Summer 1-1-2017

Distribution Changes, Genetic Population Structure, and a Novel Environmental DNA (eDNA) Detection Method for Darters (Subgenus Nothonotus) in the Upper Ohio River Watershed

Anthony S. Honick
Duquesne University

DISTRIBUTION CHANGES, GENETIC POPULATION STRUCTURE, AND A NOVEL ENVIRONMENTAL DNA (eDNA) DETECTION METHOD FOR DARTERS (SUBGENUS *NOTHONOTUS*) IN THE UPPER OHIO RIVER WATERSHED

A Dissertation

Submitted to the Bayer School of Natural and Environmental Sciences

Duquesne University

In partial fulfillment of the requirements for

the degree of Doctor of Philosophy

By

Anthony S. Honick

August 2017

DISTRIBUTION CHANGES, GENETIC POPULATION STRUCTURE, AND A NOVEL ENVIRONMENTAL DNA (eDNA) DETECTION METHOD FOR DARTERS (SUBGENUS *NOTHONOTUS*) IN THE UPPER OHIO RIVER WATERSHED

By

Anthony S. Honick

Approved April 26, 2017

Dr. Brady A. Porter Associate Professor of Biological Sciences (Committee Chair)	Dr. David G. Argent Professor of Biology California University of Pennsylvania (Committee Member)
Dr. Michael I. Seaman Associate Professor of Biological Sciences (Committee Member)	Dr. John F. Stolz Professor of Biological Sciences (Committee Member)
Dr. Phillip P. Reeder Dean, Bayer School of Natural and Environmental Sciences	Dr. Joseph R. McCormick Chair, Department of Biological Sciences Associate Professor of Biology

ABSTRACT

DISTRIBUTION CHANGES, GENETIC POPULATION STRUCTURE, AND A NOVEL ENVIRONMENTAL DNA (eDNA) DETECTION METHOD FOR DARTERS (SUBGENUS *NOTHONOTUS*) IN THE UPPER OHIO RIVER WATERSHED

By

Anthony S. Honick

August 2017

Dissertation supervised by Dr. Brady A. Porter

In the upper Ohio River watershed three species of small-bodied benthic fish the Bluebreast Darter, *Etheostoma* (*Nothonotus*) *camurum* (Cope), the Tippecanoe Darter, *Etheostoma* (*Nothonotus*) *tippecanoe* Jordan and Evermann, and the Spotted Darter, *Etheostoma* (*Nothonotus*) *maculatum* Kirtland previously existed in disjunct distributions due to poor water quality and habitat degradation. Signs of recovery indicated that these species were moving from areas of refugia into the deeper mainstem waters of the Allegheny and Ohio rivers and expanding their distributions. To provide information for the proper conservation management of these species this dissertation was divided into three stages: 1) distribution records were updated by performing State-mandated electrified-benthic trawling and compiling as many historic and contemporary records as possible, 2) the genetic diversity and genetic population structure of *E. camurum* was assessed using six polymorphic microsatellite loci, and 3) environmental DNA

(eDNA) methods with species detection from water samples via fragment analysis were developed to assist current survey methods which are costly, time consuming, and may be harmful to the fish. The surveys and compilation of data showed that *E. camurum* and *E. tippecanoe* are utilizing deeper habitat (than previously reported) in the tailwaters of the navigational lock and dam system, and have nearly continuous distributions from the upper Allegheny River downstream into the Ohio River. *Etheostoma maculatum* showed a less robust expansion and a more limited use of the tailwater habitat. The genetic assessment of *E. camurum* indicated high genetic diversity within their populations with no evident signs of isolation or inbreeding. The genetic population structure of *E. camurum* was weak indicating that the navigational lock and dam system was not strongly influencing gene flow between the populations. In addition, there were signs of a newly advancing population. With eDNA methodologies, a protocol was developed that successfully detected *E. tippecanoe* eDNA from water samples taken from the Allegheny and Kiskiminetas rivers and Deer Creek in Harmarville, PA. A second set of PCR primers were developed that have the potential to detect all three focal species using eDNA from water samples.

DEDICATION

I would like to dedicate this dissertation to my family who helped me in so many ways that I can't even begin to thank them enough. I especially want to thank my wife, Sandi, and my two wonderful sons, TJ and Craig, all of whom sacrificed many things so that I could complete this journey. Without my parents, William and MaryAnn, and all of their support, none of this would have been possible.

