Assessing the Linkage Between Aquatic Biodiversity and Water Chemistry in the Stones River Watershed

By

Jacqueline Williams

A thesis presented to the Honors College of Middle Tennessee State University in partial fulfillment of the requirements for graduation from the University Honors College

Spring 2022

Thesis Committee:

Dr. Cole Easson, Thesis Director

Dr. Rebecca Seipelt-Thiemann, Second Reader

Dr. Dennis Mullen, Thesis Committee Chair

Assessing the Linkage Between Aquatic Biodiversity and Water Chemistry in the Stones River Watershed

By

Jacqueline Williams

APPROVED:

____

Dr. Cole Easson, Thesis Director
Department of Biology

____

Dr. Rebecca Seipelt-Thiemann, Second Reader
Department of Biology

____

Dr. Dennis Mullen, Thesis Committee Chair
Department of Biology

____

Dr. John R. Vile, Dean
University Honors College

ACKNOWLEDGEMENTS

My primary thanks are owed to Dr. Easson. Thank you for accepting me into your lab and asking me to be on this project. You opened the door into the exciting and eternally frustrating world of science and I walked in wide-eyed, twitterpated, and naive of the journey ahead. Thank you for every opportunity you have invited me to take part in. They have changed my life. Mostly, thank you for guiding and helping me through the processes of research and data analysis when we were both just a little bit too busy. What an adventure it has been.

Thank you to the Stones River Watershed Association, MTSU, the Biology department, and the University Honors College for providing the programs, space, resources, and financial support for young scientists like myself to do research. I would not be here without these institutions.

Thank you to my friends and family who loved and believed in me throughout this journey. Your questions, prayers, texts, and calls were more uplifting than you know. To my mom, Teresa, who listened to all of my frustrations with patience and my elations with matched excitement. Your love for learning has inspired me for years. You encouraged me to pursue this degree and the honors college. You have supported me in every tangible and intangible way. This one is for you.

ABSTRACT

The Stones River Watershed is home to great aquatic diversity but is subject to anthropogenic pollution from multiple sources. Traditional biodiversity survey methods require a lot of time and expertise for the proper identification of each taxonomic group and involve highly invasive methods. A new biodiversity survey tool, environmental DNA (eDNA) metabarcoding, is a non-invasive and fast method to efficiently and accurately capture species richness within a variety of environment types. It has the ability to quickly assess an ecosystem's species composition and aid conservation efforts. In this study, we captured eDNA and measured water nutrient concentrations that are common in anthropogenic runoff. We discovered over 9000 amplicon sequence variants that corresponded to unique sequences of animals in the Stones River Watershed. While this high diversity was not correlated with water nutrients, it is a useful base of knowledge for conservation work to build on.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ……………………………………………………………iii

ABSTRACT……………………………………………………………………….iv

LIST OF TABLES…………………………………………………………………vi

LIST OF FIGURES……………………………………………………….………..v

INTRODUCTION……………………………………………………………….1

METHODS…………………………………………………………………………3

RESULTS………………………………………………………………………..7

DISCUSSION……………………………………………………………………19

REFERENCES…………………………………………………………………..24

LIST OF TABLES

Table 1. Sample site names, abbreviations, and locations ……………………………4

Table 2. Taxonomic breakdown of 9,214 ASVs………………………………………15

LIST OF FIGURES

Figure 1. All Sampling Sites on the Stones River Watershed……………………………3

Figure 2. ASVs richness across all collection sites ……………………………...… 11

Figure 3. ASV richness across river locations……………………………………….12

Figure 4. ASV richness across stream environment…………………………………12

Figure 5. Beta diversity across collection sites, river fork, and stream environment.….14

Figure 6. Relative abundance of eDNA sequences …………………………………….16

Figure 7. Prevalence of Fish genera across all sites…………………………………….17

Figure 8. Mollusca ASV prevalence across all sites ………………………..………..19

Figure 9. Phosphate concentration means across all sites…………………….…………21

Figure 10. Nitrate concentration means across all sites……..…………………………..22

Figure 11. Ammonia concentration means across all sites…………………...…………23

INTRODUCTION

The Stones River watershed consists of 11 named streams and 1 reservoir that span 5 counties in middle Tennessee. The Stones River flows into the Cumberland River watershed, which flows into the Ohio River in Kentucky, eventually joining the Mississippi River to terminate in the Gulf of Mexico. An approximated 1.2 million people live in the areas drained by the Stones River, with population hot spots in the cities of Murfreesboro and Smyrna. Land use in these areas is highly variable, including industrial, residential, farmland, and city properties, both privately and government-owned. Given the diversity of land use, nutrient inputs into the Stones River watershed are greatly varied and could influence environmental conditions that affect aquatic biodiversity. Additionally, there are several small impoundments (mills and dams) and one large impoundment (Percy Priest Dam) that may further impact aquatic species biodiversity. (Streamer, 2021)

Pollution and anthropogenic (human-driven) impacts in the Stones River Watershed are well documented including contaminants from sewage (Brown and Broughton 1981) and pharmaceuticals (Kaur et al., 2020). A 2020 publication from TDEC indicates that fish are unsafe to eat from portions of the East Fork due to their high mercury content (Tennessee, 2020). Three pollutants (nitrogen, phosphorus, and ammonia) are known to have a significant presence in anthropogenic runoff, mainly from agriculture and urban sources because they are common components in fertilizers. Although the presence of these pollutants is well documented, we have a poor grasp of how they impact the biodiversity of this important ecosystem. Amid the pollution, the Stones River remains home to a remarkable diversity of fish and other organisms that live

both in and near these aquatic environments (e.g., amphibians, reptiles, mollusks, plants, etc.; Mullen et al. 2006; Niemiller et al., 2009; TDEC 2010). A thorough census of the aquatic biodiversity is needed to better understand the impacts of these pollutants and their connection with human activities in the watershed.

When using traditional methodology, capturing accurate and thorough biodiversity estimates in aquatic systems is challenging and time-consuming. These methods often involve temporarily impounding parts of the stream followed by catching and/or stunning the organisms present to count them. Additionally, trained experts in each taxonomic group are needed for the proper identification of organisms. These combined logistical hurdles limit our ability to conduct these types of biodiversity surveys. However, a new method, called metabarcoding, offers a solution to these time-consuming, destructive, and arduous traditional surveys. This technique involves the collection of environmental DNA (eDNA) using filtration coupled with next-generation sequencing to estimate biodiversity and the relative abundance of organisms in an ecosystem (Deiner et al., 2017). This technique captures free-living microorganisms and the cells and DNA that larger organisms naturally shed into the water and, by sequencing a conserved genetic region in their mitochondria, has the potential to determine the identity of each species in a diverse and mixed community. These techniques have been applied to a wide number of biological systems including open-ocean, freshwater, and terrestrial ecosystems (Deiner et al., 2017; Easson et al 2020; Rees et al., 2014). These less-invasive techniques offer a powerful tool for capturing the biological richness of an ecosystem and can even capture rare, small, or low abundance taxa that may be missed in

more traditional survey techniques. Overall, metabarcoding of eDNA offers a powerful tool for understanding biological richness in the Stones River Watershed.

In order to assess overall species diversity and potential correlations between water chemistry and species diversity, we measured the water chemistry parameters, nitrate, ammonia, and phosphate, and we collected eDNA samples on filter membranes at 14 sites along the watershed.

METHODS

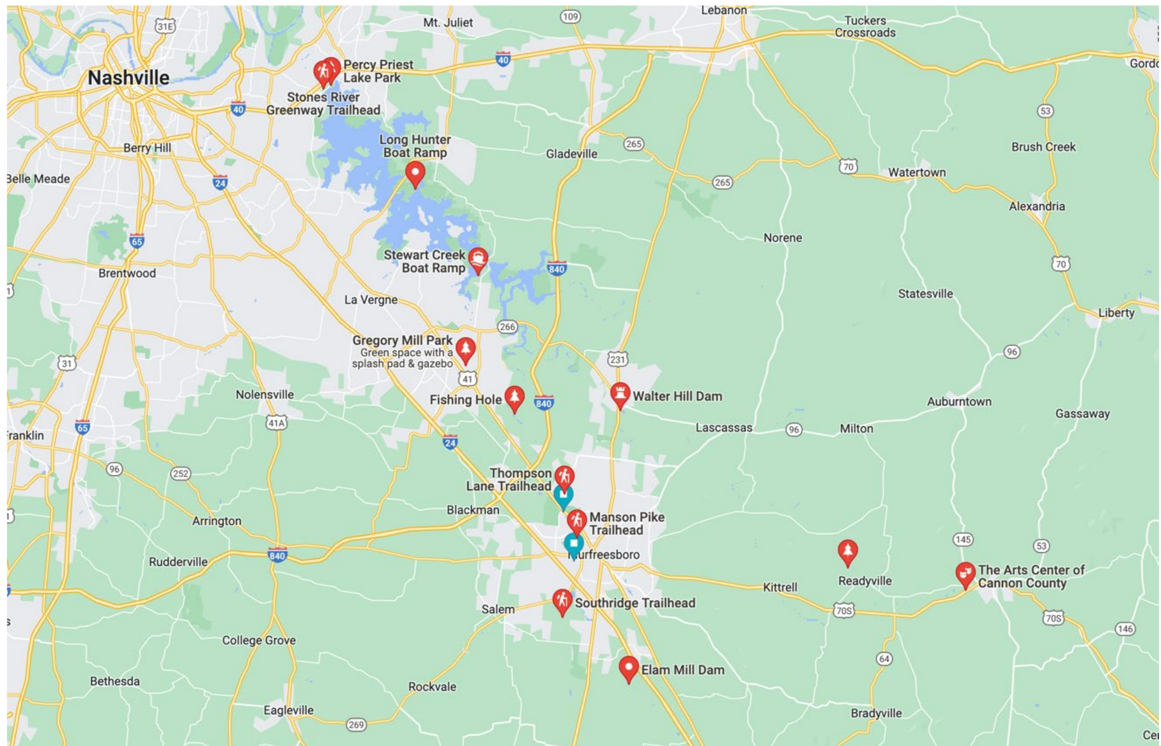

Figure 1: All Sampling sites on the Stones River Watershed. Image Source: Google Maps

Table 1. Sampling site names, acronyms, and locations.

