VALIDATION AND IMPLEMENTATION OF DNA METABARCODING FOR SEA LAMPREY DIETARY ANALYSIS

By

Conor O'Kane

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Fisheries and Wildlife – Master of Science

2024

ABSTRACT

Genetic dietary analysis has been a rapidly growing area of study due to several advantages it holds over conventional methods, such as enhanced taxonomic resolution and the ability to detect rare or degraded prey items. In this thesis, DNA metabarcoding is applied to investigate the dietary composition of parasitic, hematophagous sea lamprey (*Petromyzon marinus*) within the Great Lakes region. This approach aims to enhance our understanding of sea lamprey feeding habits by addressing limitations with traditional diet assessment methods, ultimately contributing to more informed management efforts. The first objective was to design a blocking primer that selectively suppresses amplification of sea lamprey DNA at the 12S rRNA gene region, allowing for clearer observations of host fish DNA detections. With a successful blocking primer, the second objective was to assess the influence of different environmental and biological variables on the retention of host fish DNA in sea lamprey digestive tracts within a controlled setting. Specifically, various temperatures and post-feeding fasting periods were examined in experimental aquaria, along with the ability to detect multiple hosts after sea lamprey had consecutively fed on different species. Results demonstrated that host DNA could remain detectable in sea lamprey digestive tracts for up to 30 days at temperatures between 5-15°C and still produce sequence reads from feedings on multiple host species. In the third objective, DNA metabarcoding was applied to wild-caught sea lamprey samples to assess the applicability of this technique in the field. Both adult and juvenile parasitic sea lamprey were collected from Lakes Huron, Superior, and Champlain during 2022 and 2023, with findings indicating potential sources of dietary differences among lakes and life stages for Great Lakes sea lamprey. Together, these results underscore the utility of DNA metabarcoding in detecting and distinguishing prey taxa, with valuable applications towards sea lamprey management strategies in the Great Lakes.

ACKNOWLEDGEMENTS

This project was funded by the Great Lakes Fisheries Commission. A special thank you to my committee members Dr. John Robinson, Dr. Kim Scribner, Dr. Nick Johnson, and Dr. Weiming Li. Thank you to Dr. Jeannette Kanefsky for guidance with laboratory protocols, in addition to all other members of the Scribner Lab and Li Lab for their feedback and advice. Thank you to all of those involved in sea lamprey collections including Bill Mattes and the Great Lakes Indian Fish and Wildlife Commission; Matt Symbal, Brad Young and U.S. Fish and Wildlife Service; Gene Mensch, Shawn Seppanen and the Keweenaw Bay Indian Community; Ryan Booth and Fisheries and Oceans Canada; and Dr. Ellen Marsden and the University of Vermont. A special thank you to Tyler Bruning, Emma Wickert, and all members of the Hammond Bay Biological Station (U.S. Geological Survey) for their contributions to this project.

TABLE OF CONTENTS

CHAPTER 1: DEVELOPMENT OF A BLOCKING PRIMER TO ENABLE DIETARY DNA METABARCODING ANALYSIS IN SEA LAMPREY

ABSTRACT

Since the establishment of the invasive sea lamprey (*Petromyzon marinus*) in the Great Lakes during the mid-20th century, extensive management efforts have been aimed at reducing their negative impact on native fishes. Despite a significant reduction in population size using several control methods, uncertainties remain concerning the damage caused by sea lamprey predation on Great Lakes fish populations. While conventional dietary assessments are hindered by the hematophagous nature of sea lamprey, DNA metabarcoding offers a promising alternative by identifying prey species DNA from sea lamprey digestive samples. This method has been used for dietary analyses in a wide variety of species, including lampreys; however, initial assessments using 12S polymerase chain reaction (PCR) primers designed to amplify vertebrate taxa indicated a high presence of sea lamprey DNA per sample. To minimize sea lamprey DNA co-amplification, I designed and tested eight blocking primers for their ability to suppress the amplification of sea lamprey DNA sequences during PCR while allowing amplification of host species DNA. This approach allows for the use of a single marker to amplify a taxonomically diverse suite of host fish species, in contrast to previous studies that used multiple primer pairs designed for specific host families (e.g. Salmonidae, Cyprinidae, Catostomidae), potentially missing rare taxa. Variations among blocking primers included altering base pair length, end sequence modification, and purification method. Samples with different sea lamprey-to-host DNA ratios were subjected to gel electrophoresis, quantitative PCR, and DNA metabarcoding to assess the ability of each blocking primer to reduce the number of sea lamprey sequence reads while maintaining host species sequence reads. Among blocking primers tested, all performed well with versions that included a C3 spacer and HPLC purification demonstrating the highest

effectiveness. Results demonstrate that the single blocking primers evaluated are a reliable method of sea lamprey dietary analysis for amplifying a taxonomically diverse range of host fish species. These experimentally validated methods lay the foundation for future research on the feeding habits and impacts of sea lamprey in the Great Lakes and their native range

INTRODUCTION

The invasion of the parasitic sea lamprey (*Petromyzon marinus*) into the Great Lakes during the mid-20th century and resulting impacts to both native fish communities and angling activity required international management efforts (Coble et al., 1990; Lawrie, 1970). Following the initial invasion and establishment, sea lamprey attacks and overfishing resulted in the loss of over 95% of the lake trout (*Salvelinus namaycush*) stock, and precipitous numerical declines to whitefish (*Coregonus clupeaformis*), walleye (*Sander vitreus*), and other economically and ecologically valuable fisheries (Smith and Tibbles, 1980). In response, the Great Lakes Fishery Commission (GLFC) was established in 1954 and tasked with implementing control programs designed to reduce sea lamprey population abundance within the Great Lakes, focusing on methods such as lampricides, sterile-male releases, physical barriers, and traps (Meyer and Schnick, 1983; Robinson et al., 2021). Control efforts were largely successful, with lampricide and barriers representing the bulk of these efforts, reducing sea lamprey population levels in the Great Lakes to ~10% of their previous peak (Heinrich et al., 2003; Robinson et al. 2021).

Uncertainty remains about the damage that sea lamprey cause to Great Lakes fisheries. Given that sea lampreys are hematophagous, feeding mostly on the blood of their prey, conventional dietary assessment methods are not applicable. Physical components of the diet, such as bones, shells, and other hard structures, are not present in sea lamprey digestive systems, requiring current damage assessments from sea lamprey in the Great Lakes to rely on inspecting

wounds or marks that lamprey leave on their hosts after feeding (Ebener et al., 2003; King Jr., 1980). However, interpretations of marking data are limited, as the focus of current protocols is mostly limited to lake trout, not the full Great Lakes fish community (Firkus et al. 2021; Treska et al. 2021). Captures of parasitic-stage sea lampreys are often the result of bycatch where anglers or vessels were targeting the lamprey's host fish (typically lake trout), caught either with a lamprey still attached or visible wound markings. As such, samples for dietary analysis are not derived from randomly selected sea lamprey within the lake, rather they are gathered from already captured host fish. A host fish must also survive an attack to be considered in the wounding assessment, as deceased fish tend to sink and are therefore unable to be captured for data collection (Adams et al., 2021; Bergstedt and Schneider, 1988). Biochemical methods such as stable isotope analysis and fatty acid profiles have been previously used to circumvent these capture biases in wounding data assessments (Happel et al., 2017; Harvey et al., 2008). However, while these methods are helpful in gathering larger ecological insights such as trophic level placement, they are unable to comprehensively and compositionally characterize sea lamprey diets at the species level. Additionally, the focus on lake trout marking rates to assess ecological damage may not account for difference in host preference. Previous studies have shown sea lamprey may prefer hosts with a higher abundance (Adams and Jones 2021). Given this, a decrease in lake trout abundance may appear as successful control if sea lamprey switch to an alternative primary host fish, as marking rates in lake trout will decrease.

Recent advances in molecular biology, and the development of molecular diet analysis via DNA metabarcoding, offers an attractive alternative that addresses these shortcomings and limitations (Pompanon et al., 2012). While various electrophoretic techniques have been used for dietary analyses previously (Deagle et al., 2005; Symondson, 2002; Walrant and Loreau, 1995),

the rise of next-generation sequencing currently has led to widespread accessibility of DNA metabarcoding for dietary assessments. As reviewed in Pompanon et al. (2012), DNA metabarcoding allows for a species-specific prey designation via DNA extracted from gut contents. Primers specific to certain conserved gene regions, such as the mitochondrial 12S and 16S ribosomal RNA genes(Deagle et al. 2009; Riaz et al. 2011), allow for amplification of sequences from a wide variety of taxa. Amplified products can be subsequently aligned to databases of known sequences to provide reliable taxonomic classifications (Yang et al., 2014). This method has been widely applied for dietary studies on mammals (Berry et al., 2017; Buglione et al., 2018; Lopes et al., 2020), birds (Hacker et al., 2021; McClenaghan et al., 2019), fishes (Berry et al., 2017; Harms-Tuohy et al., 2016; Jakubavičiūtė et al., 2017), and other taxa, including the flesh-feeding Arctic lamprey (Shink et al., 2019) and sea lamprey in the Great Lakes (Johnson et al., 2021).

Johnson et al. (2021) introduced a method for identifying host species using DNA extracted from sea lamprey feces using three taxon-specific primers that individually targeted salmonids, catastomids, and cyprinids. The method was largely successful, showing that diet composition varied between sea lamprey captured in the northern basin of Lake Huron and those from a tributary of Lake Huron. However, the study was meant as a proof-of-concept, and the use of multiple taxon-specific primers limited their ability to compare relative sequence abundance of multiple hosts detected in individual sea lamprey fecal samples (Johnson et al., 2021). While beneficial for detecting a wider range of hosts, use of more conserved vertebrate primers also amplifies large amounts sea lamprey DNA, lowering the proportion of usable data. This reduction decreases the effectiveness of higher data outputs in mitigating sequence read

biases introduced by extraction and amplification stochasticity (Alberdi et al., 2018; Leray and Knowlton, 2017; Polz and Cavanaugh, 1998).

One method of counteracting the amplification of sea lamprey DNA is to apply a blocking primer (Vestheim and Jarman, 2008). The inclusion of an effective blocking primer during the PCR process can significantly suppress predator DNA amplification while increasing the relative amount of amplified prey fragments in dietary studies (Vestheim et al., 2011). Blocking primers are also capable of blocking amplification of prey sequences, potentially limiting benefits (Piñol et al. 2015). Unlike universal primers, which are designed to anneal broadly to various taxonomic groups, blocking primers are designed to attach to a specific target. Amplification prevention then occurs either via 1) annealing inhibition, where the blocking primer binding site overlaps with the universal primer binding site and the physical presence of the blocker prevents annealing, or 2) elongation arrest, where the blocking primer attaches downstream and physically prevents non-target sequence elongation (Vestheim et al. 2011). Once attached to their specific target, the physical presence of the blocking primer prevents amplification of the sequence during PCR. As such, if a blocking primer were designed with the proper specificity to only anneal to sea lamprey sequences, it would suppress the overall amplification of sea lamprey sequences within the sample and still allow for the amplification of host species sequences. This method has been successfully applied in other DNA metabarcoding dietary studies (Jakubavičiūtė et al., 2017; Leray et al., 2013; Su et al., 2018).

This study focused on design and testing of blocking primers to determine effectiveness for amplification suppression of the sea lamprey 12S rRNA gene region, while allowing amplification of host species DNA. An annealing inhibiting blocking primer design was selected over an elongation arrest blocker given its expected higher efficiency (Vestheim et al., 2011). I

designed eight blocking primers, representing each possible combination of three primer design features (base pair (bp) length, end sequence modification, and purification method). I further tested blocking primer effectiveness using three primer evaluation methods (visualization of conventional PCR amplification products, quantitative PCR, and high-throughput metabarcoding sequencing of single species), each applied to single and mixed-species templates in mock communities, as well as dietary samples of wild-caught adult sea lamprey.

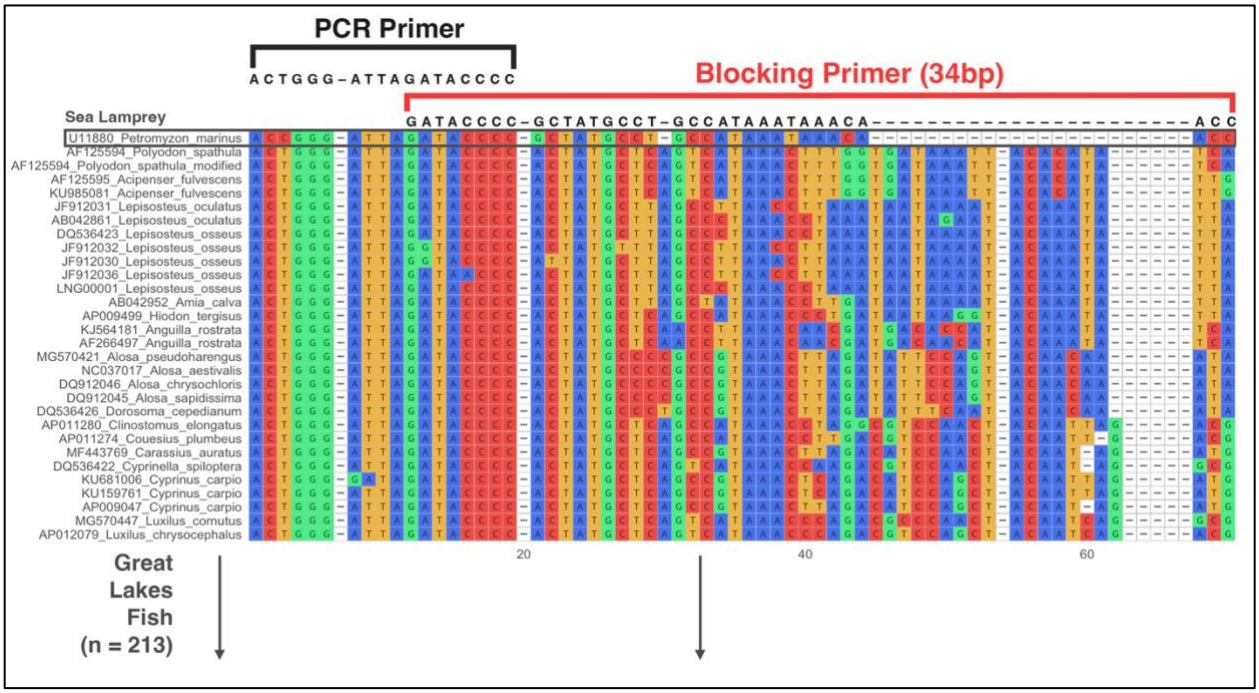
METHODS

Blocking Primer Development

A collection of 213 sequences was obtained from the NCBI GenBank database (Sayers et al. 2022) ranging from 89 to 107bp for the 12S mitochondrial rRNA gene region for 149 Great Lakes fish species, targeting the same segment as used by Riaz et al. (2011; 12S-V5 primer set). Multi-sequence alignments were created using MEGA (v. 6.0; Koichiro et al. 2013) for comparison of these sequences to the same gene region in sea lamprey. Additionally, the forward 12S-V5 primer (Riaz et al. 2011) sequence was appended to the 5' end of the sequence alignment to allow for the consideration of an annealing inhibiting blocking primer design.

Variation in the sequence length can affect primer species specificity and annealing temperature (Vestheim et al., 2011). For this study, primer lengths of 34bp and 36bp were selected to include a 23bp gap that was noted in the alignment between sea lamprey and other Great Lakes fishes beginning at bp 42 (including the primer binding region) in the sea lamprey sequence (Figure 1.1). Targeting this gap should enhance primer specificity for sea lamprey sequences relative to prey fish DNA sequences.

Figure 1.1. Visualization for the overlap of a 34bp blocking primer with the universal 12S PCR primer. The first sequence is the target sequence (sea lamprey), while the following sequences represent the first 30 sequences of Great Lakes fish from the constructed sequence database of 213 sequences. The first 75 nucleotides of each sequence are shown, with total length being 156 nucleotides (including gaps and primers).



Two end modifications and two purification methods were selected for comparison of blocking primer effectiveness: a C3 spacer (three hydrocarbons added at the 3' end) and a 3' inverted dT (reverse-linked nucleotide). End modifications aim to decrease the possibility of amplification from the blocking primer itself, and while a 3' C3 spacer has often been used (Homma et al., 2022; Nelson et al., 2017; Robeson II et al., 2018), other end modifications such as an 3' inverted dT have been successful in similar dietary studies (Egizi et al., 2013). Both additions are standard with most oligonucleotide suppliers and offer to improve blocking primer efficiency by inhibiting both DNA polymerase extension and 3' exonuclease degradation (Egizi et al., 2013; Liu et al., 2019).

For primer purification methods, desalting is typically used by most vendors, while High-Performance Liquid Chromatography (HPLC) purification is recommended for usage with

blocking primers for superior binding efficiency (Vestheim et al., 2011). Eight blocking primers were developed, representing all combinations of these variables (purification methods, end modifications, and sequence length variation; Table 1.1). The eight blocking primers evaluated in this study were synthesized and purified by Integrated DNA Technologies (Coralville, IA).

Table 1.1. Table of the eight blocking primers designed and tested in this study. The primer name, length (bp), end modification, purification method, 5'-3' sequence, and melting temperature (T_m) for each is given.

Blocking Primer	Length (bp)	End Modification	Purification Method	Sequence (5'–3')	Tm*
Blocker 1	36	C3 spacer	Desalted	GATACCCCGCTATGCCTGCCATAAATAAACAACCGT/3SpC3/	65.8
Blocker 2	36	C3 spacer	HPLC	GATACCCCGCTATGCCTGCCATAAATAAACAACCGT/3SpC3/	65.8
Blocker 3	36	Inverted dT	Desalted	GATACCCCGCTATGCCTGCCATAAATAAACAACCGT/3InvdT/	65.8
Blocker 4	36	Inverted dT	HPLC	GATACCCCGCTATGCCTGCCATAAATAAACAACCGT/3InvdT/	65.8
Blocker 5	34	C3 spacer	Desalted	GATACCCCGCTATGCCTGCCATAAATAAACAACC/3SpC3/	63.8
Blocker 6	34	C3 spacer	HPLC	GATACCCCGCTATGCCTGCCATAAATAAACAACC/3SpC3/	63.8
Blocker 7	34	Inverted dT	Desalted	GATACCCCGCTATGCCTGCCATAAATAAACAACC/3InvdT/	63.8
Blocker 8	34	Inverted dT	HPLC	GATACCCCGCTATGCCTGCCATAAATAAACAACC/3InvdT/	63.8

*Tm calculated using nearest neighbor method

Primer Evaluation: Conventional PCR

An initial assessment of blocking primer effectiveness was conducted using conventional PCR and gel electrophoresis to visualize amplified products. Single-species DNA samples from multiple individual sea lamprey and two Great Lakes native host fish species (lake trout and walleye) were diluted to 5ng/μL. Lake trout was selected due the current understanding that the species constitutes a large component of sea lamprey diets. Walleye was selected given its potential as a host and higher sequence similarity to sea lamprey within the region of interest. Each of the eight blocking primers were then added to one of the six single-species DNA samples for 48 total samples that included both a blocking primer and DNA from either sea lamprey, lake trout, or walleye. For each of the six single-species samples, a no-blocking primer

sample was included as a positive amplification control. Additionally, each blocking primer was included in a no DNA, PCR negative amplification control. Blocking primers were tested at a relative concentration of 10:1 to unmodified 12S primers (12S-V5; Riaz et al. 2011), following recommendations from Vestheim et al. (2011).

PCR was performed in a 15 μL reaction volume with 1.5 μL of 10X AmpliTaq Gold PCR Buffer II (Applied Biosystems, Waltham, MA), 0.36 μL of dNTPs (10mM), 1.2 μL of $MgCl_2$ (25mM), 0.75 μL of BSA (20 mg/mL), 0.8 μL of unmodified 12S F and R primers (10 μM) and blocking primers (100 μM), 6.54 μL of Millipore water (UV treated), 0.25 μL of AmpliTaq Gold (5U/μL) and 2 μL of template DNA (20 ng/μL) for all species. Positive control samples replaced blocking primer additions with Millipore water. Negative control samples replaced DNA template with water. Thermal conditions for PCR were as follows: 10 min at 95°C (1x); 30s at 95°C, 30s at 57°C, 45s at 72°C (40x); and 5min at 72°C. Following amplification, 4 μL of PCR product and 2.5 μL of glycerol loading dye was run on a 1% agarose gel with GelRed stain (0.5x; Biotium, Fremont, CA). Gels were photographed under UV light in a Labnet Enduro GDS II imaging system (Labnet HQ, Edison, NJ).

Primer Evaluation: Quantitative PCR

To gain insight into PCR amplification efficiency of mixed species samples, three mock communities were set up with varying DNA concentration ratios of sea lamprey and either lake trout or white sucker (M1, M4, and M5; Table 1.2). Alongside lake trout, white sucker is a known host fish of wild sea lamprey and was thus included in this analysis. All DNA samples of sea lamprey and host fish were diluted to 5 ng/μL prior to mixing. M1 was an equal ratio of sea lamprey DNA and lake trout DNA (1:1) assembled from samples of known DNA concentrations. M4 and M5 contained skewed 9:1 ratios of sea lamprey DNA to lake trout and white sucker

DNA, respectively. Testing both uniform and skewed lamprey-to-host DNA ratios allowed for comparisons of DNA amplification at equal concentrations, along with comparisons those more closely aligned with wild-caught samples from preliminary analyses (where sea lamprey sequences comprised ~90% of all sequence reads). Additionally, single-species DNA samples were again used, with two replicates of sea lamprey, lake trout, white sucker, and walleye each being incorporated into the analysis. All single-species samples were diluted to 5 ng/μL for consistency with the mixed-species samples. A final wild-caught sample, HP3, was also included. This sample contained extracted DNA from the gut contents of a parasitic-stage sea lamprey caught in Lake Huron. For this sample and all wild-caught samples in this study, DNA extractions followed the protocol used by Johnson et al. (2021) with the gMax Mini Genomic DNA Kit (IBI Scientific, Dubuque, IA).

Table 1.2. Table of DNA samples used for quantitative PCR analyses, showing sample name, species DNA included in the sample, and details of which individual for single-species samples, lamprey-to-host DNA concentration ratios for mixed-species samples, or capture data for the wild-caught sample (HP3).

Sample	Species	Details
SL1	Sea Lamprey	Individual 1
SL2	Sea Lamprey	Individual 2
LT1	Lake Trout	Individual 1
LT2	Lake Trout	Individual 2
WAE1	Walleye	Individual 1
WAE2	Walleye	Individual 2
WS1	White Sucker	Individual 1
WS2	White Sucker	Individual 2
M1	Sea Lamprey, Lake Trout	50:50 Lamprey:Host
M4	Sea Lamprey, Lake Trout	90:10 Lamprey:Host
M5	Sea Lamprey, White Sucker	90:10 Lamprey:Host
HP3	Wild-Caught	Huron Parasitic

Each the 12 DNA samples were subjected to quantitative (q)PCR with each blocking primer and in a no-blocking-primer control. A no-DNA control sample was also analyzed for each blocking primer. qPCR was performed in a 20 µL reaction volume including 10 µL of 2X Forget-Me-Not EvaGreen Master mix (low ROX; Biotium, Fremont, CA), 0.8 µL of forward and reverse primers at 10 µM and blocking primers at 100 µM (10x the concentration of 12S primers), 5.6 µL of sterile water, and 2 µL of template DNA. A no-template control (NTC) was included through the addition of sterile water in place of template DNA. Thermal cycling took place on a QuantStudio 6 (Applied Biosystems, Waltham, MA) and conditions were as follows: 2 min at 95°C (1x); 5s at 95°C, 10s at 57°C, and 20s at 72°C (40x; imaging at extension step). Amplification plots were used to determine cycle threshold (Ct) values, which were used to compare differences in relative DNA concentrations among samples with and without blocking primers. The degree of suppression was then calculated by assuming a doubling of DNA per cycle and using the difference in cycle thresholds as an exponent of two. Amplified products were also subjected to melt curve analyses immediately following completion of the final cycle to compare melting temperature (*Tm*) peaks of products from single-species samples and mixed-species samples with and without blocking primers.

DNA Metabarcoding

Mixed-species samples were also subjected to DNA metabarcoding tests to provide species-specific sequence read counts. Mock community samples M1 (1:1 sea lamprey:lake trout DNA ratio), M4 (9:1 sea lamprey:lake trout DNA ratio), and M5 (9:1 sea lamprey:white sucker DNA ratio) from the qPCR analyses along with three parasitic-stage sea lamprey from Lake Huron (HP3, HP5, and HP15) and an adult sea lamprey captured in a trap during spawning migration from Lake Champlain (CA14) were included in metabarcoding tests. Reaction

volumes were 15 µL with the same reagent concentrations as the conventional PCR analyses, again substituting UV-treated Millipore water in place of DNA template for NTC samples and including no-blocking-primer controls for all DNA samples.

PCR followed the same conditions as the conventional PCR analysis, and an additional set of samples was run with 25 cycles instead of 40 cycles. This test assessed whether reducing the number of PCR cycles would prevent amplification of sea lamprey sequences in later cycles. However, as a higher PCR cycle number should increase overall ratios between target and non-target DNA (Vestheim et al. 2011) along with testing capacity being limited, only two blocking primers from previous qPCR tests (Blocker 2 and 6) and an NTC were selected for the additional 25-cycle run while all eight blocking primers were tested in 40-cycle PCR. This selection was made based on the high performance of these blocking primers in previous analyses (see above). Relative sequence read counts for all members of mock communities and wild-caught samples were evaluated for both cycle numbers.

A secondary PCR was run after samples were barcoded with i7 (1 µL; 10 µM) and i5 (2 µL; 5 µM) index primers along with 2X Qiagen Plus MM (5 µL; #206152) to allow for demultiplexing of sequencing reads to their corresponding samples. Barcoding PCR conditions were as follows: 15 min at 95°C (1x); 10s at 95°C, 30s at 65°C, 30s at 72°C (10x); and 5min at 72°C. PCR products were then pooled and bead size selected at 0.5x and 1.2x bead concentrations to remove long and short sequences, respectively. Pooled libraries were diluted and analyzed via TapeStation (assay High Sensitivity D100 ScreenTape) for sequence length confirmation before being sequenced on a 300 cycle Illumina MiSeq lane (v2 Micro; 2 x 150bp paired end) at Michigan State University's Research Technology Support Facility.

Processing of sequencing data into relevant feature tables was performed using the mothur software package (version 1.48.0; Schloss et al., 2009). Sequence reads were demultiplexed, assembled into contigs with the allowance of two mismatches from overlapping paired-end reads, and subsequently trimmed of primer sequences. Trimming was specified to retain only the overlapping regions.

Following the creation of contigs and continuing through the mothur pipeline, a series of commands were used in the preprocessing and analysis of the sequence data (see APPENDIX B: SUPPLEMENTARY MATERIALS). Initially, sequences were summarized and subsequently screened to filter out reads exceeding 107 bp or large homopolymers. Unique sequences were generated and counted before being aligned to the reference database, then pre-clustered (0 differences allowed within clusters). Chimeric reads were identified and removed. Classifications were defined for the sequences with a confidence score cutoff of 80% or higher. A distance matrix was then calculated with a cutoff of 0.03, and sequences were clustered based on a distance cutoff of 0.01. Finally, sequences were grouped into operational taxonomic units (OTUs) based upon a 99% similarity threshold using our 12S rDNA Great Lakes fish database (see APPENDIX B: SUPPLEMENTARY MATERIALS). Sequence outputs were then cleaned and organized into a community matrix table with count numbers by OTU (see APPENDIX B: SUPPLEMENTARY MATERIALS), from which further data analyses and visualizations were constructed.

A series of paired t-tests were used to statistically compare the effectiveness of each blocking primer. Sea lamprey read counts across the seven dietary samples for each blocking primer were compared to sea lamprey read counts of unblocked samples. This provided a basis for determining whether mean read counts of sea lamprey DNA decreased with the inclusion of

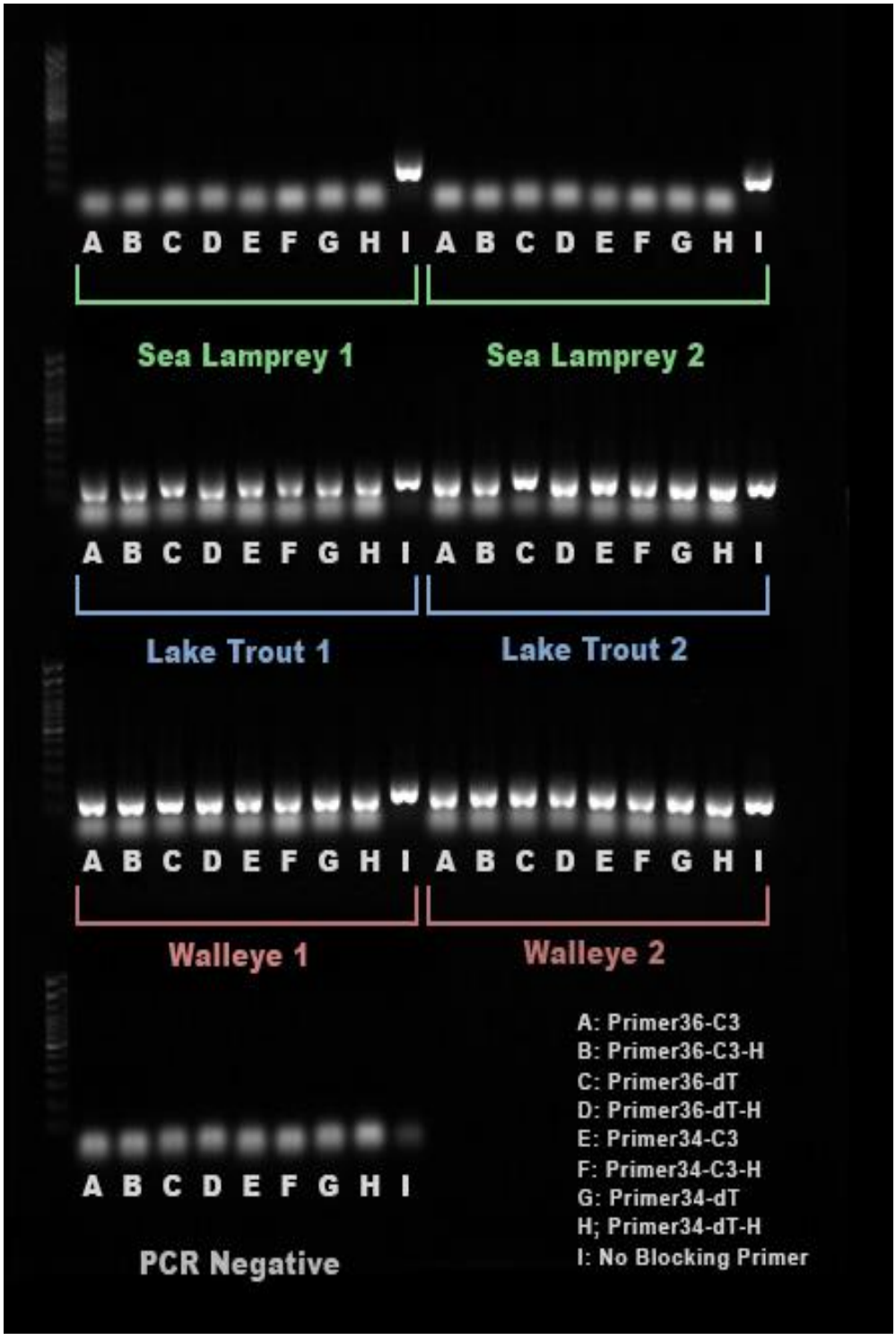
individual blocking primers. Similarly, lake trout read counts across dietary samples with each blocker were compared to lake trout read counts in unblocked samples. For lake trout sequence read comparisons, samples M5 and CA14 were not included, as the M5 sample focused on white sucker as host DNA and CA14 was an outlier wild-caught sample with very low amounts of host DNA. Blocking primers that significantly reduced sea lamprey sequence reads while not significantly reducing host fish sequence reads were deemed effective.

RESULTS

Primer Evaluation: Gel Electrophoresis

PCR amplicon band presence was used to identify successful sample amplification (Figure 1.2). No samples that included a blocking primer showed a visible band of the expected size, while both control samples without a blocking primer showed a band. Results indicated that amplification of sea lamprey DNA was suppressed in all blocking primer samples. For walleye and lake trout replicates, all samples both with and without blocking primers showed a band, indicating no visually identifiable inhibition of host species amplification. The PCR negative samples showed no bands of the expected fragment size in any sample.

Figure 1.2. Visualization of results for gel electrophoresis primer evaluation method for both replicates of sea lamprey, lake trout, and walleye. A-H indicate which primer was included in the sample, and I indicates the sample without a blocking primer. PCR negatives contained no DNA.

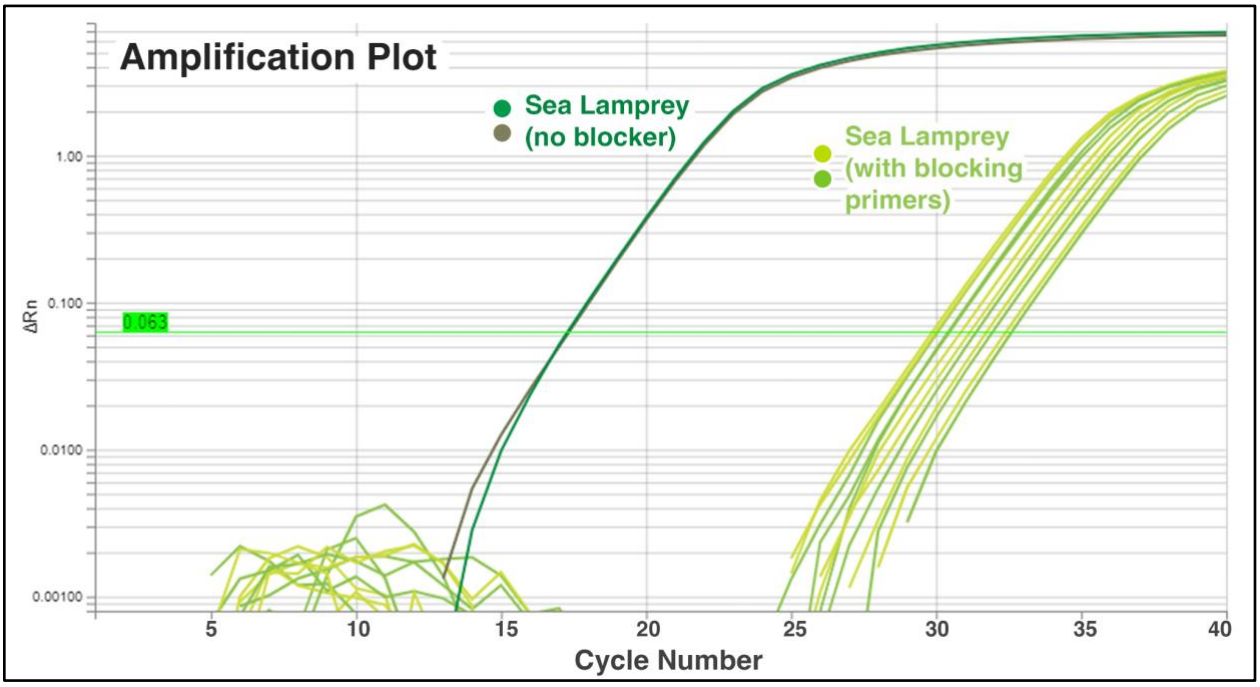


While target DNA bands were easily visible, the presence of primer dimers was also noted in all samples that included a blocking primer, including negative controls (Figure 1.2). Bands well below the target length were considered as primer dimers (typically between 10-20bp), as similar results have been found in previous blocking primer analyses (Liu et al., 2019). Presence of these primer dimers prompted the use of a Sage Science BluePippin size selection step for future analyses, which was able to fully reduce primer dimer presence (data not shown).

Primer Evaluation: Quantitative PCR

Amplification plots were used to compare samples that only contained DNA from one species. Sea lamprey DNA amplification was compared in PCR reactions with and without blocking primers to compare PCR amplification profiles based on changes in Ct values as a quantitative measure of the degree of amplification suppression (Figure 1.3). In sea lamprey samples that did not include a blocking primer, the mean Ct value across both replicates was 17.4 cycles (sd = 0.02). When blocking primers were included, the mean Ct value across replicates was 31.12 cycles (sd = 0.11), a mean Ct difference of 13.72 cycles or an average of $2^{13.72}$ = 13,494x suppression of sea lamprey amplification in reactions with a blocking primer. The best-performing blocking primer in this comparison was Blocker 6 for both replicates (Ct values of 32.65 and 32.48 cycles). Single-species samples of lake trout, walleye, and white sucker DNA were also compared between samples that did and did not contain blocking primers to ensure there was no inhibition of amplification from target species. For lake trout, the mean Ct value difference between samples with and without a blocking primer was 1.28 cycles (sd = 0.17), indicating minimal suppression of 2.43x. Walleye samples had a mean Ct value difference of 3.37 cycles (sd = 0.22), resulting in a mean suppression of 10.3x. White sucker samples had the lowest mean Ct value difference at 0.59 cycles (sd = 0.11), an overall mean suppression of 1.5x.

Figure 1.3. Amplification plot showing results for the qPCR primer evaluation illustrating the Ct differences between blocker and no-blocker reactions for sea lamprey samples. Dark green and gray lines towards the left side represent both sea lamprey samples with no blocking primers included in the reaction, while lighter green lines to the right represent the same sea lamprey samples with each of the eight blocking primers included. The horizontal green line represents the threshold used to determine Ct values for each reaction.

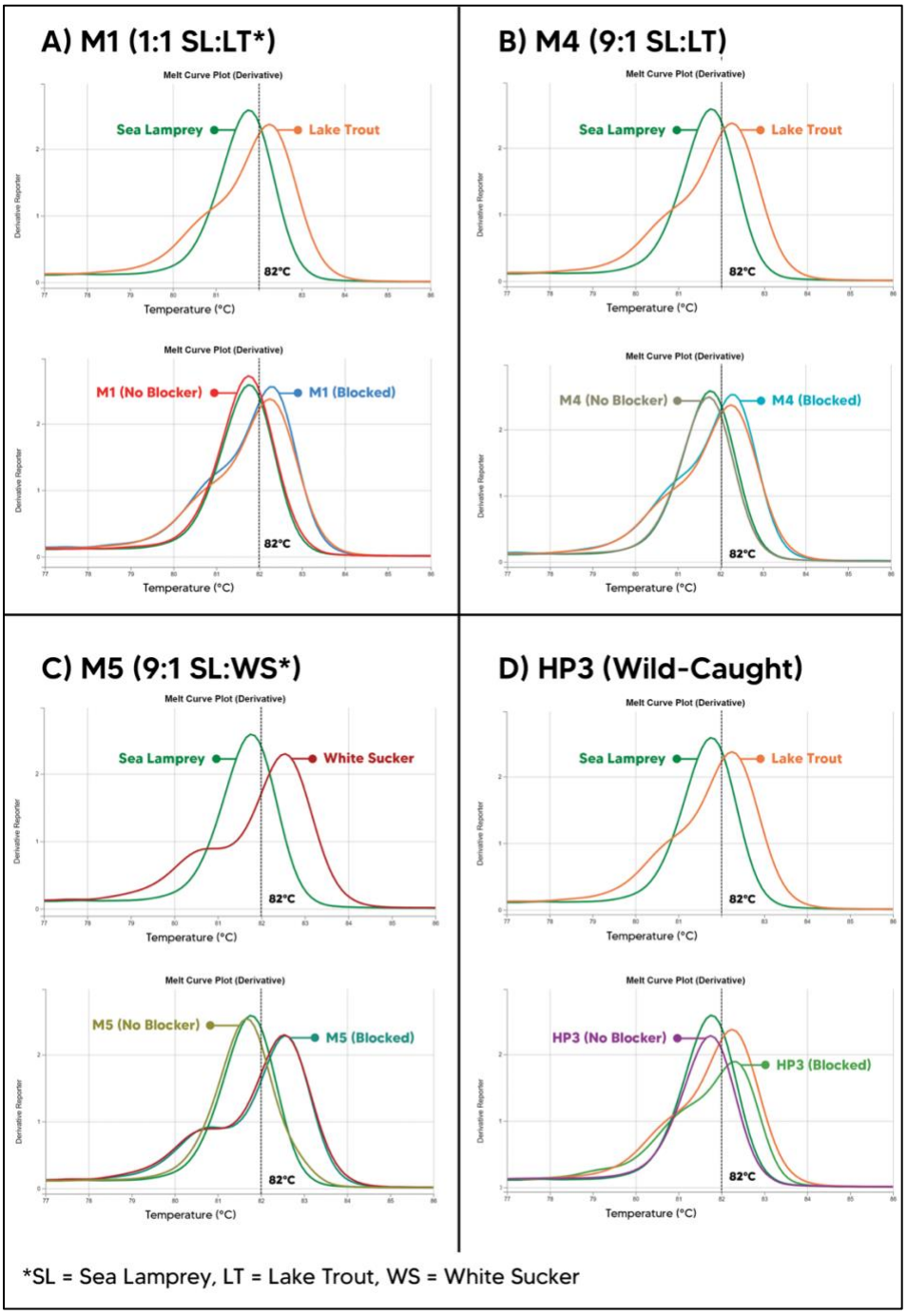


Melt curve plots were created to examine the difference in melting temperature (T_m) between mixed-species, host-only, and sea lamprey-only samples with and without blocking primers. Together, both sea lamprey sample replicates with no blocking primers generated a mean T_m of 81.77°C (sd = 0.03), establishing a sea lamprey baseline for the comparison of mixed-species samples when a blocking primer was included. Similar baselines were established for host species (lake trout, white sucker, and walleye) to identify T_m shifts. For all four tested mixed samples, when a blocking primer was included, the T_m shifted from the sea lamprey baseline and towards the T_m of the associated host species. Additionally, all HP3 wild-caught samples displayed a shift away from the sea lamprey baseline T_m and towards the lake trout

baseline T_m with the inclusion of a blocking primer, indicating a higher presence of amplified lake trout DNA.