ACKNOWLEDGEMENTS

I want to thank Brady Porter for his guidance and patience as I prolonged my journey by completing not one, but two internships. I truly appreciated his empathy when things got tough, and I am grateful for all of the skills he passed on: give a man a fish and feed him for a day, but teach him how to electrofish and feed him for life! Thanks also to my other committee members: Mike Seaman, John Stolz, and Dave Argent who provided helpful advice along the way. Lastly, I owe a huge amount of gratitude to all of the faculty, staff, and students (particularly those in the Porter Lab) in the Biology Department who helped me through my classes and with my research.

I would also like to thank Doug Locy (Aquatic Systems, Inc.) and Mike Koryak (U.S. Army Corps of Engineers, retired) for assistance with the electrified-benthic surveys, and all the individuals who helped with tributary electrofishing. I would also like to thank Elizabeth Dakin for help with sampling, providing comments on my manuscript, and providing technical advice at all stages. Thanks to all the agencies that supplied historic survey data including: Pennsylvania Natural Heritage Program, the Pennsylvania Fish and Boat Commission (Doug Fischer), Pennsylvania Department of Environmental Protection (Rick Spear), the Ohio Environmental Protection Agency, and the Ohio River Valley Water Sanitation Commission. I would also like to thank the U.S. Geological Survey for providing funds to the Pennsylvania Water Resources Research Center that funded my electrified-benthic trawling surveys and the North American Native Fishes Association for awarding me a student grant that helped fund eDNA primer screening.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF FIGURES, Continued

Page

CHAPTER ONE

Expanded Distributions of Three *Etheostoma* Darters (subgenus *Nothonotus*) within the Upper Ohio River Watershed

Within the upper Ohio River watershed, three *Etheostoma* darter species in the subgenus *Nothonotus* have been documented in disjunct populations and were listed as threatened or endangered in Ohio and Pennsylvania. Tailwater habitat below navigation lock and dam (L/D) installations has been shown to contain diverse darter assemblages. Because *Etheostoma camurum* (Bluebreast Darter), *Etheostoma maculatum* (Spotted Darter), and *Etheostoma tippecanoe* (Tippecanoe Darter) often live in similar habitats, I hypothesized that all three were occupying tailwater habitat below navigational L/Ds. Electrified-benthic trawling verified *E. camurum* and *E. tippecanoe* below eight L/D installations and at depths ranging from 1.4 to 4.5 m and 1.4 to 5.9 m, respectively. *Etheostoma maculatum* was only found below one L/D. In the Ohio River, benthic trawling documented *E. camurum* and *E. tippecanoe* utilizing habitat located within deposition zones and areas above and below islands. Analysis of contemporary and historic distribution data shows that *E. camurum* and *E. tippecanoe* now span large sections of the river, but the range of *E. maculatum* is more limited and warrants close monitoring. Our study confirms the effectiveness of utilizing benthic trawling in non-wadeable rivers to survey for benthic species such as river-inhabiting darters.

Contributions

Chapter one of this dissertation has been accepted for publication as of 04 April 2017, in the peer-reviewed journal of the *Northeastern Naturalist* 2017 24(2):209–234. The following provides details of each author's contributions:

Anthony S. Honick

- Co-authored the grant that was obtained from U.S. Geological Survey and administered through the Pennsylvania Water Resources Research Center
- Completed two seasons of field work to gather data
- Tracked down and processed all contemporary and historic data from Pennsylvania
    - Quality control checked all data (OH and PA) and converted it into a uniform format for GIS application
- Created all table and figures
- Drafted, revised, and submitted final version of manuscript for publication

Brian J. Zimmerman

- Provided all site data – contemporary and historic from Ohio
- Queried data from the Ohio State University Museum of Biological Diversity – Fish database
- Processed Ohio data into an organized summary by species
- Provided comments on the first draft and then provided a substantial amount of revision to the final version to satisfy a reviewer's comments about information needing an Ohio-based perspective

Jay R. Stauffer Jr.