Site Name:	Acronym	Location on Watershed	Latitude/Longitude
Above Spillway	AS	Percy Priest Lake	36°09'29.1"N 86°36'57.3"W
Below Spillway	BS	Stones River	36°09'28.5"N 86°37'11.6"W
Broad Street Greenway Trailhead	BT	West Fork Stones River	35°52'35.9"N 86°25'20.9"W
Elam Mill	EM	Middle Fork Stones River	35°45'36.1"N 86°22'11.2"W
Goochie Ford	GF	East Fork Stones River	35°50'18.0"N 86°11'20.9"W
Gregory Mill Park	GM	Stewart Creek	35°58'16.3"N 86°30'14.7"W
Long Hunter State Park Boat Ramp	LH	Percy Priest Lake	36°05'23.3"N 86°32'43.9"W
Manson Pike Greenway Trailhead	MP	West Fork Stones River	35°51'22.9"N 86°24'45.3"W
Molloy Lane	ML	West Fork Stones River	35°50'32.8"N 86°24'54.2"W
Nice Mill Dam	NM	West Fork Stones River	35°56'27.8"N 86°27'51.3"W
Southridge Greenway Trailhead	ST	West Fork Stones River	35°48'19.0"N 86°25'29.0"W
Stewart Creek Boat Ramp	SC	Percy Priest Lake	36°01'58.8"N 86°29'33.1"W
Thompson Lane Greenway Trailhead	TL	West Fork Stones River	35°53'16.9"N 86°25'23.8"W
Walter Hill Dam	WH	East Fork Stones River	35°56'31.5"N 86°22'38.6"W
Woodberry Bridge	WB	East Fork Stones River	35°49'23.9"N 86°05'24.4"W

Field Sampling

Field sampling of eDNA and water chemistry was conducted from August 2021 to November 2021 across 14 sites on the watershed. All 14 sites were sampled twice with the exception of BS, AS, and MP, which were sampled once (Table 1). Landcover around the sites was highly variable. Most notable land uses were a landfill near Walter Hill

Dam, a sod farm near Elam Mill, a cow pasture near Goochie Ford, a golf course next to Broad Street Trailhead and Thompson Lane Trailhead sites, and Nashville Shores amusement park next to Above and Below spillway sites. Many sites were within or downstream from major cities. Molloy Lane and Manson Pike Trailhead on the west fork were within Murfreesboro while Broad Street Trailhead and Thompson Lane were downstream from the major city area. Gregory Mill and Stewart Creek were downstream from Smyrna. Sites varied from lentic to lotic. Lentic (lake) sites included Above Spillway, Below Spillway, Long Hunter Boat Ramp, and Stewart Creek Boat Ramp. All other sites were lotic (river). (Table 1, Figure 1)

eDNA collection:

At each of the sites listed in Table 1, eDNA was collected using a portable peristaltic pump fitted with an in-line filtration system consisting of a prefilter to exclude large debris and a 0.22-micron pore size filter for capturing eDNA. Filtered water was collected in a 2 L food-grade container and used for water nutrient analysis. Total volume filtered through each 0.22-micron filter was recorded, ranging from 0.15 L to 2.2 L, as well as the total volume filtered during each visit, ranging from 0.5 L to 4.5 L (Table 1). Once clogged, each filter was removed from the housing, folded with forceps, placed in a cryovial with RNAlater, and stored in an onsite cooler with ice packs. The cryovials were taken back to the lab and frozen at -20ºC. Aseptic techniques were observed during each visit to limit contamination of eDNA samples. All waste was collected to be disposed of properly in the laboratory at MTSU.

Nutrient Measurement:

Nutrient concentrations of phosphorus, ammonia, and nitrogen were measured using a HACH colorimeter following powder pillow reagent protocol methods 8048 (Phosphate: orthophosphate), 8155 (Ammonia: salicylate & cyanurate reagents), and 8039 (Nitrate: nitrites and nitrates through cadmium reduction). Nutrient measurements were done onsite with filtered water collected in the food-grade container. Wastewater was collected for proper disposal in the lab. Each nutrient test was read twice and recorded as a range for that site, e.g. nitrate: 1.1 to 1.2 mg/L.

Laboratory Processing and Sequencing of eDNA

DNA Extraction:

Filters stored in RNAlater in the -20ºC freezer were thawed and approximately one-third was sectioned using aseptic techniques. The remaining two-thirds of the filter was put back in the cryovial with RNAlater and refrozen at -20ºC. The one-third portion of the filter was cut into smaller pieces to aid the extraction of DNA. The Qiagen DNeasy PowerLyzer PowerSoil Kit (Qiagen) was used and the provided protocol was followed (Supplementary file 1)with adjustments to the following steps: 2. Added 20 µl of Proteinase K. 4. Ran the homogenizer for 10 minutes at 9.0 RPM. 8. Transferred all of the supernatant rather than 600 µl. 10. Transferred all of the supernatants rather than 750 µl. 11. Added approximately 1,000 µl rather than 1,200 µl. 17. Incubated the samples with C6 solution at room temperature for 10 minutes before centrifuging. DNA was stored at -20ºC until samples were used for PCR.

Preparing DNA for Sequencing:

We followed the detailed protocol outlined in Leray et al. (2016) to prepare DNA samples for sequencing. We chose 2 to 4 samples from each site visit. After extraction, sample amplification was verified using PCR and the following reaction recipe: 12.5 µl of Platinum HotStart MasterMix (Invitrogen), 9.5 µl of nuclease-free water, 1 µl of forward primer (mlCOIint), 1 µl of reverse primer (jgHCO), and 1 µl of DNA for a total sample volume of 25 µl. Thermal cycler conditions followed Leray et al (2016). Amplification of the ~313 base pair (bp) COI gene was verified through gel electrophoresis on an agarose gel at 60V for 45 minutes. After confirming amplification, the reaction was doubled for a total volume of 50 µl and the above thermal cycler conditions were repeated. PCR duplicates were pooled and bead purified following Leray et al. (2016). DNA concentration was quantified using the Qubit fluorometer and then each sample was diluted to 10 nM with nuclease-free water. Samples were pooled using 10 µl of each normalized sample. This was followed by end repair, bead cleaning, a-tailing, ligation of y-adapters, and double bead cleaning protocols as written in Leray et al (2016). At this stage, sample pools with different Illumina adapters were diluted to 4nM and pooled. Samples were loaded onto a V2 chemistry 2x250bp flow cell kit for the Illumina MiSeq. Before sequencing, a 15% PhiX standard was added for increased sequence diversity.

Data Analysis

Sequences were demultiplexed using unique combinatorial barcodes into separate forward and reverse fastq files with the program cutadapt (version 4.0). Primer sequences of 26 bp were then removed from the 5' end of each sequence using cutadapt (Martin 2011). Demultiplexed sequences were then processed using the DADA2 pipeline

in R studio (version 334; Callahan et al. 2016). The DADA2 pipeline in R was used to verify sequence quality and assign amplicon sequence variants (ASVs). In conjunction with the DADA2 pipeline, we used the Midori database (Leray et al. 2018; http://reference-midori.info/server.php) for the taxonomy assignment of ASVs. Statistical analysis of eDNA data was restricted to amplicons that returned confirmed taxonomic IDs at the phylum level and were ±10 base pairs (bp) of the target amplicon size, 313 bp.