To visualize these shifts, melt curves displaying T_m peaks were plotted to demonstrate which DNA was being primarily amplified in mixed-species samples when compared to the single-species baselines (Figure 1.4). In all four mixed samples, the shape and peak of the melt curves for the unblocked sample closely resembled the melt curves for the sea lamprey sample. For all blocked samples the shape and peak of the melt curves closely resembled that of the host species. This indicates amplification of primarily sea lamprey DNA when no blocking primer is included, and amplification of primarily host fish DNA when a blocking primer is included for every mixed sample. It should be noted that the "blocked" melt curves in Figure 1.4 only include those of samples with Blocker 6. All blockers showed similar effects (data not shown).

Figure 1.4. Melt curve plots showing results for the qPCR primer evaluation method for M1 (1:1 sea lamprey:lake trout DNA ratio), M4 (9:1 sea lamprey:lake trout), M5 (9:1 sea lamprey:white sucker), and HP3 (wild-caught Lake Huron parasitic-stage sea lamprey) samples with Blocker 6. Top figures in each quadrant (A, B, C, and D) indicate baseline melt curves for sea lamprey and host species amplicons (all single-species samples with no blocking primer). Bottom figures in each quadrant overlay the sample melt curves with and without a blocking primer to illustrate the shift in melting temperature (T_m). Note: all blockers were tested and showed similar results, but only Blocker 6 is shown here for clarity.



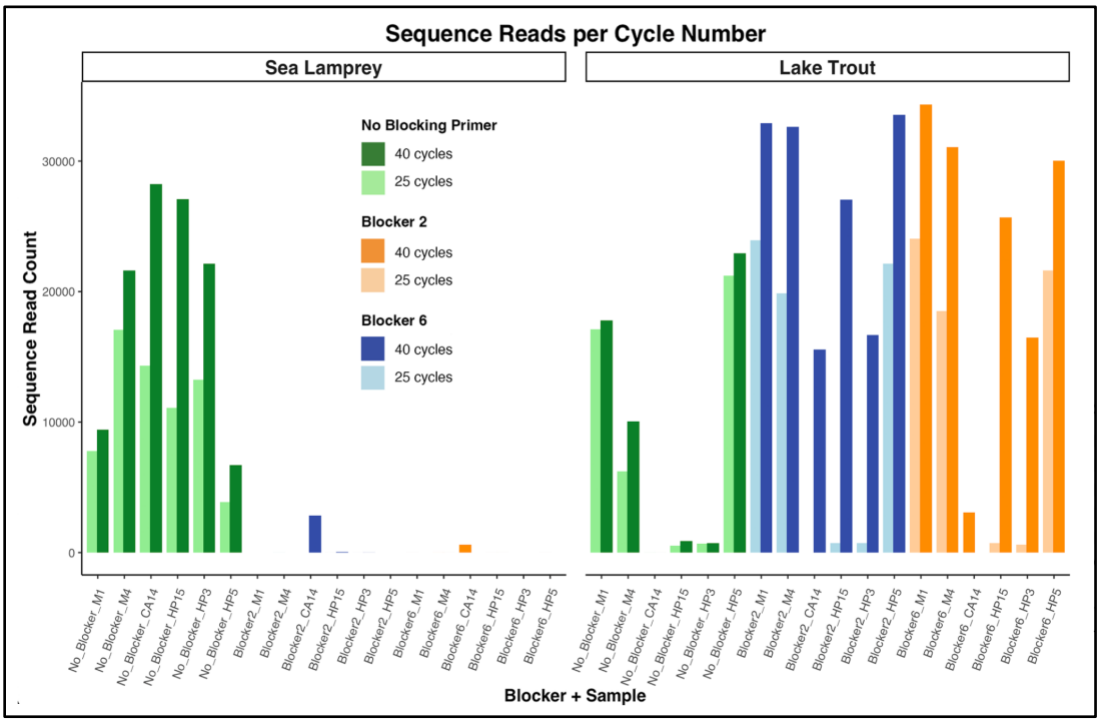
DNA Metabarcoding: OTU Classifications

Initial analysis of DNA metabarcoding results via inspection of raw OTU sequence read counts from mothur analyses shows a total of 23 OTUs were established with at least one sequence read count. Of these 23 OTUs, the top four in terms of read count (*Salvelinus namaycush*, *Catostomus commersonii*, Petromyzontidae unclassified, and Salmonidae unclassified) accounted for 96.76% of the total sequence reads (Appendix A: Figure 1.1). As the primary goal of this study is to compare the effectiveness of different blocking primers and not an in-depth investigation into the dietary composition of wild-caught sea lamprey, only these first four OTUs were used for blocking primer comparisons. Some OTUs with smaller proportional read counts could likely be grouped with these larger OTUs, such as categorizing *Salvelinus unclassified* as lake trout given the high quantities of known lake trout DNA in mock communities and observable attachment of wild-caught sea lamprey to lake trout hosts. However, to avoid additional assumptions and given the low impact these groupings have on the results, none of the unclassified, lower-proportioned OTUs were grouped with any of the top four OTUs. The only assumption made in terms of "unclassified" OTUs was that, given the prevalence of known sea lamprey DNA in both mock communities and wild-caught samples, Petromyzontidae unclassified was considered to represent sea lamprey (*Petromyzon marinus*) reads. Assignments of sea lamprey sequences at the family level were expected, as the 12S sequence for sea lamprey in our reference database differs from three other lamprey species (American brook lamprey, *Lampetra appendix*; northern brook lamprey, *Icthyomyzon fossor*; and silver lamprey, *Icthyomyzon unicuspis*) by a single base pair.

DNA Metabarcoding: Cycle Number Comparisons

Sequence read counts were visualized to compare differences between sea lamprey and lake trout community outputs based on 25-cycle and 40-cycle PCR runs (Figure 1.5). With sea lamprey, amplification suppression was high from both 25 and 40 cycles, indicating no benefit of cycle number. However, a noticeable improvement was made in the 40-cycle PCR concerning the amount of host read amplification. In all samples, host reads were higher from 40-cycle PCR runs. In sample CA14, a wild-caught adult sample from Lake Champlain, no lake trout reads were detected for either blocking primer in their respective 25-cycle runs; however, both 40-cycle runs with each blocking primer detected lake trout sequences. Sequence reads for the M5 sample with white sucker showed similar results.

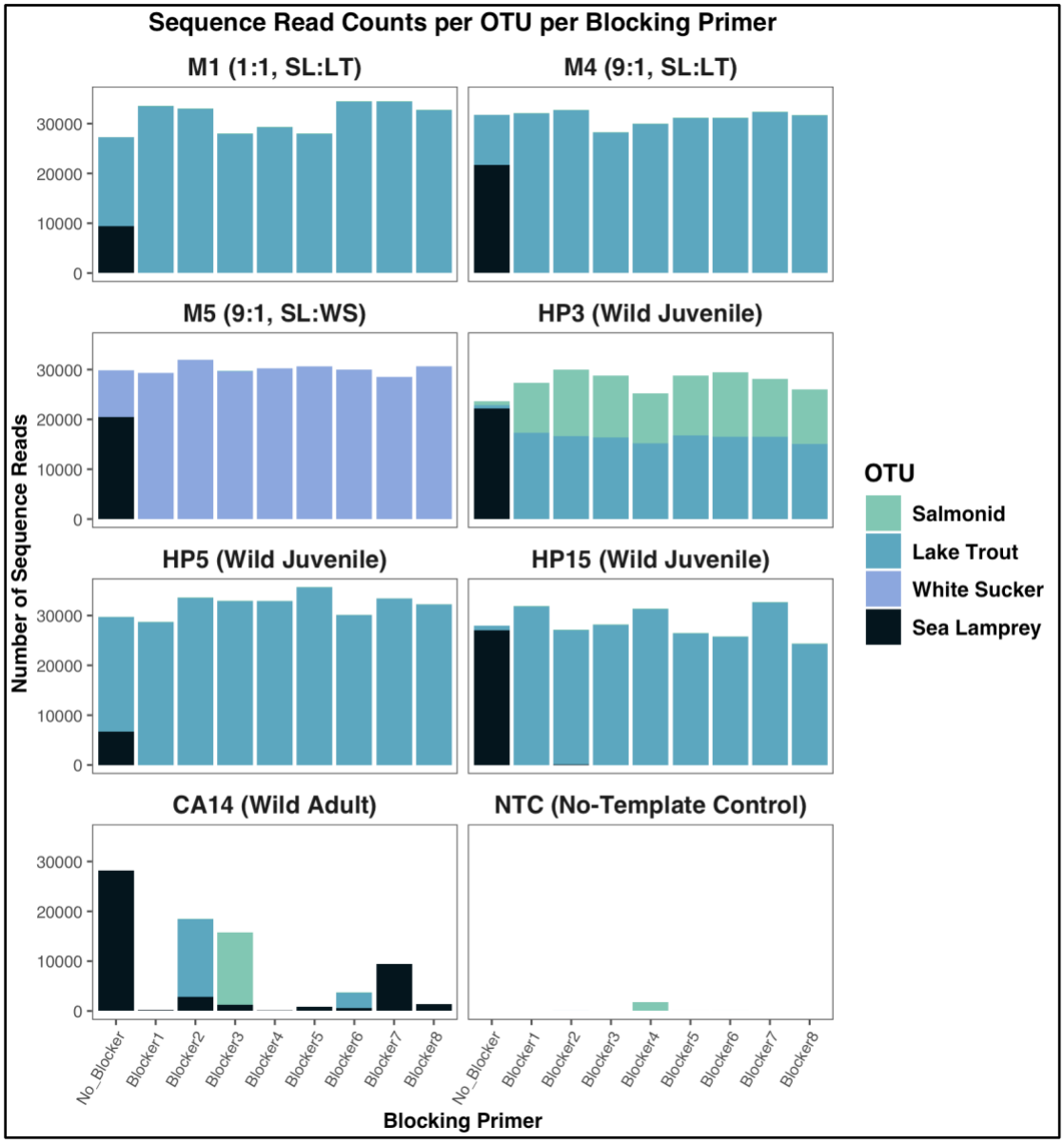
Figure 1.5. Sequence read counts for sea lamprey and lake trout from mixed-species samples in 25 and 40-cycle PCR runs with Blocker 2, Blocker 6, or no blocking primer included. Green columns indicate samples with no blocking primer, orange columns indicate Blocker 2 samples, and blue columns indicate Blocker 6 samples. Lake trout reads are consistently higher in 40-cycle samples, with effective suppression of sea lamprey reads across both 25 and 40-cycle samples.



DNA Metabarcoding: Blocking Primer Sequence Reads

All 40-cycle samples were compared using the distribution of sequence reads per OTU with each blocking primer (Figure 1.6). A sample with no blocking primer (No_Blocker) was used as a baseline for comparisons.

Figure 1.6. Sequence read count per OTU comparisons between mixed-species samples. The first column in each panel represents that sample with no blocking primer included. NTC samples did not include DNA. High levels of host sequence amplification and lamprey sequence suppression are seen across all blocking primers, with CA14 providing insights into the most effective primers as only Blocker 2 and Blocker 6 detected species-specific host reads.



In all three mock communities (M1, M4, and M5), the number of sea lamprey reads with the inclusion of any blocking primer decreased an average of 99.99% across all mock communities, showing the efficacy of all tested blocking primers. Similarly, in Huron Parasitic (HP) wild-caught samples, even with high proportions of sea lamprey in samples without a blocking primer, all blocking primers were extremely effective in suppressing sea lamprey amplification averaging a reduction of 99.98% of sequence reads while still allowing for the amplification of host fish.

The only DNA sample to show clear discrepancies in blocking primer effectiveness was CA14, a wild-caught spawning adult from Lake Champlain. Across all tested blocking primers, only three suppressed sea lamprey amplification and allowed for host amplification in this sample. Of these three, only Blockers 2 and 6 registered species-level identification (*Salvelinus namaycush*) of the host, with Blocker 3 only able to provide family-level identifications (Salmonidae) of host sequences. Given that no contamination of sequences was detected in the NTC samples for Blockers 2 and 6, these host reads were designated as true positives. Between both Blocker 2 and 6, the proportion of host reads to total reads was comparable (84.6% and 83.2%, respectively).

Paired t-tests statistically demonstrated the differences in read counts between samples that included blocking primers and samples that did not include blocking primers (Table 1.3). All blocking primers significantly reduced sea lamprey read counts in dietary samples ($p < 0.001$), and significantly increased lake trout read counts ($p < 0.05$). Additionally, t-values were all > 6.0 for sea lamprey read count comparisons and all < -4.0 for lake trout read comparisons, emphasizing the effectiveness of all blocking primers. Full sequence read data for both lake trout

and sea lamprey for all samples (except M5 in lake trout comparisons) is included in the supplemental material (Appendix A: Figure S1.1).

Table 1.3. Paired t-tests comparing read counts for samples that contain one of the eight blocking primers to samples that do not contain a blocking primer for both sea lamprey and lake trout.

Sea Lamprey Read Count Comparison

Blocker	Mean (Unblocked)	Mean (Blocked)	p-value
Blocker 1	19387.29	35.57	0.0008
Blocker 2	19387.29	409.00	0.0007
Blocker 3	19387.29	187.71	0.0007
Blocker 4	19387.29	11.29	0.0008
Blocker 5	19387.29	114.43	0.0008
Blocker 6	19387.29	88.57	0.0008
Blocker 7	19387.29	1360.57	0.0006
Blocker 8	19387.29	194.71	0.0007

Lake Trout Read Count Comparison

Blocker	Mean (Unblocked)	Mean (Blocked)	p-value
Blocker 1	8728	23866.83	0.0204
Blocker 2	8728	26386.50	0.0006
Blocker 3	8728	22226.17	0.0154
Blocker 4	8728	23057.67	0.0185
Blocker 5	8728	22948.50	0.0113
Blocker 6	8728	23439.50	0.0072
Blocker 7	8728	24838.50	0.0142
Blocker 8	8728	22606.00	0.0105

DISCUSSION

DNA metabarcoding can be a powerful tool for dietary analysis, particularly in the case of sea lamprey where it could be used to overcome hurdles posed by traditional methods. Current approaches to monitoring damage inflicted by sea lamprey have limitations in providing a complete picture, including failure to capture deceased prey, relying on targeted host capture,

and subjectivity in wound assessments prevent a comprehensive understanding. DNA metabarcoding could address some of these limitations by providing a time-specific, location-specific and species-specific dietary profile. In this study, I used three molecular methods (gel electrophoresis, quantitate PCR, and DNA metabarcoding) to evaluate the effectiveness of several blocking primers for isolating and identifying prey species from sea lamprey digestive samples. All blocking primers tested were deemed effective in suppressing the amplification of 12S mitochondrial rDNA sea lamprey sequences while retaining amplification of prey sequences, although two blocking primers employing C3 spacer end modifications and HPLC purification exhibited the highest efficiency and enhanced species-specific detection of prey. These results help to lay the groundwork for future research that will provide more extensive insights into Great Lakes sea lamprey diets and their impact on native fishes.

The effectiveness of blocking primers to suppress amplification of a non-target species has been investigated previously, including in the context of dietary composition studies (Leray et al. 2013; Pertoldi et al. 2021; Homma et al. 2022). While some studies present sample richness comparisons of blocking primers against other predator DNA-inhibiting techniques such as peptide nucleic acid (PNA) clamps or restriction enzymes, these studies often lack comprehensive comparisons of blocking primer variables and efficiencies (Lefèvre et al. 2020; Taerum et al. 2020). Often, end modifications such as C3 spacers and dT inversions are not fully contrasted, and a single end modification is chosen for all included blockers within the study (Huggins et al. 2020; Rojahn et al. 2021). Previous studies also often lack other forms of comparison such as variation in purification methods or the inclusion of mock communities with known DNA concentrations of specific species (De Barba et al. 2014; Toju and Baba 2018; Pertoldi et al. 2021). In our study, conventional PCR and gel electrophoresis provided a rapid,

but limited evaluation of the candidate blocking primers. Quantitative PCR was also rapid, and allowed us to evaluate the degree of amplification in single species samples, but interpretations of melt curves for mixed-template samples can be more subjective (particularly when predator and prey amplicons are more similar than in this study). Finally, DNA metabarcoding provided a substantially more detailed view of performance in mock communities and wild-caught sea lamprey samples. By including multiple methods of analysis in this study, stronger conclusions about the efficacy of these blocking primers could be drawn, therefore bolstering arguments for their use in future molecular diet analysis in sea lamprey.

This study drew from previous work by Johnson et al. (2021) and their research on applying DNA metabarcoding to sea lamprey fecal samples, but made improvements to field applications of this technique. In particular, this study assessed the use of a universal 12S rRNA vertebrate primer. In Johnson et al. (2021), three marker designs were used to specifically target either the Salmonidae, Cyprinidae, or Catostomidae families. While the specificity can reduce amplification of predator species, this ultimately reduces the capability to detect rare host species from other families. Successful amplification of host species with a universal vertebrate primer, as shown in this study, provides a basis for molecular diet analysis in larger, lake-wide studies of sea lamprey dietary compositions with the ability to detect rare taxa. Another previous study on Arctic lamprey diets also used universal vertebrate primers (Shink et al. 2019), emphasizing the benefit of this approach for broader improvements to dietary assessment across lamprey species and ranges. While sea lamprey are invasive within the Great Lakes, they are threatened by various environmental stressors within their native range in the North Atlantic (Guo et al. 2017). Additionally, the focus on sea lamprey control within the Great Lakes has shifted attention away from native lamprey species within the region and their conservation needs (Lucas et al. 2021).

Alongside sea lamprey, the reference database used in this study also contained 12S rRNA gene sequences for three Great Lakes native lamprey: American brook lamprey, northern brook lamprey, and silver lamprey. Base pair differences at 12S between native lamprey and sea lamprey were minimal, with American brook lamprey having a single nucleotide difference and both other native lamprey having no differences within the blocking primer annealing region. This taxonomic similarity implies these blocking primers may be applicable within dietary assessments for conservation purposes, both with other parasitic native lamprey species (i.e., silver and chestnut lamprey) and with sea lamprey across their native range.

While all blocking primers were deemed successful in the DNA metabarcoding dataset, the CA14 wild-caught adult sample allowed for a magnified look at the effectiveness of all eight blocking primers, where only two were able to amplify species-specific host sequences within the sample. This adult was captured during upstream spawning migration and thus was unlikely to have fed recently at the time of capture as the juvenile feeding stage of its life cycle had concluded (Beamish 1980). As such, it is likely that lower amounts of dietary DNA were present, resulting in the noticeably higher sea lamprey-to-host sequence read ratio in the unblocked sample. When blockers were then included in the CA14 sample, only two were able to generate species-specific host reads. These two blocking primers, Blocker 2 and Blocker 6, both contained a C3 spacer and were HPLC purified, with the only difference being the base pair length (36bp and 34bp, respectively). With the ratio of sea lamprey sequences to host sequences from these two blockers in the CA14 sample being fairly comparable, it is likely that either of these two blocking primers will suffice for future analyses that employ this tool for sea lamprey dietary analyses.

An additional point of interest from the CA14 sample is the implication of sea lamprey dietary assessments from migrating adults. These fish are typically captured via traps and barriers, creating a non-selective collection method (Miehls et al. 2020). Implementing DNA metabarcoding to analyze the feeding patterns from adult lamprey captured this way bypasses previous biases from the capture of juvenile sea lamprey, where host fish must first be targeted, captured, and assessed for attached lamprey or wound markings from a previous lamprey attack. Additionally, this would allow for host fish who did not survive a lamprey attack to still be included in analyses, as the DNA from that fish may still be present within the lamprey digestive tract whether the host survived the attack or not.

While the selected mock communities and wild-caught samples in this study allowed for the determination of an effective blocking primer, this study could have benefitted from additional adult samples. With future objectives in mind and the promising implementation of adult sea lamprey dietary assessments, additional adult samples would have bolstered the feasibility of this approach. From a metabarcoding perspective, juvenile lamprey that have been immediately removed from a captured host and frozen for further analysis may harbor proportionately larger amounts of DNA from that immediate host. Even with a molecular feeding history, hosts which the lamprey fed on previously may escape detection due to this imbalance. Adult samples, in contrast, would not suffer from this potential bias as physical removal from an immediate host is not required. While CA14 did allow for the verification and ultimate determination of an effective blocking primer in this study, it would be beneficial to examine blocking primer efficacy across multiple adult samples to confirm the applicability of this approach.

These results are encouraging for future research into sea lamprey dietary composition such as experimental analyses of sea lamprey in controlled settings and in-situ investigations for compiling a more in-depth look into sea lamprey feeding patterns across the Great Lakes. Nonetheless, additional logistical questions need to be addressed prior to broad application of the molecular diet analysis methods used here and in Johnson et al. (2021). Specifically, experimental studies that explore how environmental variables such as temperature and fasting period impact sequence read count are needed. With a better understanding of the effects of these variables on sequence reads and host detections, DNA metabarcoding would be available to be applied throughout the Great Lakes to improve our understanding of prey preference, damage to important Great Lakes fisheries, and ultimately, sea lamprey control and improvements to restoration of lampreys worldwide.

LITERATURE CITED

Adams, J.V., Jones, M.L., Bence, J.R., 2021. Using simulation to understand annual sea lamprey marking rates on lake trout. J. Gt. Lakes Res., Supplement on Sea Lamprey International Symposium III (SLIS III) 47, S628–S638. https://doi.org/10.1016/j.jglr.2020.08.008

Alberdi, A., Aizpurua, O., Gilbert, M.T.P., Bohmann, K., 2018. Scrutinizing key steps for reliable metabarcoding of environmental samples. Methods Ecol. Evol. 9, 134–147. https://doi.org/10.1111/2041-210X.12849

Beamish FWH. 1980. Biology of the North American Anadromous Sea Lamprey, Petromyzon marinus. Can J Fish Aquat Sci. 37(11):1924–1943. doi:10.1139/f80-233.

Bergstedt, R.A., Schneider, C.P., 1988. Assessment of Sea Lamprey (Petromyzon marinus) Predation by Recovery of Dead Lake Trout (Salvelinus namaycush) from Lake Ontario, 1982–85. Can. J. Fish. Aquat. Sci. 45, 1406–1410. https://doi.org/10.1139/f88-164

Berry, T.E., Osterrieder, S.K., Murray, D.C., Coghlan, M.L., Richardson, A.J., Grealy, A.K., Stat, M., Bejder, L., Bunce, M., 2017. DNA metabarcoding for diet analysis and biodiversity: A case study using the endangered Australian sea lion (Neophoca cinerea). Ecol. Evol. 7, 5435–5453. https://doi.org/10.1002/ece3.3123

Buglione, M., Maselli, V., Rippa, D., de Filippo, G., Trapanese, M., Fulgione, D., 2018. A pilot study on the application of DNA metabarcoding for non-invasive diet analysis in the Italian hare. Mamm. Biol. 88, 31–42. https://doi.org/10.1016/j.mambio.2017.10.010

Coble, D.W., Bruesewitz, R.E., Fratt, T.W., Scheirer, J.W., 1990. Lake Trout, Sea Lampreys, and Overfishing in the Upper Great Lakes: A Review and Reanalysis. Trans. Am. Fish. Soc. 119, 985–995. https://doi.org/10.1577/1548-8659(1990)119<0985:LTSLAO>2.3.CO;2

Deagle, B.E., Tollit, D.J., Jarman, S.N., Hindell, M.A., Trites, A.W., Gales, N.J., 2005. Molecular scatology as a tool to study diet: analysis of prey DNA in scats from captive Steller sea lions. Mol. Ecol. 14, 1831–1842. https://doi.org/10.1111/j.1365-294X.2005.02531.x

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(2):306–323. doi:10.1111/1755-0998.12188.

Ebener, M.P., Bence, J.R., Bergstedt, R.A., Mullett, K.M., 2003. Classifying Sea Lamprey Marks on Great Lakes Lake Trout: Observer Agreement, Evidence on Healing Times between Classes, and Recommendations for Reporting of Marking Statistics. J. Gt. Lakes Res., Sea Lamprey International Symposium (SLIS II) 29, 283–296. https://doi.org/10.1016/S0380-1330(03)70494-8

Egizi, A., Healy, S.P., Fonseca, D.M., 2013. Rapid blood meal scoring in anthropophilic Aedes albopictus and application of PCR blocking to avoid pseudogenes. Infect. Genet. Evol. 16, 122–128. https://doi.org/10.1016/j.meegid.2013.01.008

Ellington, A., Pollard, J.D., 2001. Introduction to the synthesis and purification of oligonucleotides. Curr. Protoc. Nucleic Acid Chem. Appendix 3, Appendix 3C. https://doi.org/10.1002/0471142700.nca03cs00

Firkus, T.J., Murphy, C.A., Adams, J.V., Treska, T.J., Fischer, G., 2021. Assessing the assumptions of classification agreement, accuracy, and predictable healing time of sea lamprey wounds on lake trout. J. Gt. Lakes Res., Supplement on Sea Lamprey International Symposium III (SLIS III) 47, S368–S377. https://doi.org/10.1016/j.jglr.2020.07.016

Hacker, C.E., Hoenig, B.D., Wu, L., Cong, W., Yu, J., Dai, Y., Li, Y., Li, J., Xue, Y., Zhang, Yu, Ji, Y., Cao, H., Li, D., Zhang, Yuguang, Janecka, J.E., 2021. Use of DNA metabarcoding of bird pellets in understanding raptor diet on the Qinghai-Tibetan Plateau of China. Avian Res. 12, 42. https://doi.org/10.1186/s40657-021-00276-3

Happel, A., Rinchard, J., Czesny, S., 2017. Variability in sea lamprey fatty acid profiles indicates a range of host species utilization in Lake Michigan. J. Gt. Lakes Res. 43, 182–188. https://doi.org/10.1016/j.jglr.2016.10.010

Harms-Tuohy, C.A., Schizas, N.V., Appeldoorn, R.S., 2016. Use of DNA metabarcoding for stomach content analysis in the invasive lionfish Pterois volitans in Puerto Rico. Mar. Ecol. Prog. Ser. 558, 181–191. https://doi.org/10.3354/meps11738

Harvey, C.J., Ebener, M.P., White, C.K., 2008. Spatial and Ontogenetic Variability of Sea Lamprey Diets in Lake Superior. J. Gt. Lakes Res. 34, 434–449. https://doi.org/10.3394/0380-1330(2008)34[434:SAOVOS]2.0.CO;2

Heinrich, J.W., Mullett, K.M., Hansen, M.J., Adams, J.V., Klar, G.T., Johnson, D.A., Christie, G.C., Young, R.J., 2003. Sea Lamprey Abundance and Management in Lake Superior, 1957 to 1999. J. Gt. Lakes Res., Sea Lamprey International Symposium (SLIS II) 29, 566–583. https://doi.org/10.1016/S0380-1330(03)70517-6

Homma, C., Inokuchi, D., Nakamura, Y., Uy, W.H., Ohnishi, K., Yamaguchi, H., Adachi, M., 2022. Effectiveness of blocking primers and a peptide nucleic acid (PNA) clamp for 18S metabarcoding dietary analysis of herbivorous fish. PLOS ONE 17, e0266268. https://doi.org/10.1371/journal.pone.0266268

Huggins LG, Koehler AV, Schunack B, Inpankaew T, Traub RJ. 2020. A Host-Specific Blocking Primer Combined with Optimal DNA Extraction Improves the Detection Capability of a Metabarcoding Protocol for Canine Vector-Borne Bacteria. Pathogens. 9(4):258. doi:10.3390/pathogens9040258.

Jakubavičiūtė, E., Bergström, U., Eklöf, J.S., Haenel, Q., Bourlat, S.J., 2017. DNA metabarcoding reveals diverse diet of the three-spined stickleback in a coastal ecosystem. PLOS ONE 12, e0186929. https://doi.org/10.1371/journal.pone.0186929

Johnson, N.S., Lewandoski, S.A., Merkes, C., 2021. Assessment of sea lamprey (Petromyzon marinus) diet using DNA metabarcoding of feces. Ecol. Indic. 125, 107605. https://doi.org/10.1016/j.ecolind.2021.107605

King Jr., E.L., 1980. Classification of Sea Lamprey (Petromyzon marinus) Attack Marks on Great Lakes Lake Trout (Salvelinus namaycush). Can. J. Fish. Aquat. Sci. 37, 1989–2006. https://doi.org/10.1139/f80-240

Koichiro Tamura, Glen Stecher, Daniel Peterson, Alan Filipski, and Sudhir Kumar. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution:30 2725-2729.

Lawrie, A.H., 1970. The Sea Lamprey in the Great Lakes. Trans. Am. Fish. Soc. 99, 766–775. https://doi.org/10.1577/1548-8659(1970)99<766:TSLITG>2.0.CO;2

Leray, M., Agudelo, N., Mills, S.C., Meyer, C.P., 2013. Effectiveness of Annealing Blocking Primers versus Restriction Enzymes for Characterization of Generalist Diets: Unexpected Prey Revealed in the Gut Contents of Two Coral Reef Fish Species. PLOS ONE 8, e58076. https://doi.org/10.1371/journal.pone.0058076

Leray, M., Knowlton, N., 2017. Random sampling causes the low reproducibility of rare eukaryotic OTUs in Illumina COI metabarcoding. PeerJ 5, e3006. https://doi.org/10.7717/peerj.3006

Liu, C., Qi, R.-J., Jiang, J.-Z., Zhang, M.-Q., Wang, J.-Y., 2019. Development of a Blocking Primer to Inhibit the PCR Amplification of the 18S rDNA Sequences of Litopenaeus vannamei and Its Efficacy in Crassostrea hongkongensis. Front. Microbiol. 10, 830. https://doi.org/10.3389/fmicb.2019.00830

Lopes, C.M., De Barba, M., Boyer, F., Mercier, C., Galiano, D., Kubiak, B.B., Maestri, R., da Silva Filho, P.J.S., Gielly, L., Coissac, E., de Freitas, T.R.O., Taberlet, P., 2020. Ecological specialization and niche overlap of subterranean rodents inferred from DNA metabarcoding diet analysis. Mol. Ecol. 29, 3143–3153. https://doi.org/10.1111/mec.15549

McClenaghan, B., Nol, E., Kerr, K.C.R., 2019. DNA metabarcoding reveals the broad and flexible diet of a declining aerial insectivore. The Auk 136, uky003. https://doi.org/10.1093/auk/uky003

Meyer, F.P., Schnick, R.A., 1983. Sea Lamprey Control Techniques: Past, Present, and Future. J. Gt. Lakes Res. 9, 354–358. https://doi.org/10.1016/S0380-1330(83)71906-4

Miehls S, Sullivan P, Twohey M, Barber J, McDonald R. 2020. The future of barriers and trapping methods in the sea lamprey (Petromyzon marinus) control program in the Laurentian Great Lakes. Rev Fish Biol Fish. 30(1):1–24. doi:10.1007/s11160-019-09587-7.

Nelson, E.J.H., Holden, J., Eves, R., Tufts, B., 2017. Comparison of diets for Largemouth and Smallmouth Bass in Eastern Lake Ontario using DNA barcoding and stable isotope analysis. PLOS ONE 12, e0181914. https://doi.org/10.1371/journal.pone.0181914

Pertoldi C, Schmidt JB, Thomsen PM, Nielsen LB, de Jonge N, Iacolina L, Muro F, Nielsen KT, Pagh S, Lauridsen TL, et al. 2021. Comparing DNA metabarcoding with faecal analysis for diet determination of the Eurasian otter (Lutra lutra) in Vejlerne, Denmark. Mammal Res. 66(1):115–122. doi:10.1007/s13364-020-00552-5.

Piñol J, Mir G, Gomez-Polo P, Agustí N. 2015. Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. Mol Ecol Resour. 15(4):819–830. doi:10.1111/1755-0998.12355.

Polz, M.F., Cavanaugh, C.M., 1998. Bias in Template-to-Product Ratios in Multitemplate PCR. Appl. Environ. Microbiol. 64, 3724–3730.

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.

Riaz T, Shehzad W, Viari A, Pompanon F, Taberlet P, Coissac E. 2011. ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. Nucleic Acids Res. 39(21):e145. doi:10.1093/nar/gkr732.

Robeson II, M.S., Khanipov, K., Golovko, G., Wisely, S.M., White, M.D., Bodenchuck, M., Smyser, T.J., Fofanov, Y., Fierer, N., Piaggio, A.J., 2018. Assessing the utility of metabarcoding for diet analyses of the omnivorous wild pig (Sus scrofa). Ecol. Evol. 8, 185–196. https://doi.org/10.1002/ece3.3638

Robinson, K.F., Miehls, S.M., Siefkes, M.J., 2021. Understanding sea lamprey abundances in the Great Lakes prior to broad implementation of sea lamprey control. J. Gt. Lakes Res., Supplement on Sea Lamprey International Symposium III (SLIS III) 47, S328–S334. https://doi.org/10.1016/j.jglr.2021.04.002

Rojahn J, Gleeson DM, Furlan E, Haeusler T, Bylemans J. 2021. Improving the detection of rare native fish species in environmental DNA metabarcoding surveys. Aquat Conserv Mar Freshw Ecosyst. 31(4):990–997. doi:10.1002/aqc.3514.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., Weber, C.F., 2009. Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Appl. Environ. Microbiol. 75, 7537–7541. https://doi.org/10.1128/AEM.01541-09

Shink, K.G., Sutton, T.M., Murphy, J.M., López, J.A., 2019. Utilizing DNA metabarcoding to characterize the diet of marine-phase Arctic lamprey (Lethenteron camtschaticum) in the

eastern Bering Sea. Can. J. Fish. Aquat. Sci. 76, 1993–2002. https://doi.org/10.1139/cjfas-2018-0299

Smith, B.R., Tibbles, J.J., 1980. Sea Lamprey (Petromyzon marinus) in Lakes Huron, Michigan, and Superior: History of Invasion and Control, 1936–78. Can. J. Fish. Aquat. Sci. 37, 1780–1801. https://doi.org/10.1139/f80-222

Su, M., Liu, H., Liang, X., Gui, L., Zhang, J., 2018. Dietary Analysis of Marine Fish Species: Enhancing the Detection of Prey-Specific DNA Sequences via High-Throughput Sequencing Using Blocking Primers. Estuaries Coasts 41, 560–571. https://doi.org/10.1007/s12237-017-0279-1

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

Toju H, Baba YG. 2018. DNA metabarcoding of spiders, insects, and springtails for exploring potential linkage between above- and below-ground food webs. Zool Lett. 4(1):4. doi:10.1186/s40851-018-0088-9.

Vestheim, H., Deagle, B.E., Jarman, S.N., 2011. Application of Blocking Oligonucleotides to Improve Signal-to-Noise Ratio in a PCR, in: Park, D.J. (Ed.), PCR Protocols, Methods in Molecular Biology. Humana Press, Totowa, NJ, pp. 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, 12. https://doi.org/10.1186/1742-9994-5-12

Walrant, A., Loreau, M., 1995. Comparison of Iso-enzyme Electrophoresis and Gut Content Examination for Determining the Natural Diet of the Groundbeetle Species Abax ater (Coleoptera: Carabidae). Entomol. Gen. 253–259. https://doi.org/10.1127/entom.gen/19/1995/253

Yang, L., Tan, Z., Wang, D., Xue, L., Guan, M., Huang, T., Li, R., 2014. Species identification through mitochondrial rRNA genetic analysis. Sci. Rep. 4, 4089. https://doi.org/10.1038/srep04089

APPENDIX A: SUPPLEMENTARY FIGURES AND TABLES

Figure S1.1. Total sequence read counts per operational taxonomic unit (OTU). For this study, only the top four OTUs were used, as they represented over 96% of the total sequence reads and included the known DNA from both mock community samples and single-specie samples. The additional OTUs were likely detections from the wild-caught samples (HP3, HP5, HP15, and CA14) used in this study, as well as possibly being from contamination.

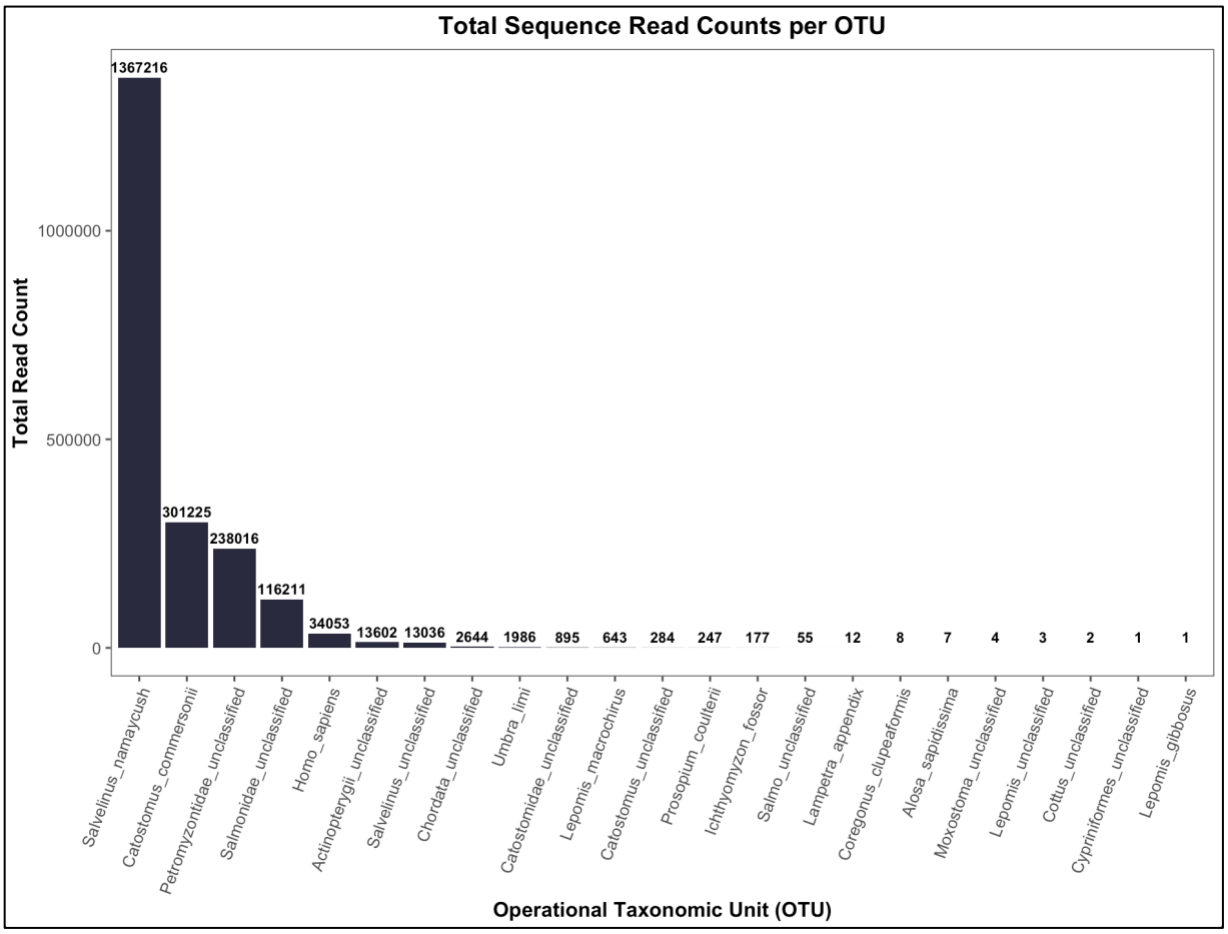


Table S1.1. Total sequence reads for lake trout (top) and sea lamprey (bottom) for all blocking primers and controls with each sample.

Lake Trout Sequence Reads

Sample	Unblocked	Blocker 1	Blocker 2	Blocker 3	Blocker 4	Blocker 5	Blocker 6	Blocker 7	Blocker 8
CA14	1	2	15563	4	2	0	3063	1	0
HP15	891	31764	27026	28104	31264	26392	25667	32604	24249
HP3	715	17312	16679	16338	15124	16742	16468	16457	15026
HP5	22915	28638	33531	32896	32817	35572	30021	33338	32155
M1	17786	33462	32911	27914	29213	27962	34345	34337	32637
M4	10060	32023	32609	28101	29926	31023	31073	32294	31569

Sea Lamprey Sequence Reads

Sample	Unblocked	Blocker 1	Blocker 2	Blocker 3	Blocker 4	Blocker 5	Blocker 6	Blocker 7	Blocker 8
CA14	28238	246	2819	1305	76	785	617	9492	1352
HP15	27068	0	37	1	1	3	1	15	2
HP3	22131	2	7	4	0	10	0	10	7
HP5	6701	1	0	0	0	0	0	0	0
M1	9429	0	0	0	0	0	0	2	0
M4	21633	0	0	4	2	3	1	1	2
M5	20511	0	0	0	0	0	1	4	0

APPENDIX B: SUPPLEMENTARY MATERIALS

Bioinformatic filtering script (mothur):

https://github.com/okaneco1/mothur/blob/main/mothur_batch_script_submission1A.txt

Blocking primer data and analysis scripts (R):

https://github.com/okaneco1/blockingprimers

CHAPTER 2: LONG-TERM DNA RETENTION IN SEA LAMPREY DIGESTIVE TRACTS: INSIGHTS FROM CONTROLLED FEEDING EXPERIMENTS

ABSTRACT

The sea lamprey (*Petromyzon marinus*), a non-native species in the Laurentian Great Lakes, has significantly impacted native fish communities and commercial fisheries, requiring population suppression efforts. While traditional control methods such as lampricides and barriers have reduced sea lamprey population abundance, questions remain regarding sea lamprey dietary composition given the potential for biases in current assessments from wounding observations in commercially important species and uncertainties in host fish mortality estimates. Recent advances in molecular technology offer a promising method of sea lamprey dietary assessment. DNA metabarcoding, which enables species-specific identification of taxonomically diverse prey items from gut contents and/or fecal samples, has proven effective in many taxa, including hematophagous species such as Artic lamprey (*Lethenteron camtschaticum*) and sea lamprey. However, studies on the effects of environmental and dietary factors on DNA retention within the digestive tract are limited, particularly within hematophagous species. Sea lamprey from the field may be exposed to different water temperatures or may have different times since last feeding. To understand the effects these factors may have on DNA retention, I used controlled feeding experiments to investigate the impacts of fasting period and water temperature on the detectability of host DNA within sea lamprey digestive tracts. Additionally, I evaluated the utility of metabarcoding for identifying multiple host species from consecutive feedings. Results indicate that DNA from hosts can be detected for up to 30 days post-feeding, with both sequence read counts and detection probability decreasing with time. Temperature effects on the detection of DNA were dependent upon fasting periods where the decrease in detection probability as a function of days fasting was slower at 15°C than 5°C. Host-switching trials

indicated multiple previous host species could be detected from a single lamprey. Findings provide valuable insights for refining dietary analysis protocols for wild-caught sea lamprey in the Great Lakes.

