Factors Affecting DNA Barcoding Classification Accuracy for Piscine Prey: An Experimental Assessment with Invasive Lionfish

Kristen A. Dahl[1,5*], Andrew Fields[2], Alison Robertson[3,4], David S. Portnoy[2], Alex Grieme[2,6], Grant Lockridge[4], and William F. Patterson III[1]

[1] *Fisheries and Aquatic Sciences Program, School of Forest, Fisheries, & Geomatics Sciences, Institute of Food and Agriculture Sciences, University of Florida, 7922 NW 71st Street, Gainesville, FL 32653, USA*

[2]*Department of Life Sciences, Texas A&M University Corpus Christi, 6300 Ocean Drive, Corpus Christi, TX 78412, USA*

[3] *School of Marine and Environmental Sciences, University of South Alabama, 600 Clinic Drive, Mobile, AL 36688, USA*

[4] *Dauphin Island Sea Lab, 101 Bienville Boulevard, Dauphin Island, AL 36528, USA*

[5]*Current Affiliation: Pacific Islands Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 1845 Wasp Blvd., Honolulu, HI 96818, USA*

[6]*Current Affiliation: Edward Via College of Osteopathic Medicine-Louisiana Campus, 4408 Bon Aire Drive, Monroe, LA 71203, USA*

[*]Corresponding author: Email: kristen.dahl@noaa.gov | ORCID: 0000-0003-4939-1084

Keywords: qPCR, molecular diet analysis, COI, self-DNA, universal primers

CRediT author statement: Kristen Dahl: Conceptualization, Methodology, Formal analysis, Visualization, Writing – Original draft preparation; Andrew Fields: Methodology, Investigation, Formal analysis, Writing – Reviewing and editing; Alison Robertson: Conceptualization, Funding acquisition, Methodology, Investigation, Resources, Supervision, Writing – Reviewing and editing; David Portnoy: Conceptualization, Methodology, Funding acquisition, Resources, Supervision, Writing – Reviewing and editing; Alex Grieme: Methodology, Investigation; Grant Lockridge: Methodology, Investigation; William Patterson: Conceptualization, Methodology, Funding acquisition, Writing – Reviewing and editing.

Abstract

The use of DNA barcoding in fish diet studies is becoming more widespread, but the effect of prey digestion on barcoding accuracy has been poorly studied. We conducted a series of controlled feeding experiments, with red lionfish (*Pterois volitans*) as predators and Gulf killifish (*Fundulus grandi*s) as prey, designed to test the effect of digestion time (12 – 42 h), prey item (sample) preservation (inside versus outside stomach), and sample preparation (washed versus unwashed) on DNA barcoding classification accuracy. Competitive interactions occurred between lionfish and killifish DNA during PCR amplification when universal COI primers were utilized; thus, primers were designed that matched the same priming site in both killifish and lionfish DNA. The proportion of amplified killifish DNA relative to lionfish DNA per sample was estimated using multiplex quantitative PCR (qPCR) and shown to decrease with longer digestion times, especially when samples were retained within lionfish stomachs for preservation. The effects of digestion and preservation method on amplifiable DNA resulted in significant effects on barcoding accuracy, resulting in erroneous identification of the predator (i.e., self-DNA) for nearly 25% of samples. Removing prey from predator stomachs prior to preservation significantly increased the proportion of amplified killifish DNA from samples, and enhanced barcoding accuracy. Overall, study results have important implications for the probability of false negatives for prey items, as well as false positives for self-DNA, when utilizing barcoding to characterize piscine diet. Results also highlight the need to conduct more method validation research given the increasing prevalence of DNA barcoding as a means to identify specific prey items in fish diet studies.

1. Introduction

Empirical diet data are widely used to characterize and quantify trophic dynamics of predator-prey interactions and are key to understanding the ecology of organisms and the structure of food webs (Paine, 1980; Sheppard and Harwood, 2005). In many aquatic and marine ecosystems, the inability to directly observe foraging behavior and trophic interactions complicates quantitative characterization of food web structure and function (Sousa et al., 2019). Visual identification of stomach contents (i.e., gut content analysis) has traditionally been used to examine trophic relationships in fishes (Hyslop, 1980), but suffers well-recognized biases due to the effects of digestion, whereby digested prey items cannot be identified visually to sufficient taxonomic resolution (Dahl et al., 2017; Haywood, 1995; Sheppard and Harwood, 2005; Ward et al., 2005). The ability to identify prey items visually often depends on factors such as prey morphology, predator digestion rate, and method of preservation of stomach contents (Carreon-Martinez et al., 2011; Legler et al., 2010). Therefore, accurate non-visual methods are often needed to obtain high-resolution diet data that are needed for understanding complex trophic interactions in marine ecosystems.

Fortunately, a range of techniques have been developed to identify organisms genetically when visual identification is problematic or impossible (Hebert et al., 2003; Symondson, 2002). In vertebrates, the most widely used approach is genetic barcoding, which utilizes species-specific markers such as 16sRNA, 18sRNA, or ITS. However, the most common approach for diet studies utilizes a ~655 base pair (bp) region of the mitochondrially-encoded cytochrome *c* oxidase subunit I (COI) gene to conduct DNA barcoding (Hebert et al., 2005; Ivanova et al., 2007). COI is amplified and sequenced using universal primers, and the species (or lowest possible taxon) is identified by direct comparison to voucher sequences in reference libraries,

such as the National Center for Biotechnology Information (NCBI) or Barcode of Life (BOLD) databases (Deagle et al., 2006; Frézal and Leblois, 2008; Kroon et al., 2020; Ward et al., 2009). DNA barcoding has proven to be an effective tool for investigating biodiversity, food safety, and illegal wildlife trade (Dawnay et al., 2007; Frézal and Leblois, 2008; Hebert et al., 2005, 2003; Nicolè et al., 2012; Valentini et al., 2009), and increasingly is used to characterize predator-prey interactions in terrestrial and aquatic systems by barcoding specific prey items from stomach contents (Sousa et al., 2019; Symondson and Harwood, 2014).

While DNA barcoding can be utilized to identify individual prey items collected from stomach contents at high taxonomic resolution, DNA degradation, polymerase chain reaction (PCR) sensitivity, and the use of universal primers have the potential to negatively impact its efficacy. PCR allows DNA amplification from extraordinarily small amounts of template, but when conducted with universal primers, which are meant to generally work for broad ranges of taxa (e.g., all fishes; Ivanova et al., 2007; Ward et al., 2005), may impede accurate identification of prey items (Jo et al., 2014; Sheppard and Harwood, 2005). Given that prey items are inherently mixed with the DNA of other consumed organisms inside the stomach, as well as the consumer's own DNA, small amounts of competing or contaminating DNA may be co-amplified and hinder sequencing and classification of the partially digested prey items of interest (Vestheim and Jarman, 2008). PCR amplification of a prey species' DNA is not only a function of the abundance of its DNA versus other DNA present in a given sample, but also the quality of DNA present (Gonzalez et al., 2012; Knebelsberger and Stöger, 2012). The process of digestion actively degrades DNA over time, and successful PCR amplification of full-length (i.e., ~ 650 bp) COI barcode sequences becomes more challenging for moderately to highly digested prey (Carreon-Martinez et al., 2011; Hajibabaei et al., 2006; Meusnier et al., 2008). In addition to

the degradative effects of digestion, sample handling, preservation, and processing methods, which vary widely among studies, also have the potential to affect DNA integrity (King et al., 2008; Stein et al., 2013) and subsequent results.

A potential issue arises when a prey sample is identified via DNA barcoding as the same species as the consumer (i.e., self-DNA with respect to the consumer). In such situations, it is difficult with COI alone to distinguish cannibalism from contamination with the consumer's DNA (Jo et al., 2014; O'Rorke et al., 2012; Sheppard and Harwood, 2005). Authors of trophic studies employing DNA barcoding have regularly reported the amplification of self-DNA among prey items for a range of taxa (Aguilar et al., 2017; Arroyave and Stiassny, 2014; Bartley et al., 2015; Braid et al., 2012; Moran et al., 2015; Valdez-Moreno et al., 2012), but few reported approaches to distinguish contamination from cannibalism. One exception occurred following a recent study that employed DNA barcoding to identify partially digested fish prey of invasive red lionfish (Dahl et al., 2017), for which microsatellite genotyping was used to distinguish cannibalism from contamination with the consumer's DNA in cases when self-DNA was identified (Dahl et al. 2018). In that study, 42% of self-DNA occurrences were confirmed as cannibalism, while the others could have been due to contamination by predator tissue during sample dissection and processing, or resulted from advanced degradation of prey DNA, leaving only lionfish consumer DNA present to be amplified.

The effect of prey digestion on identification success has rarely been investigated in fish barcoding studies (but see Carreon-Martinez et al., 2011 and Moran et al. 2015). Furthermore, the effect of different sample processing and preservation methods (e.g., storing prey inside versus outside of the consumer's stomach) on DNA barcoding accuracy remains uncharacterized. In this study, a series of controlled digestion experiments were conducted, with red lionfish

(*Pterois volitans;* hereafter lionfish) as predatory consumers and Gulf killifish (*Fundulus grandi*s; hereafter killifish) as prey, which were designed to test the effect of digestion time, preservation method (inside versus outside lionfish stomachs), and sample preparation method (washed versus unwashed) on classification accuracy of DNA barcoding. The concentration of amplified killifish and lionfish DNA for each sample was measured with multiplex quantitative PCR (qPCR), and the effect of digestion time, sample preservation, and sample preparation on the proportion of DNA belonging to killifish prey (relative to lionfish predators) amplified in samples was tested. Results are interpreted in the context of best practices for field collection, sample storage, and DNA extraction methods to optimize the quantity and quality of prey DNA available for amplification, and thus the accuracy of downstream molecular analyses used to identify prey.