After assigning taxonomy, alpha and beta diversity metrics were calculated. Since the relative abundance of eDNA does not necessarily reflect organism abundance for all taxa, we constrained our analyses to presence/absence metrics. All statistical analysis was done in R (R Core Team) with the vegan (version 2.5-7; Oksanen et al. 2020) and phyloseq (McMurdie and Holmes 2013) packages. Organism richness was calculated for each sample and compared across sampling sites, river branches, and stream environments with an analysis of variance (ANOVA) once normality was confirmed. ASV composition (beta diversity) was compared across sampling sites, river branches, and stream environments using a permuted multivariate ANOVA (PERMANOVA). Water chemistry data were analyzed compared across sites using a ANOVA in R and these data were then compared to ASV richness using a correlation analysis and ASV composition using a canonical correspondence analysis. While the eDNA data offer a wealth of data for exploration, we elected to further explore two ecologically important groups in our dataset, Acteroptergii and Mollusca. Within each of these groups, we assessed differences in alpha and beta diversity across sampling sites, river branches, and stream environments. Figures were generated in R using the vegan,

ggplot2, and phyloseq packages (McMurdie and Holmes 2013; Wickman 2016; Oksanen et al. 2020)

RESULTS

Sequence quality filtering and processing

The sequencing run generated 10,888,102 high-quality sequence reads for 82 eDNA samples. Samples were demultiplexed using the combinatorial barcodes outlined in Leray et al. (2016) using the program cutadapt (cutadapt -e 0.15 --no-indels -g barcode1.fasta -G barcode2.fasta -o {name1}-{name2}.R1.fastq.gz -p {name1}-{name2}.R2.fastq.gz srw1_S1_L001_R1_001.fastq.gz srw1_S1_L001_R2_001.fastq.gz). This method recorded the 6 base pair barcodes at the 5' end of the forward and reverse reads for each sequence and using the barcode labels in the loaded fasta files (barcode1.fasta & barcode2.fasta) assigned names to each sequence based on the combination of the two barcodes (e.g. mlCOIT1_jgHCOT1 - barcode tag 1 for forward and reverse reads, respectively). Barcode combinations were then translated to sample names manually in the output directory. The barcode was followed by a 26 base pair primer on the 5' end of each sequence, which was removed using cutadapt (cutadapt -u 26 -o trimmed.fastq input.fastq). The results of these steps were 164 demultiplexed sequence files (forward and reverse fastq files for each sample) with the barcode (6 bp) and primer (26 bp) removed. The total length of each sequence was thus 218bp.

The trimmed sequences were input into R (R Core Team) and run through the DADA2 pipeline (Callahan et al. 2016) following default parameters for quality filtering, error rate calculation, dereplication, chimera removal, and inference of amplicon

sequence variants (ASVs). The sequence processing protocol can be found at https://benjjneb.github.io/dada2/tutorial.html. These parameters resulted in 52,395 ASVs in the dataset of 82 samples. After calculation of ASVs, sequence length was examined; sequences that were substantially longer (> 320 bp) or shorter (< 50 bp) were removed from the dataset before taxonomic assignment. Taxonomic assignment was accomplished using the MIDORI server (Leray et al. 2018; http://reference-midori.info/server.php) using the RDP classifier (Wang et al. 2007). The unique COI database was used with a confidence cutoff of 0.6. The output taxonomy thus only showed taxonomic assignments with > 60% confidence. We further filtered the dataset to include only sequences with a positive taxonomic assignment to the phylum level, which resulted in a total of 9,214 ASVs across 82 samples.

Alpha diversity

Estimates of ASV richness across ranged 34-788 ASVs in a single sample (34 - GM2 and 788 - WH2). Mean ASV richness was 270 ±170 ASVs and median richness was 231±170 ASVs. We detected 4 outliers in the dataset based on values exceeding the maximum richness values (601 ASVs) calculated for the whole dataset (GF3P2-781ASVs, LH2P2-661ASVs, NM3P2-699ASVs, WH2P2-788ASVs). Since these samples represent outliers, they were excluded from richness statistics. Comparisons of ASV richness among collection sites showed significant differences among sites (ANOVA, df = 14, F = 2.30, P = 0.013), but few pairwise differences were observed (Tukey HSD P < 0.05; Figure 2). ASV richness across river fork also displayed significant differences (ANOVA, df = 3, F = 4.596, P = 0.005). Samples collected on the east fork of the Stones river contained higher ASV richness than those collected in

Stewarts Creek, Percy Priest Reservoir, and the west fork of the Stones River (Tukey HSD $P < 0.05$; Figure 3). Clear differences in richness were observed between lotic and lentic environments (ANOVA, $df = 1$, $F = 8.13$, $P = 0.006$), with lotic environments exhibiting higher ASV richness, (Figure 4).

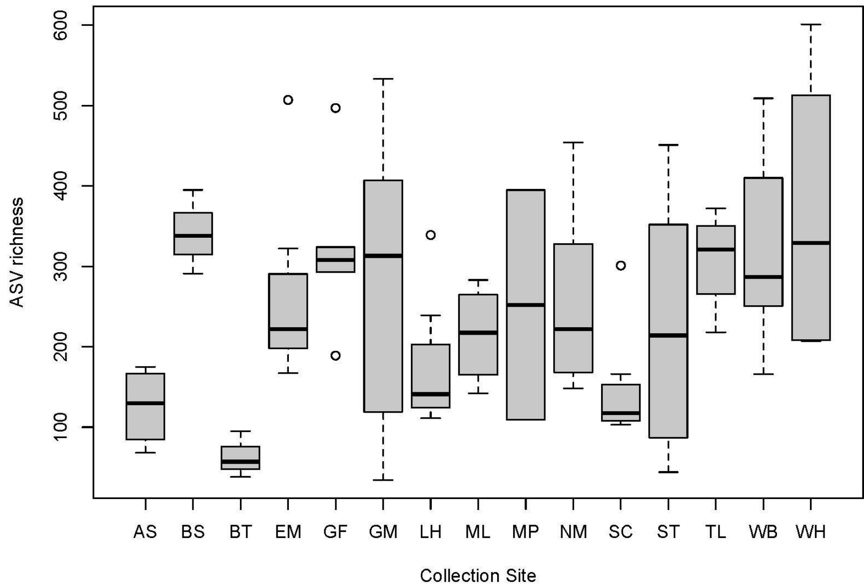

Figure 2: ASV richness across collection sites. Boxes represent first and third quartiles, dark midline within boxes represent the median, whiskers represent possible max and min values, and dots represent outliers. Site key: AS: Above Spillway, BS: Below Spillway, BT: Broad Streat Trailhead, EM: Elam Mill, GF: Goochie Ford, GM: Gregory Mill, LH: Long Hunter, ML: Molly Lane, MP: Manson Pike, NM: Nice Mill, SC: Stewart Creek, ST: Southridge Trailhead, TL: Thompson Lane Trailhead, WB: Woodberry Bridge, WH: Walter Hill Dam.

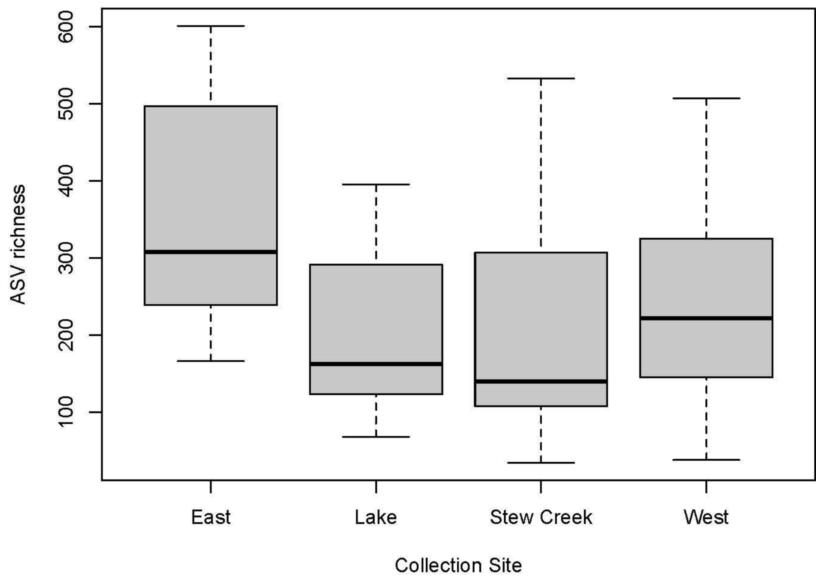

Figure 3: ASV richness across river locations. Boxes represent first and third quartiles, dark midline within boxes represent the median, and whiskers represent possible max and min values. East = east fork of Stones River; Lake = Percy Priest Reservoir; Stew Creek = Stewart Creek; West = west fork of Stones River

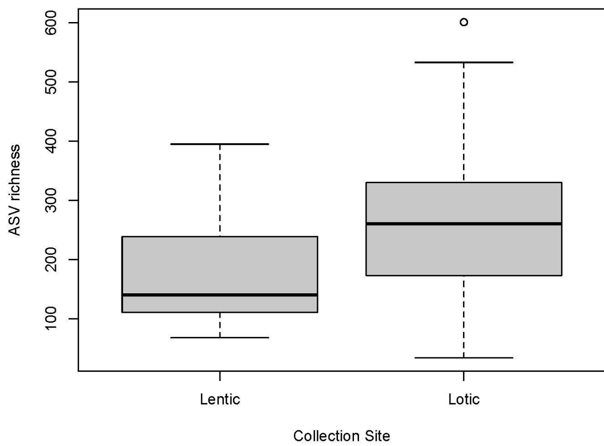

Figure 4: ASV richness across stream environments. Boxes represent first and third quartiles, dark midline within boxes represent the median, whiskers represent possible max and min values, and dots represent outliers. A total of 4 sites were in lentic environments and 11 were in lotic sites.

Beta diversity

Comparisons of dissimilarity across samples showed significant ASV composition differences due to three factors: river forks (PERMANOVA; df = 3, F = 6.34, R2 = 0.18, P = 0.001), among collection sites (PERMANOVA; df = 14, F = 3.17, R2 = 0.42, P = 0.001), and environment type (lotic vs lentic; PERMANOVA, df = 1, F = 13.342, R2 = 0.15, P = 0.001). Comparisons that included environment type, collection site, and river fork indicated a variance overlap, with all variance explained by river fork and environment type included in the 42% explained by collection site. While considerable site variation existed in our collections, a major singular driver of dissimilarity in composition appears to be environment type with lotic (river) and lentic (lake) environments composed of distinct ASVs (Figure 5).