- Provided a portion of the PA data being housed by the Penn State University Museum
- Provided comments on an early draft of the manuscript

David G. Argent

- Provided data from his personal collection records for PA
- Provided extensive comments at all stages of manuscript preparation

Brady A. Porter

- Provided field assistance and monetary support from a grant we co-authored
- Guided manuscript preparation
- Provided comments throughout manuscript preparation

Expanded Distributions of Three *Etheostoma* Darters (Subgenus *Nothonotus*) within the Upper Ohio River Watershed

Anthony S. Honick[1,*], Brian J. Zimmerman[2], Jay R. Stauffer Jr.[3], David G. Argent[4], and Brady A. Porter[1]

Abstract - Within the upper Ohio River watershed, 3 *Etheostoma* darter species in the subgenus *Nothonotus* have been documented in disjunct populations and were listed as threatened or endangered in Ohio and Pennsylvania. Tailwater habitat below navigation lock and dam (L/D) installations has been shown to contain diverse darter assemblages. *Etheostoma camurum* (Bluebreast Darter), *E. maculatum* (Spotted Darter), and *E. tippecanoe* (Tippecanoe Darter) often live in similar habitats; thus, we hypothesized that all 3 were occupying tailwater habitat below navigational L/Ds. Electrified benthic trawling verified Bluebreast Darter and Tippecanoe Darter below 8 L/D installations and at water depths varying from 1.4 m to 4.5 m and 1.4 m to 5.9 m, respectively. Spotted Darter was only found below 1 L/D. In the Ohio River, benthic trawling documented Bluebreast Darter and Tippecanoe Darter utilizing habitat located within deposition zones and areas above and below islands. Analysis of contemporary and historic distribution data shows that Bluebreast Darter and Tippecanoe Darter now span large sections of the river, but the range of Spotted Darter is more limited and warrants close monitoring. Our study confirms the effectiveness of utilizing benthic trawling in non-wadeable rivers to survey for benthic species such as river-inhabiting darters.

Introduction

Etheostoma (*Nothonotus*) *camurum* (Cope) (Bluebreast Darter) was described from the headwaters of the Cumberland River in Tennessee (Cope 1870) and is known to have variable population sizes (Page 1983, Trautman 1981) and a disjunct distribution in the upper Allegheny drainage (PA, NY); Cheat, Little Kanawha, and Elk river drainages (WV); Walhonding and Scioto drainages (OH); Wabash drainage (IN, IL); Cumberland drainage (KY, TN); Licking and upper Kentucky drainages (KY); and Duck, Elk, and upper Tennessee drainages (TN, AL, VA) (see Supplemental File 1 available online at http://www.eaglehill.us/NENAonline/suppl-files/n24-2-N1537-Honick-s1, and, for BioOne subscribers, at http://dx.doi.org/10.1656/N1537.s1). Bluebreast Darter habitat is reported to consist of moderate to swift riffles, raceways, and runs of moderate- to large-sized clear streams, and

[1]Bayer School of Natural and Environmental Sciences, Duquesne University, Pittsburgh, PA 15282. [2]School of Environment and Natural Resources, Ohio State University Museum of Biodiversity, Columbus, OH 43212. [3]Ecosystem Science and Management, The Pennsylvania State University, University Park, PA 16802. [4]Department of Biological and Environmental Sciences, California University of Pennsylvania, California, PA 15419 [*]Corresponding author - ashonick@gmail.com.

Manuscript Editor: David Halliwell

rivers running over silt-free boulders, cobble, and gravel at depths of 0.5 m to 1.5 m (Boschung et al. 2004, Etnier and Starnes 1993, Stauffer et al. 1995, Trautman 1981).

Etheostoma (*Nothonotus*) *tippecanoe* Jordan and Evermann (Tippecanoe Darter) was described from the Tippecanoe River at Marshland, IN (Jordan and Evermann 1890). This species is known to have dramatic year-to-year variation in population sizes (Trautman 1981), and Stauffer (2016) noted that populations in French Creek, PA, cycled every 3 years. Tippecanoe Darters have disjunct distributions in the upper Allegheny drainage (PA); Elk and Little Kanawha rivers (WV); lower Muskingum River, Walhonding River and the Scioto River drainage (OH); East Fork White River and upper Wabash River drainage (IN); Licking River, Kentucky River drainage and Green River (KY); and Big South Fork, Red Stones, and Harpeth rivers (TN) (see Supplemental File 1 available online at http://www.eaglehill.us/NENAonline/suppl-files/n24-2-N1537-Honick-s1, and, for BioOne subscribers, at http://dx.doi.org/10.1656/N1537.s1). Tippecanoe Darters inhabit riffles of medium to large rivers with slow to moderate currents and substrates of clean, fine gravel, sand, and cobble (Cooper 1983, Etnier and Starnes 1993, Trautman 1981).