2. Methods

2.1. Fish Collection and Acclimation

Lionfish were collected using SCUBA gear in summer 2017 and spring 2018 from Perdido Pass, Alabama on bridge rubble (30°02' N, 87°33' W) and an artificial reef (30°02' N, 87°39' W) at ~29 m depth. Collections were conducted by two paired-diver teams, with each team diving only once on each sampling date to minimize transit time for captured fish. Divers used hand nets with monofilament mesh to capture fish, then grasped lionfish by the operculum, inverted the net, and placed them inside a custom holding container constructed from a 35-L plastic basket with a zippered neoprene cover. This allowed quick access to the basket while minimizing the danger of venomous lionfish spines to divers, stress on lionfish, and escape of previously captured individuals. Dive teams collected 3-5 fish per dive (up to 15 per trip) and

then slowly ascended to the surface with the plastic basket containing lionfish. Once on the surface, fish were immediately vented with an 18-gauge hypodermic needle to relieve the effects of barotrauma (Collins et al. 1999). After venting, lionfish were placed in aerated holding tanks containing water with salinity and temperature that matched the seawater at collection. Fish were transported to the Dauphin Island Sea Lab wetlab facility for stabilization and acclimation. Fish were collected, handled, and ultimately euthanized according to the Dauphin Island Sea Lab IACUC protocol 1640293.

Lionfish were stabilized in a 3800-L recirculating community tank following live transport. Aeration maintained O_2 saturation at or near 100% for the duration of the acclimation and experimental periods. Water quality was tested regularly for ammonia and nitrate using test kits (LaMotte, Chestertown, USA). If ammonia was detected, a 10% water change was performed. Lionfish were monitored daily until normal swimming behavior was established (~1 week) and then were acclimated in the community tank for an additional 3-4 weeks prior to experimentation. During this period, lionfish were maintained on a diet of live brown shrimp, *Farfantepenaeus aztecus*, until 2 weeks prior to feeding experiments. Across the course of study, 80% of lionfish survived collection and the acclimation period.

Killifish were collected from baited traps in marshes around Dauphin Island, Alabama (30°15' N, 88°07' W) and maintained in an aerated bucket during transport to the wetlab facility located 5 km away. Killifish were sorted by size (target: 60-90 mm TL) and then placed in a community tank (200 L) maintained at a salinity of 10 ppt at 24-26 °C with daily assessment of water quality. Killifish were acclimated for a period of 10 days prior to feeding experiments. Fish were not fed for 5 days to adapt to the new environment and were thereafter fed a

commercial fish flake (Tetramin® Tropical Flake, Tetra, Morris Plains, NJ), every 2-3 days to satiation.

2.2. Digestion Experiments

Lionfish were transferred to individual 70-L aquaria for digestion experiments. Aquaria contained sand, PVC tubes, and bricks for environmental enrichment. Food was withheld to allow gut clearance and acclimation to new housing. Due to a reliance on wild-caught *P. volitans,* size could not be controlled, thus, small (>80 g), medium (80-120 g), and large (>120 g) individuals were distributed evenly among treatments. Salinity and water temperature (°C) were recorded during digestion experiments, which were conducted in September 2017 and July 2018.

Digestion experiments were conducted at 25 °C. The target sample size was 6 lionfish randomly selected for each digestion period, with periods being 12, 24, 30, 36, and 42 h. At the start of digestion trials, lionfish were fed two Gulf killifish. Killifish were visually confirmed as consumed within 5 min of being placed in the tanks. Any killifish not consumed during this period were removed from tanks and euthanized. At designated time points after killifish ingestion, individual lionfish were euthanized by rapidly severing the spinal cord. Each fish was blotted dry, weighed (nearest g) and measured (total length, mm) before dissection. Lionfish sex was determined visually by inspecting the gonads. Stomachs were dissected from euthanized lionfish. One killifish was aseptically removed from each stomach and preserved in a plastic vial containing 95% EtOH (i.e., outside stomach), while the other killifish was left inside the stomach and the entire stomach preserved inside in a plastic vial containing 95% EtOH. All samples were stored in a freezer (≤20 °C), then shipped on dry ice to the Marine Genomic Laboratory at Texas A&M University – Corpus Christi (Corpus Christi, Texas) for all genetic analyses.

2.3. Gut Content DNA Extraction and Amplification

Qualitative digestion scores were assigned for each sample following Aguilar et al. (2017) as: 1) lightly digested = easily identified with scales/skin mostly intact; 2) moderately digested = retaining most morphological characteristics, such as fins, skull, and possibly viscera, skin or scales absent or barely present; 3) severely digested = head, fins, skin, or scales absent with pieces of tissue encasing spinal column or loose tissue; and, 4) nearly fully digested.

Two small (~1 mm^3, 15-25 mg) plugs of muscle tissue were excised from each sample using sterile razor blades. Washed samples were excised from the interior of the killifish and rinsed with DI water, while unwashed samples were excised from the exterior surface of the killifish and were not rinsed prior to DNA extraction. DNA was extracted from each sample with DNeasy Blood & Tissue extraction kits (Qiagen, Valencia, California). DNA standards for qPCR optimization were extracted from muscle tissue collected from lionfish and killifish which had not been fed to lionfish using a Mag-Bind Blood and Tissue DNA HDQ 96 kit (Omega Bio-tek, Norcross, Georgia). All extracts were stored at 4 °C until qPCR and DNA barcoding analyses were performed.

2.4. Quantitative PCR Design and Assay

Quantitative PCR (qPCR) was employed to estimate the concentration of COI DNA amplified from each sample (Heid et al., 1996). Otherwise similar to conventional PCR, multiplex qPCR also includes probes, which emit fluorescence as the reaction proceeds. The amount of fluorescence emitted during reactions allows estimation of the concentration of the templates of interest, where the intensity of fluorescence above background level (threshold

cycle number, or Ct) is directly correlated with initial template quantity. Multiplexing was used to simultaneously quantify killifish and lionfish DNA present in each partially digested sample. This approach required species-specific probes with different fluorescent labels associated with each target species. COI DNA sequences from lionfish (n = 11) and killifish (n = 5) were downloaded from GenBank, aligned, and a consensus sequence for each species assembled using Clustal Omega (version 1.2.4). Primers and probes were then designed using Primer3 (Version 0.4.0; http://bioinfo.ut.ee/primer3-0.4.0/), and OligoAnalyzer (idtdna.com; Table 1).

The amount of amplifiable DNA in each sample was estimated with an AccuBlue High Sensitivity dsDNA Quantification Assay (Biotium, Fremont, California) or a Qubit dsDNA HS Assay (Thermo Scientific, Waltham, Massachusetts). When DNA concentration was greater than 10 ng/µL, sample extracts were diluted to a concentration between 1-10 ng/µL. The optimized qPCR assay was performed in triplicate for each sample and standard using 2x Master Mix (IDT, San Diego, California) at a 1x concentration, 0.25 µM of each primer, 0.2 µM of each probe, and 1µL of template DNA. PCR had an initial denaturing step of 95°C for 2 min, then 45 cycles of 95°C for 15 sec, then 66°C for 1 min, followed by an evaluation of fluorescence using a StepOnePlus qPCR machine (Applied Biosystems, Waltham, Massachusetts). DNA standards were used in a 10-fold dilution series from 100 ng/µl to 10^{-3} ng/µl. Within every qPCR plate, five standards of known concentration from both target species were included in triplicate to construct standard calibration curves of C_t (threshold number of cycles to produce constant fluorescence signal) versus log standard DNA concentration. In addition to standards, three plate wells contained only water as negative controls. DNA quantification for samples was interpolated from the resulting linear calibration curves.

2.5. DNA Barcoding

Universal fish primers developed by Ward et al. (2005) were initially used for DNA barcoding of prey samples, but ultimately were found to preferentially amplify lionfish DNA over killifish DNA (Table 2A). This preferential amplification was due to apparent base-pair differences in the primer sites between species. Therefore, new primers were designed to amplify a portion of the DNA barcoding region for both species. The lionfish mitochondrial genome (GenBank accession no. **KJ739816.1**) and killifish mitochondrial genome (GenBank accession no. **FJ445396.1**) were downloaded from GenBank, aligned using Clustal Omega (Sievers et al., 2011), and then visually checked. The COI barcoding region was identified and primers were designed to amplify a 364-bp segment of DNA which was sufficiently different (45 single nucleotide polymorphisms) to allow discrimination between killifish and lionfish (Table 2B). Both primers contained a single ambiguity, allowing primers to exactly match both species at the same priming site (Table 2B).

The diluted DNA used in the qPCR assay was used as template and the COI barcoding region of each sample was amplified via a 25-µL reaction using 1x Buffer, 1.5 mM $MgCl_2$, 200 mM dNTPs, 0.1 µM of each primer, 0.5 units of GoTaq (Promega, Madison, Wisconsin), and 1µL of template DNA. PCR was performed with an initial denaturation at 95°C for 2 min followed by 40 cycles of 95°C for 1 min, 55°C for 1 min, and 72°C for 1 min, and then a final extension at 72°C for 7 min. After checking for successful PCR amplification on 1% agarose gels using Gel Red (Biotium, Fremont, California), PCR products were cleaned with a 1x concentration of Mag-Bind TotalPure NGS (Omega Bio-Tek, Norcross, Georgia). Samples were Sanger-sequenced on an ABI 3730xl DNA Analyzer at Retrogen, Inc. (San Diego, California). The COI barcode sequences were analyzed to assess length and quality, and data were quality

trimmed in ApE (2.0; available at https://jorgensen.biology.utah.edu/wayned/ape) to remove ambiguous or low-quality bases and remnant primer from amplification or sequencing reactions. Sequences were queried with BLAST against the NCBI GenBank nucleotide sequence database (https://www.ncbi.nlm.nih.gov/genbank/) to identify sequences with closest match. BLAST searches returned the closest hundred sequences in the reference databases based on values of sequence similarity (e.g., % identity), and E-value, from which the closest species match was identified. If a sample did not amplify during PCR, amplification was tried with undiluted DNA samples up to two times before determining the sample had too little DNA to amplify due to advanced digestion.

2.6. Data Analyses

All statistical analyses were conducted in R (version 3.6.0, R Core Team 2016; http://www.R-project.org/). One-way analysis of variance (ANOVA) models were computed to test whether lionfish TL, tank temperature, or tank salinity differed among digestion treatments. The correlation between qualitative digestion scores and the percentage of prey mass remaining was tested with Pearson's *r* for a subset of samples for which killifish prey mass was recorded. Digestion score data violated normality and equal variance assumptions, thus a Kruskal–Wallis test was computed to test whether digestion score had a significant effect on percent killifish DNA (out of total DNA) amplified. Mean concentrations of killifish and lionfish DNA estimated with qPCR were calculated for triplicate runs for each sample, and then standardized by proportion of total DNA amplified per sample {e.g., proportion killifish DNA = [killifish DNA concentration/(killifish DNA concentration + lionfish DNA concentration)]}.