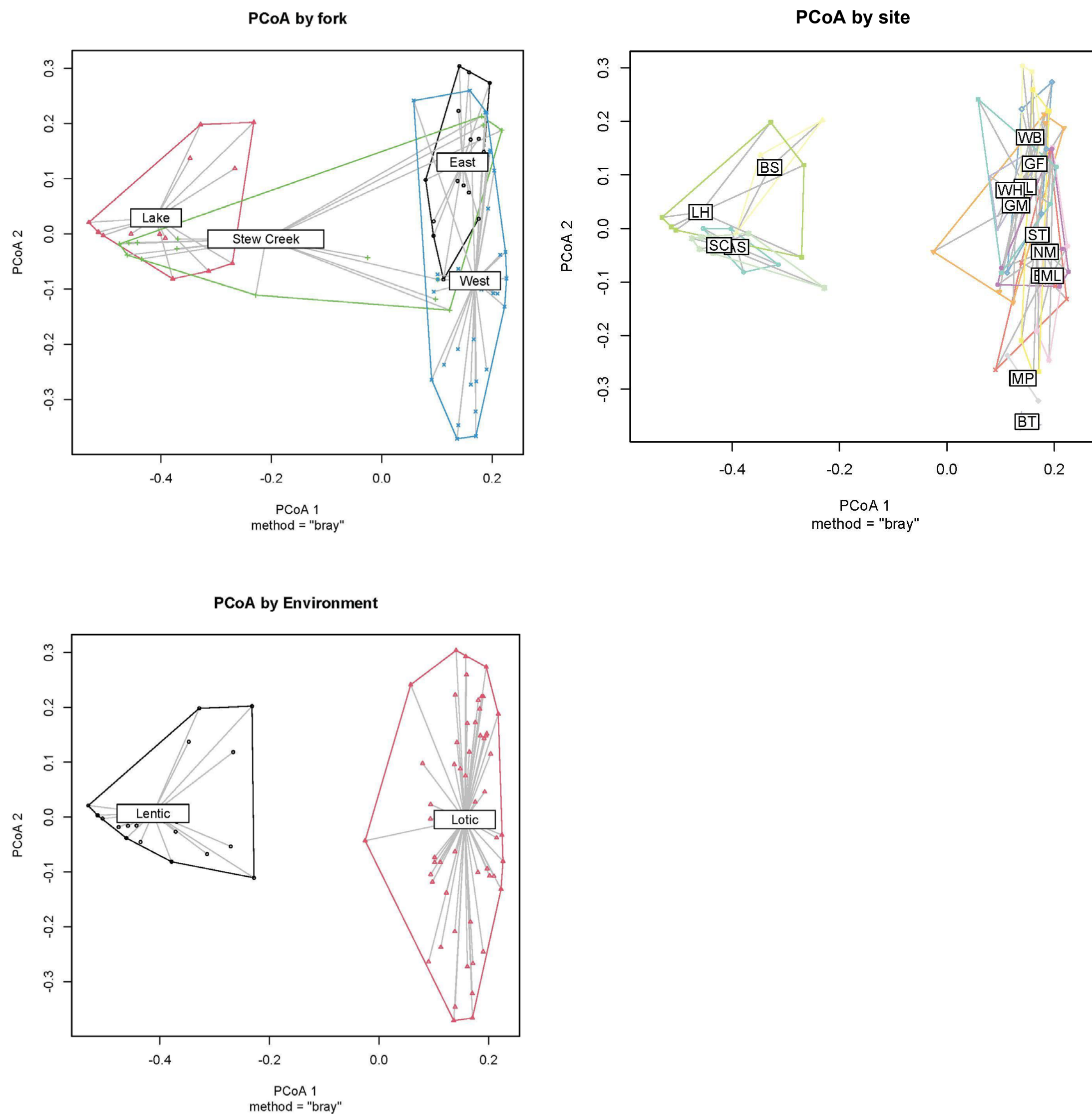

Figure 5: Beta diversity across (A) collection sites, (B) river fork, and (C) Environment type. Site key: AS: Above Spillway, BS: Below Spillway, BT: Broad Streat Trailhead, EM: Elam Mill, GF: Goochie Ford, GM: Gregory Mill, LH: Long Hunter, ML: Molly Lane, MP: Manson Pike, NM: Nice Mill, SC: Stewart Creek, ST: Southridge Trailhead, TL: Thompson Lane Trailhead, WB: Woodberry Bridge, WH: Walter Hill Dam.

The taxonomic breakdown for the 9,214 ASVs identified to at least phylum is shown in Table 2. Dominant taxa (top 100) across all samples are shown in Figure 6. These taxa include rotifers in the families Brachionidae and Synchaetidae, gastriches in the family Chaetonotidae, mollusks in the family corbiculidae, bony fish in the family

Cyprinidae, insects in the family Halimococcidae, cnidarians in the family Olindiidae, poriferans in the family Spongillidae, and several taxa belonging to Arthropoda, Chordata, Cnidaria, Insecta, Platyhelminthes, and Ploima that could not be classified to family by the COI database (Figure 6). Two major constituents of the collected eDNA across most samples were from the phyla Rotifera and Arthropoda. Beta diversity analysis revealed stark differences in lotic and lentic samples, which are driven in part by the increased presence of several rotifer ASVs.

Table 2: Taxonomic breakdown to at least phylum level of the 9,214 ASVs identified.

Phylum	Class	Number of ASVs
Annelida		179
Arthropoda		6,304
Bryozoa		10
Chordata		1541
	Actinopterygii	150
	Amphibia	2
	Aves	6
	Mammalia	17
	Reptilia	3
	Unclassified	1363
Cnidaria		572
Entoprocta		8
Gastrotricha		52
Mollusca		78
	Bivalvia	22
	Gastropoda	52
	Unclassified	4
Nematoda		27
Nemertea		2

Phylum	Class	Number of ASVs
Platyhelminthes		15
Porifera		61
Rotifera		360
Tardigrada		1

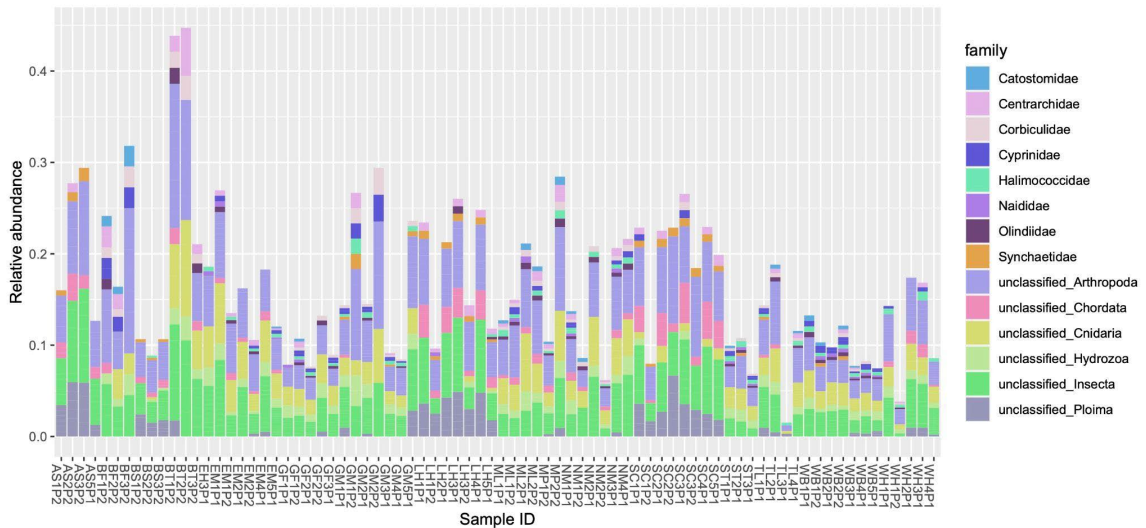

Figure 6: Stacked bar chart of relative abundance of eDNA sequences at all sampling sites. Site key: AS: Above Spillway, BS: Below Spillway, BT: Broad Streat Trailhead, EM: Elam Mill, GF: Goochie Ford, GM: Gregory Mill, LH: Long Hunter, ML: Molly Lane, MP: Manson Pike, NM: Nice Mill, SC: Stewart Creek, ST: Southridge Trailhead, TL: Thompson Lane Trailhead, WB: Woodberry Bridge, WH: Walter Hill Dam.