INTRODUCTION

One of the most devastating invaders of the Laurentian Great Lakes, the sea lamprey (*Petromyzon marinus*), is a hematophagous ectoparasite that has been the subject of international management efforts to control lethal effects on native fish communities, including commercial and sport fisheries throughout the region (Lawrie 1970; Coble et al. 1990). Following initial invasion in the early 20th century, sea lamprey contributed to the ~95% decrease in population size of lake trout (*Salvelinus namaycush)* across the Great Lakes (Smith and Tibbles 1980). Additionally, they played a role in numerical declines and resulting harvest reductions in additional species such as lake whitefish (*Coregonus clupeaformis*), walleye (*Sander vitreus*), and suckers (*Catostomus* spp. and *Moxostoma* spp.) in Lakes Huron, Michigan, and Superior (Smith and Tibbles 1980).

Control efforts since the late 1900s have successfully reduced sea lamprey populations by approximately 90% of their historical peak (Heinrich et al. 2003; Robinson et al. 2021). Numerical declines were largely attributed to barriers that block spawning migration routes and lampricides that target larval sea lamprey in streams (Hrodey et al. 2021; Siefkes et al. 2021). A key consideration for implementing control strategies in the Great Lakes is the life stage at which sea lamprey are targeted (Jones 2007). Sea lampreys begin their life as larvae in stream substrate, where they remain for 3–7 years before undergoing metamorphosis and migrating downstream to open water (Potter 1980). Subsequently, sea lamprey metamorphose to a parasitic juvenile stage and feed on medium- to large-sized fish, often mortally wounding their hosts (Kitchell 1990).

After 12-18 months residence time, sea lamprey sexually mature and leave open water and migrate into tributaries to spawn. During this stage, cessation of feeding occurs and the intestinal tract atrophies prior to gonadal development and spawning (Larsen 1980).

Due to the blood-feeding nature of sea lamprey, there are challenges associated with accurate assessments of sea lamprey diets to estimate damages incurred on the Great Lakes ecosystem. Current dietary analysis and damage estimation methods for sea lamprey utilize host fish that have been captured by fishing vessels (Hume et al. 2021; Treska et al. 2021), allowing for direct observation of attached lamprey or wound markings left by a previous attack (Lantry et al. 2015). However, these methods have limitations due to high host fish mortality (Swink 2003) and the tendency for deceased fish to sink, preventing the capture of any host fish that do not survive an attack (Bergstedt and Schneider 1988). Additionally, observing parasitic lamprey in such large lakes is a difficult task. Further biases could be introduced in the subjectivity of assessing wounding damage and the downstream effect this has on lake-wide damage estimates (Firkus et al. 2021). Variation in sea lamprey diets across the Great Lakes (e.g., Harvey et al. 2008) may violate assumptions underlying estimates of economic injury (Irwin et al. 2012), in which case it may be especially important to prioritize control efforts in places where fish stocks of greatest ecological and economical value are most at threat.

Recent advances in molecular techniques are well-suited for addressing limitations to traditional methods of diet compositional assessment (Waraniak et al. 2019). Molecular diet analysis via DNA metabarcoding is a newer, more comprehensive approach that allows multi-species dietary interrogation and enables species-specific assignments of prey items from DNA extracted from gut content samples (Pompanon et al. 2012). While these approaches require known sequences for potential prey species, extensive DNA sequence data are freely accessible

via online repositories such as GenBank (Sayers et al. 2022; https://www.ncbi.nlm.nih.gov/genbank/) and BOLD (Ratnasingham et al. 2024; https://v3.boldsystems.org/). Particularly noteworthy is the extent of fish species sequence data currently available for potential sea lamprey host species, such as 99% of Great Lakes fish species being represented in mitochondrial cytochrome oxidase I (COI) databases (Trebitz et al. 2015). Previous studies have successfully employed DNA metabarcoding to study the diets of mammals (Berry et al. 2017; Buglione et al. 2018; Lopes et al. 2020), birds (McClenaghan et al. 2019; Hacker et al. 2021), fishes (Berry et al. 2015; Harms-Tuohy et al. 2016; Jakubavičiūtė et al. 2017), and other taxa. This approach has been successful for similar applications in other hematophagous species such as leeches (Drinkwater et al. 2019), mosquitos (Reeves et al. 2018; Estrada-Franco et al. 2020), ticks (Gariepy et al. 2012), vampire bats (Bohmann et al. 2018), and, particularly, Artic lamprey (Shink et al. 2019) and sea lamprey (Johnson et al., 2021). However, broad-scale application of these methods requires a greater understanding of the factors that influence retention time and detectability of host DNA (including temperature and fasting duration).

The amplification of dietary DNA from hematophagous species has received increased attention due to recent approaches using invertebrate-derived DNA (iDNA) to amplify gene regions in vertebrates from the bloodmeals of invertebrates such as leeches (Drinkwater et al. 2019; Fahmy et al. 2019; Nguyen et al. 2021). Early research using a quantitative PCR assay to investigate the feasibility of this method found that amplifiable genetic material could remain within leech dietary systems for up to 4 months (Schnell et al. 2012). However, this timeframe may not directly transfer to dietary DNA retention in fishes, as leeches have notably slow digestive rates and may go 6 months without feeding (Sawyer 1986). While many studies have

used DNA metabarcoding to analyze the diets of fishes, including hematophagous vampire catfishes (family Vandellinae; Bonato et al. 2022), dietary composition has been the primary focus and data on the retention time of DNA within fish digestive tracts is lacking. Given the variability in degradation time of DNA in different environments (Levy-Booth et al. 2007), including fish DNA (Turner et al. 2015), understanding the persistence of host fish DNA within sea lamprey digestive tracts would better inform future field studies using this technique.

Our goal was to evaluate how fasting period and water temperature influenced DNA degradation rates in sea lamprey digestive tracts, while also examining the detectability and composition of host DNA following consecutive host feedings. This study addresses this goal through feeding trials with newly-transformed parasitic-stage sea lampreys under controlled experimental conditions. Detection of DNA may vary across temperatures given the positive influence of increased environmental temperatures on metabolic rates and subsequently higher DNA degradability (Pilliod et al. 2014). Given this, I predicted that DNA detectability in sea lamprey digestive tracts would decrease in warmer temperature groups. Furthermore, I expected longer fasting periods to allow for more DNA to pass through the digestive system, thus lowering detection rates of host DNA within these groups.

METHODS

Sea Lamprey Collections

Out-migrating recently metamorphosed (herein termed transformer) sea lamprey were collected by staff of the Great Lakes Indian Fish and Wildlife Commission and U.S. Fish and Wildlife Service and then transported to U.S. Geological Survey Hammond Bay Biological Station located in Millersburg, Michigan. Transformers were collected from Furlong Creek,

Michigan in 2022 for temperature and fasting period feeding trials, and from Marengo River, Wisconsin in 2023 for host-switching feeding trials.

Transformers were collected using drift nets during the fall while migrating downstream to the Great Lakes where they begin to feed on host fish. This life-stage was targeted to ensure that lamprey would both readily attach to a host fish in experimental settings and have no prior parasitic feeding history. All transformers collected in this fashion were maintained in aquaria in the laboratory until feeding trials began.

Temperature and Fasting Period Feeding Trials

A rack system of 18 replicate 200 L temperature-controlled tanks (0.9 m x 0.6 m x 0.5m) was established at the Hammond Bay Biological Station (USGS) to house transformers and host fish (lake trout; sourced from Sullivan Creek National Fish Hatchery). Length (mm) and weight (g) measurements of all lake trout were collected prior to their introduction into the experimental aquaria. Transformers were then weighed and measured, and a single sea lamprey and lake trout were added to each of the tanks. Each sea lamprey was allowed up to 10 days to attach and feed on the lake trout. Attachment status was noted daily to quantify duration of attachment, with the majority feeding 6-8 days. Aquaria were maintained at temperatures of either 5, 10, or 15°C. This range represents the conditions that parasitic sea lamprey likely experience in the Great Lakes and the range of temperatures tested in previous experiments (Farmer et al. 1977). After seven days feeding time, the host fish were removed from the tank.

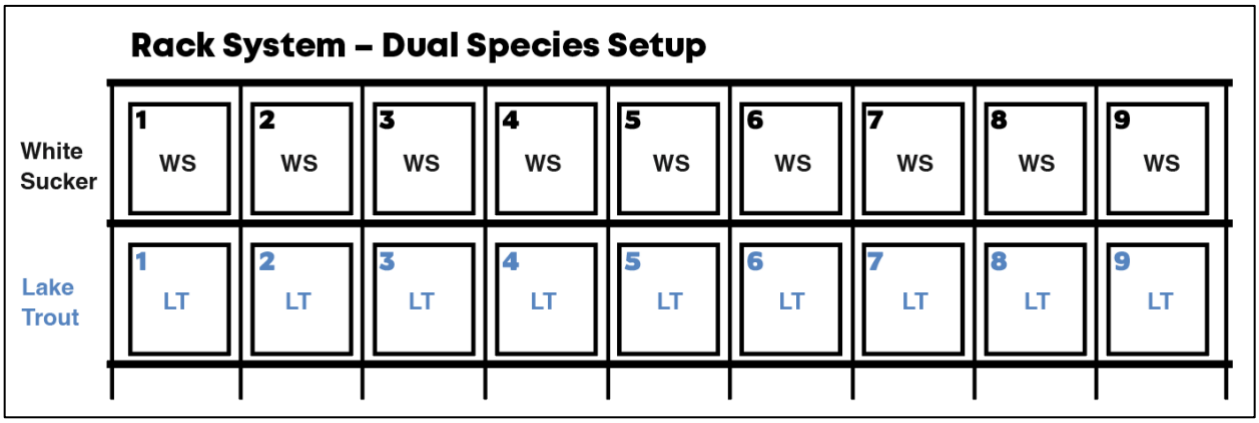
Upon removal of hosts, sea lamprey were measured and weighed and assigned to a fasting period category of either 0, 5, 10, 20, or 30 days. The process was repeated with another sea lamprey and lake trout being added to each of the aquaria until all sea lamprey had been given the opportunity to feed on a lake trout. At least five replicate sea lamprey per fasting

period and temperature combination were included, amounting to a total of 77 sea lamprey. Each of the lamprey was externally tagged in the dorsal fin with a polyethylene streamer tag (Hallprint PST12P) for identification prior to fasting. Once each lamprey completed their respective fasting periods, the sea lamprey was euthanized with an excess of tricaine methanesulfonate (MS-222) and a final length and weight measurement was taken. Sea lamprey were then frozen whole and stored at -80°C for subsequent analyses. Temperature and fasting period experiments were conducted from February to April in 2022. All animal handling activities associated with these experiments followed protocols approved by the Michigan State University's Institutional Animal Care and Use Committee (IACUC Protocol: PROTO202100285).

Host-Switching Feeding Trials

Host-switching experiments replicated many of the conditions and methods used in the temperature and fasting period feeding trials described above. Host-switching feeding trials took place at Hammond Bay Biological Station, using a naïve set of 67 transformers. In the same 18 temperature-controlled replicate aquaria, with nine lake trout (sourced from Sullivan Creek National Fish Hatchery) and nine white sucker (*Catostomus commersonii*; sourced from Michigan Wholesale Bait) evenly split with a single host fish per tank (Figure 2.1). All host fish were weighed and measured prior to introduction to experimental aquaria. Transformer sea lamprey were also weighed, measured, and placed into aquaria with each tank receiving a single lamprey. All aquaria maintained a temperature of 10°C and lamprey were allowed to feed on the host (lake trout or white sucker in this case) for up to 10 days. Visual observations were made daily to assess attachment and feeding.

Figure 2.1. Flow-through replicate tanks (200 L) used for feeding trials pairing one host species (LT = lake trout, WS = white sucker) and one recently metamorphosed sea lamprey (transformer) per tank.



Following the feeding period, host fish were swapped so that tanks with white sucker were replaced with a lake trout and tanks with lake trout were replaced with a white sucker. Each sea lamprey was then allowed to feed on the second, complementary host fish for an equivalent amount of time. Visual assessments of attachment to monitor feeding were performed daily. Between the two feeding periods, each sea lamprey was weighed and measured as proxy for feeding on the first host. Upon completion of the second feeding period, sea lamprey were removed, weighed, and grouped according to a fasting period of either 0, 5, 10, 20, 30, or 45 days. Sea lamprey were tagged with visible implant elastomer (VIE) tags for identification (see Hume et al. 2024). When lamprey completed their designated fasting times, final length and weight measurements were taken and sea lamprey were euthanized with an excess of MS-222, frozen whole, and stored at -80°C. Host-switching feeding experiments were conducted from January to March in 2023. Experiments were approved through animal handling protocols (PROTO202100285).

DNA Extraction and Amplification

After sea lamprey were thawed, the digestive tract was removed and dissected. Each sea lamprey and corresponding digestive tract was photographed before and after dissection. The intestinal tract was dissected and opened with a sterile scalpel (No. 11 blade, WSI disposable sterile). A sterile cotton swab (Dynarex 6-inch sterile cotton tipped applicator) was run along the inside length of the digestive tract, collecting as much digestive material as possible. The swab was then preserved in RNAlater (Thermo Fisher Scientific, Waltham, MA) in a 1.5 mL microcentrifuge tube, labelled, and stored at -80°C. After all sea lamprey digestive samples had been collected, tubes were sent to the Molecular Ecology Lab at Michigan State University for DNA extraction, amplification, and sequencing.

DNA extractions were performed with gMax Mini Genomic DNA Kit (IBI Scientific, Dubuque, IA), following Johnson et al. (2021). All samples were then amplified via polymerase chain reaction (PCR) using vertebrate-specific 12S-V5 primer (Riaz et al. 2011) targeting a ~140bp region of the 12S mitochondrial rRNA gene. Reactions also included a lamprey-specific blocking primer (Blocker 6), designed to suppress amplification of the 12S gene region in sea lamprey (Chapter 1). Two technical replicate PCR were included for each digestive sample.

PCR was performed in a 15 uL reaction volume with 1.5 µL of 10X AmpliTaq Gold PCR Buffer II (Thermo Fisher Scientific, Waltham, MA), 0.36 µL of dNTPs (10mM), 1.2 µL of $MgCl_2$ (25mM), 0.75 µL of BSA (20 mg/mL), 0.8 µL of forward and reverse primers (10 µM) and sea lamprey Blocker 6 blocking primer (100 µM), 6.54 µL of Millipore water (UV treated), 0.25 µL of AmpliTaq Gold (5U/µL) and 2 µL of template DNA. A negative control for each PCR plate was included and replaced DNA template with water. Thermal conditions for PCR were as follows: 10 min at 95°C (1x); 30 s at 95°C, 30 s at 57°C, 45 s at 72°C (40x); and 5 min

at 72°C (1x). Following amplification, 4 µL of PCR product and 2.5 µL of glycerol loading dye was run on a 1% agarose gel with GelRed stain (0.5x, Biotium, Fremont, CA). Gels were photographed under UV light in a Labnet Enduro GDS II imaging system (Labnet HQ, Edison, NJ) to ensure amplification.

DNA Metabarcoding and Bioinformatics

Dual index barcoding was used for all samples to allow demultiplexing of sequencing reads to their corresponding samples. The indexing PCR was conducted in a final volume of 10 uL, including 2X Qiagen Plus master mix (5 uL), i7 (1uL; 10 uM) and i5 (2uL; 5 uM) index priers, and 2 uL of template. Reaction conditions in the indexing PCR were as follows: 15 min at 95°C (1x); 10s at 95°C, 30s at 65°C, 30s at 72°C (10x); and 5min at 72°C. Dual-indexed PCR products were then pooled and bead size selected at 0.5x and 1.2x bead concentrations to remove long and short sequences, respectively. Pooled libraries were diluted and analyzed via TapeStation (assay High Sensitivity D100 ScreenTape) for sequence length confirmation before being sequenced on a 300 cycle Illumina MiSeq lane (v2 Standard; 2 x 150bp paired end) at Michigan State University's Research Technology Support Facility.

Raw sequencing data were processed using computational resources provided by the Michigan State University High-Performance Computing Center and the software package mothur (version 1.48.0; Schloss et al. 2009). A maximum of two mismatches were allowed between the barcodes and sample sequences during demultiplexing to increase tolerance for minor sequencing errors. Overlapping sections were trimmed to minimize discrepancies. Initial screening removed sequences with ambiguities and identified unique sequences. These sequences were then aligned to a reference database containing 220 sequences ranging from 89 to 107bp from the 12S mitochondrial rRNA gene region for 149 Great Lakes fish species.

Sequences outside of 85-110 bp or with more than seven homopolymers were removed, and the remaining sequences were clustered with zero differences allowed. Chimeric sequences were removed and taxonomic classification was assigned with the mothur default confidence score cutoff of 80% or higher. A distance matrix was then calculated with a cutoff of 0.03, followed by clustering with a 0.01 cutoff. A 99% similarity threshold was used to then group sequences into OTUs. This output was then cleaned and organized into a community matrix table with read counts for every OTU from each sample using R/Rstudio (v.2023.12.1+402).

Data Analysis and Statistics: Temperature and Fasting Period Trials

To determine if the weight gained (surrogate for food consumed) by sea lamprey differed across the water temperatures tested (5, 10, 15°C) an ANOVA with a post-hoc Tukey's HSD was used. I also assessed the correlation between weight gain and the duration of attachment, using the cor.test function in the *stats* R package.

Intra-specific genetic differences can result in sequence variation that cannot be fully captured by a single reference database; however, given all sea lamprey had fed only on lake trout, three OTUs with substantial representation in the sequencing data (*Salvelinus namaycush, Salvelinus* unclassified*,* and Salmonidae unclassified) were combined to represent total lake trout read counts. These separate OTUs likely occur due to minor sequencing errors that make sequences indistinguishable from other similar sequences. This issue, often due to high intraspecific variation or low interspecific variation, can result in classifications being assigned at higher taxonomic levels. To evaluate the effects of temperature and fasting period on the number of lake trout reads, linear, Poisson, negative binomial, and zero-inflated negative binomial models with predictor variables of weight gain, temperature, fasting period, and days attached (visual observation) to the host were compared against one another. These four models

were selected given certain characteristics of the data such as the presence of excess zeros and overdispersion. Results from this initial model selection with a global model including all predictor variables suggested that the negative binomial GLM provided the best fit to the data, so it was used in subsequent modeling analyses. Read counts were modeled against each predictor variable, both individually and in combination, and assessed with Akaike information criterion (AIC; Akaike 1973) for model fit using the dredge function from the *MuMIn* package in R (Kamil Bartoń 2010). Prior to fitting the models, read counts were averaged from both replicates, then rounded to the nearest whole number. This allowed for the use of data from both replicates in the same model.

Read count data from individual replicates were then translated into binary detection data. Two criteria were used to indicate a detection of a host species within a sample: 1) a minimum of 10 sequence reads of that host in the sample and 2) a relative read abundance (RRA) of that host in the sample above 1%. Positive detections were limited to samples that met both criteria. RRA was determined as the proportion of reads attributed to that host over the total number of reads in the sample (excluding human reads). These thresholds are derived from Drake et al. (2022) and are purposefully stringent. Given that all lamprey were fed known host fish, there is no investigation for rare taxa within samples and removal of true positives with smaller read counts is less likely. These thresholds serve to ultimately limit false positives that can occur from issues such as contamination and index hopping. All no-DNA control samples registered zero read counts for host fish and thus were not included in the determination of detection thresholds.

An occupancy modeling framework (MacKenzie et al. 2002) was used to estimate the detection probability of host fish in sea lamprey digestive tracts. Occupancy models, traditionally

used in ecological observation studies to account for imperfect detection, were adapted here for dietary DNA data. In our study, "occupancy" represents the presence of host fish DNA within a sea lamprey dietary tract, and "observations" consisted of the two technical PCR replicates to act as repeated detection attempts of each digestive sample. To assess the probability of detection as a function of ecological and biological variables, fasting period, water temperature, weight gain, and total days attached to the host were incorporated as covariates and each combination was compared against a null model. These variables were only added as covariates within the "observation" model. All models were evaluated using AIC, and the best fit model was selected for further interpretation. Analyses were conducted using the *unmarked* package (Fiske and Chandler 2011) in R, and detection probability was plotted as a function of the covariates from the best fit model.

Data Analysis and Statistics: Host-Switching Trials

To evaluate whether a sea lamprey's first host or second host had, on average, higher read counts, an ANOVA was performed to determine if the relationship between read count and host order was significant. Read counts were averaged between PCR technical replicates for both hosts to prevent pseudoreplication and sample size inflation.

As these trials were focused on the ability to detect a feeding history through the presence of host fish DNA in the first host, the average read counts between replicates of only the first host were used as the primary dependent variable. Independent variables included relative weight gain on host 1 compared to host 2, the species of host 1 (lake trout or white sucker), days attached to host 1, fasting period since detachment from both hosts, and the interaction between these factors. Sea lamprey that did not consecutively feed on both hosts, as inferred by weight gain during both feeding periods, were removed from the analysis. Similar to the temperature

and fasting period trials, a two-phase model selection approach was used. First, a linear, Poisson, negative binomial, and zero-inflated negative binomial model were fit to the read count data using all covariates and compared via AIC. Upon determination of the best-fit model, all combinations of predictor variables were included and compared via AIC to identify the set of predictor variables that best fit the data.

Read counts for the first host were converted into binary detections using the same thresholds as previous temperature and feeding trials. To evaluate the relationship between various ecological and biological factors on the likelihood of detecting DNA from host 1 in each sea lamprey digestive sample, an occupancy modelling framework was again set up using both PCR replicates as observations, treating positive detections as indicators of "occupancy." Covariates in this model included number of days attached to host 1, relative weight gain on host 1, fasting period, and host 1 species. All combinations of these variables, including a null model, were evaluated and compared to determine the best-fit model using AIC.

RESULTS

Temperature and Fasting Period Trials

Weight gain in sea lamprey ($n = 77$) from feeding on lake trout varied significantly ($p < 0.001$, $df = 2$, $F = 31.28$) across temperature groups as shown through an ANOVA (Figure 2.2). Specifically, sea lamprey gained significantly more weight under experimental feeding trial conditions of 15°C as compared to both 10°C and 5°C feeding conditions (Tukey's HSD; $p < 0.001$). Weight gain of sea lamprey did not differ when feeding at either 5°C or 10°C ($p = 0.625$). The number of days lamprey spent attached to their host did not correlate with weight gain ($p = 0.43$, $r = 0.09$) (Figure 2.3). Given the lack of correlation between these variables, both were included as covariates in linear modeling analyses.

Figure 2.2. Weight gain (g) of recently transformed sea lamprey from feeding on lake trout for up to 10 days as a function of experimental tank temperature (°C) group. Each point represents an individual, boxes represent the interquartile range (IQR), the heavy line in the box shows the median, and whiskers extend to the range, excluding outliers (points that exceed the IQR by 1.5 times).

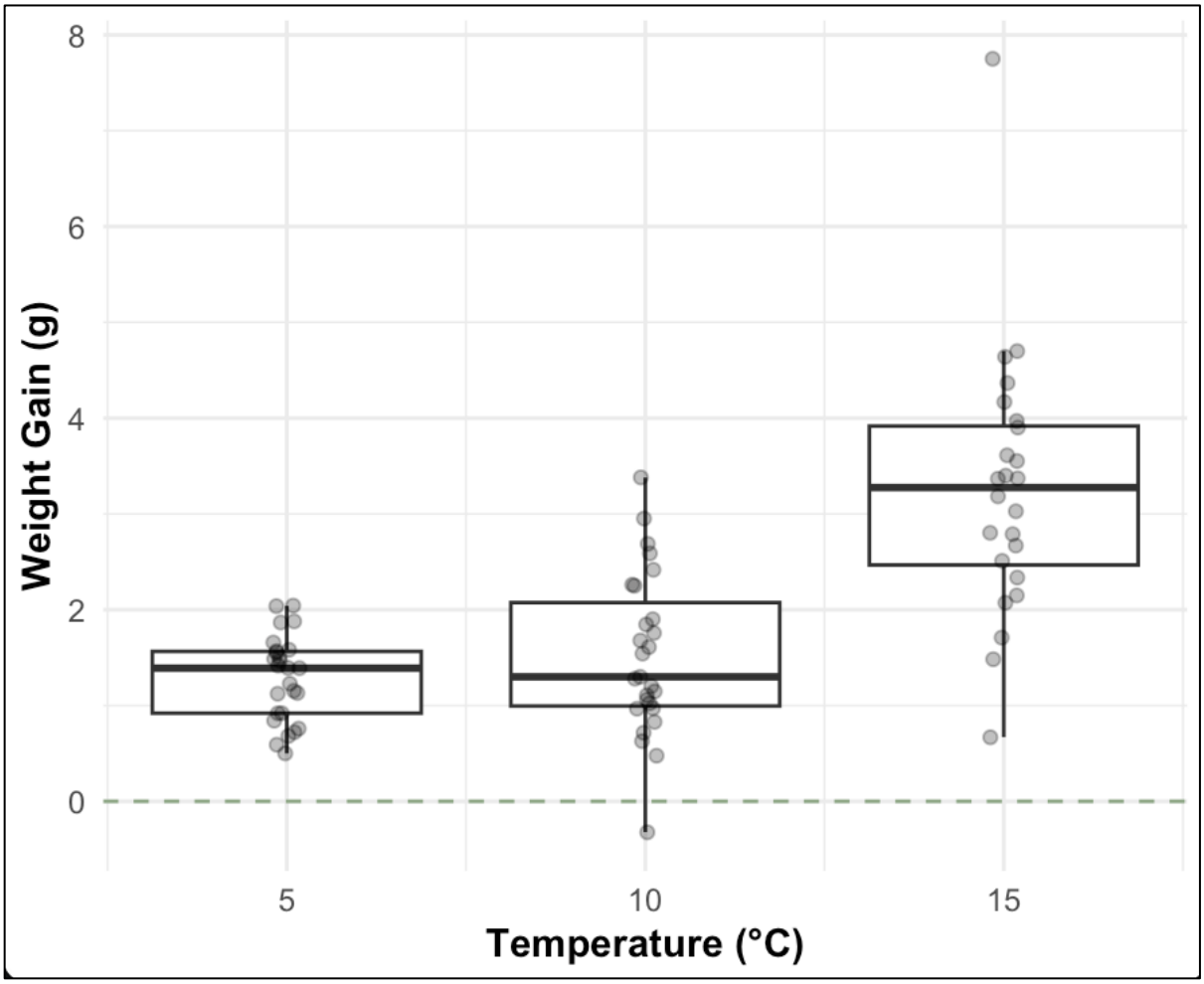
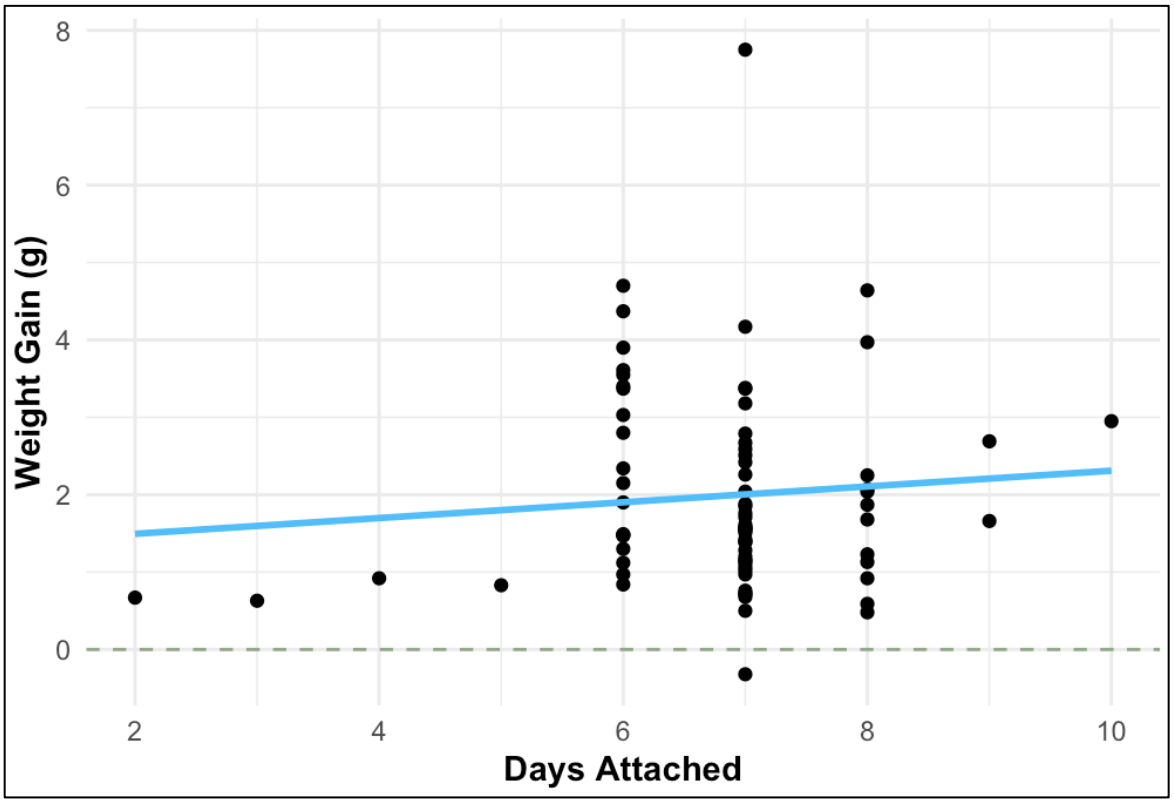


Figure 2.3. Weight gain (g) of recently transformed sea lamprey during lake trout feeding trials as a function of feeding exposure (days attached). Points represent individual sea lamprey, with a line of best linear fit (blue) to indicate the overall trend.



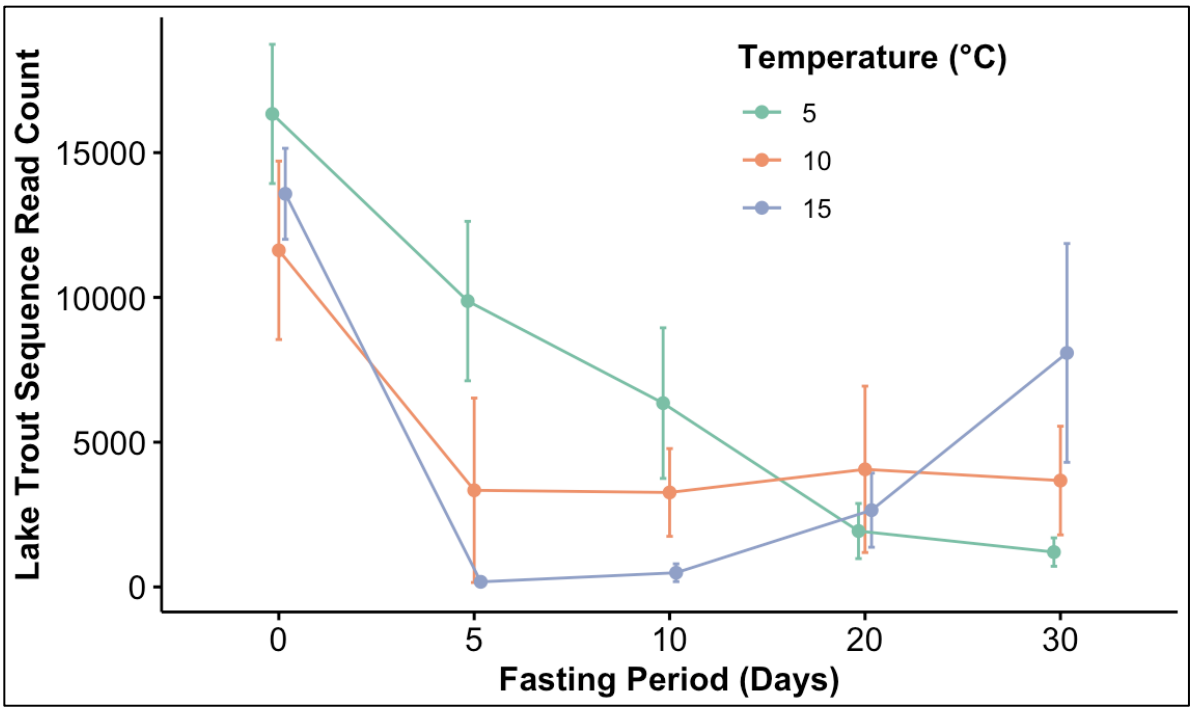
A total of 1,600,781 sequence reads were collected across 77 sea lamprey digestive samples (APPENDIX A: Figure S2.1). Read count model comparisons using AIC determined the negative binomial model was the best fit for the data. Further assessments of a negative binomial model with various predictor variable combinations indicated the model that included fasting period, temperature, and their interaction provided the best fit to the read count data (Table 2.1). Both fasting period and the interaction between fasting period and temperature exhibited statistically significant effects ($p < 0.05$). However, temperature alone did not have a statistically significant effect ($p = 0.13$), indicating that temperature influences sequence read counts primarily through its interaction with fasting period, rather than as an independent factor. Each additional fasting day was associated with an estimated 11.2% decrease in lake trout reads.

However, the positive coefficient on the interaction term ($\beta = 0.008$) suggests that this decline in read count slows at higher temperatures. This pattern can be further investigated by analyzing individual temperature groups (Figure 2.4). At 5°C, read count declined across fasting time more consistently. However, for both 10°C and 15°C, read count declines sharply after five days and remained low for the remainder of the fasting period, with the exception of an increase in read count after 30 days fasting at 15°C.

Table 2.1. Model selection table for negative binomial models comparing results of different combinations of predictor variables (days attached to host, fasting period, temperature, weight gain, and fasting period:temperature interaction) as a function of host DNA sequence read count. Models are listed in order of best fit from top to bottom, and the null model is outlined. AIC calculations indicate the model including an interaction between fasting period and temperature was best fit.

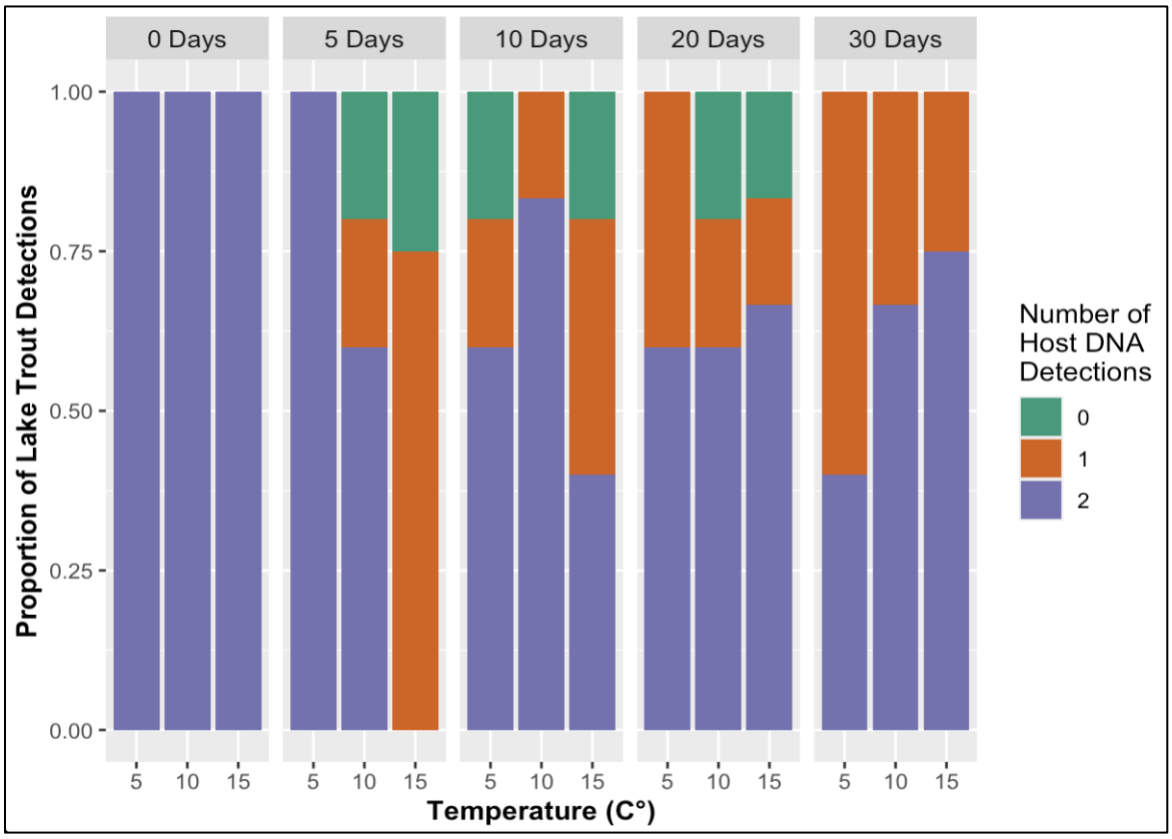
Model	df	AIC	Delta_AIC
~temp * fasting_period	5	1436.29	0.00
~temp * fasting_period + days_attached	6	1436.38	0.10
~weight_gain + temp * fasting_period	6	1436.78	0.50
~fasting_period	3	1437.21	0.93
~weight_gain + temp + fasting_period + days_attached	7	1437.32	1.03
~weight_gain + temp * fasting_period + days_attached	7	1437.32	1.03
~1	2	1441.79	5.50
~temp	3	1442.81	6.53
~days_attached	3	1443.55	7.27
~weight_gain	3	1443.78	7.50

Figure 2.4. Lake trout sequence read count described as a function of fasting period (days) and temperature (°C). Points represent the average number of lake trout reads within sea lamprey digestive samples at each fasting period/temperature combination, with whiskers indicating the range of read counts. Water temperature (°C) indicated by line color.



Host detection data indicated that across sea lamprey from all temperature and fasting period groups (n = 77), lake trout was not detected in either replicate in six samples while lake trout was only detected in a single replicate in 18 samples. Lake trout detections were highest in the zero-day fasting period group, with detections across both replicates for all samples at all three temperatures (Figure 2.5). Fasting periods of 5, 10, and 15 days each had multiple samples where lake trout was not detected in either sample, but at 30 days of fasting lake trout was detected in all samples for at least one replicate.

Figure 2.5. Proportion of host DNA (lake trout) detections between both technical PCR replicates grouped by fasting period (column headings) and temperature (°C). The combined number of host DNA detections between both replicates is indicated by color with green indicating no detections in either replicate, orange indicating a single detection in at least one replicate, and purple indicating detections in both replicates. Each bar indicates a group of transformer sea lamprey that were fed lake trout at a certain temperature (5, 10, or 15°C) and allowed to fast for a certain number of days (0, 5, 10, 20, or 30 days).



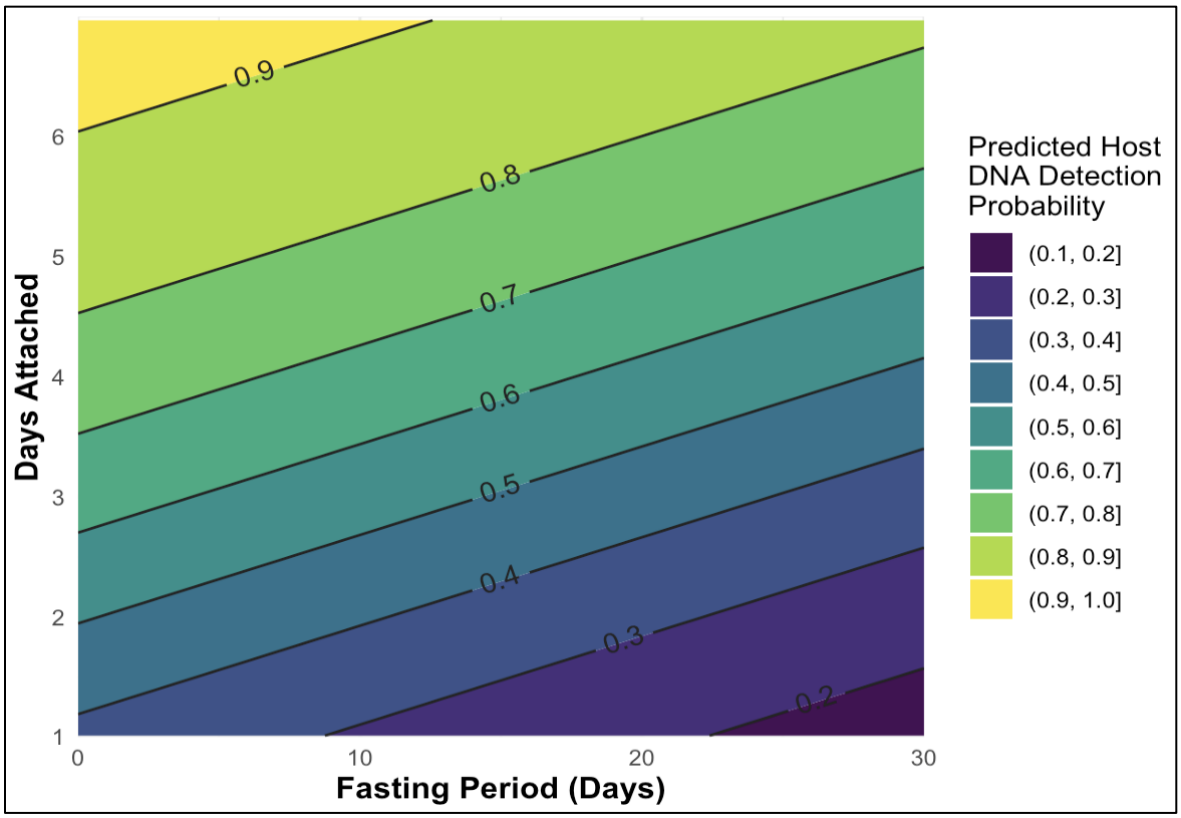
Occupancy model selection indicated several models with different variable combinations that provided a better fit than the null model (Table 2.2). Overall, the best fit model included both days attached and fasting period (AIC = 146.08). Temperature and weight gain inclusions in some models indicated slight improvements, but only when days attached and/or fasting period were included as well. A contour plot showing predicted host detection probability given days attached and fasting period showed slight decreases in the probability of detecting a host as fasting period increased and as days attached decreased (Figure 2.6). Estimated detection

probabilities were above 90% after zero days of fasting, given attachment for at least 6 days, and above 80% after 30 days of fasting given at least 6.8 days of attachment. An increase in days attached from one day to two days can increase the probability of host detection between 44.3% at zero fasting days and 59.9% at 30 fasting days. Alternatively, an increase in fasting period from zero days to one day decreases detection probability between 0.2% at seven days attached and 2.9% at one day attached.