Models were fitted to the data to predict factors that affected: 1) the proportion of killifish DNA amplified from samples as was estimated via qPCR, and 2) the probability of samples being successfully identified as killifish via barcode sequencing. Factors tested for all models were digestion time (12-42 h), preservation method (inside stomach, outside stomach), sample preparation (washed, unwashed), temperature, and lionfish TL. Initial models included interaction terms but were sequentially dropped due to non-significance. Beta regression models were computed to predict which factors affected the proportion of killifish DNA amplified in samples using the betareg package (Cribari-Neto and Zeileis, 2010; Ferrari and Cribari-Neto, 2004; Douma and Weedon, 2019). Logistic regression models were computed to predict which factors affected successful species identification via barcode sequencing with the MASS package (Venables and Ripley, 2002). For samples that amplified successfully with sufficient product for sequencing, a success (1) represented samples identified as killifish, whereas samples identified as lionfish were considered to be a failure (0). Samples that failed to provide a readable sequence were also considered a failure (0). In both beta and logistic regression analyses, candidate models were fitted with a maximum likelihood estimator (Ferrari and Cribari-Neto, 2004), and model building was performed using a backward stepwise technique. The most parsimonious models were identified by comparing Akaike Information Criterion (AIC; Akaike, 1974) values among candidate models (Burnham and Anderson, 1998). Pseudo-R^2 values were calculated to assess the extent to which the most parsimonious models explained deviance in the data. Variance inflation factors (VIFs) were calculated to test for multicollinearity among independent variables.

3. Results

3.1 Feeding Experiment

Lionfish (n = 42) collected for feeding experiments ranged in size from 166 to 370 mm TL, with a mean ± SE of 236.7 ± 8.4 mm TL. Lionfish ranged in mass from 45 to 656 g, with a mean ± SE of 190.9 ± 21.9 g. Inspection of gonads at dissection revealed that feeding experiments were conducted on 18 female and 22 male lionfish, with sex unable to be determined for 2 individuals. Feeding experiments conducted in September 2017 had mean ± SE salinity of 27.6 ± 0.2, and mean ± SE seawater temperature of 25.2 ± 0.1 °C. Feeding experiments conducted in July 2018 had mean ± SE salinity of 32.1 ± 0.2, and mean ± SE seawater temperature of 25.5 ± 0.1 °C across all experimental systems. Neither temperature (ANOVA, $F_{4,37} = 1.534$, $p = 0.213$) nor lionfish TL (ANOVA, $F_{4,37} = 2.550$, $p = 0.055$) were significantly different between digestion treatments. Salinity was significantly different between time treatments (ANOVA, $F_{4,37} = 6.177$, $p < 0.001$), driven by significant pairwise differences between the 12-h (mean: 27.4) treatment and 30-h and 42-h (mean: 32.2) treatments.

There were seven individual digestion trials in which lionfish consumed only one killifish. When this occurred, the single killifish was left in the stomach for preservation in EtOH. Additionally, one lionfish consumed no killifish, therefore, it was dissected and its stomach contents analyzed as a negative control. Advanced digestion in the 42-h treatment resulted in a lack of intact prey tissue for preservation outside the stomach, thus only samples preserved inside the stomach existed for this treatment (Table 3). Ultimately, DNA analyses were conducted on 60 killifish samples, of which 37 were preserved in EtOH inside the lionfish consumer's stomach and 23 were preserved outside the stomach prior to DNA extraction (Table

3). For each consumed killifish, one muscle tissue sample was washed and a second left unwashed prior to DNA extraction, which resulted in 120 total samples (Table 3).

Killifish prey digestive scores of 1 through 4 constituted 3.3, 41.7, 36.7, and 18.3% of samples, respectively. There was a positive correlation ($p < 0.001$; $r = 0.69$) between percent mass loss of killifish samples measured during digestion and assigned digestion score (Fig. 1).

3.2. Quantitative PCR

The amount of amplified killifish DNA versus lionfish DNA was estimated via qPCR for each sample (n = 120) in triplicate, as well as for positive (n = 60) and negative (n = 18) controls. Nine samples were analyzed in a second assay to improve low-level quantitation or to reach concentrations that were above background level. We proceeded with either the quantification estimates from the best run if one run failed, or the grand mean across runs if there was no apparent difference among runs. Seven prey samples failed to amplify detectable DNA during qPCR, despite repeated attempts at amplification. The total quantity of DNA estimated by qPCR for predator and prey combined ranged between 0.0001 and 27.07 ng/µl, with a mean ± SE of 2.23 ± 0.32 ng/µl. The negative control from an empty lionfish stomach amplified 0.93 ng/µl of lionfish DNA, and failed to amplify any killifish DNA.

There was a significant difference in percent killifish DNA among digestion scores (Kruskal-Wallace $H = 24.21$, $p < 0.001$). The proportion of killifish DNA detected via qPCR was lower at higher digestion scores, but this was more pronounced for samples preserved inside lionfish stomachs (Fig. 2A). Regardless of whether killifish tissue samples were washed or unwashed, those that were removed from lionfish stomachs prior to preserving in EtOH exhibited higher proportions of prey DNA across all digestion scores (>0.62; Fig. 2A). Across

treatment combinations, the proportion of killifish DNA amplified from samples was lowest for unwashed samples that were preserved inside lionfish stomachs (Fig. 2A).

Estimated quantities of DNA measured with qPCR revealed declines in killifish DNA, relative to lionfish DNA, as digestion time increased, and in response to preservation treatment (inside versus outside lionfish stomachs; Fig. 2B). Declines in the proportion of killifish DNA were apparent after 24 h, and were most pronounced in samples that were preserved in EtOH inside lionfish stomachs in the 36-h and 42-h treatments, for which killifish DNA proportion dropped below 0.40 for washed as well as unwashed samples (Fig. 2B). Across all treatment combinations, killifish DNA proportion was highest for unwashed samples preserved outside lionfish stomachs that had digested only for 12 h (mean ± SE: 0.81 ± 0.08) or 24 h (mean ± SE: 0.81 ± 0.05; Fig. 2B).

The final beta regression model, testing factors that affected the proportion of killifish DNA amplified from samples, included only digestion time and preservation method as independent factors, with the model explaining a relatively small fraction of the total deviance in the data (pseudo $R^2 = 0.24$; Table 4A). Digestion time had a negative effect and preserving killifish tissues outside lionfish stomachs had a positive effect on the proportion of killifish DNA that amplified in samples (Table 5A). For the model predicting the proportion of killifish DNA amplified from samples stored inside lionfish stomachs, an average decrease in the proportion of killifish DNA of 0.012 per h could be expected up to 42 h after ingestion. For samples preserved outside lionfish stomachs, the model predicted slightly lower average decreases in killifish DNA per hour of digestion time (0.010 per h). Though washing samples prior to DNA amplification often led to higher proportions of killifish DNA, sample preparation (i.e., washed versus unwashed) was not significant in the model ($p = 0.636$). Similarly, temperature ($p = 0.079$), and

lionfish TL (p = 0.547) did not significantly impact the proportion of killifish DNA detected among samples.

3.3. DNA barcoding classification success

In total, 57 samples and 1 control (lionfish tissue) initially sequenced with the universal barcoding primers (Ward et al. 2005) were identified by BLAST as lionfish, indicating a potential primer mismatch or amplification bias. An alignment revealed four base-pair differences between the killifish genome and the Ward et al. (2005) FishF1 primer, with three of these differences present in the 3' region of the primer. Contrary to this, the published lionfish genome sequence appears to have only two bases different, with only one difference located in the 3' region of the primer. Subsequently, designed primers that specifically targeted COI fragments for this study (Table 2B) yielded a PCR amplification success of 88.3% (n = 106) for all samples. The remaining samples (n = 14) could not be amplified or identified via DNA barcoding, which was attributable to either DNA degradation or low DNA concentration. Of the 106 samples that were sequenced, 82 (77.4%) were identified correctly by BLAST as killifish, 23 (21.6%) as red lionfish, and one unidentified due to poor sequence quality (Fig. 3). An additional sample which was run as a negative control (empty predator gut) was identified as lionfish.

In all but one case, killifish prey samples erroneously identified as lionfish had been preserved inside a lionfish stomach (Fig. 3). Lionfish were identified via BLAST more frequently as digestion time increased, indicating the negative impact of advanced digestion (Fig. 3B). Erroneous lionfish identifications peaked in the 36 (72.2%) and 42 h (75.0%) treatments (Fig. 3B). Sample washing prior to DNA extraction had little effect on DNA barcoding accuracy,

as nearly equal numbers of prey samples identified as lionfish were washed (n = 11) and unwashed (n = 12). Samples with a digestion score of 1 (n = 4) and 2 (n = 42) were correctly identified as killifish via DNA barcoding in >98% of cases, regardless of preservation method. Samples assigned a digestion score of 3 (n = 39), were correctly identified as killifish via DNA barcoding 69.2% the time, and samples assigned a digestion score of 4 (n = 21) were correctly identified as killifish via DNA barcoding only 47.6% of the time.

The most parsimonious logistic regression model, testing factors that affected success of prey identification, included digestion time and preservation method as independent factors, and explained a moderate proportion of deviance in the data (pseudo $R^2 = 0.40$; Table 4B). Digestion time had a negative effect on success of identification, with samples preserved outside lionfish stomachs having a higher likelihood of being correctly identified as killifish (Table 5B). For samples preserved outside lionfish stomachs, the model predicted >80% killifish identification success across all experimental time treatments (Fig. 4A). For samples preserved inside lionfish stomachs, the model predicted >80% classification success between 1 and 24 h, which then dropped to <50% success between 30 and 36 h post-ingestion, and was <10% at 42 h (Fig. 4B). Washing killifish tissue samples prior to DNA extraction had little effect on barcoding success, and was not significant in the model ($p = 0.262$). Neither temperature ($p = 0.491$) nor lionfish TL ($p = 0.581$) had a significant effect on classification success of prey samples.