Bony Fish (Class Actinopterygii)

Fish are of particular interest in conservation and recreation and we, therefore, focused on these species in our analysis. In our dataset, we detected fish belonging to 13 families and 33 genera as well as several ASVs that we were unable to classify further than class level. Prevalence of fish ASVs in the dataset varied across sites and ranged from just a single species detected (above spillway in Percy Priest reservoir) to over 70 (Woodberry bridge in Readyville, TN; Figure 7). Prevalent genera included

Campostoma (stonerollers; 12 sites), Lepomis (sunfish; all 15 sites), Etheostoma (darters; 13 sites), Pimephales (bluntnose minnows; 10 sites), Hypentelium (suckers; 10 sites), and Moxostoma (redhorses; 11 sites). ASV richness of fish was significantly different across sampling sites (ANOVA, df = 14, F = 3.5, P = 0.003) and environment types (lentic vs lotic; ANOVA; df = 1, F = 21.53, P < 0.001). The east and west forks of the Stones River did not show differences in Actinopterygii richness, and all differences were due to overall lower richness in lentic environments (Tukey HSD P < 0.05). Differences in fish ASV composition were observed across sampling sites, river forks, and environment types (PERMANOVA P < 0.01). In general, we observed a large amount of heterogeneity across sites and while lotic river sites tended to have higher fish richness, clear trends in composition were less apparent.

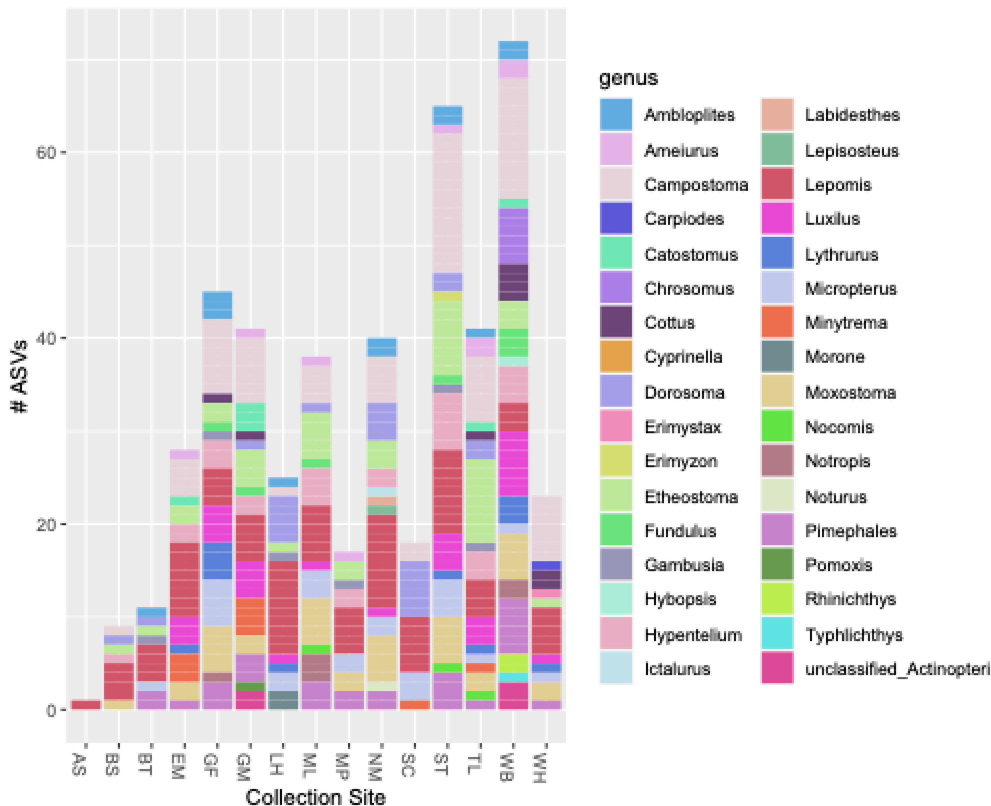

Figure 7: Prevalence of Fish genera across sites. Bars represent the number of occurrences of Actinopterygii ASVs across all samples in a site. Site key: AS: Above Spillway, BS: Below Spillway, BT: Broad Streat Trailhead, EM: Elam Mill, GF: Goochie Ford, GM: Gregory Mill, LH: Long Hunter, ML: Molly Lane, MP: Manson Pike, NM: Nice Mill, SC: Stewart Creek, ST: Southridge Trailhead, TL: Thompson Lane Trailhead, WB: Woodberry Bridge, WH: Walter Hill Dam.

Mollusca

Mollusks, especially bivalves are also of interest to conservation efforts due to the high diversity observed in the southeastern U.S. and their ecological role within aquatic environments. Across all eDNA samples, 79 Mollusca taxa were detected. These taxa belonged to the classes Bivalvia (22 ASVs) and Gastropoda (52 ASVs), with 4 ASVs only classified to phylum. (Figure 8)

Within Bivalvia, the 22 ASVs represented 3 families and 6 genera. The most prevalent genus of bivalves (observed at 14 sites) was *Corbicula* (Asian clam), which is invasive in North America. Six ASVs of freshwater mussels (family: Unionidae; genera: *Lampsilis, Medionidus, Villosa*) were observed along with 14 ASVs of freshwater clams (families: Corbiculidae & Sphaeriidae; genera: *Corbicula, Pisidum, Sphaerium*).

Within Gastropoda, we observed 5 families and 5 genera across the 53 ASVs. The most prevalent genera was *Pleurocera*, which is a common freshwater snail in eastern North America. Other genera included *Deroceras, Ferissia, Physella, Menetus, and Lithasia*. These taxa were often abundant at collection sites (Williams personal observation) where they could be observed on a rocky substrate.

Mollusca richness was highly variable across sites, and higher in both the east and west forks of the Stones River compared to the lake environments (ANOVA, df = 3, F = 6.645, P < 0.001) but was equal across the two branches (Tukey HSD P > 0.05). Despite

a lack of differences in richness across the two river branches, there were differences in mollusk composition (pairwise PERMANOVA, P = 0.05; Figure 8).

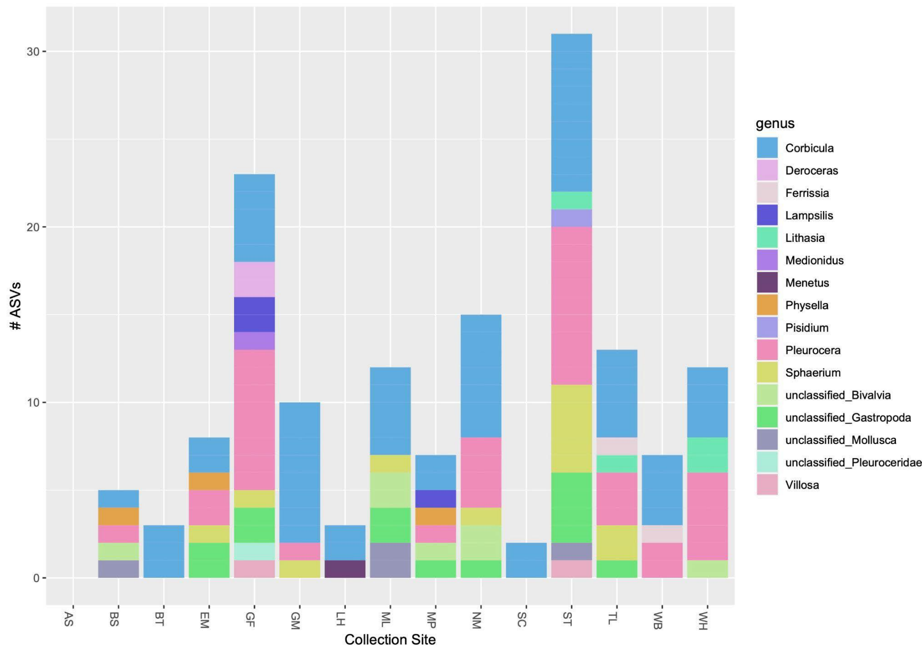

Figure 8: Stacked bar chart of Mollusca ASV prevalence across sites. Genera are represented as distinct colors while site abbreviations are listed along the x-axis. Site key: AS: Above Spillway, BS: Below Spillway, BT: Broad Streat Trailhead, EM: Elam Mill, GF: Goochie Ford, GM: Gregory Mill, LH: Long Hunter, ML: Molly Lane, MP: Manson Pike, NM: Nice Mill, SC: Stewart Creek, ST: Southridge Trailhead, TL: Thompson Lane Trailhead, WB: Woodberry Bridge, WH: Walter Hill Dam.

Nutrient concentrations

Nitrate, phosphate, and ammonia concentrations varied across sites and over time (Figure 9, 10, 11). Across all sites and sample dates, the phosphate concentrations ranged from 0.04 mg/L to 0.83 mg/L, nitrate concentrations ranged from 0.6 to 4.5 mg/L, and ammonia concentrations ranged from 0.12 mg/L to 1.09 mg/L. Sites with nutrient concentration measurements above the accurate range were left out of our calculations. Nutrient concentrations did not show a clear correlation to changing river flow (Pearson's

correlation; Nitrate: r = 0.38, P = 0.15; Ammonia: r = -0.23, P = 0.36; Phosphate: r = 0.40, P = 0.11), as some nutrient concentrations increased as river flow decreased while others decreased with decreasing river flow. Additionally, there was no difference in nutrient concentrations across the east and west branches of the Stones River (t-test, $P>0.05$).

We next investigated the relationship of the Stones River nutrient concentrations to ASV richness by measuring correlation between each nutrient concentration and mean ASV richness at each site and found no significant correlation for phosphate (r = 0.28, P = 0.30), nitrate (r = 0.2, P = 0.47), or ammonia (r = 0.17, P = 0.55).