Table 2.2. Results from occupancy model selection for trials comparing the relationship between the number of host fish sequence reads in sea lamprey digestive samples with temperature, fasting period, days attached to host, and weight gain as independent variables. Both fasting period and number of days of host attachment were included in the model of best fit. Models are listed in order of best fit from top to bottom, and the null model is outlined.

Model	nPars	AIC	delta	AICwt
~fasting_period + days_attached	4	146.083	0.000	0.386
~fasting_period + days_attached + weight_gain	5	147.233	1.150	0.217
~days_attached + weight_gain	4	148.549	2.466	0.113
~temp + fasting_period + days_attached	6	149.401	3.318	0.074
~fasting_period	3	150.299	4.216	0.047
~temp + fasting_period + days_attached + weight_gain	7	150.751	4.668	0.037
~1	2	151.494	5.411	0.026
~fasting_period + weight_gain	4	151.805	5.722	0.022
~temp + days_attached + weight_gain ~ 1	6	152.263	6.180	0.018
~temp * fasting_period	7	152.377	6.294	0.017
~temp + fasting_period	5	152.683	6.600	0.014
~weight_gain	3	153.338	7.255	0.010
~temp	4	153.813	7.730	0.008
~temp + fasting_period + weight_gain	6	154.672	8.589	0.005
~temp + weight_gain	5	155.468	9.385	0.004
~days_attached	3	157.240	11.157	0.001
~temp + days_attached	5	159.192	13.109	0.001

Figure 2.6. Predicted host DNA detection probability contour plot with the number of days a sea lamprey is attached to its host against fasting period (days). Contour lines show probability ranges in intervals of 0.1 with purple to yellow indicating lower to higher detection probability, respectively.



Host Switching Trials

A total of 1,206,927 sequence reads averaged between both PCR replicates were collected across 67 sea lamprey digestive samples (Figure S2.1). Read counts for the two host species (averaged across replicates) were compared via an ANOVA for sea lamprey that fed on both hosts (n = 37), indicating significantly higher read count averages in host 2 over host 1 (p = 0.015). Mean read count for the first host among all lamprey was 2705 (sd = 4460), while mean read count for the second host was 5568 (sd = 6341). A linear, Poisson, negative binomial, and zero-inflated negative binomial model were then compared via AIC, with the negative binomial showing the highest fit. Further model selection using AICc comparing all combinations of

predictor variables found the model with an interaction between the fasting periods since feeding on the two hosts to be as good of a fit as the null model (AIC = 611.5 for both), suggesting that relative weight gain on the first host, species of the first host, fasting days, and days attached to the first host did not significantly improve the model (Table 2.3).

Table 2.3. Model selection table for negative binomial models comparing results of different combinations of predictor variables (days attached to host 1, fasting days since host 2, fasting days since host 1, host 1 species, relative weight gain, and the interaction of fasting days since host 1 and host 2) as a function of host 1 DNA detections. Models are listed in order of best fit from top to bottom, and the null model is outlined. AICc calculations indicated no substantial improvements from any combination of predictors over the null model (outlined).

Model	DF	AICc	Delta	Weight
~fasting_days_host2 * fasting_days_host1	4	611.46	0	0.16
~1	2	611.48	0.02	0.16
~days_attached_host1	3	612.14	0.68	0.12
~fasting_days_host1	3	613.72	2.26	0.05
~fasting_days_host2	3	613.72	2.26	0.05
~relative_weight_gain	3	613.75	2.29	0.05
~host1_species	3	613.76	2.3	0.05
~days_attached_host1 + host1_species	4	614.41	2.95	0.04
~days_attached_host1 + fasting_days_host2	4	614.51	3.05	0.04
~days_attached_host1 + fasting_days_host1	4	614.51	3.05	0.04
~fasting_days_host2 + relative_weight_gain	4	616.09	4.63	0.02
~fasting_days_host1 + relative_weight_gain	4	616.09	4.63	0.02
~fasting_days_host1 + host1_species	4	616.11	4.65	0.02
~fasting_days_host2 + host1_species	4	616.11	4.65	0.02
~host1_species + relative_weight_gain	4	616.15	4.69	0.02
~days_attached_host1 + fasting_days_host1 + host1_species	5	616.9	5.44	0.01
~days_attached_host1 + fasting_days_host2 + host1_species	5	616.9	5.44	0.01
~fasting_days_host1 + host1_species + relative_weight_gain	5	618.59	7.13	0
~fasting_days_host2 + host1_species + relative_weight_gain	5	618.59	7.13	0

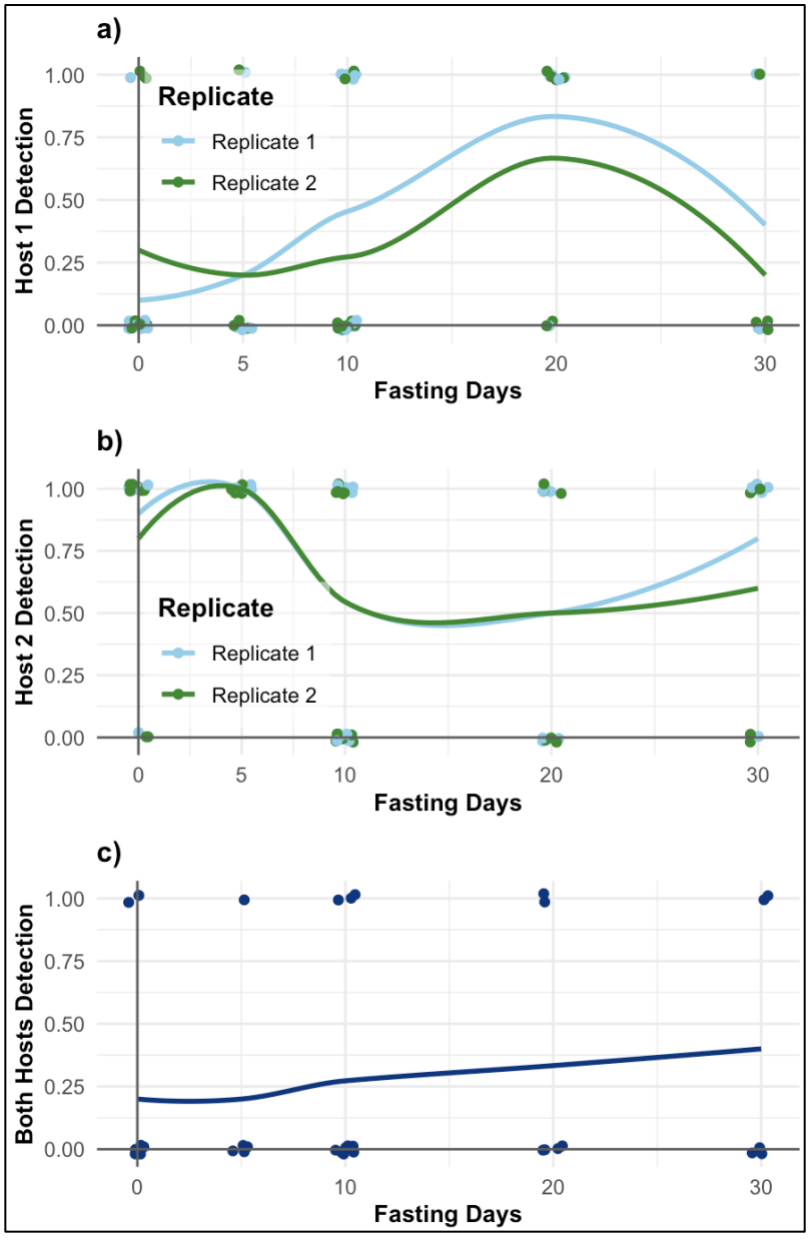
As with the previous experiment, read counts were translated into detections when total read count for a host exceeded 10 reads and the RRA of that host was at least 1% of total sample read count. All negative samples had zero total read counts, and thus were not considered for the

establishment of detection thresholds. Among all lamprey that fed on both host fish (n = 37 with positive weight gain during both feeding periods), 43.2% registered DNA detections of the first host in at least one replicate for all fasting periods. Both 30-day and 45-day fasting periods also registered detections of host 1, with two detections (n = 5) and one detection (n = 5), respectively (Table 2.4). Besides the single 15-day fasting lamprey, the highest fasting proportion of detections occurred during the 20-day fasting period group with six total detections (n = 7). This increase in host 1 detections at the 20-day fasting mark was reflected with a decrease in host 2 detections between 10 and 20 days (Figure 2.7).

Table 2.4. Summary table of host 1 detections in multi-species trials where two hosts were fed on categorized by fasting day. Total detections indicates the number of positive host DNA detections for samples at each fasting period category. Detection proportion indicates the percentage of digestive samples at each fasting day category that registered host DNA detections.

Fasting Days	Total Detections	Samples	Detection Proportion
0	3	10	0.30
5	1	5	0.20
10	5	11	0.46
20	5	6	0.83
30	2	5	0.40

Figure 2.7. Charts a and b show DNA detection outcomes for host 1 and host 2, respectively, across different fasting days. Each point represents either a host detection (1) or non-detection (0) from either PCR replicate, with blue indicating Replicate 1 and green indicating Replicate 2. LOESS curves are used to illustrate trends in detection likelihood over time. Chart **c** shows detection outcomes for both hosts, where a positive detection (1) indicates the presence of DNA from both hosts in either replicate. LOESS curves were again used to illustrate the change in the likelihood of detecting both hosts across two replicates over time.



Among all considered models for the host 1 detection occupancy model, no models indicated relationships that were a better fit than the null model as determined by AIC (Table

2.5). Univariate models with days attached, relative weight gain, and fasting days indicated the closest fit to the null model with delta AIC values < 1.00. Using the null model, the estimated detection probability of host 1 DNA across all sea lamprey that fed on both hosts was 76.9%.

Table 2.5. Results of occupancy model selection for the first 20 models comparing the relationship of host 1 DNA detections to several predictor variables including days attached to host 1, relative weight gain on host 1, fasting days since both hosts, and host 1 species. Models are ordered from top to bottom based on best fit using AIC value, and the null model is outlined. Results indicate the null model provides the best fit over either individual or combined predictor variables.

Model	DF	AIC	Delta	AICwt
~1	2	84.103	0	0.059
~rel_weight_gain * host1_species	5	84.193	0.09	0.056
~rel_weight_gain + host1_species + rel_weight_gain * host1_species	5	84.193	0.09	0.056
~days_attached_1	3	84.555	0.452	0.047
~rel_weight_gain	3	85.041	0.938	0.037
~fasting_days	3	85.065	0.962	0.036
~fasting_days_1	3	85.065	0.962	0.036
~host1_species + days_attached_1	4	85.135	1.032	0.035
~host1_species	3	85.37	1.267	0.031
~rel_weight_gain + days_attached_1	4	85.376	1.273	0.031
~days_attached_2	3	85.422	1.319	0.03
~fasting_days + days_attached_1	4	85.81	1.707	0.025
~fasting_days_1 + days_attached_1	4	85.81	1.707	0.025
~host1_species + days_attached_2	4	85.98	1.877	0.023
~days_attached_1 + days_attached_2	4	86.016	1.913	0.023
~rel_weight_gain + host1_species	4	86.057	1.955	0.022
~rel_weight_gain + host1_species + days_attached_1	5	86.281	2.178	0.02
~host1_species + days_attached_1 + days_attached_2	5	86.342	2.239	0.019
~fasting_days_1 + host1_species	4	86.701	2.598	0.016

DISCUSSION

Host DNA Detectability Across Environmental Conditions

Experimental data from this study demonstrated that host fish DNA is detectable using DNA metabarcoding from sea lamprey digestive tract samples across a range of environmental temperatures (5-15°C) consistent with periods of Great Lakes occupancy and fasting periods as long as 30 days. Further, multiple host species were detected when sea lamprey fed on different host fish species consecutively. Previous studies (Shink et al. 2019; Johnson et al. 2021) have already conducted genetic studies to identify lamprey diet composition. This study importantly quantified the effects of temperature, fasting periods, and multiple host feedings which had yet to be investigated. Given the range of water temperatures sea lamprey experience during periods of Great Lakes residence and greatest lethal impact on host fishes, their suspension of feeding during adult migration, and their capacity to feed on multiple species, it was important to understand the impact of these factors on intestinal DNA detectability for potential use in studies spanning the Great Lakes. Results indicate that not only is it feasible to identify the last host of a sea lamprey caught, but we can also identify previous feedings on a different species.

Increased temperatures have been noted to have effects on metabolic rate and the rate of DNA degradation, particularly in the detectability of eDNA (Pilliod et al. 2014; Strickler et al. 2015; Ruppert et al. 2019). As such, I expected models to indicate a significant relationship between temperature and host DNA read count. However, only the interaction between temperature and fasting period was significant, suggesting that the relationship between temperature and host read count varies according to fasting duration rather than exhibiting a consistent overall effect. A potential reason for this dependency is highlighted by the increase in weight gain seen within the 15°C group (Figure 2.2). Higher temperatures likely increase

consumption rates which may have compensated for potential effects on DNA degradation from increased temperatures or metabolic rates.

A negative relationship was noted between host read count and fasting period using a negative binomial model. However, this decrease is not consistent as a steeper decline occurred after the initial five days of fasting before leveling out, and remaining lower through the remaining 25 days (Figure 2.4). The increase in the 15°C group after 30 days fasting is difficult to explain, as two out of four samples demonstrated this increase, removing the potential of a single outlier. Mean weight gain for 15°C at 30 fasting days (1.94; sd = 1.0) was similar to both 5°C (1.40; sd = 0.44) and 10°C (1.53; sd = 0.65) at 30 fasting days, reducing the potential that these lamprey simply exhibited increased feeding on their hosts. It is possible that the low sample size (n = 4) of the 15°C/30-day fast group allowed for exaggerated differences in read counts between groups. Other explanations may involve changes in digestive physiology that result from increased water temperatures such as absorption rate changes, pH level increases, or shifts in digestive enzyme activity (Volkoff and Rønnestad 2020). It is possible that cellular breakdown due to heat stress within the digestive tract releases residual host DNA. Animal studies have demonstrated the uptake of small amounts of dietary DNA into intestinal epithelial cells (Rizzi et al. 2012). When coupled with the possibility of systemic intestinal damage that can occur in fish due to heat stress (Yang et al. 2022), a second wave of host DNA that had been taken up by lamprey intestinal cells may be released into the digestive tract as a result of cell shedding. Regardless of underlying function, our results highlight the applicability of this technique in the field, given that host DNA can still be detectable even at higher temperatures of 15°C and over longer fasting periods of 30 days. Furthermore, five additional lamprey during the host-switching trials were fasted for 45 days after feeding on the first host. While they were not

included in host-switching trial analysis as they did not feed on their secondary host, DNA was still detectable after 45 days of fasting in one out of the five lamprey, indicating the potential for DNA detections beyond 30 days of fasting.

Detection Thresholds

Detection thresholds have been used as a way to limit false positives that can occur from low-abundance errors, whether using a minimum sequence copy thresholds (MSCTs) or sample-based methods such as relative read abundance (RRA) thresholds (Alberdi et al. 2018; Ando et al. 2018). However, this method is hindered by its arbitrary nature, limited effect on errors that exceed thresholds, and potential to remove true positives below set thresholds (Deagle et al. 2019; Kelly et al. 2019; Littleford-Colquhoun et al. 2022). Given the goal of this study was to evaluate whether one of only two species was present in the sample, a more conservative approach was taken by setting both a strict read count threshold and an RRA threshold. Drake et al. (2022) outlines the complementary nature of including a sample-based threshold with a taxon contamination threshold, typically determined by the read counts per OTU in negative samples. However, all negative sample sequences were removed during bioinformatic filtering, so a common MSCT of 10 read counts was employed alongside a 1% RRA threshold (see Deagle et al. 2019; Drake et al. 2022). In our dataset, the majority (75%) of detections after 30 fasting days were based on at least 500 sequence reads, suggesting that a moderate increase in detection thresholds would still allow for host identification after an extended period of fasting.

Stochasticity within the PCR process and sequencing errors are both prevalent issues for metabarcoding studies (Alberdi et al. 2018; Dopheide et al. 2019). Common procedures for mitigating these issues involve the inclusion of biological and/or technical replicates (Ando et al. 2020). Given the small size of transformer sea lamprey and limited ability to retrieve ample

digestive material, biological replicates were difficult to obtain. As such, two technical PCR replicates were included per sea lamprey digestive sample. Both restrictive and additive approaches have been taken for the interpretation of technical replicates in metabarcoding studies (De Barba et al. 2014; Leray and Knowlton 2015; Alberdi et al. 2019), where restrictive methods only retain sequences detected in many or all replicates and additive methods allow for detections in any replicate. With these tradeoffs in mind, I plotted both methods through a bar chart representing the proportion of detections between both replicates (Figure 2.3). While a restrictive strategy may be appropriate in the context of this study, when applying this technique in the field to evaluate sea lamprey dietary composition the detection of rare taxa becomes more relevant. An interest in rare detections increases the negative aspects of the restrictive filter and its potential to overlook rare host species. One way to balance these tradeoffs is to utilize a relaxed approach as recommended by Alberdi et al. (2019), where detections are needed in 2 of 3 replicates. While this requires additional costs for at least three technical replicates, this may be the most effective method of balancing false positives and detections of rare taxa when applying DNA metabarcoding to sea lamprey dietary studies in the field.

Occupancy Modeling and Replicate Strategies

Another benefit of incorporating technical replication is that it allows for analysis of the data in occupancy modeling framework, which can produce estimates of detection probability from presence/absence data (MacKenzie et al. 2002). This method has been previously used in metabarcoding studies to account for various forms of uncertainty (Ficetola et al. 2015; Doi et al. 2019; Fukaya et al. 2022). An added benefit of particular use for this study is the ability to incorporate different covariates to evaluate their influence on detection probability. In the case of the temperature/fasting period trials, significant relationships between host detections and both

fasting period and days attached allowed for detection probability estimates that incorporated these variables. When considering only fasting period, a slight decrease over time is noticeable, but detection probabilities still remain relatively high after extended periods of time (with point estimates of detection probability between 91.6% at zero days fasted to 76.8% at 30 days fasted; Figure 2.8). Estimates of the detection probability of the assay can also be used to determine the number of replicates required for reliable detection of a host species, depending on the time of collection for lamprey field studies. For instance, with a 76.8% chance of detection at 30 days, you would need three replicates for the probability of detection to exceed 95%, as shown:

$$1 - (1 - 0.768)^2 = 0.946 \qquad (2\ replicates)$$

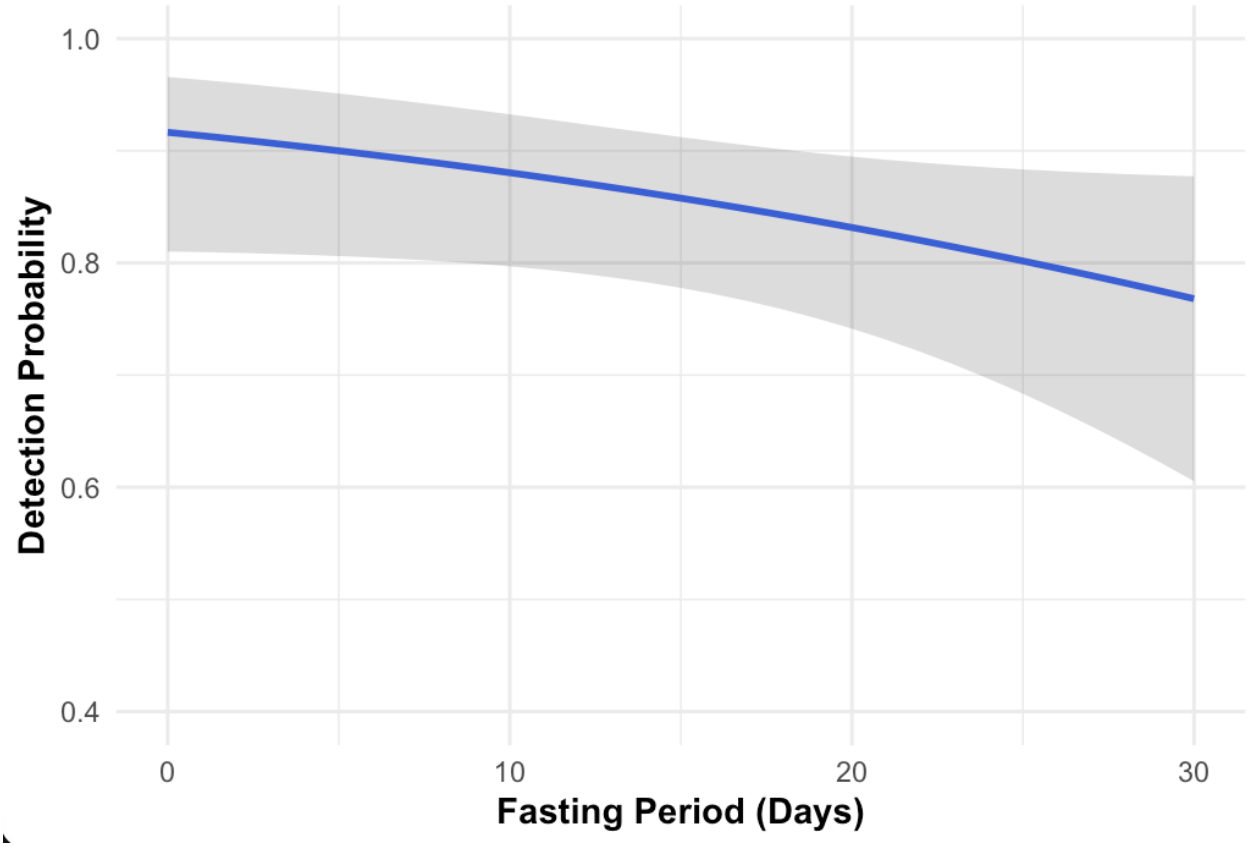
$$1 - (1 - 0.768)^3 = 0.988 \qquad (3\ replicates)$$

However, if field collections were earlier in the season, then only two replicates may be sufficient to exceed a 95% detection probability. As shown in the contour plot for both fasting period and days attached **(**Figure 2.6), additional replicates may also be helpful if sea lamprey have spent less time attached to their hosts. This assessment, however, is likely less applicable to field studies as the time a given lamprey was attached to its previous host is unknown. Instead, it highlights the relationship that a lamprey spending longer periods of time attached to its host likely indicates an increase in feeding as reflected by the increased detectability of host DNA in the digestive tract.

When an occupancy framework was applied to the host-switching trial data, influences from environmental factors on host detection were less noticeable. It is possible that low sample size had an impact, as any lamprey that did not feed on two different host species consecutively were not included. Nonetheless, the model can still be used to estimate a detection probability of 76.9% for host 1. While not as high as detection probabilities for low fasting period times in the

first trials, this probability is comparable to detecting the most recent host from lamprey that had fasted for 30 days. Borrowing from previous concepts of determining replicate numbers, detection of a feeding history is possible but would benefit from additional replicates as host DNA in the digestive tract from previous feedings is likely diminished.

Figure 2.8. Probability of detecting host DNA in sea lamprey digestive tracts as a function of the number of fasting days post-feeding. Results derived from an occupancy model describing host detection probability as a function of fasting period (days) in the temperature and fasting period trials. Grey shading indicates 95% confidence interval.



While occupancy models were unable to determine strong relationships between covariates and host detections in the host-switching trials, visual assessments may offer an alternative perspective. When binary detections were plotted against fasting days, a complementary pattern seemed to arise between detections of host 1 and host 2 during the fasting

period (Figure 2.7a and 2.7b). After zero and five fasting days, host 1 detections were relatively uncommon (~25%) compared to host 2 detections (~95%). However, a steady increase in host 1 detections occurred up until a peak around 20 days at ~70% between both replicates, while host 2 detections decrease across a similar timeline. Final 30-day detections for host 1 and 2 are again mirrored with a decrease in detections for host 1 and an increase in detections for host 2. A possible explanation may be the timeframe for DNA to pass through the lamprey digestive system. With lamprey that were sampled at either zero or five fasting days, host 2 DNA may be far more abundant within the digestive tract due to the recency of feeding. This may result in host 2 sequences producing the majority of total sequence reads, leading to a lower fraction of host 1 reads. After around 10-20 fasting days, host 2 DNA presence may have decreased from initial post-feeding amounts (as in single-species feeding trials; Figure 2.4), allowing for a more balanced proportion of sequence reads between both hosts. Following this period, the remaining host 1 DNA may be exiting the digestive tract, leading to higher amounts of host 2 detections. This may suggest a peak of around 20 days where previous host DNA detections are most likely as the overabundance of immediate host DNA has passed through the system. This balance between host 1 and host 2 DNA over time is demonstrated in the relatively consistent proportion of samples that detected both hosts in at least one replicate over the course of the fasting period (Figure 2.7c). However, samples sizes in this study (37 total sea lamprey fed on both hosts) limited our ability to fully characterize trends in detections of multiple hosts over time. More samples are likely needed to improve our understanding of these interactions. Encouragingly, detections of host 1 are still present well after a second host has been parasitized, providing confidence for future applications of this technique.

Implications for Field Studies and Dietary Analysis in Sea Lamprey

Ultimately, these investigations are most useful for directing methods and data interpretations for future field studies evaluating sea lamprey diet composition and host preference. Temperature and fasting period trials with a single host species indicated that, although there are slight negative impacts, longer fasting timeframes and higher temperatures do not eliminate the ability to identify host species via DNA metabarcoding. Detection thresholds were set at fairly conservative levels to prevent false positives, but the need to detect rare hosts in field studies may prompt the use of less conservative detection thresholds, if reducing false negatives is a higher priority. Replicate detection analyses showed that the inclusion of technical PCR replicates can be an important tool for mitigating false negatives as well. If resources allow, at least three technical replicates may be preferable for use in molecular diet analysis of sea lamprey. This aligns with the replication needed for at least 95% detection probability of previous hosts at higher fasting timeframes, along with allowing for a relaxed approach to detection interpretations (presence in 2/3 replicates) as recommended by Alberdi et al. (2019). This is beneficial for adult lamprey, as individuals may have undergone longer fasting periods prior to capture, and host detection may be less consistent at fasting periods of 30 days or more. For juvenile parasitic lamprey, direct removal from their immediate hosts allows for easy identification, although this likely results in recovery bias for that host's DNA given feeding recency. As the detection proportion of host 1 at zero fasting days was 36.4% (n = 11), an increase in replicates should help to address biases associated with an overabundance of immediate host DNA. When considering environmental temperature, the two lowest read count ranges were both 5 and 10 fasting day periods at 15°C (Figure 2.4). This may suggest a precautionary approach to either include more replicates or implement more relaxed detection

thresholds when water temperature is higher, although the overall relationship between read count and temperature remained insignificant. Additionally, read counts at 20 fasting days were consistent with 5°C and 10°C temperatures and read counts at 30 days were higher than both other temperatures. Since it's impossible to determine when captured adult sea lamprey last fed or began migration, those with varying fasting periods may be captured together, complicating efforts to counteract sequencing biases within samples that are sequenced simultaneously.

The methods demonstrated in this study are encouraging for both juvenile and adult sea lamprey dietary studies. Thousands of juvenile and adult sea lamprey are captured every year from either commercial fishing bycatch or trappings at 29 index streams across the Great Lakes for control efforts and population estimates (Adams et al. 2021; Hume et al. 2021). This provides the opportunity for robust analyses of sea lamprey dietary composition and preferences across the Great Lakes. Given the ability to detect previous hosts, there is the potential for gaining insights on host preference in parasitic juveniles. Additionally, sampling biases from commercial fishing vessels targeting certain host fish can be bypassed with random trappings of adult sea lamprey given the ability to detect hosts after 30 fasting days. This provides the potential to help understand the hosts sea lamprey fed on prior to spawning migration. Given the slower decline in read counts across fasting days in the 5°C group, it is possible that early-season collections could provide better results due to both cooler water temperatures and limited fasting periods.

LITERATURE CITED

Adams JV, Barber JM, Bravener GA, Lewandoski SA. 2021. Quantifying Great Lakes sea lamprey populations using an index of adults. Journal of Great Lakes Research. 47:S335–S346. doi:10.1016/j.jglr.2021.04.009.

Adams JV, Jones ML. 2021. Evidence of host switching: Sea lampreys disproportionately attack Chinook salmon when lake trout abundance is low in Lake Ontario. Journal of Great Lakes Research. 47:S604–S611. doi:10.1016/j.jglr.2020.03.003.

Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika*, *60*(2), 255-265.

Alberdi A, Aizpurua O, Bohmann K, Gopalakrishnan S, Lynggaard C, Nielsen M, Gilbert MTP. 2019. Promises and pitfalls of using high-throughput sequencing for diet analysis. Molecular Ecology Resources. 19(2):327–348. doi:10.1111/1755-0998.12960.

Alberdi A, Aizpurua O, Gilbert MTP, Bohmann K. 2018. Scrutinizing key steps for reliable metabarcoding of environmental samples. Methods in Ecology and Evolution. 9(1):134–147. doi:10.1111/2041-210X.12849.

Ando H, Fujii C, Kawanabe M, Ao Y, Inoue T, Takenaka A. 2018. Evaluation of plant contamination in metabarcoding diet analysis of a herbivore. Sci Rep. 8(1):15563. doi:10.1038/s41598-018-32845-w.

Ando H, Mukai H, Komura T, Dewi T, Ando M, Isagi Y. 2020. Methodological trends and perspectives of animal dietary studies by noninvasive fecal DNA metabarcoding. Environmental DNA. 2(4):391–406. doi:10.1002/edn3.117.

Beamish FWH. 1980. Biology of the North American Anadromous Sea Lamprey, Petromyzon marinus. Can J Fish Aquat Sci. 37(11):1924–1943. doi:10.1139/f80-233.

Bence JR, Bergstedt RA, Christie GC, Cochran PA, Ebener MP, Koonce JF, Rutter MA, Swink WD. 2003. Sea Lamprey (Petromyzon marinus) Parasite-host Interactions in the Great Lakes. Journal of Great Lakes Research. 29:253–282. doi:10.1016/S0380-1330(03)70493-6.

Bergstedt RA, Schneider CP. 1988. Assessment of Sea Lamprey (Petromyzon marinus) Predation by Recovery of Dead Lake Trout (Salvelinus namaycush) from Lake Ontario, 1982–85. Can J Fish Aquat Sci. 45(8):1406–1410. doi:10.1139/f88-164.

Berry O, Bulman C, Bunce M, Coghlan M, Murray DC, Ward RD. 2015. Comparison of morphological and DNA metabarcoding analyses of diets in exploited marine fishes. Marine Ecology Progress Series. 540:167–181. doi:10.3354/meps11524.

Berry TE, Osterrieder SK, Murray DC, Coghlan ML, Richardson AJ, Grealy AK, Stat M, Bejder L, Bunce M. 2017. DNA metabarcoding for diet analysis and biodiversity: A case study using the endangered Australian sea lion (Neophoca cinerea). Ecology and Evolution. 7(14):5435–5453. doi:10.1002/ece3.3123.

Bohmann K, Gopalakrishnan S, Nielsen M, Nielsen L dos SB, Jones G, Streicker DG, Gilbert MTP. 2018. Using DNA metabarcoding for simultaneous inference of common vampire bat diet and population structure. Molecular Ecology Resources. 18(5):1050–1063. doi:10.1111/1755-0998.12891.

Bonato KO, Silva PC, Carvalho FR, Malabarba LR. 2022. Trophic interactions of vampire catfishes (Siluriformes: Vandelliinae) revealed by metabarcoding analysis of stomach contents. Freshwater Biology. 67(3):542–548. doi:10.1111/fwb.13861.

Buglione M, Maselli V, Rippa D, de Filippo G, Trapanese M, Fulgione D. 2018. A pilot study on the application of DNA metabarcoding for non-invasive diet analysis in the Italian hare. Mammalian Biology. 88:31–42. doi:10.1016/j.mambio.2017.10.010.

Coble DW, Bruesewitz RE, Fratt TW, Scheirer JW. 1990. Lake Trout, Sea Lampreys, and Overfishing in the Upper Great Lakes: A Review and Reanalysis. Transactions of the American Fisheries Society. 119(6):985–995. doi:10.1577/1548-8659(1990)119<0985:LTSLAO>2.3.CO;2.

Crawford SS, Canada NRC. 2001. Salmonine Introductions to the Laurentian Great Lakes: An Historical Review and Evaluation of Ecological Effects. NRC Research Press.

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. Molecular Ecology Resources. 14(2):306–323. doi:10.1111/1755-0998.12188.

Deagle BE, Kirkwood R, Jarman SN. 2009. Analysis of Australian fur seal diet by pyrosequencing prey DNA in faeces. Molecular Ecology. 18(9):2022–2038. doi:10.1111/j.1365-294X.2009.04158.x.

Deagle BE, Thomas AC, McInnes JC, Clarke LJ, Vesterinen EJ, Clare EL, Kartzinel TR, Eveson JP. 2019. Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? Molecular Ecology. 28(2):391–406. doi:10.1111/mec.14734.

Doi H, Fukaya K, Oka S, Sato K, Kondoh M, Miya M. 2019. Evaluation of detection probabilities at the water-filtering and initial PCR steps in environmental DNA metabarcoding using a multispecies site occupancy model. Sci Rep. 9(1):3581. doi:10.1038/s41598-019-40233-1.

Dopheide A, Xie D, Buckley TR, Drummond AJ, Newcomb RD. 2019. Impacts of DNA extraction and PCR on DNA metabarcoding estimates of soil biodiversity. Methods in Ecology and Evolution. 10(1):120–133. doi:10.1111/2041-210X.13086.

Drake LE, Cuff JP, Young RE, Marchbank A, Chadwick EA, Symondson WOC. 2022. An assessment of minimum sequence copy thresholds for identifying and reducing the prevalence of artefacts in dietary metabarcoding data. Methods in Ecology and Evolution. 13(3):694–710. doi:10.1111/2041-210X.13780.

Drinkwater R, Schnell IB, Bohmann K, Bernard H, Veron G, Clare E, Gilbert MTP, Rossiter SJ. 2019. Using metabarcoding to compare the suitability of two blood-feeding leech species for sampling mammalian diversity in North Borneo. Molecular Ecology Resources. 19(1):105–117. doi:10.1111/1755-0998.12943.

Elbrecht V, Leese F. 2015. Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer Bias and Biomass—Sequence Relationships with an Innovative Metabarcoding Protocol. PLOS ONE. 10(7):e0130324. doi:10.1371/journal.pone.0130324.

Ellen Marsden J, Chipman BD, Nashett LJ, Anderson JK, Bouffard W, Durfey L, Gersmehl JE, Schoch WF, Staats NR, Zerrenner A. 2003. Sea Lamprey Control in Lake Champlain. Journal of Great Lakes Research. 29:655–676. doi:10.1016/S0380-1330(03)70522-X.

Estrada-Franco JG, Fernández-Santos NA, Adebiyi AA, López-López M de J, Aguilar-Durán JA, Hernández-Triana LM, Prosser SWJ, Hebert PDN, Fooks AR, Hamer GL, et al. 2020. Vertebrate-Aedes aegypti and Culex quinquefasciatus (Diptera)-arbovirus transmission networks: Non-human feeding revealed by meta-barcoding and next-generation sequencing. PLOS Neglected Tropical Diseases. 14(12):e0008867. doi:10.1371/journal.pntd.0008867.

Fahmy M, Ravelomanantsoa NAF, Youssef S, Hekkala E, Siddall M. 2019. Biological inventory of Ranomafana National Park tetrapods using leech-derived iDNA. Eur J Wildl Res. 65(5):1–13. doi:10.1007/s10344-019-1305-3.

Farmer GJ, Beamish FWH, Lett PF. 1977. Influence of Water Temperature on the Growth Rate of the Landlocked Sea Lamprey (Petromyzon marinus) and the Associated Rate of Host Mortality. J Fish Res Bd Can. 34(9):1373–1378. doi:10.1139/f77-197.

Ficetola GF, Pansu J, Bonin A, Coissac E, Giguet-Covex C, De Barba M, Gielly L, Lopes CM, Boyer F, Pompanon F, et al. 2015. Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. Molecular Ecology Resources. 15(3):543–556. doi:10.1111/1755-0998.12338.

Firkus TJ, Murphy CA, Adams JV, Treska TJ, Fischer G. 2021. Assessing the assumptions of classification agreement, accuracy, and predictable healing time of sea lamprey wounds on lake trout. Journal of Great Lakes Research. 47:S368–S377. doi:10.1016/j.jglr.2020.07.016.

Fiske I, Chandler R. 2011. unmarked: An R Package for Fitting Hierarchical Models of Wildlife Occurrence and Abundance. Journal of Statistical Software. 43:1–23. doi:10.18637/jss.v043.i10.

Ford MJ, Hempelmann J, Hanson MB, Ayres KL, Baird RW, Emmons CK, Lundin JI, Schorr GS, Wasser SK, Park LK. 2016. Estimation of a Killer Whale (Orcinus orca) Population's Diet Using Sequencing Analysis of DNA from Feces. PLOS ONE. 11(1):e0144956. doi:10.1371/journal.pone.0144956.

Fukaya K, Kondo NI, Matsuzaki SS, Kadoya T. 2022. Multispecies site occupancy modelling and study design for spatially replicated environmental DNA metabarcoding. Methods in Ecology and Evolution. 13(1):183–193. doi:10.1111/2041-210X.13732.

Gariepy TD, Lindsay R, Ogden N, Gregory TR. 2012. Identifying the last supper: utility of the DNA barcode library for bloodmeal identification in ticks. Mol Ecol Resour. 12(4):646–652. doi:10.1111/j.1755-0998.2012.03140.x.

Guo Z, Andreou D, Britton JR. 2017. Sea Lamprey Petromyzon marinus Biology and Management Across Their Native and Invasive Ranges: Promoting Conservation by Knowledge Transfer. Reviews in Fisheries Science & Aquaculture. 25(1):84–99. doi:10.1080/23308249.2016.1233166.

Hacker CE, Hoenig BD, Wu L, Cong W, Yu J, Dai Y, Li Y, Li J, Xue Y, Zhang Yu, et al. 2021. Use of DNA metabarcoding of bird pellets in understanding raptor diet on the Qinghai-Tibetan Plateau of China. Avian Research. 12(1):42. doi:10.1186/s40657-021-00276-3.

Happel A, Rinchard J, Czesny S. 2017. Variability in sea lamprey fatty acid profiles indicates a range of host species utilization in Lake Michigan. Journal of Great Lakes Research. 43(1):182–188. doi:10.1016/j.jglr.2016.10.010.

Harms-Tuohy CA, Schizas NV, Appeldoorn RS. 2016. Use of DNA metabarcoding for stomach content analysis in the invasive lionfish Pterois volitans in Puerto Rico. Marine Ecology Progress Series. 558:181–191. doi:10.3354/meps11738.

Harvey CJ, Ebener MP, White CK. 2008. Spatial and Ontogenetic Variability of Sea Lamprey Diets in Lake Superior. Journal of Great Lakes Research. 34(3):434–449. doi:10.3394/0380-1330(2008)34[434:SAOVOS]2.0.CO;2.

Heinrich JW, Mullett KM, Hansen MJ, Adams JV, Klar GT, Johnson DA, Christie GC, Young RJ. 2003. Sea Lamprey Abundance and Management in Lake Superior, 1957 to 1999. Journal of Great Lakes Research. 29:566–583. doi:10.1016/S0380-1330(03)70517-6.

Homma C, Inokuchi D, Nakamura Y, Uy WH, Ohnishi K, Yamaguchi H, Adachi M. 2022. Effectiveness of blocking primers and a peptide nucleic acid (PNA) clamp for 18S metabarcoding dietary analysis of herbivorous fish. Kumar R, editor. PLoS ONE. 17(4):e0266268. doi:10.1371/journal.pone.0266268.

Hrodey PJ, Lewandoski SA, Sullivan WP, Barber JM, Mann KA, Paudel B, Symbal MJ. 2021. Evolution of the sea lamprey control barrier program: The importance of lowermost barriers☆. Journal of Great Lakes Research. 47:S285–S296. doi:10.1016/j.jglr.2021.10.006.

Hubbs CL, Pope TEB. 1937. The Spread of the Sea Lamprey Through the Great Lakes. Transactions of the American Fisheries Society. 66(1):172–176. doi:10.1577/1548-8659(1936)66[172:TSOTSL]2.0.CO;2.

Huggins LG, Koehler AV, Schunack B, Inpankaew T, Traub RJ. 2020. A Host-Specific Blocking Primer Combined with Optimal DNA Extraction Improves the Detection Capability of a Metabarcoding Protocol for Canine Vector-Borne Bacteria. Pathogens. 9(4):258. doi:10.3390/pathogens9040258.

Hume JB, Bennis S, Bruning T, Docker MF, Good S, Lampman R, Rinchard J, Searcy T, Wilkie MP, Johnson NS. 2024. Evaluation of Larval Sea Lamprey Petromyzon marinus Growth in the Laboratory: Influence of Temperature and Diet. Aquaculture Research. 2024(1):5547340. doi:10.1155/2024/5547340.