The logistic regression, predicting DNA barcoding success based on percent killifish DNA amplified in qPCR samples, was statistically significant ($p < 0.001$, $R^2 = 0.80$). All samples for which amplified DNA was >51% killifish were correctly identified as killifish via barcoding (Fig. 5). The model fit to the experimental data also predicted a 50% probability of correctly identifying prey via barcoding when killifish DNA constituted ≥42% of the extracted DNA and a

99% correctly identification when killifish DNA constituted ≥62% of the extracted DNA in the multiplex (Fig. 5).

4. Discussion

Results of this study demonstrate that the degree of digestion has a significant effect on the percent of prey DNA amplified from a sample. However, DNA barcoding was largely successful in correctly identifying killifish, regardless of the degree of digestion, provided that samples had been removed from lionfish stomachs prior to preservation in EtOH. For samples that were left inside lionfish stomachs prior to being preserved in EtOH, the probability of erroneously identifying the samples as lionfish (i.e., amplifying and sequencing self-DNA) increased substantially with digestion time, related to the fact that prey DNA quality and concentration diminished as digestion progressed.

4.1. Effect of Digestion and Preservation Method

It was anticipated that the classification accuracy of barcode sequencing would be dictated by how much killifish DNA was present in a sample relative to lionfish DNA. Relative DNA concentrations measured with qPCR suggest that prey items experienced the highest levels of DNA degradation when they were digested in the living predator's stomach longer than 24 h, and when they were stored within the predator stomach, alongside chyme and digestive enzymes. At levels of advanced digestion, DNA of the prey species was degraded such that a template of sufficient length and/or quality was lacking, thus the COI barcode could not be properly amplified. In those samples, the consumers' DNA, presumably of higher quality, was instead preferentially amplified and the samples were subsequently identified as the predator.

Alternatively, samples in which more than half of the total amplified DNA came from the prey (relative to predator DNA) were correctly identified. Models fit using experimental data confirmed a relationship between relative DNA quantity (prey:predator) and DNA barcoding accuracy, and indicate that when prey DNA was greater than 60% of the amplified DNA in a given sample, there was nearly 100% probability of correctly identifying the prey.

Sample handling and preservation practices are not standardized among studies utilizing DNA barcoding to identify the partially digested prey items in fishes. For instance, prey samples may be extracted subsequent to the whole preservation of predators (Valdez-Moreno et al., 2012), preserved inside the predator gut alongside other consumed prey, chyme, and digestive enzymes (Carreon-Martinez et al., 2011; Dahl et al., 2017; Oyafuso et al., 2016), or rinsed and preserved separately from the predator stomach upon dissection (Aguilar et al., 2017). Experimental results presented here indicate that suboptimal sample preservation can lead to erroneous identifications, and strongly suggest that sampling protocols are an important consideration when DNA barcoding will be used to identify specific items of piscine predators. Removing prey from predator stomachs when samples are initially processed can increase the ratio of amplified prey to predator DNA and minimize the chance of a false-positive for self-DNA when using barcode sequencing to identify prey samples. This modification is supported by accurate DNA barcoding of killifish prey in the current study across all digestion times (i.e., up to 42 h post-ingestion) when samples were preserved and stored outside lionfish stomachs. A similar effect is likely if multiple prey species in different stages of digestion are present in a predator's gut, and the removal and individual storage of prey items should reduce the possibility of cross-contamination (Shokralla et al., 2010), though this was not tested by Shokralla et al. (2010) or in this study. Though storage time in preservative (EtOH) was not tested as a factor in

here, extended periods of time (> 6 months) can also negatively impact the quantity of DNA amplified during PCR (Bisanti et al., 2009; Stein et al., 2013), especially when samples are not regularly curated. Thus, best practices should involve removing prey items and preserving them separately, with DNA extractions performed as soon as possible after preservation to avoid introducing spurious results into DNA barcoding. All of these suggested modifications should be easy to incorporate into existing sampling protocols.

4.2. False Positives for Self-DNA

The results of controlled experiments revealed a concerning degree of false-positive or self-DNA identifications made when using barcode sequencing for sample identification. While authors of diet studies applying DNA barcoding frequently report the amplification of self-DNA among prey items, there is no consensus on how to handle such results (Jo et al., 2014; Sheppard and Harwood, 2005). By itself, DNA barcoding often cannot discriminate between COI sequences resulting from true cannibalism versus false positives, yet self-DNA identifications are often attributed to contamination (Bartley et al., 2015; Dahl et al., 2017; Moran et al., 2015), and inferred to be true instances of cannibalism in other situations (Arroyave and Stiassny, 2014; Braid et al., 2012; Côté et al., 2013; Oyafuso et al., 2016; Valdez-Moreno et al., 2012). Thus, it is important to recognize the impacts of different sample processing and preservation practices used among studies upon results, especially those that report self-DNA. Specifically, studies in which consumer stomachs are preserved with prey items still inside, especially for extended periods, may have erroneous identification of cannibalism caused by amplifying predator DNA, or cross contamination from other prey items.

The issue of self-DNA amplification frequently arises in the application of DNA metabarcoding, which combines universal primers and high-throughput next-generation sequencing to amplify and identify mixtures of DNA originating from multiple species in gut, fecal, or environmental samples (Pompanon et al., 2012; Ruppert et al., 2019; Shokralla et al., 2012; Sousa et al., 2019). When used to investigate trophic relationships, concomitant amplification of abundant, high-quality predator DNA regularly compromises detection of scarcer and degraded prey DNA (Deagle et al., 2006; Piñol et al., 2014). One approach to combat this potential bias is the incorporation of blocking oligonucleotides (oligos), which are designed to bind to predator DNA and prevent its amplification during PCR (Vestheim et al., 2011). However, careful consideration of the likelihood of cannibalism occuring must be given when deciding whether to incorporate blocking oligos into PCR. Additionally, sufficient genetic divergence must exist between predator and prey because blocking oligos may prevent amplification of closely related fish prey from piscine predators, which could confound taxonomic inference from barcoding results (Krehenwinkel et al., 2017). Clearly, guidance is needed on practices that may appropriately account for cannibalism in taxa, such as fishes, where this is a common occurrence (Pereira et al., 2017; Polis, 1981), while reducing the potential for amplifying self-DNA.

4.3. Primer Considerations

The relationships observed between the relative concentration of amplifiable prey DNA and barcoding accuracy herein were made under controlled settings using study-specific primers, and, therefore, there were only two possible outcomes. Nevertheless, our findings have important implications for future studies utilizing barcoding approaches to identify fish prey taxa.

Optimized upstream processes (e.g., prey handling, storage, etc.) discussed above may lessen the likelihood of contamination, but other processes inherent to the use of universal primers may hinder prey identification (Cristescu, 2014). The aim in using universal primers is to amplify DNA from a broad range of potential taxa, thus they are commonly used in trophic studies that employ DNA barcode sequencing due to a lack of *a priori* knowledge of the predator's diet (Valentini et al., 2009). However, base-pair mismatches at primer binding sites can result in the preferential amplification of some species over others during PCR (De Barba et al., 2014), and in extreme cases may completely obscure the presence of certain prey species. Primer bias, though not an original focus in this study, was apparent as universal fish primers developed by Ward et al. (2005) were found to preferentially amplify lionfish DNA over killifish DNA 100% of the time, regardless of DNA proportions due to differences in complementarity between the primers and priming sites in DNA of the two species. If this situation occurred during a field study, certain prey species might not be detected, reducing the efficacy of DNA barcoding to accurately characterize diet. Ward et al. (2005) reported that five of 211 species failed to amplify using their universal primers, across varied groups, and including congeners of species which amplified successfully. Thus, primer specificity to all potential prey should be assessed when possible by aligning available sequences from global repositories (e.g., GenBank, BOLD) against the universal primers of choice (Wilcox et al., 2013) or by testing amplification of DNA isolated directly from potential prey species collected from the target ecosystem.

Due to the large size of the COI sequence (~650 bp), it has been suggested that primers should preferentially target shorter DNA fragments to increase the probability of detecting and recovering highly degraded DNA from samples preserved for extended periods (Meusnier et al., 2008) or inside the gut of predators (Deagle et al., 2006), as they will recover a more reliable and

complete prey spectrum (Leray et al., 2013). Partial, or mini, barcodes within the full-length barcode, such as the ones used in this study, have been shown to be effective in identifying specimens with degraded DNA (Leray et al., 2013; Meusnier et al., 2008; Shokralla et al., 2015), and may be a potential avenue for overcoming problems of low amplification success and high contamination by non-prey DNA (Pompanon et al., 2012). Furthermore, the use of a partial (364 bp) COI barcode for identification of killifish from feeding trials likely resulted in conservative estimates of erroneous identifications because full-length barcodes from degraded (i.e., digested) DNA would be less likely to amplify over that of more abundant high-quality lionfish DNA. Thus, for studies amplifying the full COI barcode, especially with universal primers, predator-biased amplification may be far higher than observed in this study.

4.4. Samples Benefitting from Barcoding

Few researchers have investigated the accuracy of DNA barcoding in identifying highly degraded prey items (Carreon-Martinez et al., 2011), and data presented here are the first that relate qualitative digestion scores with estimated quantities of amplified prey (and predator) DNA and barcoding accuracy. Similar relationships examined over a greater diversity of predator-prey systems may prove to be informative, given that researchers will not know how long prey items have been digesting inside the gut of field-collected specimens, or what percentage of prey tissue remains. Digestion score was a reasonable predictor of the percent killifish DNA amplified from samples, but even highly digested prey items were successfully barcoded as killifish if they had been preserved in EtOH outside of the consuming lionfish's stomach. Following 12 to 24 h digestion, all killifish prey remained visually identifiable to at least family level and this corresponded to low (e.g., 1, 2) digestion scores, high quantities of

prey DNA, and high probability of accurate prey identification via BLAST. At the highest (42 h) level of digestion, prey were the least likely to be visually identifiable and, therefore, benefitted most from molecular methods, but this was also when DNA barcoding became less reliable. Perhaps a comparative weighting strategy could be applied to stomach contents analyses that use barcode sequencing, such that a higher weight is placed on results obtained from less digested samples (higher quality DNA) as compared to more digested samples (lower quality DNA), so that predation on prey with lower confidence of accurate identification is not overestimated (King et al., 2008).