Canonical correspondence analysis compared species compositions with nutrient concentrations to investigate potential relationships. A significant relationship was detected between species composition and all three nutrient measurements ($P< 0.01$). Despite this significant result, the overall impact on species composition was minimal with all three variables together only explaining < 6% of the total variance in species composition across sampling sites. Thus, these results indicate a weak affect of nutrient concentrations that likely only correlated with changes in a few members of the overall community.

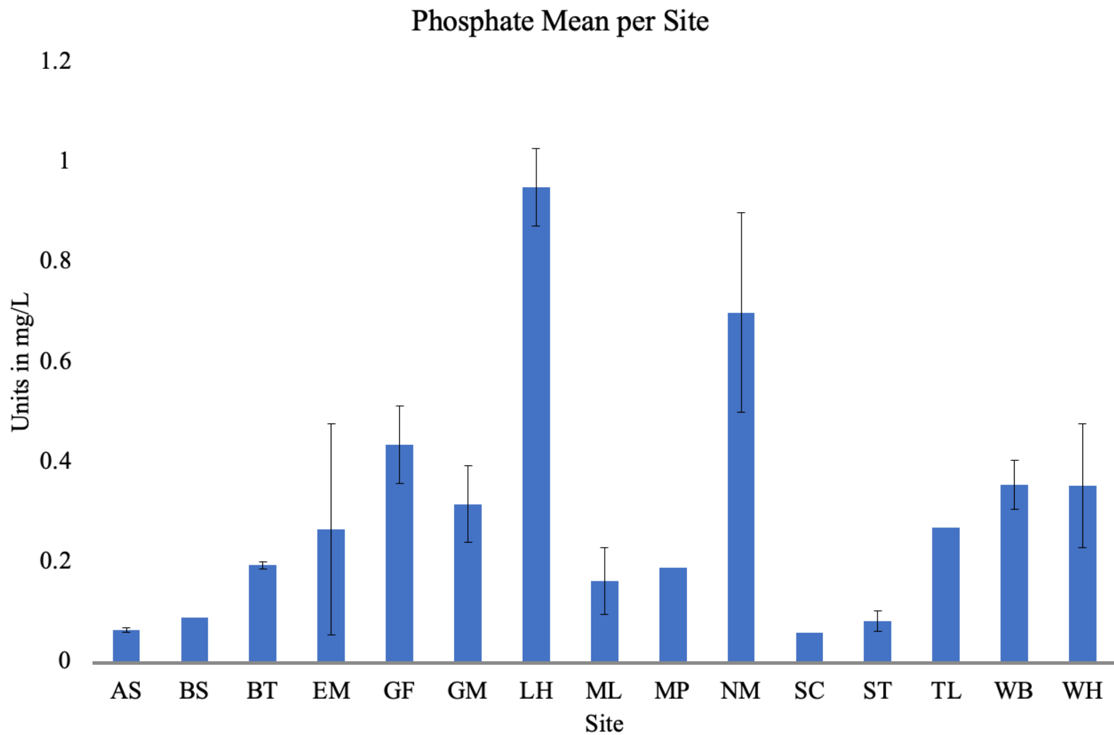

Figure 9: Phosphate means in mg/L per site sampled. Error bars represent the standard deviation of the sample means. Sites without standard deviation bars represent sites where only one nutrient measurement was taken or the nutrient measurements were equal and therefore did not have a standard deviation. Site key: AS: Above Spillway, BS: Below Spillway, BT: Broad Streat Trailhead, EM: Elam Mill, GF: Goochie Ford, GM: Gregory Mill, LH: Long Hunter, ML: Molly Lane, MP: Manson Pike, NM: Nice Mill, SC: Stewart Creek, ST: Southridge Trailhead, TL: Thompson Lane Trailhead, WB: Woodberry Bridge, WH: Walter Hill Dam.

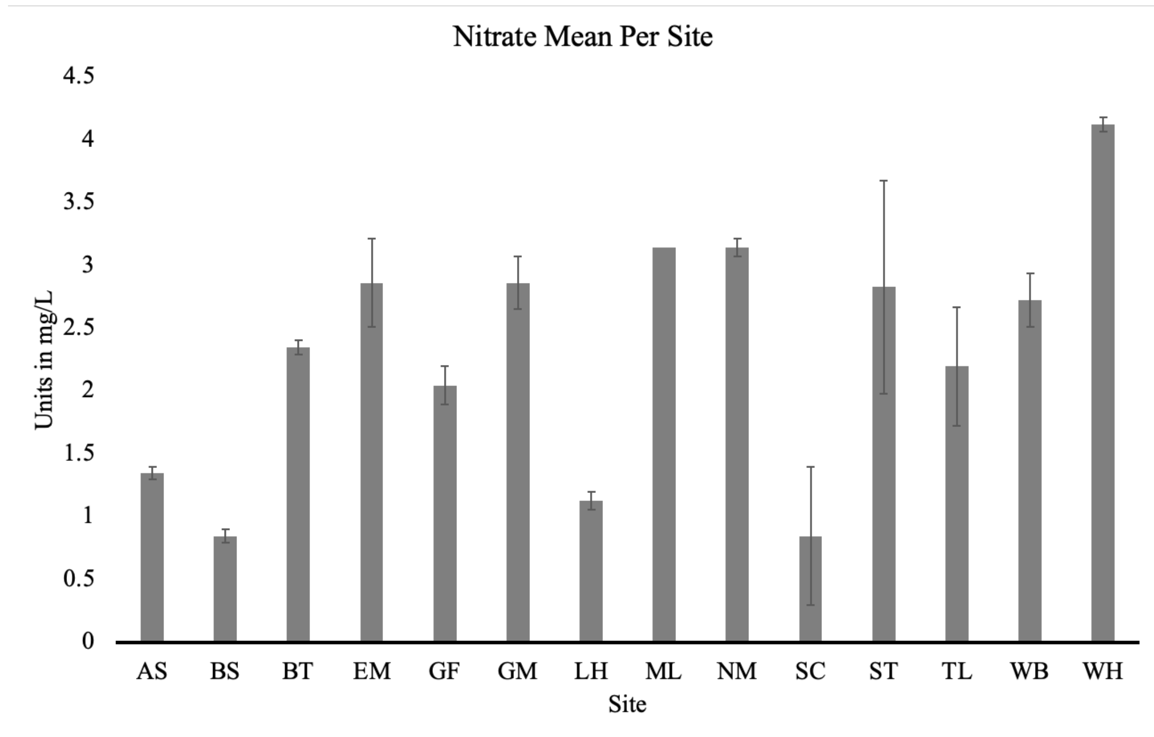

Figure 10: Nitrate means in mg/L per site sampled. Error bars represent the standard deviation of the sample means. Sites without standard deviation bars represent sites where only one nutrient measurement was taken or the nutrient measurements were equal and therefore did not have a standard deviation. Manson Pike (MP) did not have a measurement for nitrate and was omitted from the graph. Site key: AS: Above Spillway, BS: Below Spillway, BT: Broad Streat Trailhead, EM: Elam Mill, GF: Goochie Ford, GM: Gregory Mill, LH: Long Hunter, ML: Molly Lane, NM: Nice Mill, SC: Stewart Creek, ST: Southridge Trailhead, TL: Thompson Lane Trailhead, WB: Woodberry Bridge, WH: Walter Hill Dam.