Hume JB, Bravener GA, Flinn S, Johnson NS. 2021. What can commercial fishery data in the Great Lakes reveal about juvenile sea lamprey (*Petromyzon marinus*) ecology and management? Journal of Great Lakes Research. 47:S590–S603. doi:10.1016/j.jglr.2021.03.023.

Irwin BJ, Liu W, Bence JR, Jones ML. 2012. Defining Economic Injury Levels for Sea Lamprey Control in the Great Lakes Basin. North American Journal of Fisheries Management. 32(4):760–771. doi:10.1080/02755947.2012.685140.

Iverson SJ, Field C, Don Bowen W, Blanchard W. 2004. Quantitative Fatty Acid Signature Analysis: A New Method of Estimating Predator Diets. Ecological Monographs. 74(2):211–235. doi:10.1890/02-4105.

Jakubavičiūtė E, Bergström U, Eklöf JS, Haenel Q, Bourlat SJ. 2017. DNA metabarcoding reveals diverse diet of the three-spined stickleback in a coastal ecosystem. PLOS ONE. 12(10):e0186929. doi:10.1371/journal.pone.0186929.

Johnson NS, Lewandoski SA, Merkes C. 2021. Assessment of sea lamprey (Petromyzon marinus) diet using DNA metabarcoding of feces. Ecological Indicators. 125:107605. doi:10.1016/j.ecolind.2021.107605.

Jones ML. 2007. Toward Improved Assessment of Sea Lamprey Population Dynamics in Support of Cost-effective Sea Lamprey Management. Journal of Great Lakes Research. 33:35–47. doi:10.3394/0380-1330(2007)33[35:TIAOSL]2.0.CO;2.

Jusino MA, Banik MT, Palmer JM, Wray AK, Xiao L, Pelton E, Barber JR, Kawahara AY, Gratton C, Peery MZ, et al. 2019. An improved method for utilizing high-throughput amplicon sequencing to determine the diets of insectivorous animals. Molecular Ecology Resources. 19(1):176–190. doi:10.1111/1755-0998.12951.

Kamil Bartoń. 2010. MuMIn: Multi-Model Inference. :1.47.5. doi:10.32614/CRAN.package.MuMIn. [accessed 2024 Sep 20]. https://CRAN.R-project.org/package=MuMIn.

Kelly RP, Shelton AO, Gallego R. 2019. Understanding PCR Processes to Draw Meaningful Conclusions from Environmental DNA Studies. Sci Rep. 9(1):12133. doi:10.1038/s41598-019-48546-x.

Kitchell JF. 1990. The Scope for Mortality Caused by Sea Lamprey. Transactions of the American Fisheries Society. 119(4):642–648. doi:10.1577/1548-8659(1990)119<0642:TSFMCB>2.3.CO;2.

Lantry B, Adams J, Christie G, Schaner T, Bowlby J, Keir M, Lantry J, Sullivan P, Bishop D, Treska T, et al. 2015. Sea lamprey mark type, marking rate, and parasite–host relationships for

lake trout and other species in Lake Ontario. Journal of Great Lakes Research. 41(1):266–279. doi:10.1016/j.jglr.2014.12.013.

Larsen LO. 1980. Physiology of Adult Lampreys, with Special Regard to Natural Starvation, Reproduction, and Death after Spawning. Can J Fish Aquat Sci. 37(11):1762–1779. doi:10.1139/f80-221.

Lawrie AH. 1970. The Sea Lamprey in the Great Lakes. Transactions of the American Fisheries Society. 99(4):766–775. doi:10.1577/1548-8659(1970)99<766:TSLITG>2.0.CO;2.

Lefèvre E, Gardner CM, Gunsch CK. 2020. A novel PCR-clamping assay reducing plant host DNA amplification significantly improves prokaryotic endo-microbiome community characterization. FEMS Microbiology Ecology. 96(7):fiaa110. doi:10.1093/femsec/fiaa110.

Leray M, Agudelo N, Mills SC, Meyer CP. 2013. Effectiveness of Annealing Blocking Primers versus Restriction Enzymes for Characterization of Generalist Diets: Unexpected Prey Revealed in the Gut Contents of Two Coral Reef Fish Species. PLOS ONE. 8(4):e58076. doi:10.1371/journal.pone.0058076.

Leray M, Knowlton N. 2015. DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. Proceedings of the National Academy of Sciences. 112(7):2076–2081. doi:10.1073/pnas.1424997112.

Levy-Booth DJ, Campbell RG, Gulden RH, Hart MM, Powell JR, Klironomos JN, Pauls KP, Swanton CJ, Trevors JT, Dunfield KE. 2007. Cycling of extracellular DNA in the soil environment. Soil Biology and Biochemistry. 39(12):2977–2991. doi:10.1016/j.soilbio.2007.06.020.

Littleford-Colquhoun BL, Freeman PT, Sackett VI, Tulloss CV, McGarvey LM, Geremia C, Kartzinel TR. 2022. The precautionary principle and dietary DNA metabarcoding: Commonly used abundance thresholds change ecological interpretation. Molecular Ecology. 31(6):1615–1626. doi:10.1111/mec.16352.

Lopes CM, De Barba M, Boyer F, Mercier C, Galiano D, Kubiak BB, Maestri R, da Silva Filho PJS, Gielly L, Coissac E, et al. 2020. Ecological specialization and niche overlap of subterranean rodents inferred from DNA metabarcoding diet analysis. Molecular Ecology. 29(16):3143–3153. doi:10.1111/mec.15549.

Lucas MC, Hume JB, Almeida PR, Aronsuu K, Habit E, Silva S, Wang CJ, Zampatti B. 2021. Emerging conservation initiatives for lampreys: Research challenges and opportunities. Journal of Great Lakes Research. 47:S690–S703. doi:10.1016/j.jglr.2020.06.004.

MacKenzie DI, Nichols JD, Lachman GB, Droege S, Andrew Royle J, Langtimm CA. 2002. Estimating Site Occupancy Rates When Detection Probabilities Are Less Than One. Ecology. 83(8):2248–2255. doi:10.1890/0012-9658(2002)083[2248:ESORWD]2.0.CO;2.

McClenaghan B, Nol E, Kerr KCR. 2019. DNA metabarcoding reveals the broad and flexible diet of a declining aerial insectivore. The Auk. 136(1):uky003. doi:10.1093/auk/uky003.

McInnes JC, Alderman R, Deagle BE, Lea M-A, Raymond B, Jarman SN. 2017. Optimised scat collection protocols for dietary DNA metabarcoding in vertebrates. Methods in Ecology and Evolution. 8(2):192–202. doi:10.1111/2041-210X.12677.

Miehls S, Sullivan P, Twohey M, Barber J, McDonald R. 2020. The future of barriers and trapping methods in the sea lamprey (Petromyzon marinus) control program in the Laurentian Great Lakes. Rev Fish Biol Fisheries. 30(1):1–24. doi:10.1007/s11160-019-09587-7.

Nguyen TV, Tilker A, Nguyen A, Hörig L, Axtner J, Schmidt A, Le M, Nguyen AHQ, Rawson BM, Wilting A, et al. 2021. Using terrestrial leeches to assess the genetic diversity of an elusive species: The Annamite striped rabbit Nesolagus timminsi. Environmental DNA. 3(4):780–791. doi:10.1002/edn3.182.

Oksanen J. 2010. Vegan : community ecology package. http://vegan.r-forge.r-project.org/. [accessed 2024 Oct 9]. https://cir.nii.ac.jp/crid/1570291225091856896.

Pertoldi C, Schmidt JB, Thomsen PM, Nielsen LB, de Jonge N, Iacolina L, Muro F, Nielsen KT, Pagh S, Lauridsen TL, et al. 2021. Comparing DNA metabarcoding with faecal analysis for diet determination of the Eurasian otter (Lutra lutra) in Vejlerne, Denmark. Mamm Res. 66(1):115–122. doi:10.1007/s13364-020-00552-5.

Pilliod DS, Goldberg CS, Arkle RS, Waits LP. 2014. Factors influencing detection of eDNA from a stream-dwelling amphibian. Molecular Ecology Resources. 14(1):109–116. doi:10.1111/1755-0998.12159.

Piñol J, Mir G, Gomez-Polo P, Agustí N. 2015. Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. Molecular Ecology Resources. 15(4):819–830. doi:10.1111/1755-0998.12355.

Potter IC. 1980. Ecology of Larval and Metamorphosing Lampreys. Can J Fish Aquat Sci. 37(11):1641–1657. doi:10.1139/f80-212.

Pukk L, Kanefsky J, Heathman AL, Weise EM, Nathan LR, Herbst SJ, Sard NM, Scribner KT, Robinson JD. 2021. eDNA metabarcoding in lakes to quantify influences of landscape features and human activity on aquatic invasive species prevalence and fish community diversity. Diversity and Distributions. 27(10):2016–2031. doi:10.1111/ddi.13370.

Ratnasingham S, Wei C, Chan D, Agda J, Agda J, Ballesteros-Mejia L, Ait Boutou H, El Bastami Z M, Ma E, Manjunath R, Rea D, Ho C, Telfer A, McKeowan J, Rahulan M, Steinke C, Dorsheimer J, Milton M, Hebert PDN (2024). BOLD v4: A Centralized Bioinformatics Platform for DNA-Based Biodiversity Data. In DNA Barcoding: Methods and Protocols, pp. 403-441. Chapter 26. New York, NY: Springer US, 2024.

Reeves LE, Gillett-Kaufman JL, Kawahara AY, Kaufman PE. 2018. Barcoding blood meals: New vertebrate-specific primer sets for assigning taxonomic identities to host DNA from mosquito blood meals. PLOS Neglected Tropical Diseases. 12(8):e0006767. doi:10.1371/journal.pntd.0006767.

Riaz T, Shehzad W, Viari A, Pompanon F, Taberlet P, Coissac E. 2011. ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. Nucleic Acids Research. 39(21):e145. doi:10.1093/nar/gkr732.

Rizzi A, Raddadi N, Sorlini C, Nordgrd L, Nielsen KM, Daffonchio D. 2012. The Stability and Degradation of Dietary DNA in the Gastrointestinal Tract of Mammals: Implications for Horizontal Gene Transfer and the Biosafety of GMOs. Critical Reviews in Food Science and Nutrition. 52(2):142–161. doi:10.1080/10408398.2010.499480.

Robinson KF, Miehls SM, Siefkes MJ. 2021. Understanding sea lamprey abundances in the Great Lakes prior to broad implementation of sea lamprey control. Journal of Great Lakes Research. 47:S328–S334. doi:10.1016/j.jglr.2021.04.002.

Rojahn J, Gleeson DM, Furlan E, Haeusler T, Bylemans J. 2021. Improving the detection of rare native fish species in environmental DNA metabarcoding surveys. Aquatic Conservation: Marine and Freshwater Ecosystems. 31(4):990–997. doi:10.1002/aqc.3514.

Ruppert KM, Kline RJ, Rahman MS. 2019. Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. Global Ecology and Conservation. 17:e00547. doi:10.1016/j.gecco.2019.e00547.

Rutter MA, Bence JR. 2003. An Improved Method to Estimate Sea Lamprey Wounding Rate on Hosts with Application to Lake Trout in Lake Huron. Journal of Great Lakes Research. 29:320–331. doi:10.1016/S0380-1330(03)70497-3.

Sard NM, Herbst SJ, Nathan L, Uhrig G, Kanefsky J, Robinson JD, Scribner KT. 2019. Comparison of fish detections, community diversity, and relative abundance using environmental DNA metabarcoding and traditional gears. Environmental DNA. 1(4):368–384. doi:10.1002/edn3.38.

Sayers EW, Cavanaugh M, Clark K, Pruitt KD, Schoch CL, Sherry ST, Karsch-Mizrachi I. 2022. GenBank. Nucleic Acids Research. 50(D1):D161–D164. doi:10.1093/nar/gkab1135.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, et al. 2009. Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Applied and Environmental Microbiology. 75(23):7537–7541. doi:10.1128/AEM.01541-09.

Schmidt SN, Olden JD, Solomon CT, Zanden MJV. 2007. Quantitative Approaches to the Analysis of Stable Isotope Food Web Data. Ecology. 88(11):2793–2802. doi:10.1890/07-0121.1.

Schnell IB, Thomsen PF, Wilkinson N, Rasmussen M, Jensen LRD, Willerslev E, Bertelsen MF, Gilbert MTP. 2012. Screening mammal biodiversity using DNA from leeches. Current Biology. 22(8):R262–R263. doi:10.1016/j.cub.2012.02.058.

Shink KG, Sutton TM, Murphy JM, López JA. 2019. Utilizing DNA metabarcoding to characterize the diet of marine-phase Arctic lamprey (*Lethenteron camtschaticum*) in the eastern Bering Sea. Can J Fish Aquat Sci. 76(11):1993–2002. doi:10.1139/cjfas-2018-0299.

Siefkes MJ, Johnson NS, Muir AM. 2021. A renewed philosophy about supplemental sea lamprey controls. Journal of Great Lakes Research. 47:S742–S752. doi:10.1016/j.jglr.2021.03.013.

Smith BR, Tibbles JJ. 1980. Sea Lamprey (Petromyzon marinus) in Lakes Huron, Michigan, and Superior: History of Invasion and Control, 1936–78. Can J Fish Aquat Sci. 37(11):1780–1801. doi:10.1139/f80-222.

Strickler KM, Fremier AK, Goldberg CS. 2015. Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. Biological Conservation. 183:85–92. doi:10.1016/j.biocon.2014.11.038.

Swink WD. 2003. Host Selection and Lethality of Attacks by Sea Lampreys (Petromyzon marinus) in Laboratory Studies. Journal of Great Lakes Research. 29:307–319. doi:10.1016/S0380-1330(03)70496-1.

Taerum SJ, Steven B, Gage DJ, Triplett LR. 2020. Validation of a PNA Clamping Method for Reducing Host DNA Amplification and Increasing Eukaryotic Diversity in Rhizosphere Microbiome Studies. Phytobiomes Journal. 4(4):291–302. doi:10.1094/PBIOMES-05-20-0040-TA.

Thomas AC, Deagle BE, Eveson JP, Harsch CH, Trites AW. 2016. Quantitative DNA metabarcoding: improved estimates of species proportional biomass using correction factors derived from control material. Molecular Ecology Resources. 16(3):714–726. doi:10.1111/1755-0998.12490.

Toju H, Baba YG. 2018. DNA metabarcoding of spiders, insects, and springtails for exploring potential linkage between above- and below-ground food webs. Zoological Lett. 4(1):4. doi:10.1186/s40851-018-0088-9.

Trebitz AS, Hoffman JC, Grant GW, Billehus TM, Pilgrim EM. 2015. Potential for DNA-based identification of Great Lakes fauna: match and mismatch between taxa inventories and DNA barcode libraries. Sci Rep. 5(1):12162. doi:10.1038/srep12162.

Treska TJ, Ebener MP, Christie GC, Adams JV, Siefkes MJ. 2021. Setting and tracking suppression targets for sea lampreys in the Great Lakes. Journal of Great Lakes Research. 47:S357–S367. doi:10.1016/j.jglr.2021.10.007.

Turner CR, Uy KL, Everhart RC. 2015. Fish environmental DNA is more concentrated in aquatic sediments than surface water. Biological Conservation. 183:93–102. doi:10.1016/j.biocon.2014.11.017.

Vestheim H, Deagle BE, Jarman SN. 2011. Application of Blocking Oligonucleotides to Improve Signal-to-Noise Ratio in a PCR. In: Park DJ, editor. PCR Protocols. Totowa, NJ:

Humana Press. (Methods in Molecular Biology). p. 265–274. [accessed 2023 Jan 27]. https://doi.org/10.1007/978-1-60761-944-4_19.

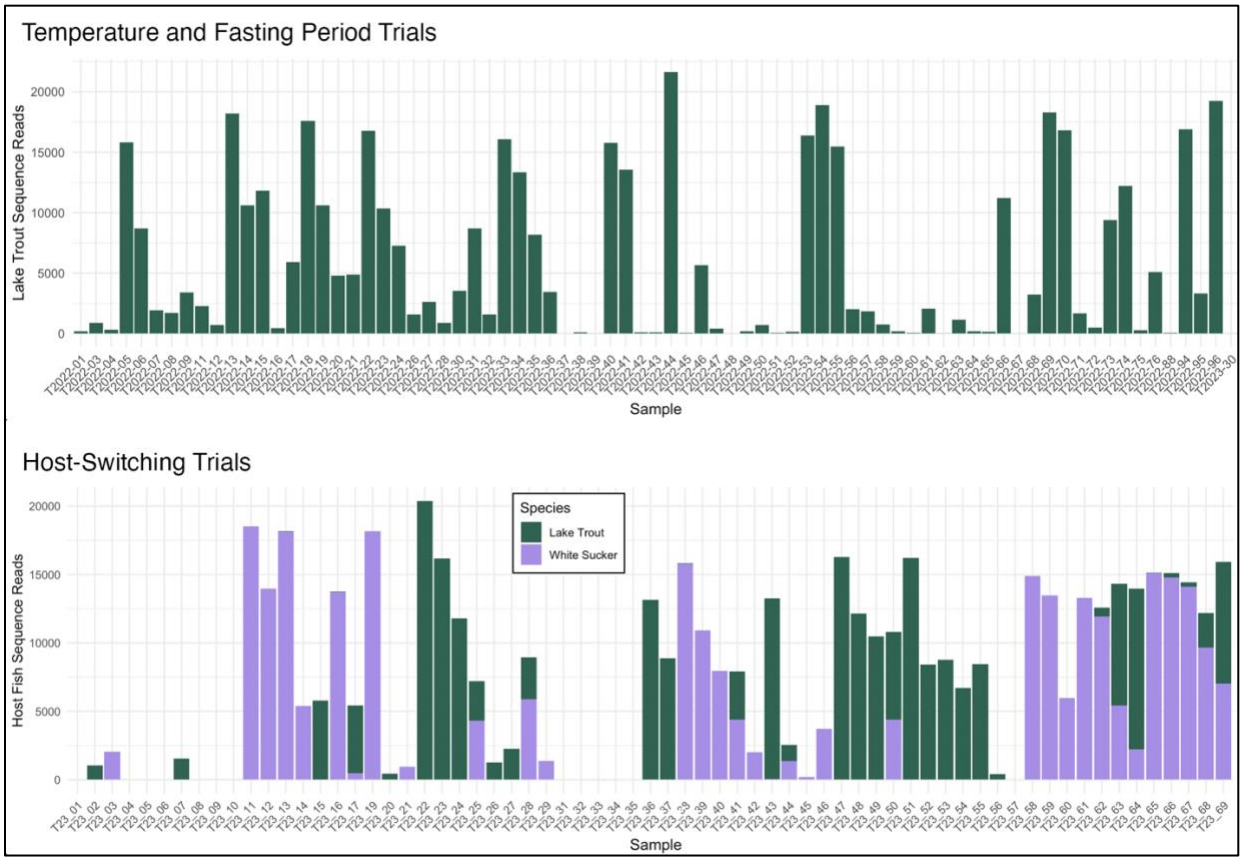
Volkoff H, Rønnestad I. 2020. Effects of temperature on feeding and digestive processes in fish. Temperature. 7(4):307–320. doi:10.1080/23328940.2020.1765950.

Waraniak JM, Marsh TL, Scribner KT. 2019. 18S rRNA metabarcoding diet analysis of a predatory fish community across seasonal changes in prey availability. Ecology and Evolution. 9(3):1410–1430. doi:10.1002/ece3.4857.

Yang S, Zhang C, Xu W, Li D, Feng Y, Wu J, Luo W, Du X, Du Z, Huang X. 2022. Heat Stress Decreases Intestinal Physiological Function and Facilitates the Proliferation of Harmful Intestinal Microbiota in Sturgeons. Front Microbiol. 13. doi:10.3389/fmicb.2022.755369. [accessed 2024 Sep 3]. https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2022.755369/full.

APPENDIX A: SUPPLEMENTARY FIGURES

Figure S2.1. Total sequence reads for host DNA across all samples for both temperature/fasting period trials (top) and host-switching trials (bottom). In the host-switching trials plot, as two hosts were used, host species for sequence reads is indicated by color with green representing lake trout and purple representing white sucker.



APPENDIX B: SUPPLEMENTARY MATERIALS

Data and statistical analyses (Temperature and Feeding Trials):

https://github.com/okaneco1/SL_experimental1

Data and statistical analyses (Host-Switching Trials):

https://github.com/okaneco1/SL_experimental2

CHAPTER 3: NEW INSIGHTS INTO LANDLOCKED PARASITIC SEA LAMPREY DIETS USING DNA METABARCODING

ABSTRACT

Following their invasion into the Laurentian Great Lakes, sea lamprey (*Petromyzon marinus*) have caused extensive ecological damage to native fisheries and fish communities. While control efforts have been largely successful in reducing overall abundance, traditional dietary assessments have had limited ability to identify sea lamprey host species. Further, current host wound assessments used to estimate overall damage to Great Lakes fisheries rely on assessments of sea-lamprey wounded lake trout that have survived previous lamprey attacks, resulting in potential limitations of our understanding of sea lamprey diets and damage. This study aims to increase our knowledge of sea lamprey diets by incorporating DNA metabarcoding to analyze the diets of blood-feeding sea lamprey across lakes (Huron, Superior and Champlain), life stages (adult and parasitic juveniles), and years (2022 and 2023). Sea lamprey were collected across each lake and year, and dietary DNA was analyzed via metabarcoding with the 12S rRNA gene region. To improve robustness of metabarcoding results, a sea lamprey-specific blocking primer was used to suppress amplification of sea lamprey sequences. Our results indicated that whitefish (Coregoninae) and lake trout (*Salvelinus namaycush*) comprised the majority of sea lamprey hosts across all samples, accounting for 42.1% and 34.3% of total sequence reads. Pronounced differences in dietary composition occurred in several specific subsets, including differences among lakes in 2023 parasitic samples and differences between stages of the life cycle in 2022 samples from Lake Huron. While most of the variance in dietary composition was unexplained, PERMANOVA analyses revealed that lake and life stage explained a larger fraction of the variance among samples (7.9% and 7.2%, respectively) than year (3.3%). These observed differences across lakes and life stages have implications for management strategies, as

targeted control efforts that focus on areas of higher ecological damage may allow for more efficient allocation of finite management resources. Furthermore, future molecular diet analysis studies across the Great Lakes could help to improve the accuracy of damage assessments.

INTRODUCTION

The sea lamprey (*Petromyzon marinus*) is a hematophagous ectoparasite that has inflicted profound ecological and economic damage as an invader of the Laurentian Great Lakes (Hubbs and Pope 1937; Lawrie 1970). As such, the Great Lakes Fishery Commission (GLFC) was instated as a bi-national organization between the United States and Canada focused on controlling the spread and damage caused by sea lamprey within the region (Robinson et al. 2021). While these efforts have been largely successful in reducing population sizes (Heinrich et al. 2003), certain challenges and limitations remain in gathering a comprehensive understanding of sea lamprey diets due to their blood-feeding nature.

Sea lamprey are primarily controlled to reduce damage inflicted upon valuable fisheries. Sea lamprey-induced damage is assessed annually within each of the Great Lakes by monitoring wounds on captured lake trout (*Salvelinus namaycush*) and estimating a wounding rate (Rutter and Bence 2003) and developing an annual assessment of adult sea lamprey abundance (Adams et al., 2021). From these estimates, metrics such as juvenile (feeding-stage) sea lamprey population size and overall ecological damage can be assessed. Historical wounding rate data also contribute to the establishment of lake-specific adult abundance targets in the sea lamprey control program. However, sea lamprey wounding surveys rely heavily on lake trout selective fisheries or assessments, which can complicate interpretation of wounding rate data. For example, there is evidence that host preferences may change as the relative abundance of prey species changes (Adams and Jones 2021), potentially resulting in skewed estimates of juvenile

sea lamprey abundance and commercial fisheries impacts (Hume et al. 2021). Additional challenges are involved when relying solely on captured fish that have either been previously attacked by sea lamprey or have sea lamprey currently attached. When estimating wounding rates and host preference, host fish that did not survive the attack sink to the bottom of the lake, thus preventing them from being captured and ultimately included in damage estimates (Bergstedt and Schneider 1988). This issue is exacerbated by the high mortality rate of sea lamprey attacks (Swink 2003).

A comprehensive understanding of sea lamprey diets would help provide a more accurate assessment of sea lamprey impacts on Great Lakes fisheries. However, traditional dietary analysis using visual identification of gut contents is not possible, given the blood diet of sea lamprey. To address this, previous studies have utilized biochemical methods such as fatty acid profiles and stable isotope analyses to investigate sea lamprey diets (Harvey et al. 2008; Happel et al. 2017). These studies have provided valuable information for understanding sea lamprey diets, such as Harvey et al. (2008) indicating that sea lamprey captured in Black Bay, Lake Superior rely more on lower trophic level species such as whitefish. However, while these methods provide information on ontogenetic dietary shifts and host variation across spatial scales, they have certain limitations as well. For example, inferences from these methods are limited to trophic-level interactions rather than supplying an understanding of species-level host identification. Additionally, for both approaches, baseline data are required for proper interpretation, which necessitates extensive collection and monitoring as the fatty acid and stable isotope signatures of prey items may change through time (Iverson et al. 2004; Schmidt et al. 2007). Regardless of these limitations, these studies have warranted further research given the

implications that sea lamprey may feed on a wide range of hosts and can vary selection over time and space.

New technological advances in genetic techniques provide an alternative method that addresses the need for further dietary assessments at the species-level, while avoiding some of the limitations associated with fatty acid profile and stable isotope analyses. DNA metabarcoding has gained traction as an effective and applicable approach to monitoring dietary composition in a variety of animal taxa including mammals (Berry et al. 2017; Buglione et al. 2018; Lopes et al. 2020), birds (McClenaghan et al. 2019; Hacker et al. 2021), fishes (Berry et al. 2015; Harms-Tuohy et al. 2016; Jakubavičiūtė et al. 2017), and other groups. This approach has been successfully applied in several hematophagous species, such as leeches (Drinkwater et al. 2019), mosquitos (Reeves et al. 2018; Estrada-Franco et al. 2020), ticks (Gariepy et al. 2012), vampire bats (Bohmann et al. 2018), and, particularly, Artic lamprey (*Lethenteron camtschaticum*; Shink et al. 2019) and sea lamprey (Johnson et al. 2021). Molecular diet analysis provides a distinct advantage in two ways, as it allows for species-level identification of prey items and dietary assessments in sea lamprey not captured while attached to a host species. These advantages were highlighted in Johnson et al. (2021), where fecal samples from Great Lakes sea lamprey were analyzed using DNA metabarcoding. Among other results, unexpected host use was found from sea lamprey in northern Lake Huron with frequent detections of white sucker (*Catostomus commersonii*) and longnose sucker (*Catostomus cotostomus*) in adult fecal samples (Johnson et al. 2021). Previous DNA metabarcoding studies also examined dietary composition of Artic lamprey, with results indicating that although salmonids were thought to be a primary food source, homogenized dietary samples rarely exhibited salmonid detections while several previously unknown food sources were discovered (Shink et al. 2019). Both of these previous

studies highlight the potential for DNA metabarcoding, particularly for lamprey, to provide novel insights into dietary compositions.

Preliminary studies using wide-scale DNA metabarcoding techniques for dietary assessments of sea lamprey within the Great Lakes region were conducted to complement field studies, including the development of a sea lamprey-specific blocking primer designed to prevent PCR amplification of sea lamprey sequences (Chapter 1) and laboratory experiments designed to assess the influence of temperature, fasting period length, and multiple host feedings on the genetic detection of hosts (Chapter 2). Insights from these studies can facilitate interpretation of results from large-scale applications of DNA metabarcoding within sea lamprey in the Great Lakes. While Johnson et al. (2021) provided a strong basis for this technique, the main objective of their study was to provide a proof-of-concept focused on a small subset of sea lamprey within northern Lake Huron over a single year. In an effort to expand upon that research, this study focuses on a larger-scale application of this technique across three lakes (Champlain, Huron, and Superior) and two years (2022 and 2023) for both adult-stage and parasitic juvenile-stage sea lamprey. These lakes differ in their fish communities (e.g. Atlantic salmon provide an important prey source for Lake Champlain sea lamprey; Marsden et al. 2003), which I expect to be reflected in dietary data, given previous studies that found differential host preferences across the Great Lakes (Harvey et al. 2008; Adams and Jones 2021). Previous analyses showed that although sea lamprey cease feeding as they enter the adult stage, detections of host DNA from dietary samples are still possible (Johnson et al. 2021; Chapter 2). This provides the opportunity for unbiased sampling of sea lamprey, as traps and barriers can non-selectively collect sea lamprey as they migrate upstream during their adult stage. The ability to detect host DNA within this life stage would provide a strong complement to ongoing dietary assessments of juvenile

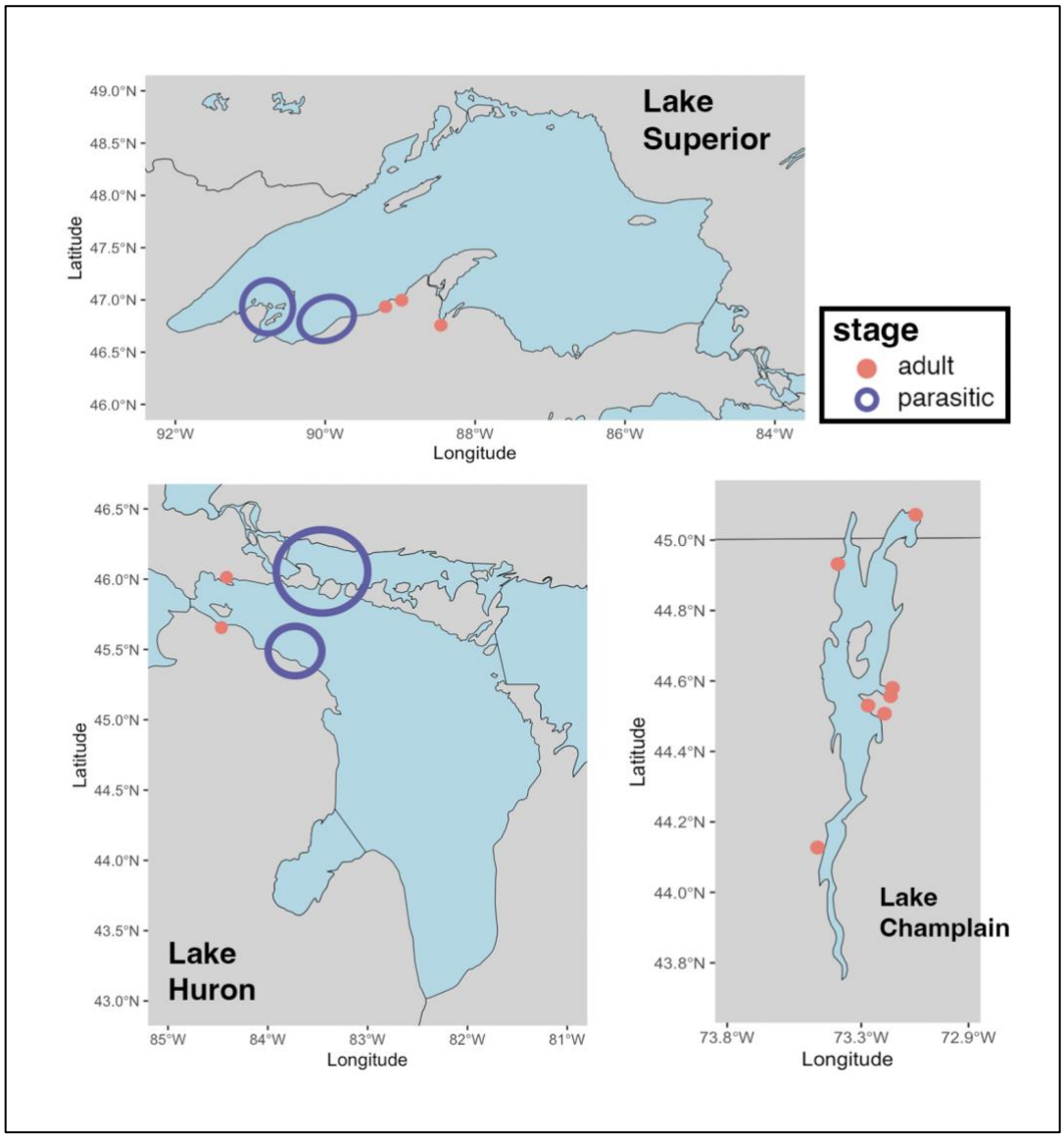
parasitic sea lamprey. Given unexpected host usage in previous lamprey DNA metabarcoding analyses, along with prior studies indicating varied host usage, I expect to see spatial and temporal variance in sea lamprey dietary composition as well as differences across adult and parasitic life stages.

METHODS

Sea Lamprey Collections

Wild-captured sea lamprey were collected from Lakes Huron, Superior, and Champlain in both 2022 and 2023. Collections included capture of both adult sea lamprey and parasitic juvenile sea lamprey. Adult sea lamprey were collected via barriers and traps in streams, while parasitic sea lamprey were removed from host fish that were targeted with fishing vessels in the basin of each lake. Data on which host fish parasitic sea lamprey were attached to was also collected from Lake Huron in 2022/2023 and Lake Superior in 2023. Stream and river collections for adults took place in the Cheboygan Rivers for Lake Huron (USFWS – Marquette Biological Station), the Misery, Falls and Firesteel rivers for Lake Superior (Keweenaw Bay Indian Community), and Mallets Creek, the Sunderland, Pond and Mullen brooks, Great Chazy and Winooski rivers, and Morpion Stream for Lake Champlain (USFWS – Lake Champlain Office; Figure 3.1).

Figure 3.1. Map of Lake Superior, Lake Huron, and Lake Champlain sea lamprey collection areas in 2022 and 2023 given available data. Capture locations for adult sea lamprey are marked with orange points, while general capture areas for parasitic sea lamprey are outlined with a purple circle. Northern parasitic collections in Lake Huron occurred during 2022 (commercial gill net fisherman) and southern parasitic collections occurred during 2023 (charter sports fisherman).



Parasitic sea lamprey collected from Lake Superior were removed from fishes captured by commercial fishermen that were targeting lake whitefish and lake trout (in coordination with the Great Lakes Indian Fish and Wildlife Commission). Parasitic sea lamprey from Lake Huron during 2022 were removed from commercial fishermen targeting lake whitefish and lake trout

(in coordination with Fisheries and Oceans Canada). During 2023, parasitic lamprey from Lake Huron were removed from fish captured by a charter recreational fisher targeting lake trout and chinook salmon. Parasitic sea lamprey from Lake Champlain were captured during trawl surveys conducted by the University of Vermont and these parasitic sea lamprey were not attached to host fishes. Following collections, all sea lamprey were frozen whole and sent to Hammond Bay Biological Station in Millersburg, MI, where they were stored at -80°C for further analysis. All sea lamprey collections were conducted in compliance with the collection permits associated with the agencies listed above.

DNA Extractions

After allowing sea lamprey to thaw, the intestinal tract was dissected and opened with a sterile scalpel (No. 11 blade, WSI disposable sterile). A sterile cotton swab (Dynarex 6-inch sterile cotton tipped applicator) was run along the length of the digestive tract to collect digestive material, preserved in RNAlater in a 1.5 mL centrifuge tube and labelled according to the stream, year and life stage of the sample. Each lamprey was photographed before and after dissection to quantify digestive material remaining in the gut. All swabs were stored at -80°C until they were prepared for extractions at the Molecular Ecology Lab at Michigan State University for further DNA analyses.

DNA extractions for all sea lamprey digestive samples used the gMax Mini Genomic DNA Kit (IBI Scientific, Dubuque, IA), following manufacturer instructions. Polymerase chain reaction (PCR) was performed using the vertebrate-specific 12S-V5 primer set (Riaz et al. 2011), which targets a ~140bp region (including primers) of the mitochondrial 12S rRNA gene. Due to the likelihood of a high presence of lamprey DNA within the samples, a lamprey-specific blocking primer (Blocker 6) was included in PCR reactions for all samples to suppress

amplification of 12S sea lamprey sequences (see Chapter 1). Two technical PCR replicates were included for each sample.

PCR reactions were conducted in a 15 µL volume with 1.5 µL of 10X AmpliTaq Gold PCR Buffer II (Applied Biosystems, Waltham, MA), 0.36 µL of dNTPs (10mM), 1.2 µL of $MgCl_2$ (25mM), 0.75 µL of BSA (20 mg/mL), 0.8 µL of forward and reverse primers (10 µM) and sea lamprey Blocker 6 blocking primer (100 µM), 6.54 µL of Milipore water (UV treated), 0.25 µL of AmpliTaq Gold (5U/µL) and 2 µL of template DNA. A negative control for each PCR plate was included by substituting DNA template with water. Thermal conditions for PCR were as follows: 10 min at 95°C (1x); 30s at 95°C, 30s at 57°C, 45s at 72°C (40x); and 5min at 72°C (1x). Following amplification, 4 µL of PCR product from a subset of samples from each PCR plate and 2.5 µL of glycerol loading dye was run on a 1% agarose gel with GelRed stain (0.5x; Biotium, Fremont, CA). Gels were photographed under UV light in a Labnet Enduro GDS II imaging system (Labnet HQ, Edison, NJ) to verify successful amplification.

Metabarcoding and Bioinformatics

Dual-index barcoding was used for all digestive samples to allow demultiplexing of sequencing reads to their corresponding samples. The indexing PCR was conducted in a final volume of 10 µL, including 2X Qiagen Plus master mix (5 µL), i7 (1µL; 10 µM) and i5 (2µL; 5 µM) index primers, and 2 µL of template. Reaction conditions in the indexing PCR were as follows: 15 min at 95°C (1x); 10s at 95°C, 30s at 65°C, 30s at 72°C (10x); and 5min at 72°C. Dual-indexed PCR products were then pooled and bead size selected at 0.5x and 1.2x bead concentrations to remove long and short sequences, respectively. Pooled libraries were diluted and analyzed via TapeStation (assay High Sensitivity D100 ScreenTape) for sequence length confirmation before being sequenced on a 300 cycle Illumina MiSeq lane (v2 Standard; 2x150

bp paired end) at Michigan State University's Research Technology Support Facility. To account or sequencing biases, technical replicates for each sample were sequenced on separate lanes.

Raw sequencing data were processed using computational resources provided by the Michigan State University High-Performance Computing Center and the software package mothur (version 1.48.0 Schloss et al. 2009) and followed the same bioinformatic protocol as Chapter 2. A maximum of two mismatches was allowed between the barcodes and sample sequences during demultiplexing to help with minor sequence errors, and sequences were trimmed to retain overlapping sections to minimize discrepancies. Initial screening removed sequences with ambiguities and identified unique sequences. These sequences were aligned to a reference database containing 220 sequences ranging from 89 to 107 bp from the 12S mitochondrial rRNA gene region for 149 Great Lakes fish species. Sequences that did not align within this range to the reference database were removed, and the remaining sequences were clustered with zero differences allowed. Chimeric sequences were filtered out before taxonomic classification with a default confidence score cutoff of 80% or higher. A distance matrix was then calculated with a cutoff of 0.03, followed by clustering with a 0.01 cutoff. A 99% similarity threshold was used to group sequences into OTUs to help differentiate closely related species. Any sequences that could not be taxonomically distinguished below the family level (e.g. Salmonidae unclassified OTUs) were subjected to BLAST (NBCI) analysis in an attempt to improve taxonomic resolution. An additional OTU for the subfamily Coregoninae was created to include all sequences with high similarity to 12S sequences from the genera *Coregonus*, *Prosopium*, and *Stenodus* to allow for differentiation of whitefish from other salmonids. Notably, there were 0-4 bp differences among six sequences representing five species across these genera in our reference database for the 12S locus. This output was then cleaned and organized into a

community matrix table with read count numbers for every OTU from each sample using R/Rstudio (v.2023.12.1+402).

Data Analysis and Statistical Methods

To incorporate both technical PCR replicates but avoid pseudoreplication, the averaged read count from both replicates was used to represent the number of sequence reads per sample. Read count data were also converted into detection data using two thresholds: total sequence read count for that OTU was at least 20 total reads and relative read abundance (RRA) was at least 1% of the total reads for that sample. Initially, the median read count value of all negative samples for all OTUs was used to represent a baseline minimum sequence read count. But as this value was only six reads, a more conservative approach was taken and 20 sequence reads was selected to represent a minimum sequence copy threshold for detection. The basis for these thresholds was derived from Drake et al. (2022).

Collection data were summarized based on the lake, life stage, year, OTU total, sample size, and weight and length measurements. Any OTUs that did not register at least one detection were removed from the total OTUs for that category. As length (mm) and weight (g) differ between juvenile and adult life stages, these variables were further analyzed for significant differences with ANOVA and subsequent Tukey's HSD post-hoc tests. All OTUs that registered at least 1,000 total sequence reads or greater than one detection were included in downstream analyses. Additionally, reads for non-prey species (e.g. *Homo sapiens* and Petromyzontidae) and ambiguous OTU classifications above the family level (e.g. Chordata unclassified) were removed, except in the case of the order-level classification of Acipenseriformes. Host data from the known, primary hosts that juvenile parasitic sea lamprey were attached to when collected were also summarized. Total detections of the primary host were compared to the total number

of samples indicated for that host species. Additionally, the top five OTUs detected in those samples (excluding unclassified OTUs at or above the family-level) were listed to assess how often juveniles had detections from multiple hosts.