4.5. Conclusions

DNA barcode sequencing is a powerful and efficient approach to identify organisms, and is becoming an integral tool to identify prey items from partially digested stomach contents of aquatic and marine predators (Barnett et al., 2010; Braid et al., 2012; Dahl et al., 2017; Oyafuso et al., 2016; Valdez-Moreno et al., 2012). However, the results presented show that preservation methodologies and degree of digestion have large impacts on the veracity of results, and thus on inferences drawn from them. Removing prey from predator stomachs prior to preservation in EtOH can increase the ratio of prey to predator DNA and improve the accuracy of DNA barcoding. Therefore, researchers applying DNA barcoding in diet studies should remove prey items from the digestive tract of the predator as soon as possible after acquiring samples. Molecular markers, as well as the primers chosen to amplify them, should be carefully considered to reduce potential sources of error stemming from the presence of predator (self) DNA, as well as DNA from other prey taxa. Partial and mini barcodes, such as those used in this study, have a greater ability to amplify degraded DNA, and may be used in future studies

specifically for samples exhibiting a high degree of degradation in tandem with blocking oligos to negate the amplification of high-quality predator DNA (Vestheim et al., 2011). In some cases, prey can be examined visually to determine that they are not the same species as the predator, however, in species where cannibalism may be occurring, as is common in piscivorous fishes, further guidance is needed on how best to handle the amplification of self-DNA via barcoding. Additional molecular techniques such as microsatellite genotyping may be necessary to disentangle cannibalism from false positives due to contamination (Dahl et al., 2018; Uriarte et al., 2019). Primer specificity for potential arrays of prey species and the predator should be considered before choosing primers, because universal primers may lead to preferential amplification of species in samples with mixed DNA. Controlled experiments should be directed towards examining the dynamics of competitive PCR interactions within other predator-prey systems using multiple prey species, as well as the frequency of erroneous barcoding identifications using different universal primers, such as those developed by Ivanova et al. (2007), on mixtures of known DNA. Continued research will improve our understanding of potential sources of error in DNA barcoding and serve to increase the accuracy the method when applied to study the trophic ecology of aquatic predators.

Acknowledgements

Funding for fish collection and fish experimental work completed by AR and GL was partially supported by the National Oceanic and Atmospheric Administration ECOHAB program through award NA11NOS4780028 (AR; publication number 992), with graduate student effort supported by NSF PIRE (1743802), and additional funding from the Florida Fish and Wildlife Conservation Commission (FWC-16188 to WFP). We thank Clayton Bennett, Alexander

Leynse, and Hunter King from the Dauphin Island Sea Lab (DISL), for their assistance in dive collections and transport of lionfish to Dauphin Island. Scott Alford (DISL) assisted with killifish collections, and Pearce Cooper assisted with sample transport from Dauphin Island. This article is archived as a publication of the Marine Genomics Laboratory at Texas A&M University–Corpus Christi (publication number 31) and in the series Genetic Studies of Marine Fishes (publication number 124)(https://www.marinegenomicslab.tamucc.edu/current-studies).

References

Aguilar, R., Ogburn, M.B., Driskell, A.C., Weigt, L.A., Groves, M.C., Hines, A.H., 2017. Gutsy genetics: Identification of digested piscine prey items in the stomach contents of sympatric native and introduced warmwater catfishes via DNA barcoding. Environ. Biol. Fishes. 100, 325–336. https://doi.org/10.1007/s10641-016-0523-8

Akaike, H., 1974. A New Look at the Statistical Model Identification: Selected Papers of Hirotugu Akaike. Springer, New York, NY. https://doi.org/10.1109/TAC.1974.1100705

Arroyave, J., Stiassny, M.L.J., 2014. DNA barcoding reveals novel insights into pterygophagy and prey selection in distichodontid fishes (Characiformes: Distichodontidae). Ecol. Evol. 4, 4534–4542. https://doi.org/10.1002/ece3.1321

Barnett, A., Redd, K.S., Frusher, S.D., Stevens, J.D., Semmens, J.M., 2010. Non-lethal method to obtain stomach samples from a large marine predator and the use of DNA analysis to improve dietary information. J. Exp. Mar. Bio. Ecol. 393, 188–192. https://doi.org/10.1016/j.jembe.2010.07.022

Bartley, T., McCann, K.S., Bartley, T.J., Braid, H.E., McCann, K.S., Lester, N.P., Shuter, B.J., Hanner, R.H., 2015. DNA barcoding increases resolution and changes structure in Canadian boreal shield lake food webs. DNA Barcodes 3, 30–43. https://doi.org/10.1515/dna-2015-0005

Bisanti, M., Ganassi, S., Mandrioli, M., 2009. Comparative analysis of various fixative solutions on insect preservation for molecular studies. Entomol. Exp. Appl. 130, 290–296. https://doi.org/10.1111/j.1570-7458.2008.00821.x

Braid, H.E., Deeds, J., DeGrasse, S.L., Wilson, J.J., Osborne, J., Hanner, R.H., 2012. Preying on commercial fisheries and accumulating paralytic shellfish toxins: A dietary analysis of

invasive *Dosidicus gigas* (Cephalopoda Ommastrephidae) stranded in Pacific Canada. Mar. Biol. 159, 25–31. https://doi.org/10.1007/s00227-011-1786-4

Burnham, K.P., Anderson, D.R., 1998. Practical use of the information-theoretic approach, in: Model Selection and Inference. Springer, New York, pp. 75–117.

Carreon-Martinez, L., Johnson, T.B., Ludsin, S.A., Heath, D.D., 2011. Utilization of stomach content DNA to determine diet diversity in piscivorous fishes. J. Fish Biol. 78, 1170–1182. https://doi.org/10.1111/j.1095-8649.2011.02925.x

Collins, M.R., McGovern, J.C., Sedberry, G.R., Meister, H.S., and Pardieck, R. 1999. Swim bladder deflation in black sea bass and vermilion snapper: potential for increasing postrelease survival. N. Am. J. Fish. Manag. 19, 828–832. https://doi.org/10.1577/1548-8675(1999)019%3C0828:SBDIBS%3E2.0.CO;2

Côté, I.M., Green, S.J., Morris, J.A., Akins, J.L., Steinke, D., 2013. Diet richness of invasive Indo-Pacific lionfish revealed by DNA barcoding. Mar. Ecol. Prog. Ser. 472, 249–256. https://doi.org/10.3354/meps09992

Cribari-Neto, F., Zeileis, A., 2010. Beta regression in R. J. Stat. Softw. https://doi.org/10.18637/jss.v034.i02

Cristescu, M.E., 2014. From barcoding single individuals to metabarcoding biological communities: Towards an integrative approach to the study of global biodiversity. Trends Ecol. Evol. 29, 566–571. https://doi.org/10.1016/j.tree.2014.08.001

Dahl, K.A., Patterson, W.F., Robertson, A., Ortmann, A.C., 2017. DNA barcoding significantly improves resolution of invasive lionfish diet in the Northern Gulf of Mexico. Biol. Invasions 19, 1917–1933. https://doi.org/10.1007/s10530-017-1407-3

Dahl, K.A., Portnoy, D.S., Hogan, J.D., Johnson, J.E., Gold, J.R., Patterson, W.F., 2018. Genotyping confirms significant cannibalism in northern Gulf of Mexico invasive red lionfish, Pterois volitans. Biol. Invasions 20, 3513–3526. https://doi.org/10.1007/s10530-018-1791-3

Dawnay, N., Ogden, R., McEwing, R., Carvalho, G.R., Thorpe, R.S., 2007. Validation of the barcoding gene COI for use in forensic genetic species identification. Forensic Sci. Int. 173, 1–6. https://doi.org/10.1016/j.forsciint.2006.09.013

De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E., Taberlet, P., 2014. DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: Application to omnivorous diet. Mol. Ecol. Resour. 14, 306–323. https://doi.org/10.1111/1755-0998.12188

Deagle, B.E., Eveson, J.P., Jarman, S.N., 2006. Quantification of damage in DNA recovered from highly degraded samples - A case study on DNA in faeces. Front. Zool. 3, 1–10. https://doi.org/10.1186/1742-9994-3-11

Douma, J.C., Weedon, J.T., 2019. Analysing continuous proportions in ecology and evolution: A practical introduction to beta and Dirichlet regression. Methods Ecol Evol. 10 (9), 1412-1430, https://doi.org/10.1111/2041-210X.13234

Ferrari, S.L.P., Cribari-Neto, F., 2004. Beta regression for modelling rates and proportions. J. Appl. Stat. 31, 799–815. https://doi.org/10.1080/0266476042000214501

Frézal, L., Leblois, R., 2008. Four years of DNA barcoding: Current advances and prospects. Infect. Genet. Evol. 8, 727–736. https://doi.org/10.1016/j.meegid.2008.05.005

Gonzalez, J.M., Portillo, M.C., Belda-Ferre, P., Mira, A., 2012. Amplification by PCR artificially reduces the proportion of the rare biosphere in microbial communities. PLoS One 7, e29973. https://doi.org/10.1371/journal.pone.0029973

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. https://doi.org/10.1111/j.1471-8286.2006.01470.x

Haywood, M.D.E., 1995. Rates at which post-larval prawns are digested by a small predatory fish and the implications for predation studies. J. Fish Biol. 47, 337–340. https://doi.org/10.1111/j.1095-8649.1995.tb01901.x

Hebert, P.D.N., Cywinska, A., Ball, S.L., deWaard, J.R., 2003. Biological identifications through DNA barcodes. Proc. R. Soc. B Biol. Sci. 270, 313–321. https://doi.org/10.1098/rspb.2002.2218

Hebert, P.D.N., Gregory, T.R., Savolainen, V., 2005. The promise of DNA barcoding for taxonomy. Syst. Biol. 54, 852–859. https://doi.org/10.1080/10635150500354886

Heid, C.A., Stevens, J., Livak, K.J., Williams, P.M., 1996. Real time quantitative PCR. Genome Res. 10, 986–994. https://doi.org/10.1101/gr.6.10.986

Hyslop, E.J., 1980. Stomach contents analysis-A review of methods and their application. J. Fish Biol. 17, 411–429. https://doi.org/10.1111/j.1095-8649.1980.tb02775.x

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. https://doi.org/10.1111/j.1471-8286.2007.01748.x

Jo, H., Gim, J.A., Jeong, K.S., Kim, H.S., Joo, G.J., 2014. Application of DNA barcoding for identification of freshwater carnivorous fish diets: Is number of prey items dependent on

size class for Micropterus salmoides? Ecol. Evol. 4, 219–229. https://doi.org/10.1002/ece3.921

King, R.A., Read, D.S., Traugott, M., Symondson, W.O.C., 2008. Molecular analysis of predation: A review of best practice for DNA-based approaches. Mol. Ecol. 17, 947–963. https://doi.org/10.1111/j.1365-294X.2007.03613.x

Knebelsberger, T., Stöger, I., 2012. DNA extraction, preservation, and amplification, in: DNA Barcodes. Humana Press, Totowa, New Jersey, pp. 311–338.