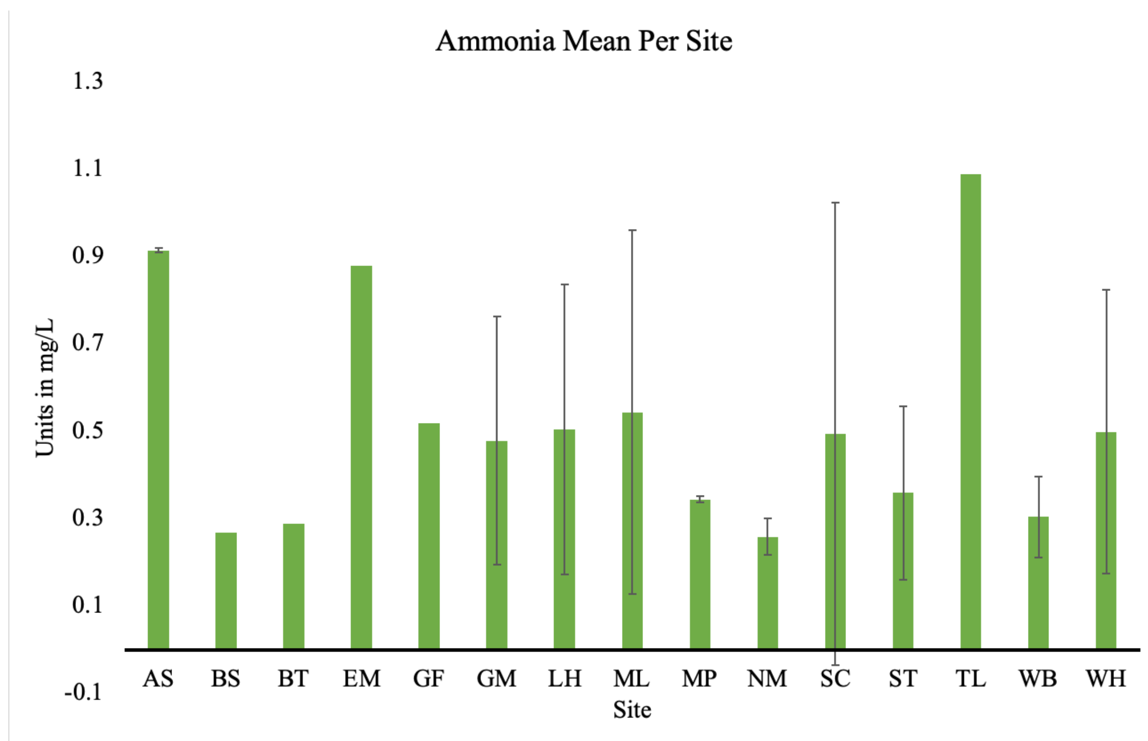

Figure 11: Ammonia means in mg/L per site sampled. Error bars represent the standard deviation of the sample means. Sites without standard deviation bars represent sites where only one nutrient measurement was taken or the nutrient measurements were equal and therefore did not have a standard deviation. Of note, Stewart Creek's (SC) standard error was larger than the site mean due to a large difference in concentration between sample dates. Site key: AS: Above Spillway, BS: Below Spillway, BT: Broad Streat Trailhead, EM: Elam Mill, GF: Goochie Ford, GM: Gregory Mill, LH: Long Hunter, ML: Molly Lane, MP: Manson Pike, NM: Nice Mill, SC: Stewart Creek, ST: Southridge Trailhead, TL: Thompson Lane Trailhead, WB: Woodberry Bridge, WH: Walter Hill Dam.

DISCUSSION

In 82 eDNA samples, we were able to distinguish 9,214 ASVs to at least the phylum level, and this high number likely speaks to high species diversity living within the Stones River Watershed. Major species richness and composition differences were observed between lotic and lentic environments, but the east and west forks of the Stones River were overall more similar with most composition differences observed at the level of collection site. Due to their relevance in conservation, we specifically focused on the

150 ASVs in class Actinopterygii and the 79 ASVs in phylum Mollusca. Within class Actinopterygii, we identified ASVs from 13 families and 33 genera including *Campostoma, Lepomis, Etheostoma, Pimephales, Hypentelium, and Moxostoma*. Within Mollusca, we were able to identify 8 families and 11 genera from the classes Bivalvia and Gastropoda, with four ASVs remaining at the phylum level. Nutrient concentrations varied greatly across sites and sampling dates. Despite this variation, we did not observe differences in nutrient concentrations between river branches or relative to river flow. Nutrient concentrations did not correlate with species richness and only was only weakly related to shifts in species composition across sampling sites in the Stones River watershed.

While species richness differed across lotic river branches, the most striking difference seemed to be between lotic and lentic environments. Both river branches displayed higher species richness than the three Percy Priest lake sites and two Stewarts creek sites, and species richness in the east fork was additionally higher than in the west fork. Lotic river sites on the east fork were surrounded by countryside (Woodberry Bridge), pasture (Goochie Ford), or a landfill (Walter Hill Dam). All east fork sites displayed a high species richness ranging from 322 (Goochie Ford) to 372 (Walter Hill) median unique ASVs. Lotic river sites on the west fork were surrounded by suburbs (Southridge Trailhead), a city (Molloy Lane and Manson Pike), a golf course (Broad Street Trailhead and Thompson Lane Trailhead), and countryside (Nice Mill). These highly variable sites on the west fork displayed a large range of species richness from 63 (Thompson Lane Trailhead) to 308 median ASVs (Broad Street Trailhead). It should be noted that Thompson Lane and Broad Street Trailhead are closer to each other than any

other sites on the west fork. Broad Street Trailhead was sampled on September 9, 2021, with river flow recorded at 117 cubic feet per second while Thompson Lane Trailhead was sampled on October 26, 2021 with river flow recorded at 98.2 cubic feet per second. Interestingly, Thompson Lane is downstream from Broad Street, and one possible explanation for this large difference in ASVs detection could be river flow or sample location within the river. All other sites had median ASVs that fell between these two sites. According to TWRA, 315 species of fish are known to occur in Tennessee's freshwater systems with at least 46 of those species occurring in Middle Tennessee (Mullen 2006). Our study was able to identify 150 ASVs of bony fish within 13 families and 33 genera with 4 ASVs not being classified beyond family. These data indicate that we were able to collect evidence for additional species of fish living in the Stones River that had not previously been recorded within the watershed (Mullen 2006). Tennessee also boasts an incredibly specious molluscan community spanning 460 total species and forms (Bogan et al. 1983). In 1998, there were 130 species of freshwater mussels that were known or had been known to be in Tennessee's waterways (bivalves; tn.gov). In 2011, review identified 38 freshwater gastropod species in Tennessee (Dillon et al. 2011) while a 2013 paper recognized 91 freshwater gastropod species (Johnson et al. 2013). We captured 78 ASVs of mollusk from Bivalvia and Gastropoda, with 4 ASVs not being classified beyond phylum. Our methods were able to capture 22 of the 130 species of bivalves that were once known to be in Tennessee rivers (tn.gov). Within Gastropoda, we captured 52 unique ASVs, 14 more possible species than previously identified in Dillon et al. (2011), though 39 fewer species than identified in Johnson et al. (2013).

Nutrient data were highly variable and seemed uninfluenced by river branch or river flow. Across all sites, there was no correlation between nutrient concentrations and species richness. This is a break from the literature on the relationship between water nutrient concentration and species richness. An *in-situ* experiment in 2020 found that eukaryotic, diatom, and bacterial communities changed in response to nutrient loading in seawater (Clark et al. 2020). A 2020 review found eDNA survey techniques effective in monitoring ecological response to harmful algal blooms (HABs) and eutrophication in terrestrial, marine, and freshwater environments (Liu et al. 2020). Most studies focus on the compositional change of microbial communities in response to nutrient concentration changes. It follows that these techniques have the potential to monitor species richness on a larger scale. Though there was no apparent relationship between species richness and nutrient concentrations, we did identify a relationship between all three nutrients and the species composition across all sites. The relationship only accounted for differences in about 6% of the overall variance in species composition. While a small amount, we believe this relationship supports our predictions that nutrient concentrations do in some way affect the organisms within the watershed.

In class Actinoptergi, we were able to identify 33 genera with the highest prevalence in Campostoma (stonerollers; 12 sites), Lepomis (sunfish; all 15 sites), Etheostoma (darters; 13 sites), Pimephales (bluntnose minnows; 10 sites), Hypentelium (suckers; 10 sites), and Moxostoma (redhorses; 11 sites). Many species were observable while in the field, including members of *Lepomis, Hypentilium, Pimephales,* and *Etheostoma* (Williams pers. obs). All genera captured through eDNA are known to inhabit Tennessee waterways and occur in Middle Tennessee (Fishbase). ASVs for

individual genera of fish exceeded the number of known species in the Stones River in several cases (Mullen 2006), which may suggest either hidden genetic diversity with the Stones River watershed or possibly the need for some adjustment of sequence clustering during eDNA bioinformatics analysis. Though composition differed across sites, there were no clear trends across branches or the river continuum. While the compositional data were less clear cut spatially, species richness differences across sites could be attributed to environmental differences, mainly lotic versus lentic environments. We detected greater species richness in the river branch sites than in the lake sites, with the east fork showing significantly more taxa than the west fork. This difference was surprising considering there were fewer sampling sites on the east fork than on the west fork and the record of pollution on the east fork. The east fork is known to have several issues with pollution, specifically with pharmaceuticals (Kaur et al. 2020) and mercury, such that fish are considered unsafe for consumption in parts of the east fork, especially around Walter Hill (TDEC 2019). Despite the pollution, the greater species richness may be due to the more rural nature of the east fork compared to the west fork, however further investigation is needed for conclusive evidence.

Within phylum Mollusca, we were able to identify 78 ASVs in Bivalvia and Gastropoda, with 4 ASVs unclassifiable past phylum. Organisms belonging to Gastopoda were observed in abundance on a rocky substrate at Southridge Trailhead. As shown in Figure 7, Southridge Trailhead showed the highest Mollusca species richness. Species richness was significantly different in lotic and lentic environments but did not differ across the lotic river sites. Within the 22 ASVs of Bivalvia, we detected 3 families and 6 genera. *Corbicula* (Asian clam), an invasive species, was the most prevalent genus, being

observed at 14 sites. This species was likely introduced to the United States in the 1920s and has since invaded nearly every watershed on the continent where it can threaten native populations of bivalves (Ferreira-Rodriguez et al. 2022). Six ASVs of freshwater mussels (family: Unionidae; genera: *Lampsilis, Medionidus, Villosa*) were observed and while these genera are common in the southeast, certain species within them are endangered (Williams et al. 1993; Lane et al. 2019). Within the 53 ASVs of Gastropoda, we observed 5 families and 5 genera. The most prevalent genus was *Pleurocera*, a common freshwater snail in eastern North America (Johnson et al. 2013). Other genera included *Deroceras, Ferissia, Physella, Menetus, and Lithasia*. These were the taxa abundant at Southridge Trailhead (Williams personal observation) and all taxa detected were known to exist within North America and Tennessee.