Principal coordinates analysis (PCoA) was performed using Bray-Curtis distances to measure differences in relative dietary composition between sea lamprey of different lakes, years, and life stages. Analyses were limited to only OTUs that produced at least 1,000 total reads. The first two principal coordinates were plotted with 95% confidence interval ellipses around centroids for each group. Lake differences were compared by categorizing sea lamprey samples into four groups containing each combination of life stage (juvenile / adult) and year (2022 / 2023). Year differences were compared by categorizing samples using each combination of lake and life stage, and life stage differences were compared via categories using each combination of lake and year. Sample sizes were included for each category, and the percentage of explained variance of the first two principal coordinates was derived from eigenvalues. Variables with the highest correlation to the first principal component were interpreted as the primary contributor to variance within each PCoA. To assess the significance of differences between groups, PERMANOVA was performed on the Bray-Curtis distances using the adonis2 function in the *vegan* R package (Oksanen 2010). Lake, year, and life stage were used as independent variables along with their interactions. Further PERMANOVAs were performed on groups that had two of either lake, year, or life stage controlled for across samples, allowing closer investigations into differences between a single variable.

To assist with explanations for differences found between sample groups, collection data were summarized for the number of samples collected throughout each season. All collections took place between April and October of their respective years, so summaries did not go beyond

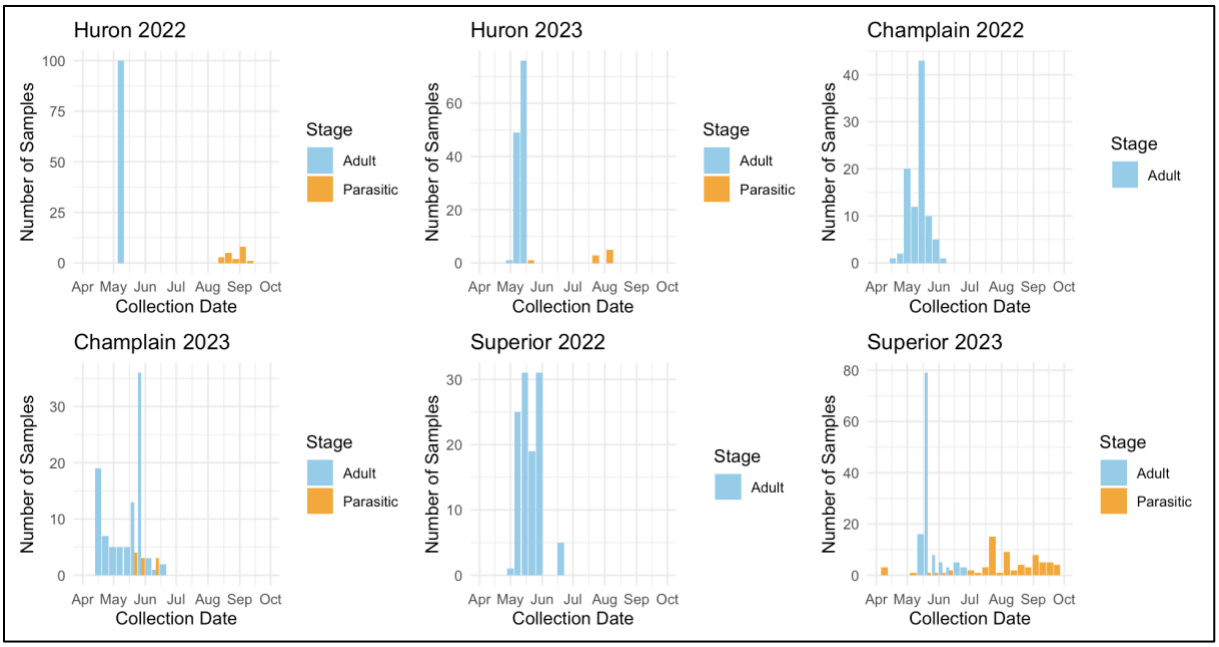
this timeframe. Collection data were available for all groups included in the PCoA and PERMANOVA analyses except for Lake Superior 2022 parasitic-stage sea lamprey.

RESULTS

Collection Data

Adult lamprey collections took place between April and June for both years across all three lakes (Figure 3.2). Collections of parasitic-stage lamprey took place during August through October for all documented subsets except for Champlain 2023, during which parasitic lamprey were collected between June and July. While parasitic lamprey were collected from Lake Superior during 2022, information on collection dates was not available and not included in analyses. All subsets with documented collection data had lamprey collected at multiple weeks throughout each period for both adult and parasitic stages, besides adult lamprey from Lake Huron in 2022, which were all collected during the same week.

Figure 3.2. Number of sea lamprey digestive samples obtained and their given collection dates for each lake and year. Adult sea lamprey samples are in blue, and parasitic sea lamprey samples are in orange. Collection dates were not available for Superior 2022 parasitic sea lamprey.



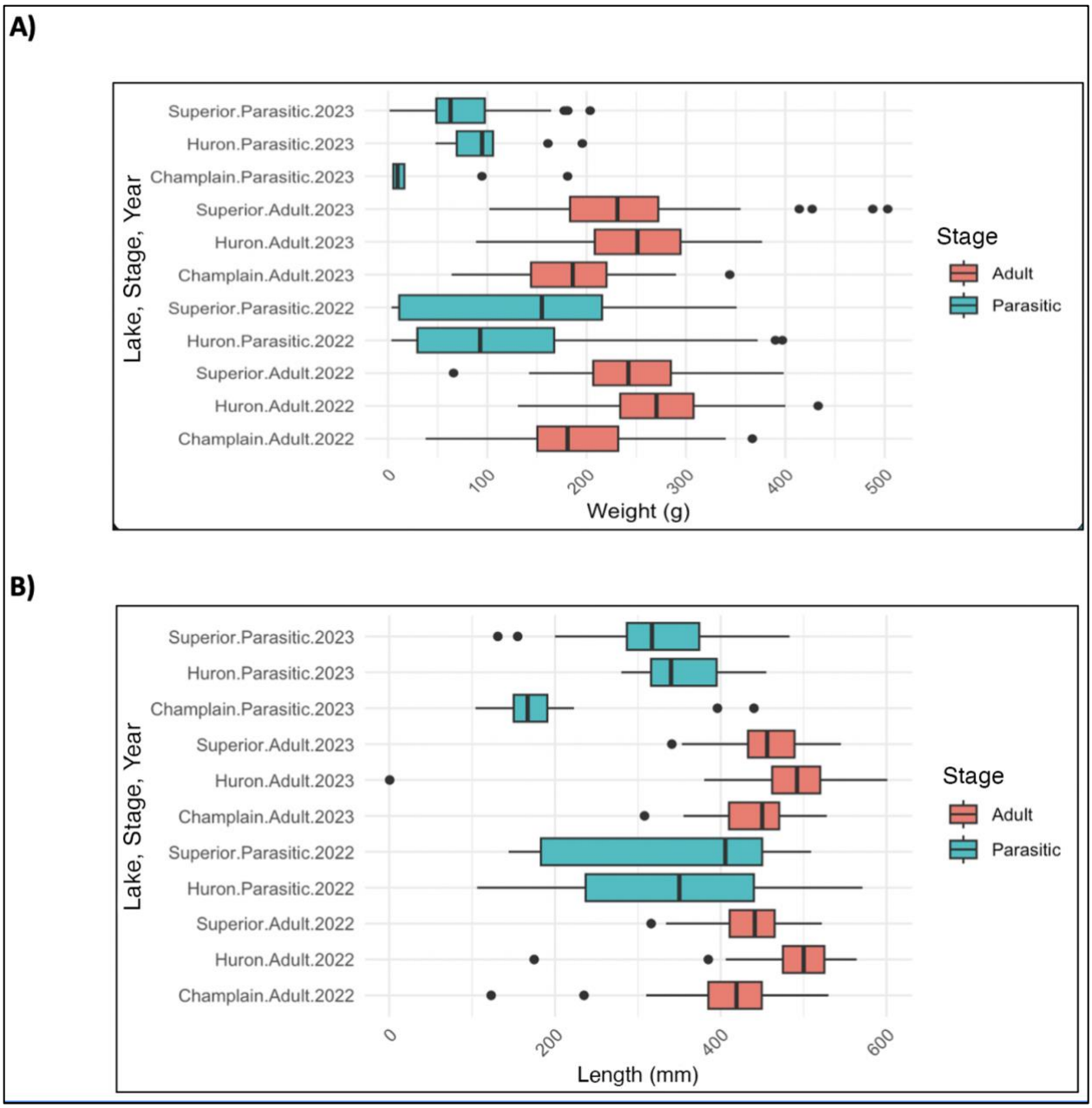
Length and Weight Comparisons

A total of 929 sea lamprey were collected and analyzed for dietary composition (Table 3.1). A significant difference was seen in both length ($p < 0.001$) and weight ($p < 0.001$) between parasitic and adult sea lamprey, with adults being an average of 111 mm longer and weighing 112 g more (Figure 3.3). Additionally, a significant difference ($p < 0.001$) was found in length measurements between years, with lamprey captured in 2023 averaging 28 mm longer than lamprey captured in 2022 when accounting for stage type. No significant differences in length were found between lakes, but weight showed lake-specific significant differences, with Lake Superior sea lamprey weighing on average 23.5 g more than Lake Champlain lamprey ($p = 0.001$) and 15.9 g more than Lake Huron lamprey ($p = 0.009$) when accounting for life stage. Weight did not differ significantly between Lake Huron and Lake Champlain sea lamprey, and there were no significant differences in weights between years.

Table 3.1. Summarization of operational taxonomic units (OTU) totals, sample size (N), weight (mean and standard deviation), and length (mean and standard deviation) for each lake-stage-year subset of sea lamprey samples.

Lake	Phase	Year	OTU Total	N	Weight (mean)	Weight (sd)	Length (mean)	Length (sd)
Champlain	Adult	2022	20	95	190.0	62.3	413	56
Huron	Adult	2022	11	65	267.4	58.8	493	55
Superior	Adult	2022	22	113	249.6	59.3	438	38
Champlain	Parasitic	2022	0	0	NA	NA	NA	NA
Huron	Parasitic	2022	15	233	119.3	103.7	344	120
Superior	Parasitic	2022	7	10	148.7	118.6	354	135
Champlain	Adult	2023	26	94	185.0	53.5	443	43
Huron	Adult	2023	16	109	247.3	65.9	486	64
Superior	Adult	2023	16	133	234.1	67.6	458	42
Champlain	Parasitic	2023	6	12	30.4	53.5	204	104
Huron	Parasitic	2023	4	9	100.1	49.6	356	56
Superior	Parasitic	2023	9	56	76.3	47.9	325	80

Figure 3.3. Weight (g) for sea lamprey at each lake-year-stage subset (A) and length (mm) of sea lamprey at each lake-year-stage (B). Boxes represent the interquartile range, with whiskers extending to include points within 1.5x that range. Points represent single outliers. Adult sea lamprey are indicated by red, parasitic sea lamprey are indicated by blue.

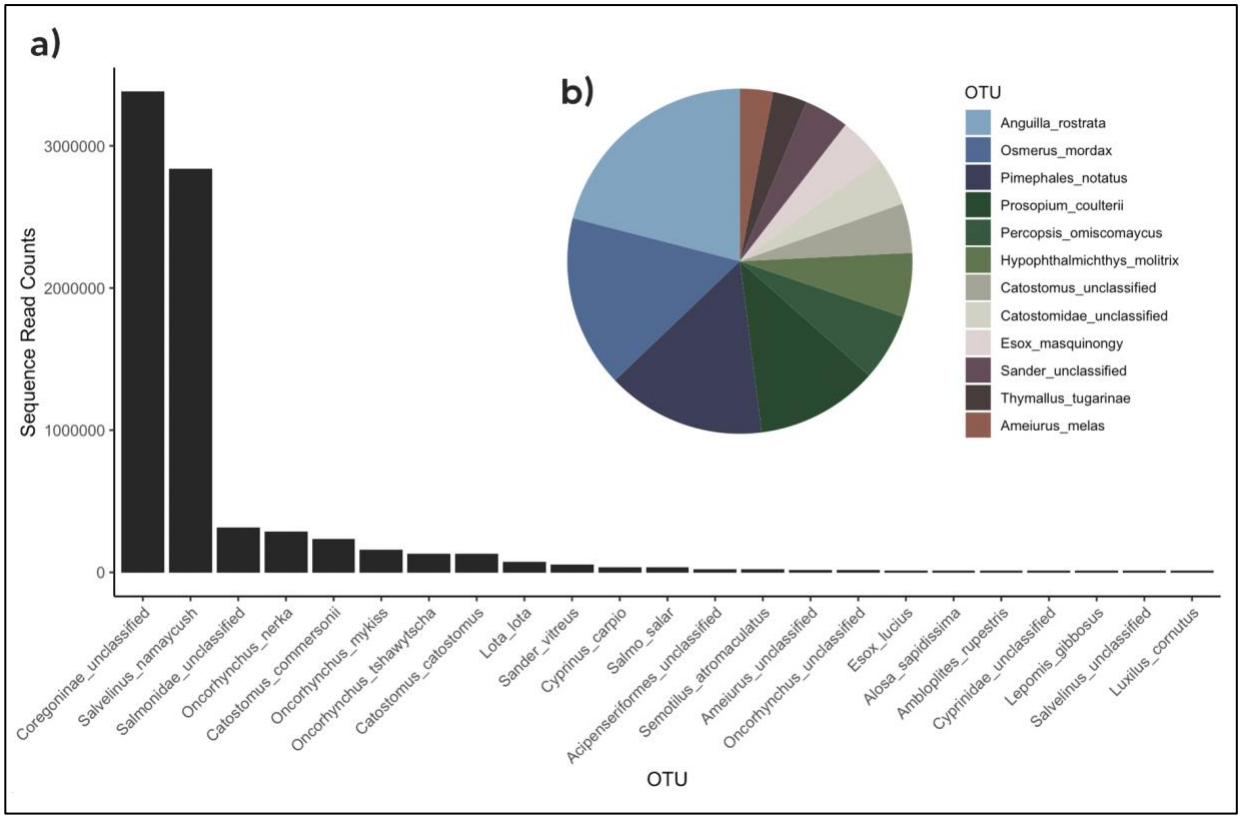


Read Count and OTU Summaries

A total of 12,473,006 sequence reads were produced from the 929 sea lamprey digestive samples analyzed. This resulted in 62 unique OTUs detected across all samples. Detected OTUs

were summarized via sequence read counts into those with at least 10,000 total reads (Figure 3.4a) and those with < 10,000 total reads but > 1,000 total reads (Figure 3.4b). The largest OTUs in terms of total sequence reads were Coregoninae unclassified (whitefish) and *Salvelinus namaycush* (lake trout), making up 42.1% and 34.3% of total sequence reads, respectively. This is likely due to the large portion of parasitic juvenile sea lamprey that were directly captured attached to either whitefish or lake trout. Many reads were indistinguishable within the family of Salmonidae (468,837 total reads), even after sequences were analyzed via BLAST, likely due to the similarity within the 12S gene region among salmonids. For the family Salmonidae in the reference database, interspecific sequence differences ranged from 0-10 bp, while intraspecific differences (for species with multiple sequences in the alignment) ranged from 1-4 bp.

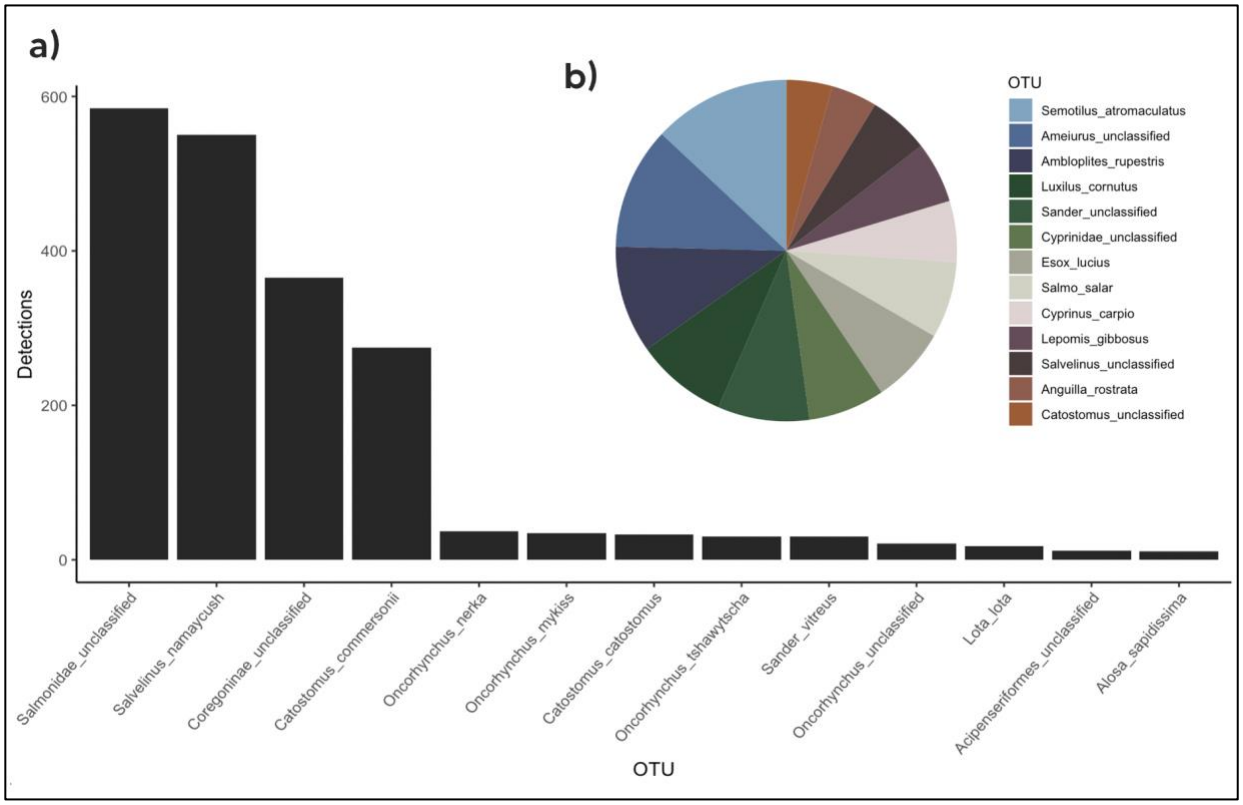
Figure 3.4. Total sequence read counts for each OTU for all samples. Bars represent OTUs that contained a total of at least 10,000 sequence reads (a), and sections of the pie chart (b) represent OTUs that contained less than 10,000 total sequence reads but above 1,000 total sequence reads. Names of each OTU are listed, with "unclassified" signifying those sequence reads could not be taxonomically distinguished within the stated taxa, and have been grouped at the genus or family level. All OTUs that were unclassified above the family level (besides Acipenseriformes) were removed.



Total OTU detections were also summarized after reads were converted to binary counts (Figure 3.5). OTUs with one or less detections (n = 25) were excluded from this figure as they are unlikely to contribute substantially to dietary compositions. The top four OTUs in terms of binary detections were Salmonidae, lake trout, whitefish, and *Catostomus commersonii* (white sucker). Salmonids were found in 63.0% of all samples, with species-specific detections of lake trout in 55.2%, whitefish in 32.7% and white sucker in 26.3%. In total, these four OTUs accounted for 85.1% of all registered detections (n = 2085). Minor detection OTUs included other salmonid species such as *Onchorhyncus mykiss* (rainbow trout), *Onchorhyncus nerka*

(sockeye salmon), and *Onchorhyncus tshawytscha* (Chinook salmon) with detections in 3.4%, 3.3%, and 2.8% of samples, respectively. Other notable species-level detections include *Catostomus catostomus* (longnose sucker), *Sander vitreus* (walleye), *Lota lota* (burbot), *Alosa sapidissima* (American shad), and *Semotilus atromaculatus* (creek chub).

Figure 3.5. Total detections for each OTU for all samples given thresholds of > 20 total sequence reads and > 1% relative read abundance within a sample. OTUs with at least 10 total detections are displayed on the bar chart, and OTUs with total detection counts between 2 and 10 displayed on the pie chart. Names of each OTU are listed, with "unclassified" signifying those sequence reads could not be taxonomically distinguished within the stated taxa, and have been grouped at the genus or family level. All OTUs that were unclassified above the family level (besides Acipenseriformes) were removed.



Detections of OTUs across all samples can be further broken down by life stage, with 320 total parasitic digestive samples and 609 adult digestive samples (Table 3.2). While both adult and parasitic samples had detections of unclassified salmonids (63.4% and 62.2%, respectively), there were differences between lake trout, Coregoninae and white sucker proportions. A higher

percentage of adult sea lamprey samples had lake trout detections at 69.1% versus only 40.3% of parasitic sea lamprey samples producing lake trout detections. Coregoninae detections were proportionally higher in parasitic samples at 57.5% detections, while only 29.7% of adult samples had detections. The largest difference was seen with white sucker, with detections in 44.7% of adult samples and less than 1% of parasitic samples at only three total detections. A sizable difference was also seen in the number of samples from both life stages that had zero detections for any OTU. Adult sea lamprey a had much higher proportion of non-detections at 23.2% when compared to parasitic sea lamprey at only 0.9%.

Table 3.2. Summary of total OTU detections (top ten OTUs by detection count) divided into adult samples and parasitic samples. OTUs represented with "unclassified" indicated a family or genus-level OTU where sequences could not be taxonomically distinguished below the given level. Percent represents the proportion of detections for the given OTU as compared to the total number of detections among samples of the same stage (adult or parasitic). Total non-detections indicate the number of samples for each stage subset that did not register any detections of any OTU.

Adult Samples (n = 609)

OTU	Detections	Percent
Salvelinus namaycush	421	69.1
Salmonidae unclassified	386	63.4
Catostomus commersonii	272	44.7
Coregoninae unclassified	181	29.7
Oncorhynchus mykiss	26	4.3
Catostomus catostomus	26	4.3
Sander vitreus	26	4.3
Oncorhynchus tshawytscha	20	3.3
Oncorhynchus nerka	13	2.1
Lota lota	11	1.8
Total Non-Detections	141	23.2

Parasitic Samples (n = 320)

OTU	Detections	Percent
Salmonidae unclassified	199	62.2
Coregoninae unclassified	184	57.5

Table 3.2 (cont'd)

Salvelinus namaycush	129	40.3
Oncorhynchus nerka	24	7.5
Oncorhynchus unclassified	13	4.1
Oncorhynchus tshawytscha	10	3.1
Oncorhynchus mykiss	8	2.5
Catostomus catostomus	7	2.2
Lota lota	7	2.2
Sander vitreus	4	1.3
Total Non-Detections	3	0.9

A total of 206 parasitic juvenile samples, all from either Lake Huron or Lake Superior, had data on the known, primary host they were attached to when collected (Table 3.3). Whitefish had the highest number of samples at 119, and a detection rate of 93.3% with 111 of those samples registering Coregoninae detections. Lake trout had a lower detection rate with 31 detections out of 48 known samples. Beyond these two species, sample sizes were lower with the 10 other taxa listed as primary hosts having 10 or less samples. While whitefish had more samples overall, the most detections of other hosts occurred in lake trout samples with 26 out of 48 samples registering detections for a host fish besides lake trout, thus implying those sea lamprey had fed on multiple hosts.

Table 3.3. Primary (known) host detections from juvenile parasitic sea lamprey from Lake Huron and Lake Superior collections. Primary host indicates the host fish that collected sea lamprey were attached too. Samples indicate how many sea lamprey were attached to that primary host, and primary host detections are the number of those lamprey where the primary host was detected. Total samples with other hosts are the number of samples within that primary host subset where hosts beyond the primary host were detected. The top five OTUs for each primary host subset are also listed (besides unclassified Salmonidae for clarity).

Primary Host	Samples	Primary Host Detections	Total Samples with Other Hosts	Top Five OTUs Among Samples
Whitefish	119	111	14	Coregoninae (111), *Oncorhynchus nerka* (9), *Salvelinus namaycush* (5), *Oncorhynchus spp.* (5), *Oncorhynchus tshawytscha* (1)
Lake Trout	48	31	26	*Salvelinus namaycush* (31), Coregoninae (18), *Oncorhynchus nerka* (3), *Catostomus commersonii* (2), *Lota lota* (2)
Pink Salmon	10	6	4	*Oncorhynchus nerka* (6), Coregoninae (3), *Oncorhynchus spp.* (3), *Oncorhynchus mykiss* (1), *Salvelinus spp.* (1)
Rainbow Trout	10	5	5	*Oncorhynchus mykiss* (5), Coregoninae (4), *Oncorhynchus nerka* (1)
Chinook Salmon	5	5	1	*Oncorhynchus tshawytscha* (5), *Salvelinus namaycush* (1)
Burbot	3	0	3	Coregoninae (2), *Oncorhynchus nerka* (1)
Herring	3	0	3	Coregoninae (3), *Salvelinus namaycush* (1)
Pickerel	3	0	3	*Salvelinus namaycush* (2), Coregoninae (1)

Table 3.3 (cont'd)

Salmon (unspecified)	2	1	2	Coregoninae (1), *Oncorhynchus tshawytscha* (1)
Atlantic Salmon	1	0	1	*Salvelinus namaycush* (1)
Sucker (unspecified)	1	1	0	*Catostomus catostomus* (1)
Walleye	1	1	1	Coregoninae (1), *Salvelinus namaycush* (1), *Catostomus commersonii* (1), *Sander vitreus* (1), *Oncorhynchus spp.* (1)

PCoA and PERMANOVA Analyses

Strong overlaps in diet were observed between adults, but much less overlap between parasitic juveniles occurred according to visual PCoA assessments of comparisons between lakes (Figure 3.6). When assessing differences between life stages (Figure 3.7), the spread of data points and corresponding ellipses for adults is broader than juveniles. This may indicate either less variability in juvenile diets or be a reflection of smaller samples sizes in certain subsets. Differences between years (Figure 3.8) are more pronounced in parasitic juveniles, although may be due to smaller sample sizes in the Champlain and Superior subsets. The largest amounts of variation were captured by the first and second principal components for all comparisons, so variables with the highest correlation to these principal components were interpreted as the primary contributors to variance within each PCoA. Dietary differences between lakes were best explained by variation in the presence of whitefish, except for adult samples from 2022, where variation was primarily associated with lake trout (PCoA1; Figure 3.6). For differences between life stages, variation was mainly associated with whitefish in both years for Lakes Huron and Superior, but differences in lake trout were the primary explanation of variance for both years of Lake Champlain samples (PCoA1; Figure 3.7). Whitefish were the most significant contributors to variance across all categories for differences between years (PCoA1; Figure 3.8).

Figure 3.6. Principal coordinate analyses (PCoA) for stage-year subset groups looking at differences between lakes. Lakes are indicated by color, with Champlain as orange, Huron as blue, and Superior as green. All plots use the first and second principal coordinates. Samples sizes of each subset are included in the titles, with each point representing a single sea lamprey sample.

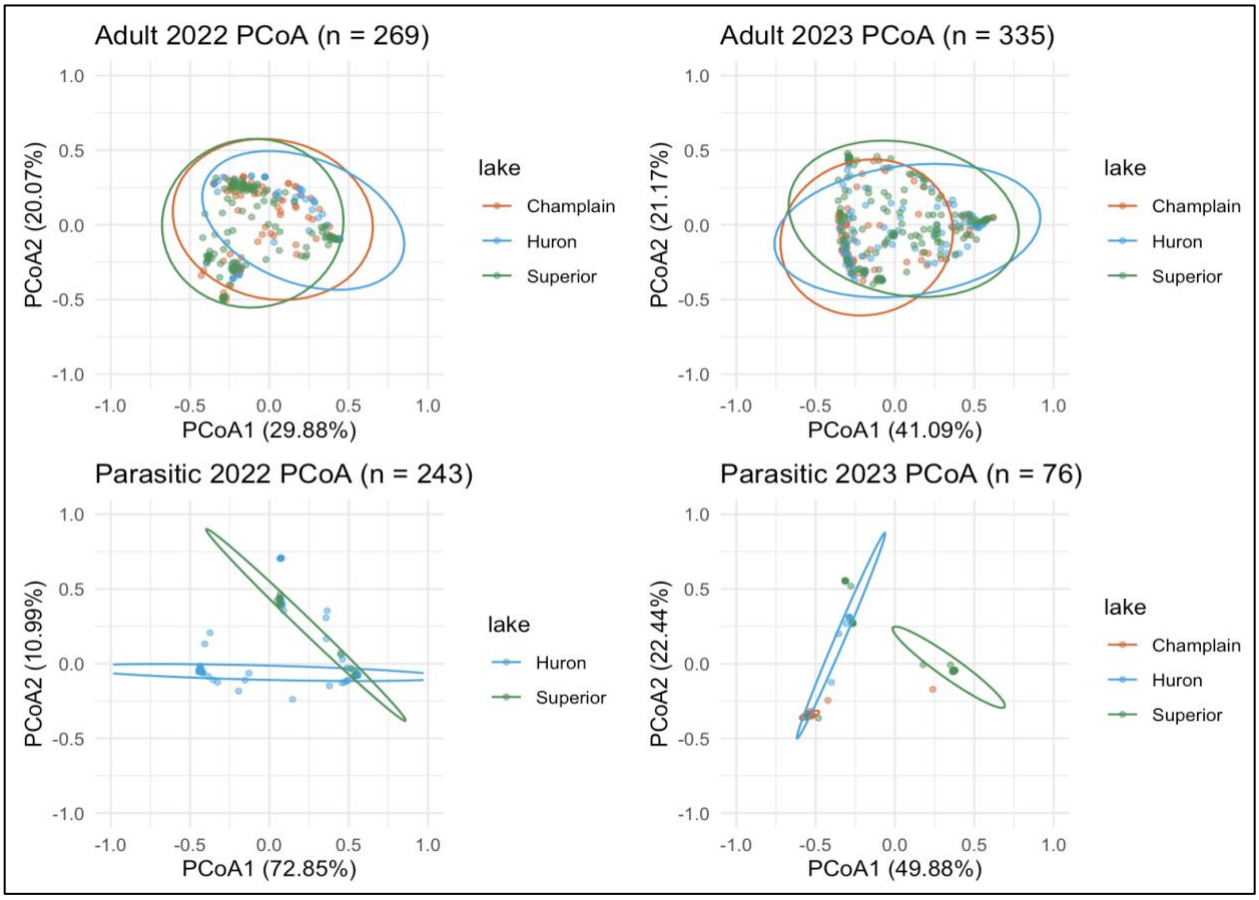


Figure 3.7. Principal coordinate analyses (PCoA) for lake-year subset groups looking at differences between stages. Stages are indicated by color, with adults as red and parasitic as blue. All plots use the first and second principal coordinates. Samples sizes of each subset are included in the titles, with each point representing a single sea lamprey sample.

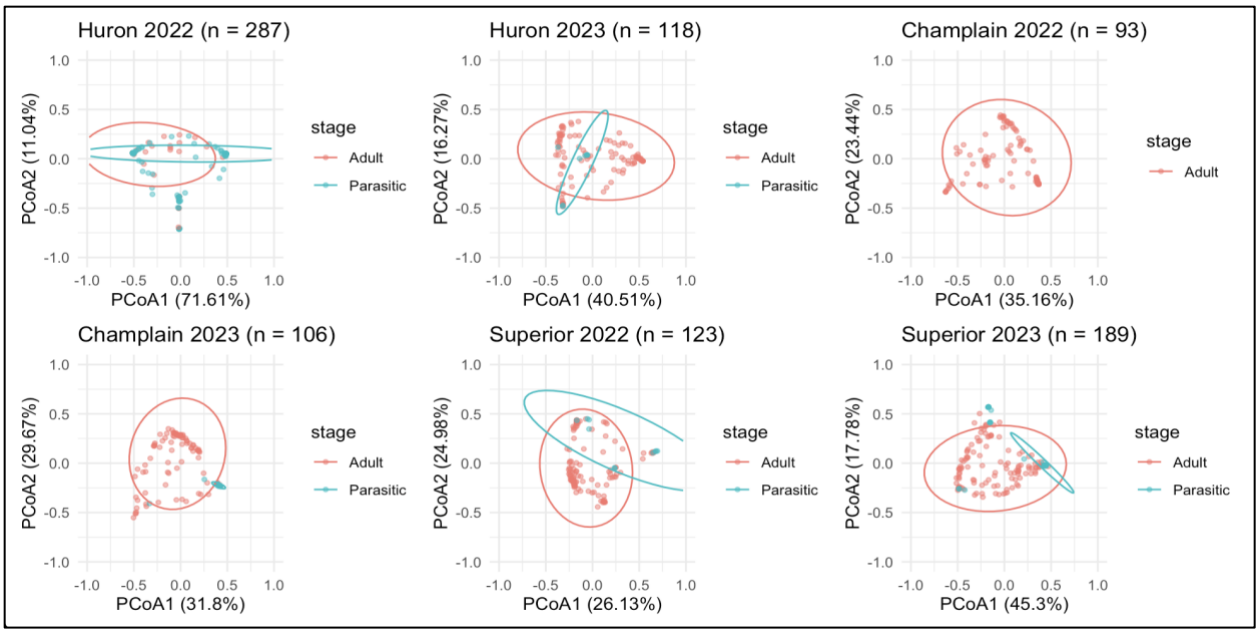
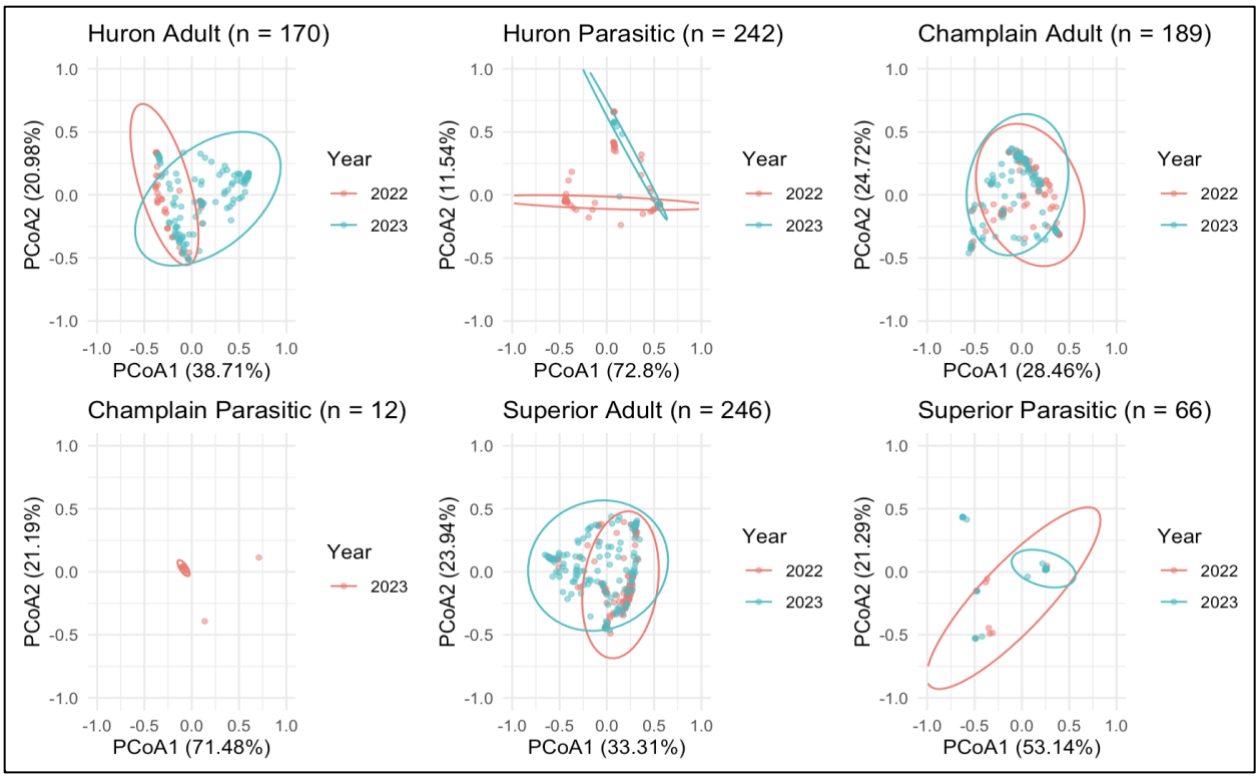


Figure 3.8. Principal coordinate analyses (PCoA) for lake-stage subset groups looking at differences between years. Years are indicated by color, with 2022 sea lamprey as red and 2023 sea lamprey as blue. All plots use the first and second principal coordinates. Samples sizes of each subset are included in the titles, with each point representing a single sea lamprey sample.



Lake, year, and life stage, along with their interactions, were significant influences on dietary composition ($p = 0.001$) according to PERMANOVA analyses (Table 3.4). Lake and life stage explained similar proportions of variance in dietary composition across samples at 7.2% and 7.9%, respectively, while year explained less variance at 3.3%, as indicated by partial R^2 values. The year:life stage interaction ($R^2 = 0.8\%$) and the three-way interaction ($R^2 = 0.8\%$) terms explained less variance. Overall, a substantial fraction of the variance in dietary composition among samples was unexplained by the independent variables included in PERMANOVA analyses. More targeted effects of these variables on dietary composition and consistency within specific subsets of data were investigated by additional PERMANOVA analyses (APPENDIX A: Table S3.1). Particular instances of high variance explanation were the

effect of lake on 2023 parasitic samples (R^2 = 35.5%) and the effect of life stage on Huron 2022 samples (R^2 = 21.9%). All effects from lake, year, and life stage on dietary composition were significant ($p < 0.05$).

Table 3.4. PERMANOVA summary looking at the overall influence (R^2) of lake, year, and life stage, along with their interactions, on dietary differences among sea lamprey samples. Analyses were conducted on the relative read counts (*individual read count/total sample read count*) of each OTU within each sample.

Variable	R2	p-value
lake	0.0720	0.001
year	0.0334	0.001
life stage	0.0794	0.001
lake:year	0.0153	0.001
lake:life stage	0.0340	0.001
year:life stage	0.0080	0.001
lake:year:life stage	0.0083	0.001

DISCUSSION

Molecular diet analysis of Great Lakes sea lamprey provided a detailed inventory of dietary compositional differences between sea lamprey caught during two years (2022, 2023), three lakes (Huron, Superior, Champlain), and two life stages (adult, parasitic). This work capitalizes on recent developments in genetic technology that have allowed for deeper insights of predator diets beyond traditional methods. While previous dietary analyses in Great Lakes sea lamprey have documented important spatial and ontogenetic differences in diets (Harvey et al. 2008; Happel et al. 2017), the lack of species-level identifications has limited our understanding of sea lamprey host preference and diet composition. My results lay the groundwork for future studies that may further resolve complications associated with sampling constraints, PCR

amplification stochasticity, and taxonomic resolution (e.g., for some members of the family Salmonidae).

Dietary Composition and Host Detections

Lake trout and whitefish have typically been thought to comprise the majority of sea lamprey diets within the Great Lakes (Harvey et al. 2008). However, recent studies involving fatty acid analysis and DNA metabarcoding have pointed at a wider variety of prey species for Great Lakes sea lamprey (Happel et al. 2017; Johnson et al. 2021). Our results support both of these suggestions, with Coregoninae (whitefish) and lake trout comprising a significant amount of total sequence reads from the study, but with 37 other OTUs represented by at least two detections across all samples. This adds further support to the notion lake trout and whitefish are critical components to sea lamprey diets in the Great Lakes, but showcases the variety of prey items attacked by sea lamprey. As sequence read counts may not directly translate into proportional dietary information for the relative abundance of each prey item (Elbrecht and Leese 2015; Piñol et al. 2015; Jusino et al. 2019; although see Sard et al. 2019), binary detection data can supplement read count data with a different perspective. From this approach, white sucker appears as a common prey item as well (detected in 26.3% of samples), supporting findings from Johnson et al. (2021). However, these common prey items appear with different detection distributions between adult and parasitic sea lamprey (Table 3.2). Unclassified Salmonidae detections were present in high proportions for both life stages. However, adult sea lamprey samples had higher proportions of lake trout and white sucker, while parasitic sea lamprey produced higher proportions of Coregoninae. This difference may be explained by the size of sea lamprey during these life stages. As parasitic-stage sea lamprey have lower length and weight on average (Figure 3.3), it is possible that parasitic sea lamprey prefer smaller species

such as those within the Coregoninae subfamily (see Harvey at el. 2008). At the adult life stage where sea lamprey are larger, this may allow for targeting large prey species, such as lake trout, or species with similar spawning migrations, such as white sucker. Differences in OTU detection proportions displayed between parasitic and adult sea lamprey provide further evidence that sea lamprey exhibit ontogenetic shifts in dietary preferences. Although, it should be noted that sampling for adults and parasites occurred at different locations and time periods.

Other species with at least 10 total detections in the dataset, including rainbow trout, longnose sucker, Chinook salmon, walleye, burbot, and sturgeon (Acipenseriformes), highlight the variety of dietary options for Great Lakes sea lamprey. Although, these species likely only have minor impacts on overall dietary composition as they were detected in less than 4% of samples. Of note is the detection of *Onchorhyncus nerka* in 3.3% of samples. While originally stocked in Lakes Huron and Ontario to aid with recreational and commercial fishing efforts, sockeye/Kokanee salmon stocking efforts have not been ongoing since the 1970s (Crawford and Canada 2001). Therefore, these detections are likely due to the similarity across *Onchorhyncus* species in the 12S rRNA gene region, with 0 bp differences within our reference dataset between *Onchorhyncus nerka* and *Onchorhyncus gorbuscha* (pink salmon), and only 2-3 bp differences between *O. nerka* and rainbow trout or Chinook salmon. As our reference database will not fully capture intraspecific variation, it is possible that other 12S variants present in these *Onchorhyncus* species match our sockeye salmon reference sequence even more closely. This idea is supported by the majority of samples with pink salmon as known hosts registering detections for *O. nerka* (Table 3.3). The genetic similarity among salmonids at 12S is highlighted with the high presence of unclassified salmonid detections across all samples, although likely does not have a large impact on the interpretation of results as total sequence reads for

unclassified salmonids (after additional classifications from our BLAST analysis) were relatively low.