Krehenwinkel, H., Kennedy, S., Pekár, S., Gillespie, R.G., 2017. A cost-efficient and simple protocol to enrich prey DNA from extractions of predatory arthropods for large-scale gut content analysis by Illumina sequencing. Methods Ecol. Evol. 8, 126–134. https://doi.org/10.1111/2041-210X.12647

Kroon, F.J., Lefèvre, C.D., Doyle, J.R., Patel, F., Milton, G., Severati, A., Kenway, M., Johansson, C.L., Schnebert, S., Thomas-Hall, P., Bonin, M.C., Cameron, D.S., Westcott, D.A., 2020. DNA-based identification of predators of the corallivorous Crown-of-Thorns Starfish (*Acanthaster* cf. *solaris*) from fish faeces and gut contents. Sci. Rep. 10, 1–14. https://doi.org/10.1038/s41598-020-65136-4

Legler, N.D., Johnson, T.B., Heath, D.D., Ludsin, S.A., 2010. Water temperature and prey size effects on the rate of digestion of larval and early juvenile fish. Trans. Am. Fish. Soc. 139, 868–875. https://doi.org/10.1577/T09-212.1

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. Front. Zool. 10, 1–14. https://doi.org/10.1186/1742-9994-10-34

Meusnier, I., Singer, G.A., Landry, J.-F., Hickey, D.A., Hebert, P.D., Hajibabaei, M., 2008. A universal DNA mini-barcode for biodiversity analysis. BMC Genomics 9, 1–4. https://doi.org/10.1186/1471-2164-9-214

Moran, Z., Orth, D.J., Schmitt, J.D., Hallerman, E.M., Aguilar, R., 2015. Effectiveness of DNA barcoding for identifying piscine prey items in stomach contents of piscivorous catfishes. Environ. Biol. Fishes 99, 161–167. https://doi.org/10.1007/s10641-015-0448-7

Nicolè, S., Negrisolo, E., Eccher, G., Mantovani, R., Patarnello, T., Erickson, D.L., Kress, W.J., Barcaccia, G., 2012. DNA barcoding as a reliable method for the authentication of commercial seafood products. Food Technol. Biotechnol. 50, 387–398.

O'Rorke, R., Lavery, S., Chow, S., Takeyama, H., Tsai, P., Beckley, L.E., Thompson, P.A., Waite, A.M., Jeffs, A.G., 2012. Determining the diet of larvae of western Rock Lobster (*Panulirus cygnus*) using high-throughput DNA sequencing techniques. PLoS One 7, e42757. https://doi.org/10.1371/journal.pone.0042757

Oyafuso, Z.S., Toonen, R.J., Franklin, E.C., 2016. Temporal and spatial trends in prey composition of wahoo *Acanthocybium solandri*: A diet analysis from the central North Pacific Ocean using visual and DNA bar-coding techniques. J. Fish Biol. 88, 1501–1523. https://doi.org/10.1111/jfb.12928

Paine, R.T., 1980. Food Webs: Linkage, Interaction Strength and Community Infrastructure. J. Anim. Ecol. 49, 667–685. https://doi.org/10.2307/4220

Pereira, L.S., Agostinho, A.A., Winemiller, K.O., 2017. Revisiting cannibalism in fishes. Rev. Fish Biol. Fish. 27, 499–513. https://doi.org/10.1007/s11160-017-9469-y

Piñol, J., San Andrés, V., 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. Mol. Ecol. Resour. 14, 18–26. https://doi.org/10.1111/1755-0998.12156

Polis, G.A., 1981. The evolution and dynamics of intraspecific predation. Annu. Rev. Ecol. Syst. 12, 225–251. https://doi.org/10.1146/annurev.es.12.110181.001301

Pompanon, F., Deagle, B.E., Symondson, W.O.C., Brown, D.S., Jarman, S.N., Taberlet, P., 2012. Who is eating what: Diet assessment using next generation sequencing. Mol. Ecol. 21, 1931–1950. https://doi.org/10.1111/j.1365-294X.2011.05403.x

Ruppert, K.M., Kline, R.J., Rahman, M.S., 2019. Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. Glob. Ecol. Conserv. 17, e00547. https://doi.org/10.1016/j.gecco.2019.e00547

Sheppard, S.K., Harwood, J.D., 2005. Advances in molecular ecology: Tracking trophic links through predator-prey food-webs. Funct. Ecol. 19, 751–762. https://doi.org/10.1111/j.1365-2435.2005.01041.x

Shokralla, S., Hellberg, R.S., Handy, S.M., King, I., Hajibabaei, M., 2015. A DNA mini-barcoding system for authentication of processed fish products. Sci. Rep. 5, 1–11. https://doi.org/10.1038/srep15894

Shokralla, S., Singer, G.A.C., Hajibabaei, M., 2010. Direct PCR amplification and sequencing of specimens' DNA from preservative ethanol. BioTechniques. 3, 305–306. https://doi.org/10.2144/000113362

Shokralla, S., Spall, J.L., Gibson, J.F., Hajibabaei, M., 2012. Next-generation sequencing technologies for environmental DNA research. Mol. Ecol. 21, 1794–1805. https://doi.org/10.1111/j.1365-294X.2012.05538.x

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539. https://doi.org/10.1038/msb.2011.75

Sousa, L.L., Silva, S.M., Xavier, R., 2019. DNA metabarcoding in diet studies: Unveiling ecological aspects in aquatic and terrestrial ecosystems. Environ. DNA. 1, 199–214. https://doi.org/10.1002/edn3.27

Stein, E.D., White, B.P., Mazor, R.D., Miller, P.E., Pilgrim, E.M., 2013. Evaluating ethanol-based sample preservation to facilitate use of DNA barcoding in routine freshwater biomonitoring programs using benthic macroinvertebrates. PLoS One. 8, e51273. https://doi.org/10.1371/journal.pone.0051273

Symondson, W.O.C., 2002. Molecular identification of prey in predator diets. Mol. Ecol. 11, 627–641. https://doi.org/10.1046/j.1365-294X.2002.01471.x

Symondson, W.O.C., Harwood, J.D., 2014. Special issue on molecular detection of trophic interactions: Unpicking the tangled bank. Mol. Ecol. 23, 3601–3604. https://doi.org/10.1111/mec.12831

Uriarte, A., Johnstone, C., Laiz-Carrión, R., García, A., Llopiz, J.K., Shiroza, A., Quintanilla, J.M., Lozano-Peral, D., Reglero, P., Alemany, F., 2019. Evidence of density-dependent cannibalism in the diet of wild Atlantic bluefin tuna larvae (*Thunnus thynnus*) of the

Balearic Sea (NW-Mediterranean). Fish. Res. 212, 63–71. https://doi.org/10.1016/j.fishres.2018.12.013

Valdez-Moreno, M., Quintal-Lizama, C., Gómez-Lozano, R., García-Rivas, M. del 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. https://doi.org/10.1371/journal.pone.0036636

Valentini, A., Pompanon, F., Taberlet, P., 2009. DNA barcoding for ecologists. Trends Ecol. Evol. 24, 110–117. https://doi.org/10.1016/j.tree.2008.09.011

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S, fourth ed. Springer, New York.

Vestheim, H., Deagle, B.E., Jarman, S.N., 2011. Application of blocking oligonucleotides to improve signal-to-noise ratio in a PCR. Methods Mol. Biol. 687, 265–274. https://doi.org/10.1007/978-1-60761-944-4_19

Vestheim, H., Jarman, S.N., 2008. Blocking primers to enhance PCR amplification of rare sequences in mixed samples – A case study on prey DNA in Antarctic krill stomachs. Front. Zool. 5, 1–11. https://doi.org/10.1186/1742-9994-5-12

Ward, R.D., Hanner, R., Hebert, P.D., 2009. The campaign to DNA barcode all fishes, FISH-BOL. J. Fish Biol. 74, 329–356. https://doi.org/10.1111/j.1095-8649.2008.02080.x

Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R., Hebert, P.D., 2005. DNA barcoding Australia's fish species. Philos. Trans. R. Soc. B Biol. Sci. 360, 1847–1857. https://doi.org/10.1098/rstb.2005.1716

Wilcox, T.M., McKelvey, K.S., Young, M.K., Jane, S.F., Lowe, W.H., Whiteley, A.R., Schwartz, M.K., 2013. Robust detection of rare species using environmental DNA: The

importance of primer specificity. PLoS One. 8, e59520.

https://doi.org/10.1371/journal.pone.0059520

Table 1. PCR primers and probes designed in this study to simultaneously amplify and quantify COI DNA of lionfish (*Pterois volitans*) and killifish (*Fundulus grandis*) from prey extracts using qPCR. For each primer, GC% indicates the percentage of guanine-cytosine content.