The use of eDNA meta-barcoding is an efficient and accurate method for conducting biodiversity surveys in a variety of ecosystems (Easson et al. 2020; Rees et al. 2014), including lentic and lotic freshwater (Deiner et al. 2016, 2017) and use of eDNA in biodiversity monitoring has increased in recent years (Liu et al. 2020). These methods are faster, less labor-intensive, and less invasive than traditional survey methods like temporary impoundments or trapping. Some studies have shown a correlation between ASV abundance and species abundance (Carraro et al. 2018), but quantitative estimates to date have only been applied to studies with target species (Carraro et al. 2018; Tilloston et al. 2018; Spear et al. 2021). With the relationship between eDNA abundance and species abundance not broadly accepted for all freshwater taxa, we have chosen not to infer species abundances from our data and to instead focus on species richness and detection. One current limitation of eDNA is the depth of the reference database of

genomic sequences. Positive identification of taxa is restricted by the availability of the sequence within the chosen database. Though eDNA metabarcoding has the potential to capture cryptic species that traditional surveys may miss, it still relies on a genetic reference database that may have limitations in some taxonomic areas. Thus, it is possible that our current dataset of positively identified ASVs is an underestimation of overall richness in these environments. As the understanding of eDNA and its implications increases, these methods have the potential to significantly streamline and boost conservation efforts. This study being the first eDNA survey of the Stones River watershed, we have just scratched the surface of what can be achieved in this watershed.

Though this study did not show a significant correlation between water nutrient concentration and species diversity, there is still an interesting story being told. We did show a connection between all three nutrient concentrations and species compositions across all sites. Though the nutrient concentration differences only account for shifts in 6% of the total taxa identified and therefore only influence a few organisms, we have shown some relationship between environmental nutrient concentrations and the organisms that live there. Further research is needed to explore and find causation this correlation. From the over 9,000 positively identified ASVs, we were able to distinguish nearly half of the known bony fish in Tennessee and several freshwater mollusks. Moreover, our study covered a relatively small area, meaning that the potential for even greater detection of species richness is possible through project upscaling. Overall, our data show a high species richness, specifically in bony fish where our ASV richness exceeds that of species richness in Tennessee (fishbase.org). Our study does not provide conclusive evidence of hidden biodiversity, but certainly generates interesting hypotheses

for future studies. This first eDNA-based survey of the Stones River Watershed provides a base for regional conservation efforts and proof the methodology is sound and shows great potential for future use. We have provided strong evidence for a great diversity of species and, albeit small, a correlation between nutrient concentrations and species composition within the Stones River Watershed.

REFERENCES

Bogan, A. & Parmalee, P. (1983) Tennessee's Rare Wildlife. Volume II: The Mollusks. Tennessee Wildlife Resource Agency.

Brown, L. C., & Broughton, C. W. (1981). A survey of pollution indicator bacteria in water wells of Rutherford County, Tennessee. Journal of the Tennessee Academy of Science, 56(3)

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nature methods*, *13*(7), 581-583.

Carraro, L., Hartikainen, H., Jokela, J., Bertuzzo, E., & Rinaldo, A. (2018). Estimating species distribution and abundance in river networks using environmental DNA. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(46), 11724–11729. https://doi.org/10.1073/pnas.1813843115

Clark, D. E., Pilditch, C. A., Pearman, J. K., Ellis, J. I., & Zaiko, A. (2020). Environmental DNA metabarcoding reveals estuarine benthic community response to nutrient enrichment - Evidence from an in-situ experiment. *Environmental pollution (Barking, Essex : 1987)*, *267*, 115472. https://doi.org/10.1016/j.envpol.2020.115472

Deiner, K., Fronhofer, E., Mächler, E. *et al.* (2016) Environmental DNA reveals that rivers are conveyer belts of biodiversity information. *Nat Commun* 7, 12544. https://doi.org/10.1038/ncomms12544

Deiner, K., Bik, H.M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D.M., de Vere, N. and Pfrender, M.E., (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular ecology*, *26*(21), pp.5872-5895.

Dillon, R. T., Jr. & M. Kohl. 2013. *The Freshwater Gastropods of Tennessee.* Internet address: http://www.fwgna.org/FWGTN

Easson, C. G., Boswell, K. M., Tucker, N., Warren, J. D., & Lopez, J. V. (2020). Combined eDNA and Acoustic Analysis Reflects Diel Vertical Migration of Mixed Consortia in the Gulf of Mexico. Frontiers in Marine Science.

Ferreira-Rodríguez, N., Gangloff, M., Shafer, G., & Atkinson, C. L. (2022). Drivers of ecosystem vulnerability to Corbicula invasions in southeastern North America. Biological Invasions, 1-12.

Froese, R. and D. Pauly. Editors. 2022. FishBase. World Wide Web electronic publication. www.fishbase.org, version (02/2022).

Jari Oksanen, F. Guillaume Blanchet, Michael Friendly, Roeland Kindt, Pierre Legendre, Dan McGlinn, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens, Eduard Szoecs and Helene Wagner (2020). vegan: Community Ecology Package. R package version 2.5-7. https://CRAN.R-project.org/package=vegan

Johnson, P.D., Bogan, A.E., Brown, K.M., Burkhead, N.M., Cordeiro, J.R., Garner, J.T., Hartfield, P.D., Lepitzki, D.A.W., Mackie, G.L., Pip, E., Tarpley, T.A., Tiemann, J.S., Whelan, N.V. and Strong, E.E. (2013), Conservation Status of Freshwater Gastropods of Canada and the United States. Fisheries, 38: 247-282. https://doi.org/10.1080/03632415.2013.785396

Kaur, R., Akuley-Amenyenu, A., Garner, K. G., & Dennis, S. O. (2020). Detection of Human Pharmaceuticals in the Surface Water of East Fork Stones River. Journal of Water Resource and Protection, 12(3), 240-259.

Lane, T. W., Hallerman, E. M., & Jones, J. W. (2019). Population genetic assessment of two critically endangered freshwater mussel species, Tennessee bean Venustaconcha trabalis and Cumberland bean Venustaconcha troostensis. *Conservation Genetics*, *20*(4), 759-779.

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.

Leray, M., Haenel, Q., & Bourlat, S. J. (2016). Preparation of Amplicon Libraries for Metabarcoding of Marine Eukaryotes Using Illumina MiSeq: The Adapter Ligation Method. *Methods in molecular biology (Clifton, N.J.)*, *1452*, 209–218. https://doi.org/10.1007/978-1-4939-3774-5_14

Leray, M., Ho, S. L., Lin, I. J., & Machida, R. J. (2018). MIDORI server: a webserver for taxonomic assignment of unknown metazoan mitochondrial-encoded sequences using a curated database. *Bioinformatics*, *34*(21), 3753-3754.

Liu, Q., Zhang, Y., Wu, H., Liu, F., Peng, W., Zhang, X., Chang, F., Xie, P., & Zhang, H. (2020). A Review and Perspective of eDNA Application to Eutrophication and HAB Control in Freshwater and Marine Ecosystems. *Microorganisms*, *8*(3), 417. https://doi.org/10.3390/microorganisms8030417

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal*, *17*(1), 10-12.

McMurdie, P. J., & Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one*, *8*(4), e61217.

Mullen, D. (2006). Fish Inventory at Stones River National Battlefield.

Niemiller, M.L., Glorioso, B.M., Nicholas, C., Phillips, J., Rader, J., Reed, E., Sykes, K.L., Todd, J., Wyckoff, G.R., Young, E.L. and Miller, B.T.,(2009). Notes on the reproduction of the Streamside Salamander, Ambystoma barbouri, from Rutherford County, Tennessee. *Southeastern Naturalist*, *8*(1), pp.37-44.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre. P., McGlin, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Henry, M., Stevens, H., Szoecs, E., Wagner, H. (2020) vegan: community Ecology Package. R package version 2.5-7

Ratnasingham, S., & Hebert, P. D. (2007). bold: The Barcode of Life Data System (http://www.barcodinglife.org). *Molecular ecology notes*, *7*(3), 355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x

Rees, H. C., Maddison, B. C., Middleditch, D. J., Patmore, J. R., & Gough, K. C. (2014). The detection of aquatic animal species using environmental DNA–a review of eDNA as a survey tool in ecology. *Journal of applied ecology*, *51*(5), 1450-1459.

Spear, M. J., Embke, H. S., Krysan, P. J., & Vander Zanden, M. J. (2021). Application of eDNA as a tool for assessing fish population abundance. *Environmental DNA*, *3*(1), 83-91.

Streamer Report. (2021) Stream Trace Detailed Report. www.usgs.gov

TDEC. (2019) Division of Water Resources – Posted Streams, Rivers & Reservoirs.

Tillotson, M. D., Kelly, R. P., Duda, J. J., Hoy, M., Kralj, J., & Quinn, T. P. (2018). Concentrations of environmental DNA (eDNA) reflect spawning salmon abundance at fine spatial and temporal scales. *Biological Conservation*, *220*, 1-11.

TWRA. (2012). The Angler's Guide to Tennessee Fish.

Tn.gov. https://www.tn.gov/twra/wildlife/fish/fresh-water-mussels-in-tennessee.html

Wang Q. et al. (2007) Naïve bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl. Environ. Microb., 73, 5261–5267.

Wickham H (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org

Williams, J. D., Warren Jr, M. L., Cummings, K. S., Harris, J. L., & Neves, R. J. (1993). Conservation status of freshwater mussels of the United States and Canada. *Fisheries*, *18*(9), 6-22.