Further investigation into the primary host detections (Table 3.3) shows differing proportions for detections of known hosts. DNA metabarcoding is an imperfect detection method, so a 93.3% detection rate of whitefish in samples where sea lamprey were attached to a whitefish is encouraging. Even though other primary hosts, such as trout and salmon species, did not replicate this high detection level, most known host species were still detected in the majority of their respective samples. An important consideration is that attachment time to these primary hosts is unknown. As Chapter 2 demonstrated, at zero fasting days, the probability of host detections can be as low as 30-40% if the sea lamprey has only been attached to its host for a single day (Figure 2.6). This may account for some instances where known hosts were not detected in the digestive sample. Additionally, these differences may be exacerbated by amplification bias. If Coregoninae were found to have a positive amplification bias at this marker, this would also help explain the relatively larger amounts other host species detections in samples where whitefish was not the primary host and relatively low amounts of other host detections in samples where whitefish was the primary host. This points toward some potential improvements for this research. It may be beneficial to include a mock community analysis in future studies to look for potential amplification biases, thus allowing for the incorporation of correction factors. Additionally, markers besides 12S may be considered, as differences in genetic similarity at other gene regions may allow for increased clarity into host DNA detections.

Dietary Variation Across Lakes and Life Stages

Lake-specific dietary variation was an important factor given interpretations of PCoA and PERMANOVA, both with overall significance for all samples and significance for specific

subsets of data. For example, when considering only parasitic sea lamprey collected in 2023, the lake from which sea lamprey were captured had a significant effect on diet, explaining 35.5% of the variance. However, the high degree of explained variance may be due to capture biases, as commercial vessels were used to target host fish and any attached sea lamprey were then collected as bycatch. With the greatest amount of variance in this subset characterized by whitefish, it is possible that targeted capture of whitefish within Lake Superior in 2023 may have contributed to the differences I observed. This also creates a lack of consistency between variance explanations for this subset, where if lake more consistently explained variance across these life stage-year categories it may offer more support to the overall explanation of variance that lake provides on the data as a whole. However, a large R^2 value of 0.355 in a single subset may inflate the overall influence of lake differences on sea lamprey diets. Nonetheless, the high percent of explained variance still suggests the possibility of lake-wide differences in host preference among sea lamprey, particularly given all life stage-year subsets and lake overall were significant factors for explaining dietary variance ($p = 0.001$).

Similar subset analyses to categorize dietary differences between adult and parasitic stage lamprey point to ontogenetic shifts in host prey use. For Huron 2022 samples, differences in life stage explained 21.9% of dietary variance. The largest driver of variance in this subset was also whitefish, again playing a larger role in the diet of parasitic-stage lamprey. A potential confounding variable for this subset is the collection of all adult-stage lamprey within the same week. However, as adult collections display less of a sampling bias due to the non-targeted nature of stream trappings, it is unlikely that this accounts for the differences seen between life stages in Huron 2022 samples. Both Superior 2023 and Champlain 2023 subsets also exhibited a high percentage of explained variance due to differences in life stage, variance explained in both

of these subsets above 12%. While Champlain 2023 parasitic sea lamprey had a smaller sample size of 24, the observation that all three subsets exhibited variance in dietary composition as a result of life stage supports the existence of ontogenetic shifts in host use, potentially due to size differences or energy requirements between adult and parasitic-stage sea lamprey. The higher consistency in which life stage differences explain dietary variance across each of these lake-year subsets, as opposed to the lower consistency in R^2 values seen across life stage-year subsets, adds support to the overall effect of life stage differences on sea lamprey diet composition (APPENDIX A: Table S3.1). Additionally, the idea of life stage-specific dietary behavior given the distinct shapes between adult and parasitic ellipses for these three subsets was supported by PCoA analyses. This coincides with previous research conducted by Harvey et al. (2008) showing that sea lamprey mass and $\delta^{13}C$ were positively correlated in Lake Superior, suggesting ontogenetic shifts in sea lamprey diets. Other studies have also examined prey preference in sea lamprey, linking host selection to prey size (Bence et al. 2003) and prey abundance (Adams and Jones 2021). Together, along with the further evidence this study provides for ontogenetic dietary shifts, it is likely that sea lamprey host selection is influenced by a variety of factors.

Temporal changes in diet composition were also supported by PERMANOVA and PCoA analyses. Year showed a significant effect overall on all sea lamprey diet samples ($p = 0.001$), although the amount of variance explained was relatively low ($R^2 = 3.3\%$). However, the Huron adult-stage subset showed a higher amount of dietary variance explained by year (15.2%). This difference was primarily driven by the higher proportion of whitefish in 2023 Huron adult samples, and a higher proportion of lake trout in 2022 Huron adult samples. As noted previously, Huron 2022 adult samples were all collected during the same week. However, unlike differences in parasitic-stage lamprey, adult-stage collections were not associated with targeting fishing

efforts, indicating a stronger probability of detecting true differences. Additionally, Huron 2023 adult lamprey were collected during the same seasonal timeframe (early May) as Huron 2022 adult lamprey, removing the potential for within-season temporal differences to shift dietary preferences. Notably, Huron 2022 adult samples had a fairly low detection rate overall (APPENDIX A: Figure S3.3), so it is likely this could have overemphasized variation. Differences in diet composition between years for Huron adults were supported by PCoA visualizations, with distinct ellipses for 2022 and 2023 samples. While other subsets comparing temporal differences showed lower percentages of explained variation (< 6%), year was still determined to have a significant effect.

Limitations on Molecular Diet Analysis

Sequence read totals can be affected by various factors such as environmental conditions, recency of feeding, and stochasticity within the PCR and sequencing processes (Alberdi et al. 2019; Dopheide et al. 2019). This creates difficulty in how to account for rare taxa, as relative sequence abundance may not accurately reflect true dietary proportions (Deagle et al. 2019). In dietary metabarcoding studies on fish using mock communities of known prey mixtures, moderate discrepancies have been noted from 5% to 60% variation in sequence read proportions relative to the control (Ford et al. 2016; Thomas et al. 2016). However, the conversion of sequences to binary detections for increased data interpretability involves selected thresholds and may overemphasize the importance of rare taxa within the diet (Deagle et al. 2019). This effect can be exacerbated in field-based studies, where the introduction of environmental DNA is possible, in addition to the possibility of contamination during PCR and sequencing (De Barba et al. 2014; McInnes et al. 2017). While sequence removal methods such as minimum sequence copy thresholds (MSCT) and RRA cutoffs can help to reduce the influence of false negatives

(Ando et al. 2018; Drake et al. 2022), our results may be more susceptible to environmental contamination if DNA in the environment leave trace amounts of non-prey DNA within the digestive tract. This may be more applicable in stream settings, where the density of fish within the system may be higher than the open basin. Further, DNA extraction and amplification took place within a lab that has previously used a variety of fish DNA, potentially leading to contamination of trace DNA. Given this, the stricter thresholds of a 20-read MSCT and a 1% RRA per sample were selected to help to eliminate these potential sources of contamination and reduce the prevalence of false positives within our samples. Additionally, using both detection data and relative sequence read count data allowed us to investigate variance between sample subsets and observe multiple perspectives on overall dietary trends for Great Lakes sea lamprey.

While this study allowed for deeper insights into sea lamprey dietary composition, limitations remain on the interpretation of these results. Given the genetic similarity of salmonids at the 12S rRNA gene region, species-level interpretations were only possible by including a BLAST analysis to go beyond the taxonomic resolution afforded by our reference database. Still, certain Coregoninae species were unable to be specified. While a whitefish category was distinguishable from other salmonids, any further classifications of *Prosopium*, *Coregonus*, or *Stenodus* was not possible with this genetic marker. It is possible that the use of a different marker, such as 16S, which has been used previously in eDNA studies targeting fish (Sard et al. 2019; Pukk et al. 2021), may allow for further classifications in this group. Similarly, a secondary metabarcoding marker may be suitable. For example, in Thomas et al. (2022), a COI "minibarcode" was used in a secondary PCR reaction to specifically quantify salmonid proportions within the diets of harbor seals, given the unreliability of the 16S marker to distinguish between steelhead and coho salmon sequences. Additional limitations include smaller

sample sizes for certain subsets of samples. For example, Superior 2022 parasitic, Huron 2023 parasitic, and Champlain 2023 parasitic samples all included fewer than 25 sea lamprey. While data were still available for these subsets and these smaller sample sizes were still incorporated into statistical analyses, the confidence with which conclusions can be made is nonetheless hindered for comparisons with limited sample sizes. Future studies would also benefit from a wider sample range across each basin. Parasitic collections were limited to smaller geographic regions within each lake. When drawing conclusions about sea lamprey dietary composition, these confined ranges may not be fully representative of sea lamprey diets in other areas of each basin. One possible direction is to focus future sampling locations near the 29 index streams across the Great Lakes region that are currently used for mark-recapture estimates (Adams et al. 2021).

Implications for Sea Lamprey Management

Results from this study document differences in dietary composition of sea lamprey within the Great Lakes on temporal and spatial scales along with ontogenetic differences. Molecular diet analysis has the ability to improve current assessments of sea lamprey diets, which provide critical information for management and control that currently rely on wounding rate estimates. Given that our results indicate potential differences in sea lamprey diets between lakes and life stages, management and damage estimates specific to each lake or life stage may be warranted. Additionally, diet compositional variation between years suggests that sea lamprey prey preferences may depend upon host fish community composition, a conclusion that supports the findings of Adams and Jones (2021). Thus, the inclusion of lake-specific fish community data such as trawling surveys may improve damage estimates and management efforts. Furthermore, the ability to detect host fish DNA in adult lamprey caught in traps during their

upstream spawning migration allows a more comprehensive overview of sea lamprey diets by avoiding capture biases from targeted host fish capture often associated with juvenile sea lamprey collections. As shown through this study, these adult-stage lamprey offer additional insights into the potential variety of lamprey diets given the high number of OTUs found in adult lamprey. This perspective complements parasitic analyses as well, given the support for varied diet preferences between life stages as suggested by previous studies such as Harvey et al (2008). Ultimately, with the continued use of DNA metabarcoding for dietary inferences in both adult and parasitic sea lamprey, further insights can be gathered to improve our understanding of sea lamprey diet composition, the accuracy of damage estimates, and the efficiency of sea lamprey control in the Great Lakes.

LITERATURE CITED

Adams JV, Barber JM, Bravener GA, Lewandoski SA. 2021. Quantifying Great Lakes sea lamprey populations using an index of adults. Journal of Great Lakes Research. 47:S335–S346. doi:10.1016/j.jglr.2021.04.009.

Adams JV, Jones ML. 2021. Evidence of host switching: Sea lampreys disproportionately attack Chinook salmon when lake trout abundance is low in Lake Ontario. Journal of Great Lakes Research. 47:S604–S611. doi:10.1016/j.jglr.2020.03.003.

Alberdi A, Aizpurua O, Bohmann K, Gopalakrishnan S, Lynggaard C, Nielsen M, Gilbert MTP. 2019. Promises and pitfalls of using high-throughput sequencing for diet analysis. Molecular Ecology Resources. 19(2):327–348. doi:10.1111/1755-0998.12960.

Alberdi A, Aizpurua O, Gilbert MTP, Bohmann K. 2018. Scrutinizing key steps for reliable metabarcoding of environmental samples. Methods in Ecology and Evolution. 9(1):134–147. doi:10.1111/2041-210X.12849.

Ando H, Fujii C, Kawanabe M, Ao Y, Inoue T, Takenaka A. 2018. Evaluation of plant contamination in metabarcoding diet analysis of a herbivore. Sci Rep. 8(1):15563. doi:10.1038/s41598-018-32845-w.

Ando H, Mukai H, Komura T, Dewi T, Ando M, Isagi Y. 2020. Methodological trends and perspectives of animal dietary studies by noninvasive fecal DNA metabarcoding. Environmental DNA. 2(4):391–406. doi:10.1002/edn3.117.

Beamish FWH. 1980. Biology of the North American Anadromous Sea Lamprey, Petromyzon marinus. Can J Fish Aquat Sci. 37(11):1924–1943. doi:10.1139/f80-233.

Bence JR, Bergstedt RA, Christie GC, Cochran PA, Ebener MP, Koonce JF, Rutter MA, Swink WD. 2003. Sea Lamprey (Petromyzon marinus) Parasite-host Interactions in the Great Lakes. Journal of Great Lakes Research. 29:253–282. doi:10.1016/S0380-1330(03)70493-6.

Bergstedt RA, Schneider CP. 1988. Assessment of Sea Lamprey (Petromyzon marinus) Predation by Recovery of Dead Lake Trout (Salvelinus namaycush) from Lake Ontario, 1982–85. Can J Fish Aquat Sci. 45(8):1406–1410. doi:10.1139/f88-164.

Berry O, Bulman C, Bunce M, Coghlan M, Murray DC, Ward RD. 2015. Comparison of morphological and DNA metabarcoding analyses of diets in exploited marine fishes. Marine Ecology Progress Series. 540:167–181. doi:10.3354/meps11524.

Berry TE, Osterrieder SK, Murray DC, Coghlan ML, Richardson AJ, Grealy AK, Stat M, Bejder L, Bunce M. 2017. DNA metabarcoding for diet analysis and biodiversity: A case study using the endangered Australian sea lion (Neophoca cinerea). Ecology and Evolution. 7(14):5435–5453. doi:10.1002/ece3.3123.

Bohmann K, Gopalakrishnan S, Nielsen M, Nielsen L dos SB, Jones G, Streicker DG, Gilbert MTP. 2018. Using DNA metabarcoding for simultaneous inference of common vampire bat diet

and population structure. Molecular Ecology Resources. 18(5):1050–1063. doi:10.1111/1755-0998.12891.

Bonato KO, Silva PC, Carvalho FR, Malabarba LR. 2022. Trophic interactions of vampire catfishes (Siluriformes: Vandelliinae) revealed by metabarcoding analysis of stomach contents. Freshwater Biology. 67(3):542–548. doi:10.1111/fwb.13861.

Buglione M, Maselli V, Rippa D, de Filippo G, Trapanese M, Fulgione D. 2018. A pilot study on the application of DNA metabarcoding for non-invasive diet analysis in the Italian hare. Mammalian Biology. 88:31–42. doi:10.1016/j.mambio.2017.10.010.

Coble DW, Bruesewitz RE, Fratt TW, Scheirer JW. 1990. Lake Trout, Sea Lampreys, and Overfishing in the Upper Great Lakes: A Review and Reanalysis. Transactions of the American Fisheries Society. 119(6):985–995. doi:10.1577/1548-8659(1990)119<0985:LTSLAO>2.3.CO;2.

Crawford SS, Canada NRC. 2001. Salmonine Introductions to the Laurentian Great Lakes: An Historical Review and Evaluation of Ecological Effects. NRC Research Press.

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. Molecular Ecology Resources. 14(2):306–323. doi:10.1111/1755-0998.12188.

Deagle BE, Kirkwood R, Jarman SN. 2009. Analysis of Australian fur seal diet by pyrosequencing prey DNA in faeces. Molecular Ecology. 18(9):2022–2038. doi:10.1111/j.1365-294X.2009.04158.x.

Deagle BE, Thomas AC, McInnes JC, Clarke LJ, Vesterinen EJ, Clare EL, Kartzinel TR, Eveson JP. 2019. Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? Molecular Ecology. 28(2):391–406. doi:10.1111/mec.14734.

Doi H, Fukaya K, Oka S, Sato K, Kondoh M, Miya M. 2019. Evaluation of detection probabilities at the water-filtering and initial PCR steps in environmental DNA metabarcoding using a multispecies site occupancy model. Sci Rep. 9(1):3581. doi:10.1038/s41598-019-40233-1.

Dopheide A, Xie D, Buckley TR, Drummond AJ, Newcomb RD. 2019. Impacts of DNA extraction and PCR on DNA metabarcoding estimates of soil biodiversity. Methods in Ecology and Evolution. 10(1):120–133. doi:10.1111/2041-210X.13086.

Drake LE, Cuff JP, Young RE, Marchbank A, Chadwick EA, Symondson WOC. 2022. An assessment of minimum sequence copy thresholds for identifying and reducing the prevalence of artefacts in dietary metabarcoding data. Methods in Ecology and Evolution. 13(3):694–710. doi:10.1111/2041-210X.13780.

Drinkwater R, Schnell IB, Bohmann K, Bernard H, Veron G, Clare E, Gilbert MTP, Rossiter SJ. 2019. Using metabarcoding to compare the suitability of two blood-feeding leech species for

sampling mammalian diversity in North Borneo. Molecular Ecology Resources. 19(1):105–117. doi:10.1111/1755-0998.12943.

Elbrecht V, Leese F. 2015. Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer Bias and Biomass—Sequence Relationships with an Innovative Metabarcoding Protocol. PLOS ONE. 10(7):e0130324. doi:10.1371/journal.pone.0130324.

Ellen Marsden J, Chipman BD, Nashett LJ, Anderson JK, Bouffard W, Durfey L, Gersmehl JE, Schoch WF, Staats NR, Zerrenner A. 2003. Sea Lamprey Control in Lake Champlain. Journal of Great Lakes Research. 29:655–676. doi:10.1016/S0380-1330(03)70522-X.

Estrada-Franco JG, Fernández-Santos NA, Adebiyi AA, López-López M de J, Aguilar-Durán JA, Hernández-Triana LM, Prosser SWJ, Hebert PDN, Fooks AR, Hamer GL, et al. 2020. Vertebrate-Aedes aegypti and Culex quinquefasciatus (Diptera)-arbovirus transmission networks: Non-human feeding revealed by meta-barcoding and next-generation sequencing. PLOS Neglected Tropical Diseases. 14(12):e0008867. doi:10.1371/journal.pntd.0008867.

Fahmy M, Ravelomanantsoa NAF, Youssef S, Hekkala E, Siddall M. 2019. Biological inventory of Ranomafana National Park tetrapods using leech-derived iDNA. Eur J Wildl Res. 65(5):1–13. doi:10.1007/s10344-019-1305-3.

Farmer GJ, Beamish FWH, Lett PF. 1977. Influence of Water Temperature on the Growth Rate of the Landlocked Sea Lamprey (Petromyzon marinus) and the Associated Rate of Host Mortality. J Fish Res Bd Can. 34(9):1373–1378. doi:10.1139/f77-197.

Ficetola GF, Pansu J, Bonin A, Coissac E, Giguet-Covex C, De Barba M, Gielly L, Lopes CM, Boyer F, Pompanon F, et al. 2015. Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. Molecular Ecology Resources. 15(3):543–556. doi:10.1111/1755-0998.12338.

Firkus TJ, Murphy CA, Adams JV, Treska TJ, Fischer G. 2021. Assessing the assumptions of classification agreement, accuracy, and predictable healing time of sea lamprey wounds on lake trout. Journal of Great Lakes Research. 47:S368–S377. doi:10.1016/j.jglr.2020.07.016.

Fiske I, Chandler R. 2011. unmarked: An R Package for Fitting Hierarchical Models of Wildlife Occurrence and Abundance. Journal of Statistical Software. 43:1–23. doi:10.18637/jss.v043.i10.

Ford MJ, Hempelmann J, Hanson MB, Ayres KL, Baird RW, Emmons CK, Lundin JI, Schorr GS, Wasser SK, Park LK. 2016. Estimation of a Killer Whale (Orcinus orca) Population's Diet Using Sequencing Analysis of DNA from Feces. PLOS ONE. 11(1):e0144956. doi:10.1371/journal.pone.0144956.

Fukaya K, Kondo NI, Matsuzaki SS, Kadoya T. 2022. Multispecies site occupancy modelling and study design for spatially replicated environmental DNA metabarcoding. Methods in Ecology and Evolution. 13(1):183–193. doi:10.1111/2041-210X.13732.

Gariepy TD, Lindsay R, Ogden N, Gregory TR. 2012. Identifying the last supper: utility of the DNA barcode library for bloodmeal identification in ticks. Mol Ecol Resour. 12(4):646–652. doi:10.1111/j.1755-0998.2012.03140.x.

Guo Z, Andreou D, Britton JR. 2017. Sea Lamprey Petromyzon marinus Biology and Management Across Their Native and Invasive Ranges: Promoting Conservation by Knowledge Transfer. Reviews in Fisheries Science & Aquaculture. 25(1):84–99. doi:10.1080/23308249.2016.1233166.

Hacker CE, Hoenig BD, Wu L, Cong W, Yu J, Dai Y, Li Y, Li J, Xue Y, Zhang Yu, et al. 2021. Use of DNA metabarcoding of bird pellets in understanding raptor diet on the Qinghai-Tibetan Plateau of China. Avian Research. 12(1):42. doi:10.1186/s40657-021-00276-3.

Happel A, Rinchard J, Czesny S. 2017. Variability in sea lamprey fatty acid profiles indicates a range of host species utilization in Lake Michigan. Journal of Great Lakes Research. 43(1):182–188. doi:10.1016/j.jglr.2016.10.010.

Harms-Tuohy CA, Schizas NV, Appeldoorn RS. 2016. Use of DNA metabarcoding for stomach content analysis in the invasive lionfish Pterois volitans in Puerto Rico. Marine Ecology Progress Series. 558:181–191. doi:10.3354/meps11738.

Harvey CJ, Ebener MP, White CK. 2008. Spatial and Ontogenetic Variability of Sea Lamprey Diets in Lake Superior. Journal of Great Lakes Research. 34(3):434–449. doi:10.3394/0380-1330(2008)34[434:SAOVOS]2.0.CO;2.

Heinrich JW, Mullett KM, Hansen MJ, Adams JV, Klar GT, Johnson DA, Christie GC, Young RJ. 2003. Sea Lamprey Abundance and Management in Lake Superior, 1957 to 1999. Journal of Great Lakes Research. 29:566–583. doi:10.1016/S0380-1330(03)70517-6.

Homma C, Inokuchi D, Nakamura Y, Uy WH, Ohnishi K, Yamaguchi H, Adachi M. 2022. Effectiveness of blocking primers and a peptide nucleic acid (PNA) clamp for 18S metabarcoding dietary analysis of herbivorous fish. Kumar R, editor. PLoS ONE. 17(4):e0266268. doi:10.1371/journal.pone.0266268.

Hrodey PJ, Lewandoski SA, Sullivan WP, Barber JM, Mann KA, Paudel B, Symbal MJ. 2021. Evolution of the sea lamprey control barrier program: The importance of lowermost barriers☆. Journal of Great Lakes Research. 47:S285–S296. doi:10.1016/j.jglr.2021.10.006.

Hubbs CL, Pope TEB. 1937. The Spread of the Sea Lamprey Through the Great Lakes. Transactions of the American Fisheries Society. 66(1):172–176. doi:10.1577/1548-8659(1936)66[172:TSOTSL]2.0.CO;2.

Huggins LG, Koehler AV, Schunack B, Inpankaew T, Traub RJ. 2020. A Host-Specific Blocking Primer Combined with Optimal DNA Extraction Improves the Detection Capability of a Metabarcoding Protocol for Canine Vector-Borne Bacteria. Pathogens. 9(4):258. doi:10.3390/pathogens9040258.

Hume JB, Bennis S, Bruning T, Docker MF, Good S, Lampman R, Rinchard J, Searcy T, Wilkie MP, Johnson NS. 2024. Evaluation of Larval Sea Lamprey Petromyzon marinus Growth in the Laboratory: Influence of Temperature and Diet. Aquaculture Research. 2024(1):5547340. doi:10.1155/2024/5547340.

Hume JB, Bravener GA, Flinn S, Johnson NS. 2021. What can commercial fishery data in the Great Lakes reveal about juvenile sea lamprey (*Petromyzon marinus*) ecology and management? Journal of Great Lakes Research. 47:S590–S603. doi:10.1016/j.jglr.2021.03.023.

Irwin BJ, Liu W, Bence JR, Jones ML. 2012. Defining Economic Injury Levels for Sea Lamprey Control in the Great Lakes Basin. North American Journal of Fisheries Management. 32(4):760–771. doi:10.1080/02755947.2012.685140.

Iverson SJ, Field C, Don Bowen W, Blanchard W. 2004. Quantitative Fatty Acid Signature Analysis: A New Method of Estimating Predator Diets. Ecological Monographs. 74(2):211–235. doi:10.1890/02-4105.

Jakubavičiūtė E, Bergström U, Eklöf JS, Haenel Q, Bourlat SJ. 2017. DNA metabarcoding reveals diverse diet of the three-spined stickleback in a coastal ecosystem. PLOS ONE. 12(10):e0186929. doi:10.1371/journal.pone.0186929.

Johnson NS, Lewandoski SA, Merkes C. 2021. Assessment of sea lamprey (Petromyzon marinus) diet using DNA metabarcoding of feces. Ecological Indicators. 125:107605. doi:10.1016/j.ecolind.2021.107605.

Jones ML. 2007. Toward Improved Assessment of Sea Lamprey Population Dynamics in Support of Cost-effective Sea Lamprey Management. Journal of Great Lakes Research. 33:35–47. doi:10.3394/0380-1330(2007)33[35:TIAOSL]2.0.CO;2.

Jusino MA, Banik MT, Palmer JM, Wray AK, Xiao L, Pelton E, Barber JR, Kawahara AY, Gratton C, Peery MZ, et al. 2019. An improved method for utilizing high-throughput amplicon sequencing to determine the diets of insectivorous animals. Molecular Ecology Resources. 19(1):176–190. doi:10.1111/1755-0998.12951.

Kamil Bartoń. 2010. MuMIn: Multi-Model Inference. :1.47.5. doi:10.32614/CRAN.package.MuMIn. [accessed 2024 Sep 20]. https://CRAN.R-project.org/package=MuMIn.

Kelly RP, Shelton AO, Gallego R. 2019. Understanding PCR Processes to Draw Meaningful Conclusions from Environmental DNA Studies. Sci Rep. 9(1):12133. doi:10.1038/s41598-019-48546-x.

Kitchell JF. 1990. The Scope for Mortality Caused by Sea Lamprey. Transactions of the American Fisheries Society. 119(4):642–648. doi:10.1577/1548-8659(1990)119<0642:TSFMCB>2.3.CO;2.

Lantry B, Adams J, Christie G, Schaner T, Bowlby J, Keir M, Lantry J, Sullivan P, Bishop D, Treska T, et al. 2015. Sea lamprey mark type, marking rate, and parasite–host relationships for

lake trout and other species in Lake Ontario. Journal of Great Lakes Research. 41(1):266–279. doi:10.1016/j.jglr.2014.12.013.

Larsen LO. 1980. Physiology of Adult Lampreys, with Special Regard to Natural Starvation, Reproduction, and Death after Spawning. Can J Fish Aquat Sci. 37(11):1762–1779. doi:10.1139/f80-221.

Lawrie AH. 1970. The Sea Lamprey in the Great Lakes. Transactions of the American Fisheries Society. 99(4):766–775. doi:10.1577/1548-8659(1970)99<766:TSLITG>2.0.CO;2.

Lefèvre E, Gardner CM, Gunsch CK. 2020. A novel PCR-clamping assay reducing plant host DNA amplification significantly improves prokaryotic endo-microbiome community characterization. FEMS Microbiology Ecology. 96(7):fiaa110. doi:10.1093/femsec/fiaa110.

Leray M, Agudelo N, Mills SC, Meyer CP. 2013. Effectiveness of Annealing Blocking Primers versus Restriction Enzymes for Characterization of Generalist Diets: Unexpected Prey Revealed in the Gut Contents of Two Coral Reef Fish Species. PLOS ONE. 8(4):e58076. doi:10.1371/journal.pone.0058076.

Leray M, Knowlton N. 2015. DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. Proceedings of the National Academy of Sciences. 112(7):2076–2081. doi:10.1073/pnas.1424997112.

Levy-Booth DJ, Campbell RG, Gulden RH, Hart MM, Powell JR, Klironomos JN, Pauls KP, Swanton CJ, Trevors JT, Dunfield KE. 2007. Cycling of extracellular DNA in the soil environment. Soil Biology and Biochemistry. 39(12):2977–2991. doi:10.1016/j.soilbio.2007.06.020.

Littleford-Colquhoun BL, Freeman PT, Sackett VI, Tulloss CV, McGarvey LM, Geremia C, Kartzinel TR. 2022. The precautionary principle and dietary DNA metabarcoding: Commonly used abundance thresholds change ecological interpretation. Molecular Ecology. 31(6):1615–1626. doi:10.1111/mec.16352.

Lopes CM, De Barba M, Boyer F, Mercier C, Galiano D, Kubiak BB, Maestri R, da Silva Filho PJS, Gielly L, Coissac E, et al. 2020. Ecological specialization and niche overlap of subterranean rodents inferred from DNA metabarcoding diet analysis. Molecular Ecology. 29(16):3143–3153. doi:10.1111/mec.15549.

Lucas MC, Hume JB, Almeida PR, Aronsuu K, Habit E, Silva S, Wang CJ, Zampatti B. 2021. Emerging conservation initiatives for lampreys: Research challenges and opportunities. Journal of Great Lakes Research. 47:S690–S703. doi:10.1016/j.jglr.2020.06.004.

MacKenzie DI, Nichols JD, Lachman GB, Droege S, Andrew Royle J, Langtimm CA. 2002. Estimating Site Occupancy Rates When Detection Probabilities Are Less Than One. Ecology. 83(8):2248–2255. doi:10.1890/0012-9658(2002)083[2248:ESORWD]2.0.CO;2.

McClenaghan B, Nol E, Kerr KCR. 2019. DNA metabarcoding reveals the broad and flexible diet of a declining aerial insectivore. The Auk. 136(1):uky003. doi:10.1093/auk/uky003.

McInnes JC, Alderman R, Deagle BE, Lea M-A, Raymond B, Jarman SN. 2017. Optimised scat collection protocols for dietary DNA metabarcoding in vertebrates. Methods in Ecology and Evolution. 8(2):192–202. doi:10.1111/2041-210X.12677.

Miehls S, Sullivan P, Twohey M, Barber J, McDonald R. 2020. The future of barriers and trapping methods in the sea lamprey (Petromyzon marinus) control program in the Laurentian Great Lakes. Rev Fish Biol Fisheries. 30(1):1–24. doi:10.1007/s11160-019-09587-7.

Nguyen TV, Tilker A, Nguyen A, Hörig L, Axtner J, Schmidt A, Le M, Nguyen AHQ, Rawson BM, Wilting A, et al. 2021. Using terrestrial leeches to assess the genetic diversity of an elusive species: The Annamite striped rabbit Nesolagus timminsi. Environmental DNA. 3(4):780–791. doi:10.1002/edn3.182.

Oksanen J. 2010. Vegan : community ecology package. http://vegan.r-forge.r-project.org/. [accessed 2024 Oct 9]. https://cir.nii.ac.jp/crid/1570291225091856896.

Pertoldi C, Schmidt JB, Thomsen PM, Nielsen LB, de Jonge N, Iacolina L, Muro F, Nielsen KT, Pagh S, Lauridsen TL, et al. 2021. Comparing DNA metabarcoding with faecal analysis for diet determination of the Eurasian otter (Lutra lutra) in Vejlerne, Denmark. Mamm Res. 66(1):115–122. doi:10.1007/s13364-020-00552-5.

Pilliod DS, Goldberg CS, Arkle RS, Waits LP. 2014. Factors influencing detection of eDNA from a stream-dwelling amphibian. Molecular Ecology Resources. 14(1):109–116. doi:10.1111/1755-0998.12159.

Piñol J, Mir G, Gomez-Polo P, Agustí N. 2015. Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. Molecular Ecology Resources. 15(4):819–830. doi:10.1111/1755-0998.12355.

Potter IC. 1980. Ecology of Larval and Metamorphosing Lampreys. Can J Fish Aquat Sci. 37(11):1641–1657. doi:10.1139/f80-212.

Pukk L, Kanefsky J, Heathman AL, Weise EM, Nathan LR, Herbst SJ, Sard NM, Scribner KT, Robinson JD. 2021. eDNA metabarcoding in lakes to quantify influences of landscape features and human activity on aquatic invasive species prevalence and fish community diversity. Diversity and Distributions. 27(10):2016–2031. doi:10.1111/ddi.13370.

Reeves LE, Gillett-Kaufman JL, Kawahara AY, Kaufman PE. 2018. Barcoding blood meals: New vertebrate-specific primer sets for assigning taxonomic identities to host DNA from mosquito blood meals. PLOS Neglected Tropical Diseases. 12(8):e0006767. doi:10.1371/journal.pntd.0006767.

Riaz T, Shehzad W, Viari A, Pompanon F, Taberlet P, Coissac E. 2011. ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. Nucleic Acids Research. 39(21):e145. doi:10.1093/nar/gkr732.

Rizzi A, Raddadi N, Sorlini C, Nordgrd L, Nielsen KM, Daffonchio D. 2012. The Stability and Degradation of Dietary DNA in the Gastrointestinal Tract of Mammals: Implications for

Horizontal Gene Transfer and the Biosafety of GMOs. Critical Reviews in Food Science and Nutrition. 52(2):142–161. doi:10.1080/10408398.2010.499480.

Robinson KF, Miehls SM, Siefkes MJ. 2021. Understanding sea lamprey abundances in the Great Lakes prior to broad implementation of sea lamprey control. Journal of Great Lakes Research. 47:S328–S334. doi:10.1016/j.jglr.2021.04.002.

Rojahn J, Gleeson DM, Furlan E, Haeusler T, Bylemans J. 2021. Improving the detection of rare native fish species in environmental DNA metabarcoding surveys. Aquatic Conservation: Marine and Freshwater Ecosystems. 31(4):990–997. doi:10.1002/aqc.3514.

Ruppert KM, Kline RJ, Rahman MS. 2019. Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. Global Ecology and Conservation. 17:e00547. doi:10.1016/j.gecco.2019.e00547.

Rutter MA, Bence JR. 2003. An Improved Method to Estimate Sea Lamprey Wounding Rate on Hosts with Application to Lake Trout in Lake Huron. Journal of Great Lakes Research. 29:320–331. doi:10.1016/S0380-1330(03)70497-3.

Sard NM, Herbst SJ, Nathan L, Uhrig G, Kanefsky J, Robinson JD, Scribner KT. 2019. Comparison of fish detections, community diversity, and relative abundance using environmental DNA metabarcoding and traditional gears. Environmental DNA. 1(4):368–384. doi:10.1002/edn3.38.

Sayers EW, Cavanaugh M, Clark K, Pruitt KD, Schoch CL, Sherry ST, Karsch-Mizrachi I. 2022. GenBank. Nucleic Acids Research. 50(D1):D161–D164. doi:10.1093/nar/gkab1135.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, et al. 2009. Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Applied and Environmental Microbiology. 75(23):7537–7541. doi:10.1128/AEM.01541-09.

Schmidt SN, Olden JD, Solomon CT, Zanden MJV. 2007. Quantitative Approaches to the Analysis of Stable Isotope Food Web Data. Ecology. 88(11):2793–2802. doi:10.1890/07-0121.1.

Schnell IB, Thomsen PF, Wilkinson N, Rasmussen M, Jensen LRD, Willerslev E, Bertelsen MF, Gilbert MTP. 2012. Screening mammal biodiversity using DNA from leeches. Current Biology. 22(8):R262–R263. doi:10.1016/j.cub.2012.02.058.

Shink KG, Sutton TM, Murphy JM, López JA. 2019. Utilizing DNA metabarcoding to characterize the diet of marine-phase Arctic lamprey (*Lethenteron camtschaticum*) in the eastern Bering Sea. Can J Fish Aquat Sci. 76(11):1993–2002. doi:10.1139/cjfas-2018-0299.

Siefkes MJ, Johnson NS, Muir AM. 2021. A renewed philosophy about supplemental sea lamprey controls. Journal of Great Lakes Research. 47:S742–S752. doi:10.1016/j.jglr.2021.03.013.

Smith BR, Tibbles JJ. 1980. Sea Lamprey (Petromyzon marinus) in Lakes Huron, Michigan, and Superior: History of Invasion and Control, 1936–78. Can J Fish Aquat Sci. 37(11):1780–1801. doi:10.1139/f80-222.

Strickler KM, Fremier AK, Goldberg CS. 2015. Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. Biological Conservation. 183:85–92. doi:10.1016/j.biocon.2014.11.038.

Swink WD. 2003. Host Selection and Lethality of Attacks by Sea Lampreys (Petromyzon marinus) in Laboratory Studies. Journal of Great Lakes Research. 29:307–319. doi:10.1016/S0380-1330(03)70496-1.

Taerum SJ, Steven B, Gage DJ, Triplett LR. 2020. Validation of a PNA Clamping Method for Reducing Host DNA Amplification and Increasing Eukaryotic Diversity in Rhizosphere Microbiome Studies. Phytobiomes Journal. 4(4):291–302. doi:10.1094/PBIOMES-05-20-0040-TA.

Thomas AC, Deagle BE, Eveson JP, Harsch CH, Trites AW. 2016. Quantitative DNA metabarcoding: improved estimates of species proportional biomass using correction factors derived from control material. Molecular Ecology Resources. 16(3):714–726. doi:10.1111/1755-0998.12490.

Toju H, Baba YG. 2018. DNA metabarcoding of spiders, insects, and springtails for exploring potential linkage between above- and below-ground food webs. Zoological Lett. 4(1):4. doi:10.1186/s40851-018-0088-9.

Trebitz AS, Hoffman JC, Grant GW, Billehus TM, Pilgrim EM. 2015. Potential for DNA-based identification of Great Lakes fauna: match and mismatch between taxa inventories and DNA barcode libraries. Sci Rep. 5(1):12162. doi:10.1038/srep12162.

Treska TJ, Ebener MP, Christie GC, Adams JV, Siefkes MJ. 2021. Setting and tracking suppression targets for sea lampreys in the Great Lakes. Journal of Great Lakes Research. 47:S357–S367. doi:10.1016/j.jglr.2021.10.007.

Turner CR, Uy KL, Everhart RC. 2015. Fish environmental DNA is more concentrated in aquatic sediments than surface water. Biological Conservation. 183:93–102. doi:10.1016/j.biocon.2014.11.017.

Vestheim H, Deagle BE, Jarman SN. 2011. Application of Blocking Oligonucleotides to Improve Signal-to-Noise Ratio in a PCR. In: Park DJ, editor. PCR Protocols. Totowa, NJ: Humana Press. (Methods in Molecular Biology). p. 265–274. [accessed 2023 Jan 27]. https://doi.org/10.1007/978-1-60761-944-4_19.

Volkoff H, Rønnestad I. 2020. Effects of temperature on feeding and digestive processes in fish. Temperature. 7(4):307–320. doi:10.1080/23328940.2020.1765950.

Yang S, Zhang C, Xu W, Li D, Feng Y, Wu J, Luo W, Du X, Du Z, Huang X. 2022. Heat Stress Decreases Intestinal Physiological Function and Facilitates the Proliferation of Harmful

Intestinal Microbiota in Sturgeons. Front Microbiol. 13. doi:10.3389/fmicb.2022.755369. [accessed 2024 Sep 3]. https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2022.755369/full.

APPENDIX A: SUPPLEMENTARY FIGURES AND TABLES

Table S3.1. PERMANOVA comparisons on sample subsets, quantifying variation associated with lake, stage, and year differences individually. Lake differences (a) compare explained variation (R^2) for the given stage-year subset along with statistical significance (p-value). Stage differences compare lake-year subsets, and year differences compare lake-stage subsets, both with explained variation and p-values.

a) Lake Differences

Stage	Year	R2	p-value
Adult	2022	0.0686	0.001
Adult	2023	0.0254	0.001
Parasitic	2022	0.0360	0.001
Parasitic	2023	0.3548	0.001

b) Stage Differences

Lake	Year	R2	p-value
Huron	2022	0.2195	0.001
Huron	2023	0.0668	0.001
Superior	2022	0.0545	0.001
Superior	2023	0.1370	0.001
Champlain	2023	0.1257	0.001

c) Year Differences

Lake	Stage	R2	p-value
Huron	Adult	0.1528	0.001
Huron	Parasitic	0.0456	0.001
Superior	Adult	0.0299	0.001
Superior	Parasitic	0.0544	0.023
Champlain	Adult	0.0173	0.007

Figure S3.1. Proportion of collected sea lamprey with positive host fish detections for all OTUs that produced at least a single detection for Lake Champlain adult samples from 2022. OTUs that are "unclassified" are groupings of sequences that were unable to taxonomically distinguished below the given taxa (either family- or genus-level).

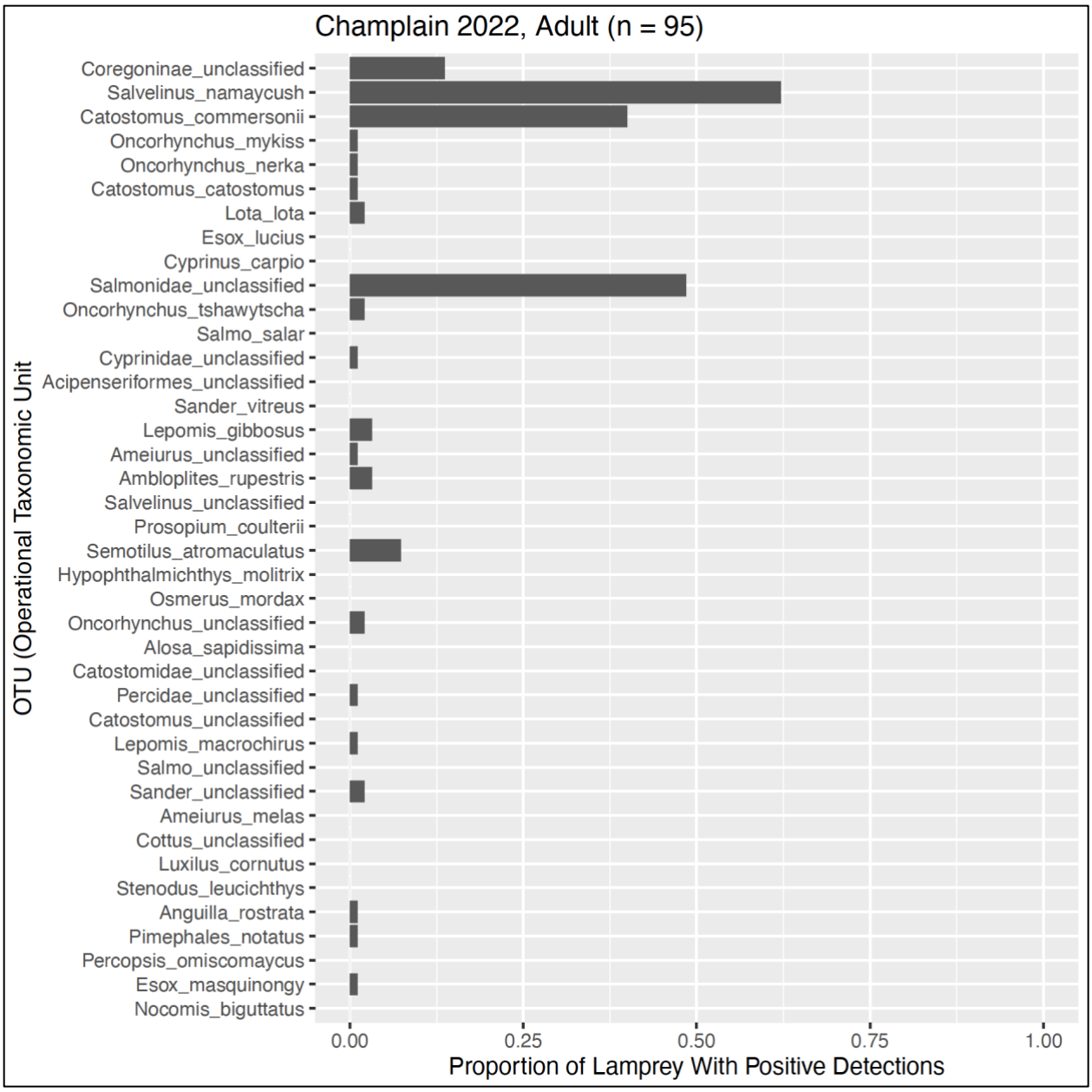


Figure S3.2. Proportion of collected sea lamprey with positive host fish detections for all OTUs that produced at least a single detection for Lake Champlain adult samples from 2023. OTUs that are "unclassified" are groupings of sequences that were unable to taxonomically distinguished below the given taxa (either family- or genus-level).