Species	Sequence type	DNA sequence	Length (bp)	GC%	Annealing temp (°C)
Lionfish	*Forward primer*	5'-GTTACAGCTCATGCTTTCGTAATAA-3'	25	36.0	53.5
	Reverse primer	5'-CGAGGAAATGCTATGTCTGGT-3'	21	47.6	54.6
	Probe	5'-ACTGGCTTATCCCGCTGATGATTGG-3'	25	52.0	61.0
Killifish	*Forward primer*	5'-TCGAGCAGAACTAAGCCAACC-3'	21	52.4	63.8
	Reverse primer	5'-ATGAAATACCAGCCAAGTGTA-3’	21	38.1	59.3
	Probe	5'-TAAACTGTTCAACCCGTCCC-3'	20	50.0	54.9

Table 2. A) Universal primers developed by Ward et al. (2005), and B) Study-specific primers designed in this study to amplify a portion of the COI barcoding region of DNA in lionfish (*Pterois volitans*) and killifish (*Fundulus grandis*). For each primer, GC% indicates the percentage of guanine-cytosine content and Tm indicates the melting point.

A)

Primer name	DNA sequence	Length (bp)	GC%	Tm (°C)
FishF1	5'-TCAACCAACCACAAAGACATTGGCAC-3'	26	46.2	58.0
FishF2	5'-TCGACTAATCATAAAGATATCGGCAC-3'	26	38.5	54.8
FishR1	5'-TAGACTTCTGGGTGGCCAAAGAATCA-3'	26	46.2	58.0
FishR2	5'-ACTTCAGGGTGACCGAAGAATCAGAA-3'	26	46.2	58.0

B)

tRNA_Tyr F	5'-TTGRTAAGAAGAGGACTTAAACC-3'	23	34.8	51.1
COI_262 R	5'-TTATACGAGGAAAWGCTATGTCTG-3’	24	37.5	52.3

Table 3. Prey sample sizes by treatment combination for qPCR and DNA barcoding analyses. Totals for digestion trials and treatments are indicated, and the grand total is in bold.

	Time digested					
Treatment	12 h	24 h	30 h	36 h	42 h	Row Total
Inside stomach, Washed	6	11	6	10	4	37
Inside stomach, Unwashed	6	11	6	10	4	37
Outside stomach, Washed	6	6	6	5	0	23
Outside stomach, Unwashed	6	6	6	5	0	23
Column Total	24	34	24	30	8	**120**

Table 4. Model selection results for: A) proportion of killifish COI DNA in prey samples estimated using qPCR, and B) DNA barcoding probability of success. The most parsimonious model of each response variable is in bold. R^2 values are McFadden's pseudo-R^2 (logistic regression) or pseudo-R^2 (beta regression).

Model (y ~ x)	Test	df	k	R^2	AIC	ΔAIC
A) *Proportion killifish DNA*						
KillifishDNA ~ Time + Preservation	**Beta regression**	**4**	**2**	**0.241**	**-40.03**	**0**
KillifishDNA ~ Time + Preservation + Washing + Temperature	Beta regression	6	4	0.254	-39.24	0.79
KillifishDNA ~ Time + Preservation + Washing	Beta regression	5	3	0.241	-38.26	0.98
KillifishDNA ~ Time + Preservation + Washing + Temperature + Lionfish TL	Beta regression	7	5	0.255	-37.63	2.4
B) *DNA barcoding classification success*						
Success ~ Time + Preservation	**Logistic regression**	**4**	**2**	**0.443**	**69.16**	**0**
Success ~ Time + Preservation + Washing	Logistic regression	5	3	0.454	69.87	0.71
Success ~ Time + Preservation + Washing + Temperature	Logistic regression	6	4	0.458	71.48	2.32
Success ~ Time + Preservation + Washing + Temperature + Lionfish TL	Logistic regression	7	5	0.461	73.17	4.01

Table 5. Results and coefficient estimates from final models for: A) proportion of killifish COI DNA and, B) DNA barcoding classification success. Coefficient estimates indicate the direction and magnitude of the effect on model response. For categorical factors, the baseline level for comparison is given in parentheses. The unit for the continuous factor time was hour.

Response	Test	Factor	df	Test statistic	p-value	Coefficient estimates
A) *Proportion killifish DNA*	Beta regression	Time	1	z = -5.27	<0.001	-0.058
		Preservation (Inside stomach)	1	z = 2.83	0.005	
		Outside stomach				0.571
B) *DNA barcoding classification success*	Logistic regression	Time	1	z = -4.20	<0.001	-0.245
		Preservation (Inside stomach)	1	z = 2.52	<0.001	
		Outside stomach				2.77

Figure 1. Correlation between percent mass loss of killifish prey measured during experimental digestion trials and visual digestion score assigned prior to DNA extraction and qPCR, tested with Pearson's *r*. Note: only a subset (n = 54) of killifish prey samples were weighed prior to and after digestion.

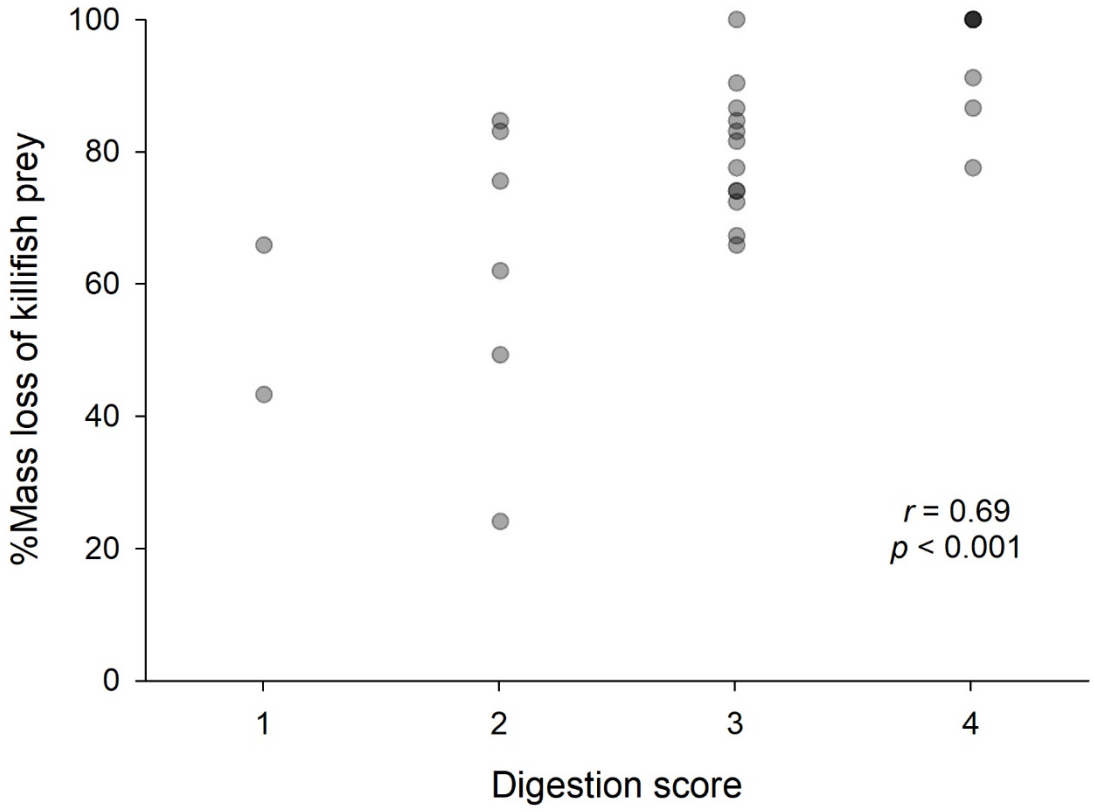
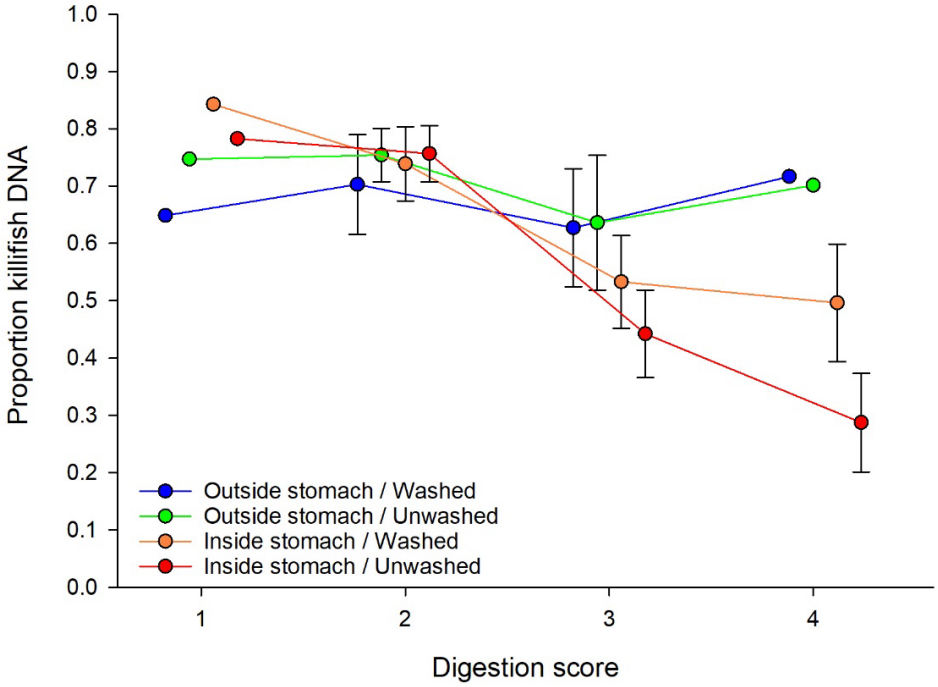


Figure 2. Mean (±SE) proportion of killifish prey DNA estimated with qPCR by: A) digestion score, and B) digestion time, across all treatments. The remaining proportion of DNA quantified (totaling 1) was that of the predator, lionfish. Note: treatment-specific means are offset along x-axes for viewing.