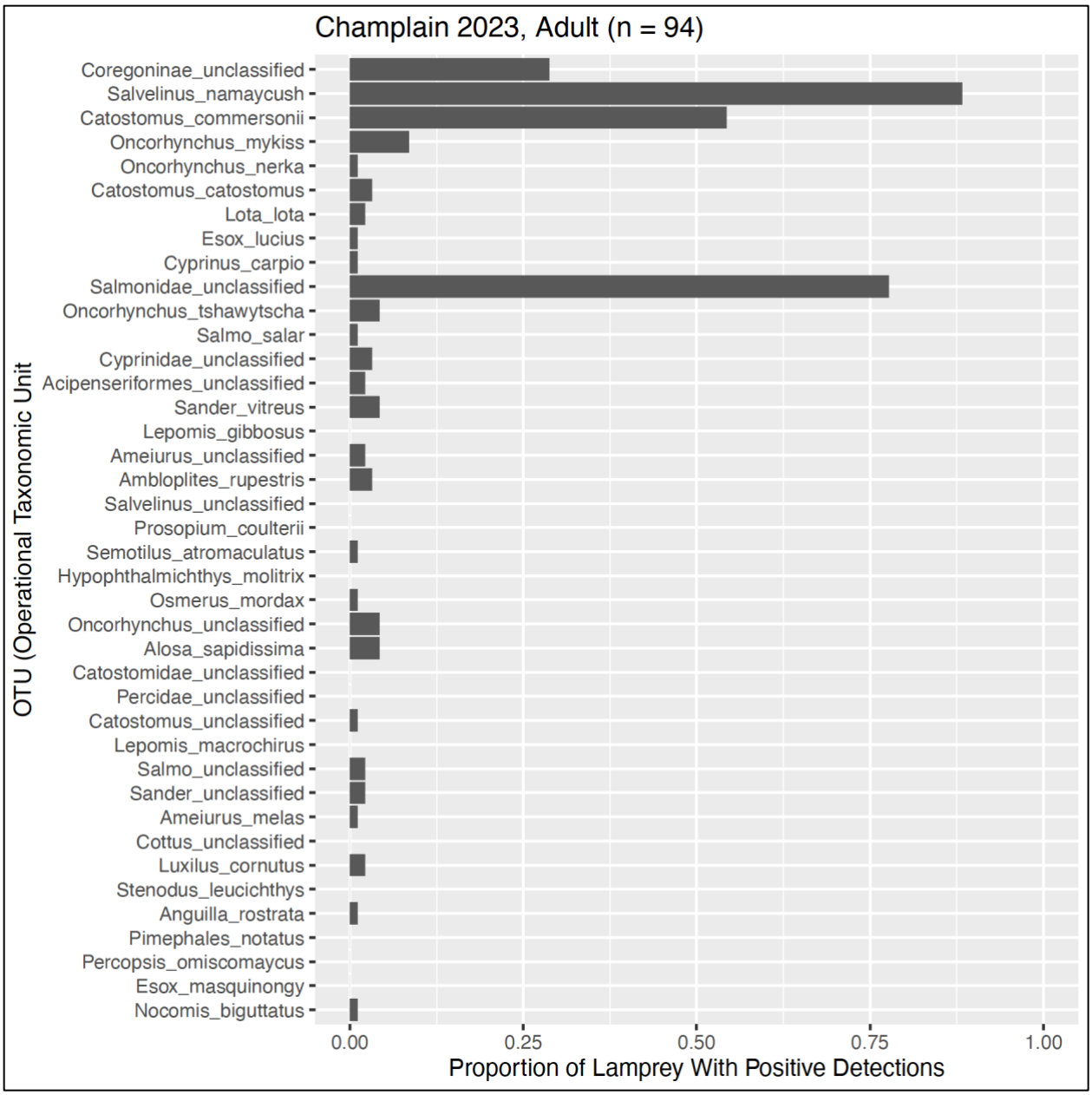


Figure S3.3. Proportion of collected sea lamprey with positive host fish detections for all OTUs that produced at least a single detection for Lake Huron adult samples from 2022. OTUs that are "unclassified" are groupings of sequences that were unable to taxonomically distinguished below the given taxa (either family- or genus-level).

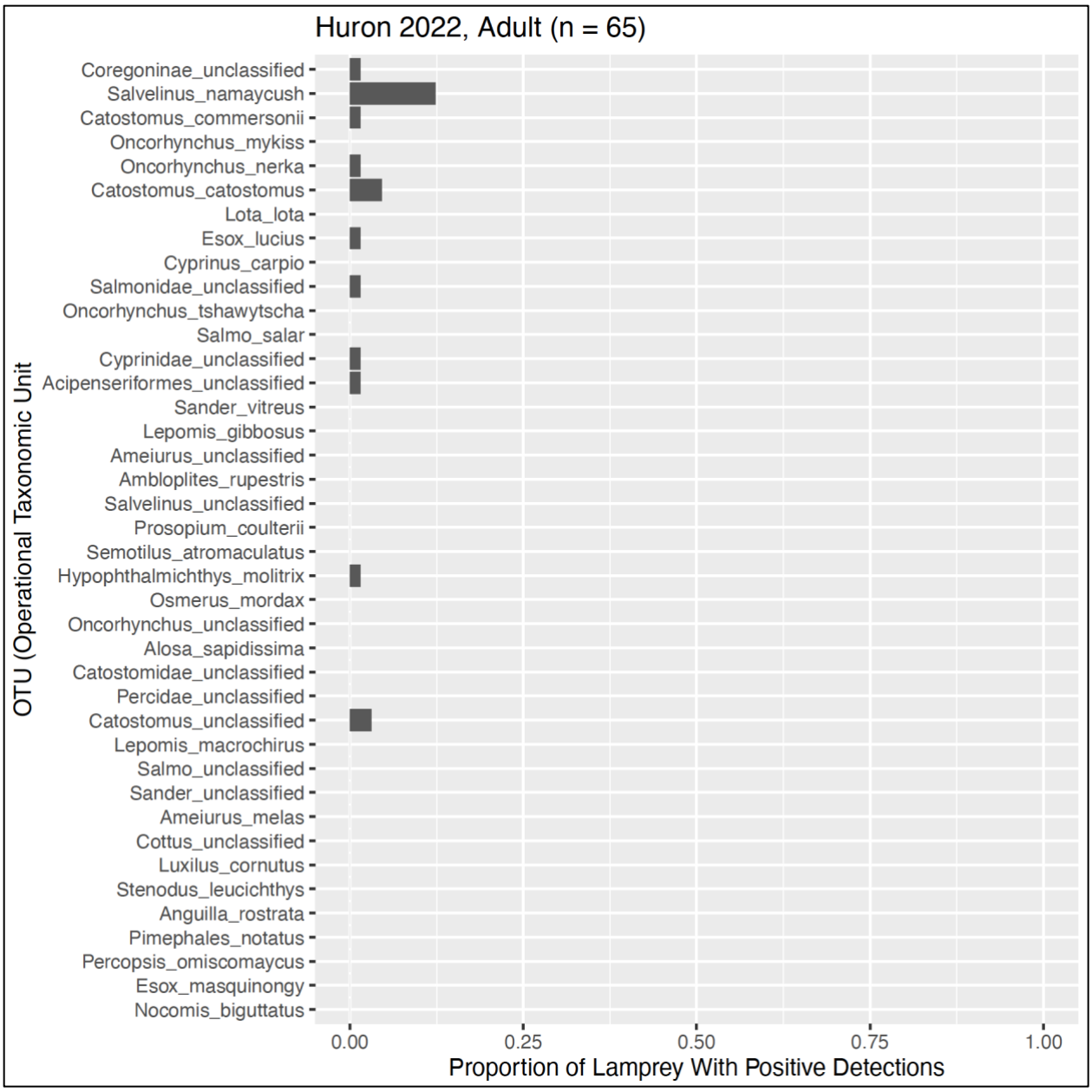


Figure S3.4. Proportion of collected sea lamprey with positive host fish detections for all OTUs that produced at least a single detection for Lake Huron adult samples from 2023. OTUs that are "unclassified" are groupings of sequences that were unable to taxonomically distinguished below the given taxa (either family- or genus-level).

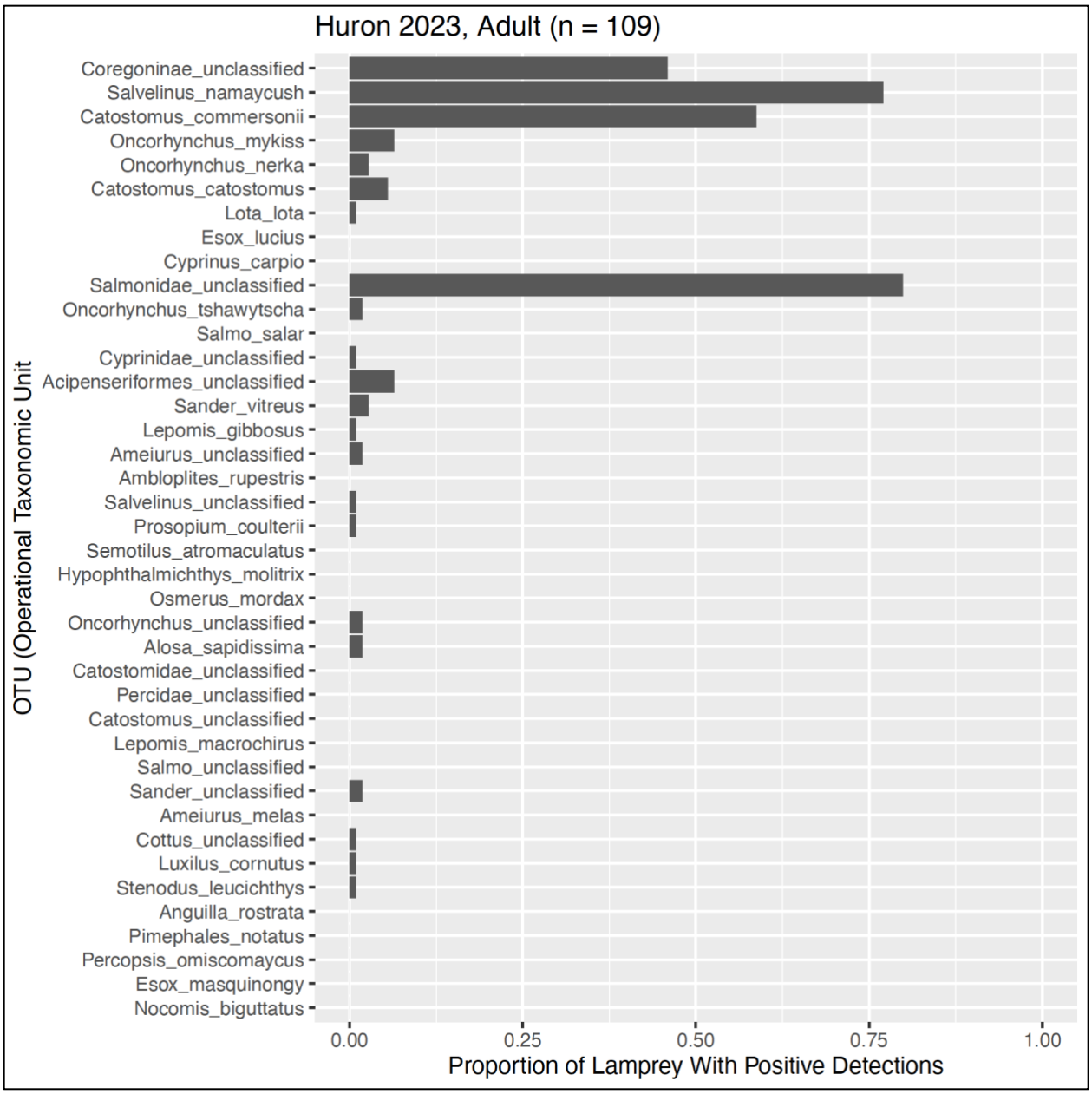


Figure S3.5. Proportion of collected sea lamprey with positive host fish detections for all OTUs that produced at least a single detection for Lake Superior adult samples from 2022. OTUs that are "unclassified" are groupings of sequences that were unable to taxonomically distinguished below the given taxa (either family- or genus-level).

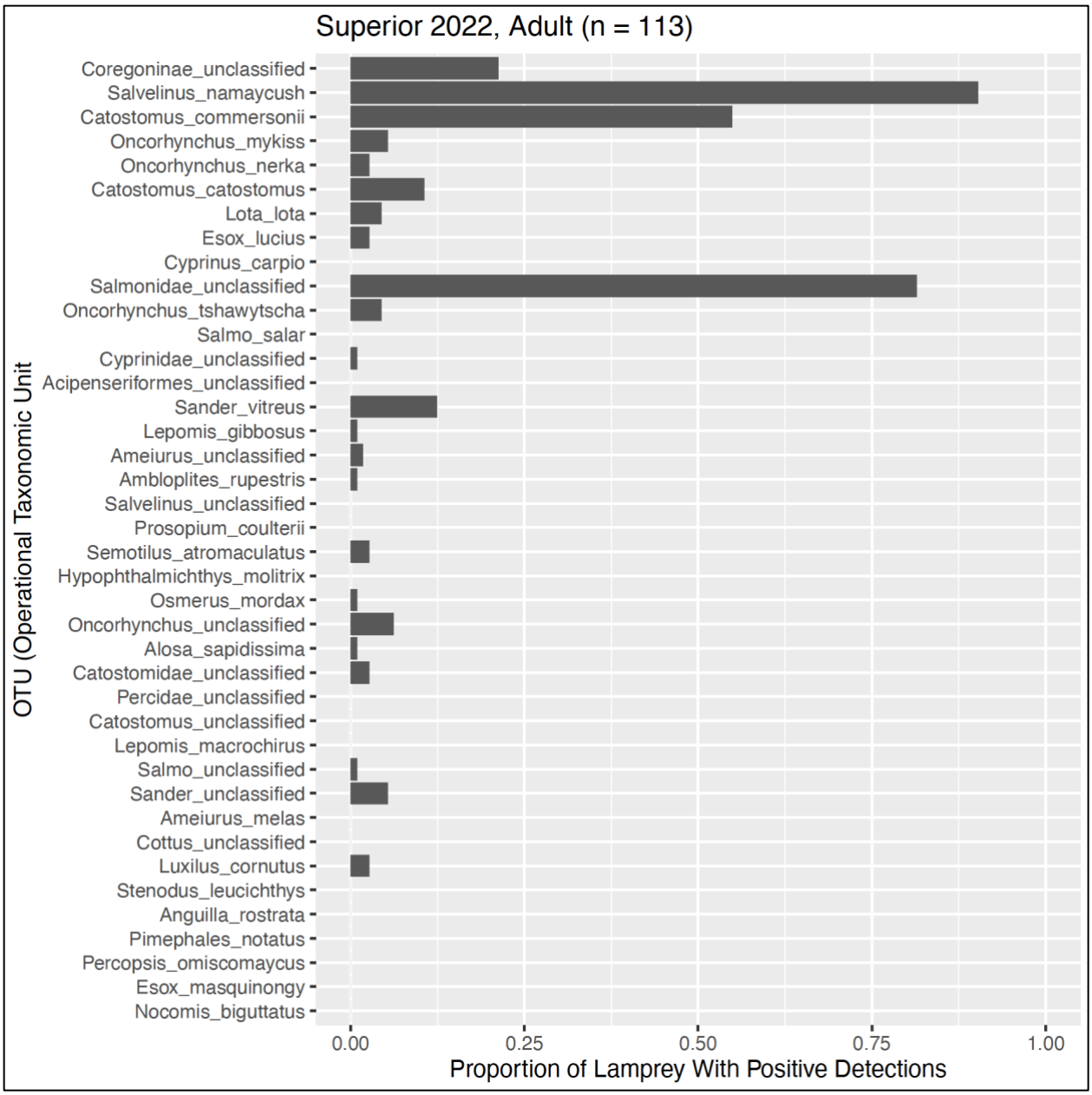


Figure S3.6. Proportion of collected sea lamprey with positive host fish detections for all OTUs that produced at least a single detection for Lake Superior adult samples from 2023. OTUs that are "unclassified" are groupings of sequences that were unable to taxonomically distinguished below the given taxa (either family- or genus-level).

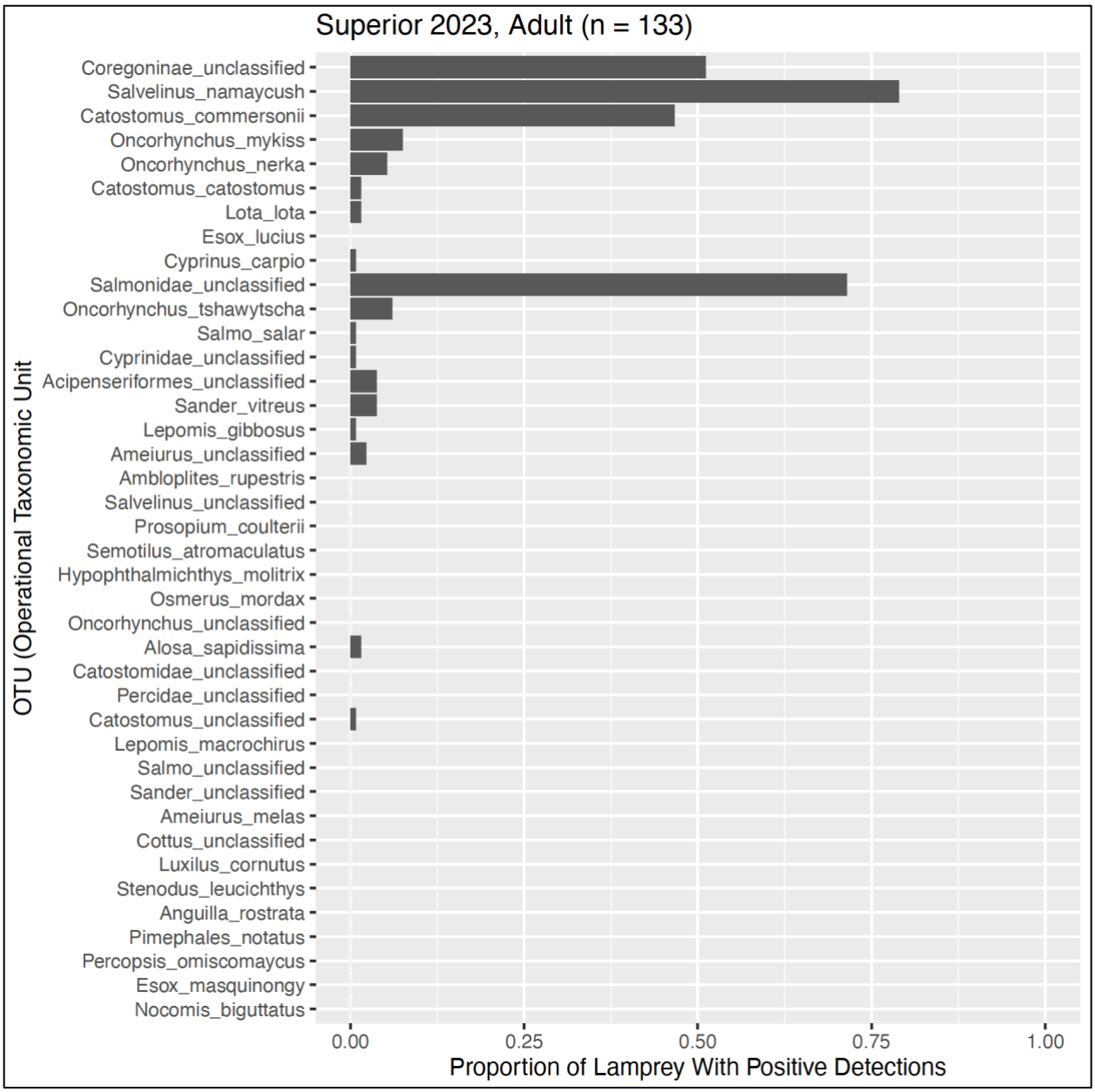


Figure S3.7. Proportion of collected sea lamprey with positive host fish detections for all OTUs that produced at least a single detection for Lake Champlain parasitic juvenile samples from 2023. OTUs that are "unclassified" are groupings of sequences that were unable to taxonomically distinguished below the given taxa (either family- or genus-level).

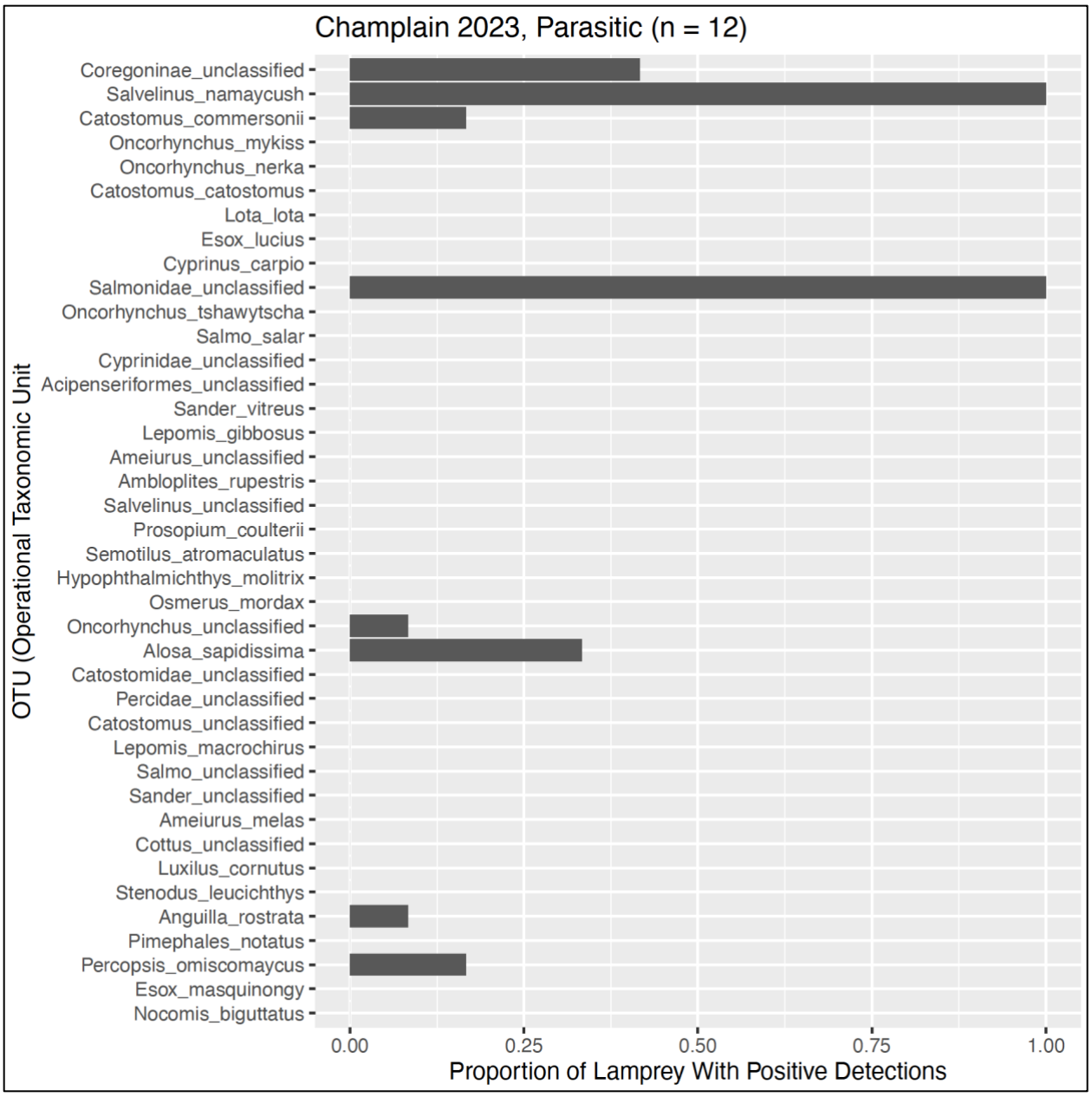


Figure S3.8. Proportion of collected sea lamprey with positive host fish detections for all OTUs that produced at least a single detection for Lake Huron parasitic juvenile samples from 2022. OTUs that are "unclassified" are groupings of sequences that were unable to taxonomically distinguished below the given taxa (either family- or genus-level).

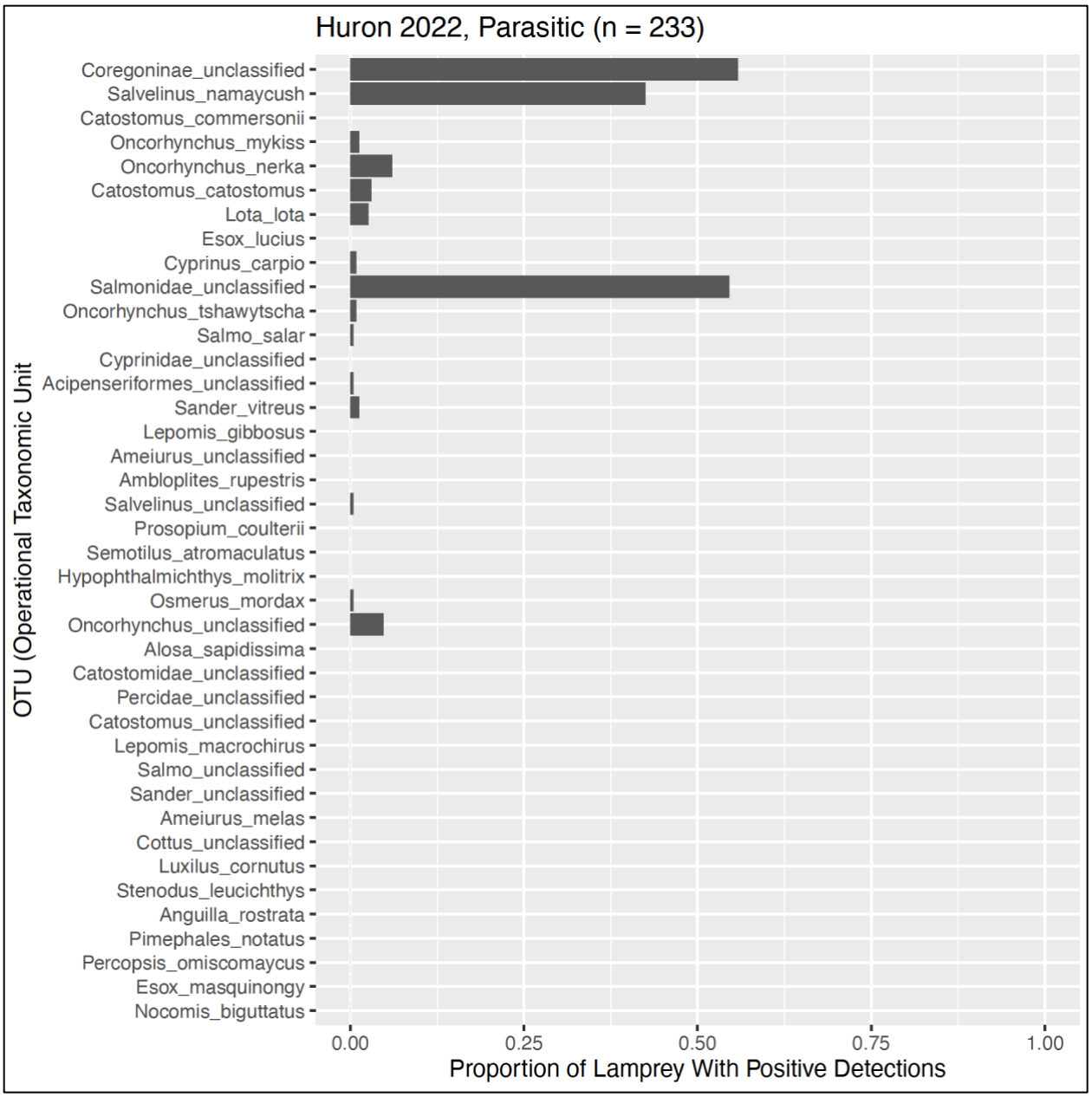


Figure S3.9. Proportion of collected sea lamprey with positive host fish detections for all OTUs that produced at least a single detection for Lake Huron parasitic juvenile samples from 2023. OTUs that are "unclassified" are groupings of sequences that were unable to taxonomically distinguished below the given taxa (either family- or genus-level).

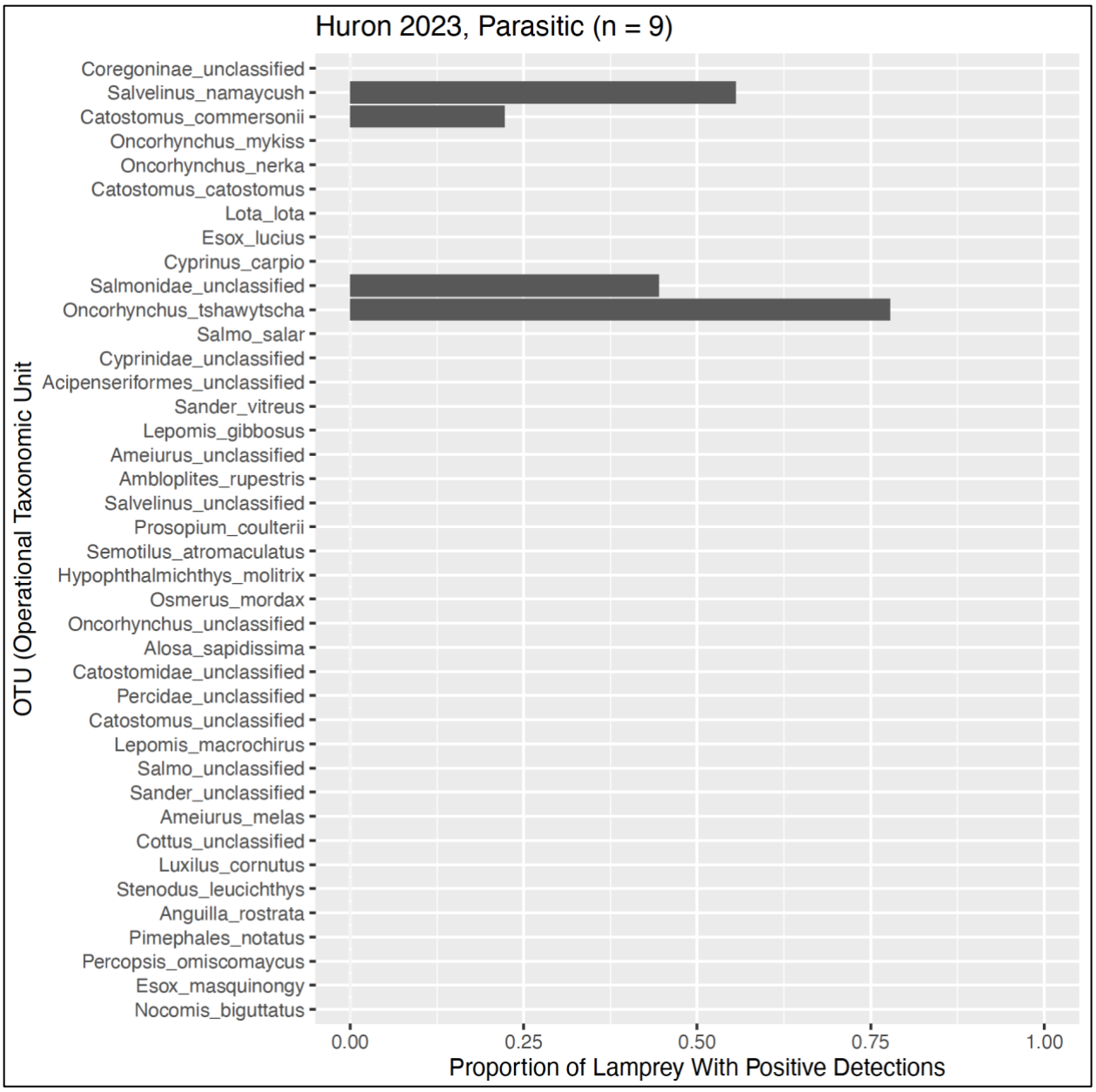


Figure S3.10. Proportion of collected sea lamprey with positive host fish detections for all OTUs that produced at least a single detection for Lake Superior parasitic juvenile samples from 2022. OTUs that are "unclassified" are groupings of sequences that were unable to taxonomically distinguished below the given taxa (either family- or genus-level).

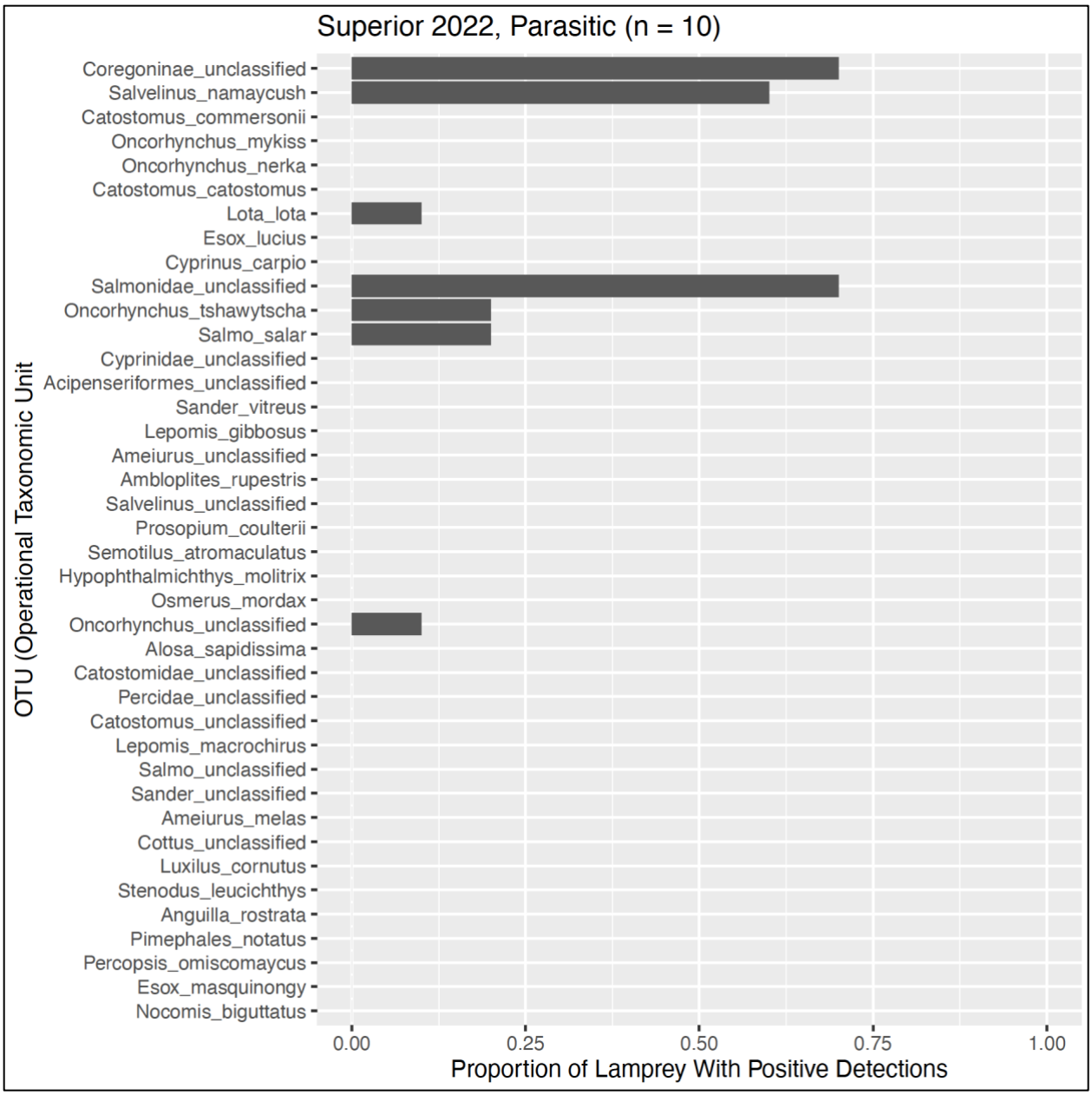
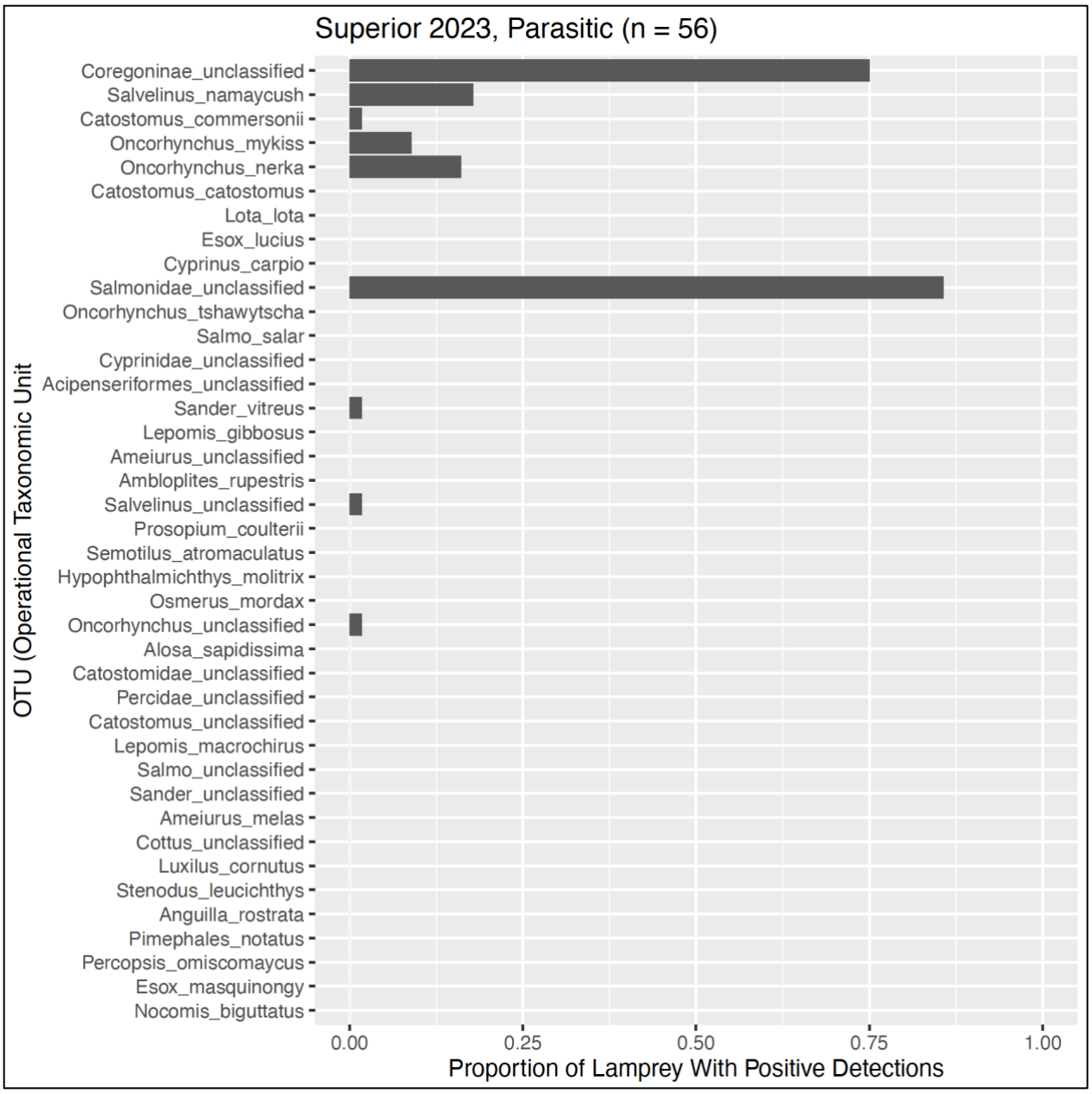


Figure S3.11. Proportion of collected sea lamprey with positive host fish detections for all OTUs that produced at least a single detection for Lake Superior parasitic juvenile samples from 2023. OTUs that are "unclassified" are groupings of sequences that were unable to taxonomically distinguished below the given taxa (either family- or genus-level).



APPENDIX B: SUPPLEMENTARY MATERIALS

Data and statistical analyses:

https://github.com/okaneco1/SL_wild_caught

Bioinformatic scripts (mothur):

https://github.com/okaneco1/mothur/tree/main