A



B

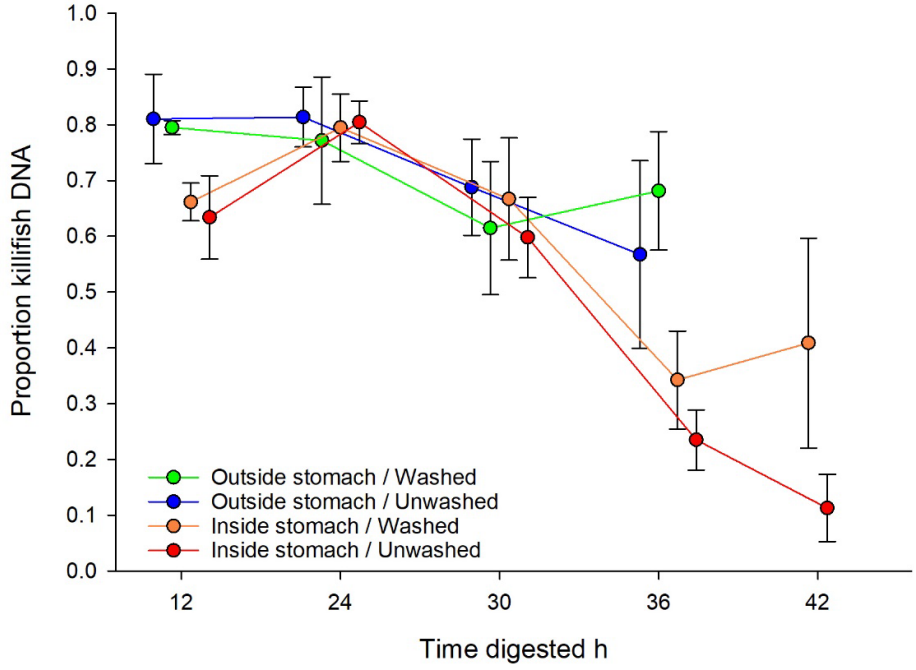


Figure 3. Percentage of samples barcoded as Gulf killifish, red lionfish, or neither for samples that: A) were removed from lionfish stomachs prior to preserving in EtOH, or B) were left inside stomachs and preserved in EtOH. Numbers above stacked bars indicate the percentage of samples successfully barcoded in each treatment.

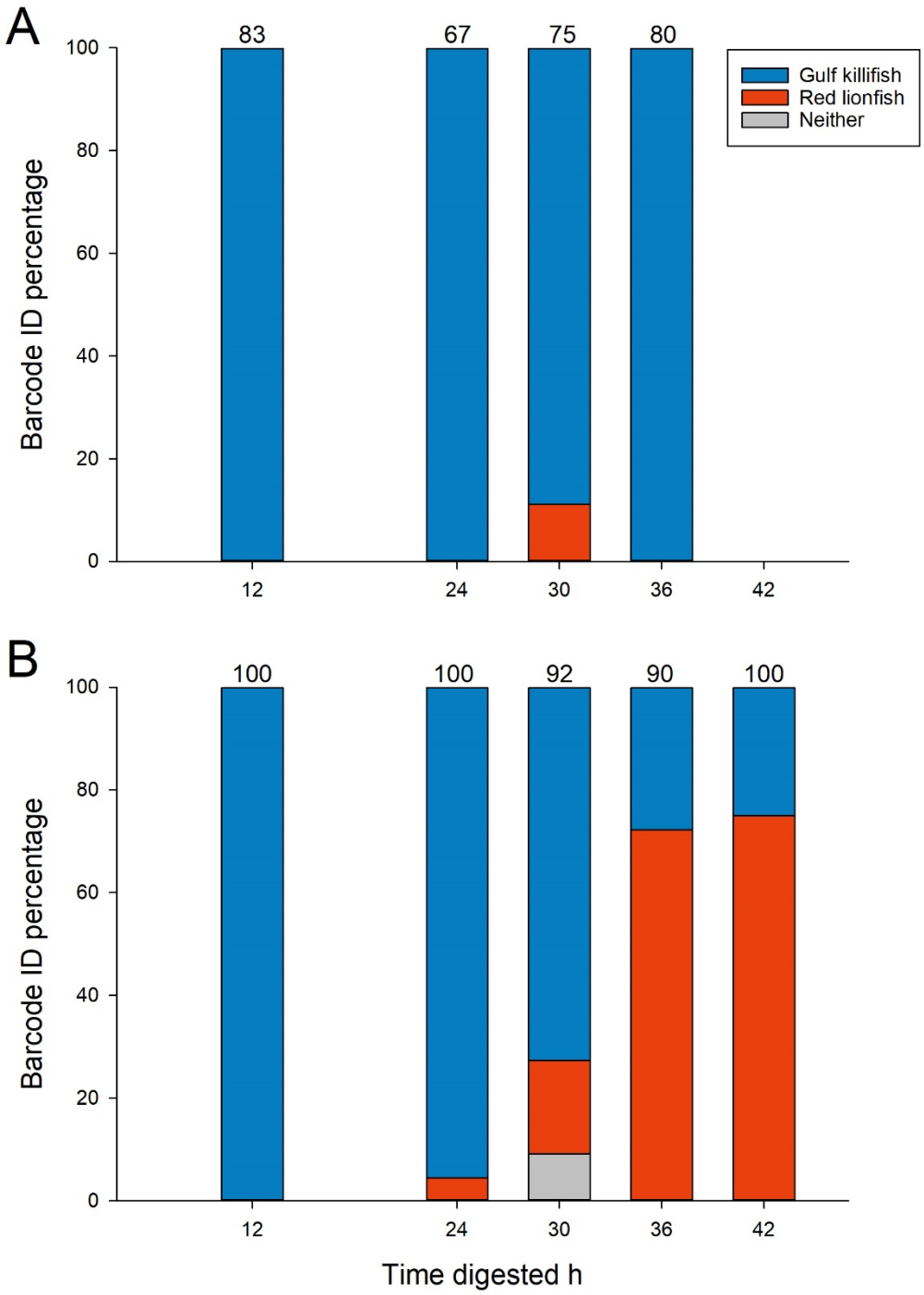


Figure 4. Relationship between probability of DNA barcoding success (i.e., 1 = correctly identifying killifish prey, 0 = erroneously identifying prey as lionfish) and time digested predicted from the logistic regression model for: A) samples that were removed from lionfish stomachs prior to preserving in EtOH, or B) samples that were left inside lionfish stomachs and preserved in EtOH. Model fits are plotted along with 95% CIs, observation symbol size scaled to sample size. Horizontal reference lines denote 50% probability of success.

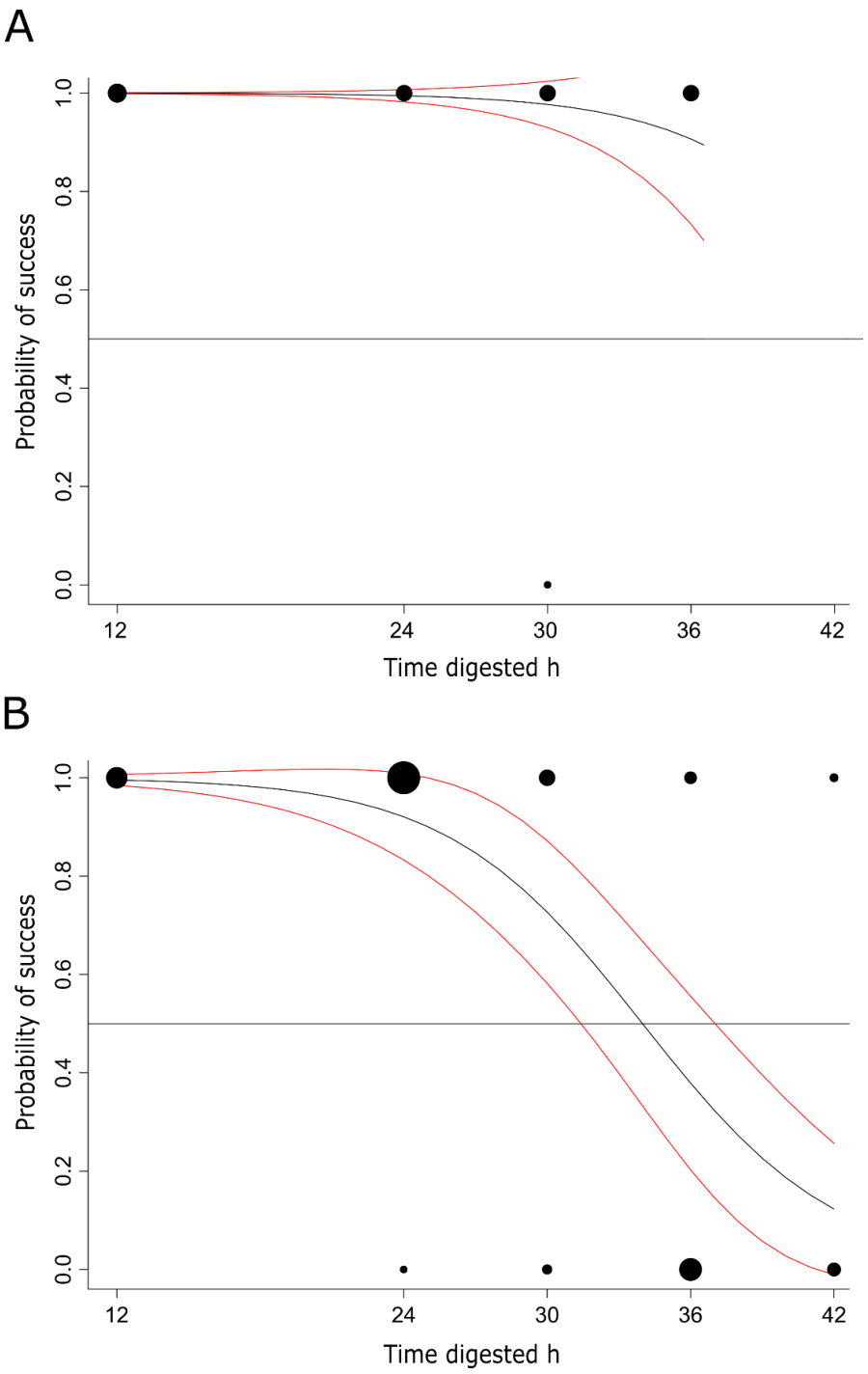


Figure 5. Relationship between DNA barcoding success (i.e., 1 = correctly identifying killifish prey, 0 = erroneously identifying prey as lionfish) and percent killifish prey DNA from experimental data (gray circles). Reference lines (red dashed) denote percentage of killifish DNA above which a 50% probability of correct DNA barcoding is predicted to occur.