Etheostoma (*Nothonotus*) *maculatum* Kirtland (Spotted Darter) was described from the Mahoning River near Youngstown, OH (Kirtland 1840), but that population was extirpated by pollution from a steel mill sometime in the mid-1850s (Trautman 1981). Historically, the Spotted Darter has been found in low population densities over a few disjunct localities in the Beaver River (now extirpated) and upper Allegheny drainage (PA and NY), Elk River (WV), Walhonding River and the Scioto River drainage (OH), Tippecanoe River (IN), and the Green River (KY) (see Supplemental File 1 available online at http://www.eaglehill.us/NENAonline/suppl-files/n24-2-N1537-Honick-s1, and, for BioOne subscribers, at http://dx.doi.org/10.1656/N1537.s1). Of these 3 species, the Spotted Darter is less broadly distributed (Kuehne and Barbour 1983, Page 1983); the largest number of remaining populations occur in the upper Allegheny River drainage of Pennsylvania and Big Darby Creek in Ohio (see Supplemental File 2 available online at http://www.eaglehill.us/NENAonline/suppl-files/n24-2-N1537-Honick-s2, and, for BioOne subscribers, at http://dx.doi.org/10.1656/N1537.s2). Spotted Darters inhabit swift riffles in medium to large streams associated with large cobble and boulder substrates (Zorach and Raney 1967).

Until recently, all 3 species were listed as either threatened or endangered by the Ohio Division of Wildlife (ODNR 2015, Ohio Revised Code 2015) and the Pennsylvania Fish and Boat Commission (Pennsylvania Bulletin 1999). In the most recent compilation of imperiled North American freshwater fishes by the American Fisheries Society, Jelks et al. (2008) listed the Tippecanoe Darter as vulnerable and the Spotted Darter as threatened and declining.

Even though it is generally accepted that the nation's waterways have experienced improved water quality conditions since implementation of the Clean Water Act (1972), nationwide assessments by Brown and Froemke (2012) and the US Environmental Protection Agency (2009) indicate that the nation's water resources are experiencing increased stress from nonpoint-source pollution. Jelks et al. (2008) corroborated these claims and reported that imperilment of inland fishes

had substantially increased since the last assessment completed by the American Fisheries Society in 1989. In contrast, and on a regional scale, surveys since 2003 in Pennsylvania (Argent and Kimmel 2010; Freedman et al. 2009a; Howell 2007; Koryak et al. 2009, 2011) and assessments from Yoder et al. (2005) and the Ohio EPA (OEPA 2016) have found that the fish communities in non-wadeable rivers of the upper Ohio River watershed were recovering. Regardless, there remains a pressing need to accurately document the return of imperiled fishes from refugia (e.g., French Creek, PA, and Big Darby Creek, OH) and track their distributional changes for future assessment of imperilment. In this study, we documented the changes in the distribution of these 3 focal darter species that have been increasing in occurrence outside of their known refugia. Several factors have contributed to elucidating these changes: (1) in Ohio, extensive routine monitoring using rigorous boat-electrofishing protocols and trawling have documented changes throughout the mainstem Ohio River and (2) in Pennsylvania, historically there may have been less-rigorous routine sampling efforts, but there has been a recent switch in sampling protocols to include benthic and electrified benthic trawling in non-wadeable rivers and within lock and dam (L/D) tailwaters. Efforts in Pennsylvania and Ohio have demonstrated the extent to which these 3 darter species now occupy non-wadeable rivers in the upper Ohio River watershed. In addition, because previous work indicated the importance of L/D tailwater habitat to benthic riverine fish species (Argent and Kimmel 2014, Freedman et al. 2009a, Koryak et al. 2009), we hypothesized that these darters in Pennsylvania were occupying tailwater habitat below L/D installations. In summary, we performed electrified benthic trawling surveys and compiled contemporary and historic data from multiple sources in Ohio and Pennsylvania to re-assess the darters' current distributions. These data have increased our understanding of the focal species' distributions within non-wadeable rivers and provided a summary of regional distribution changes that are imperative to documenting recovery since the delisting of Bluebreast Darter (in Ohio and Pennsylvania) and Tippecanoe Darter and Spotted Darter (in Pennsylvania).

Study Area

We analyzed contemporary and historic fish-survey data collected from rivers and streams in the upper Ohio River watershed of Pennsylvania and Ohio. Target water bodies included the mainstem rivers and tributaries of the Ohio River from river kilometer (rkm) 790.0 (the Ohio/Indiana border) upstream to Pittsburgh, PA (rkm 0), the Allegheny River from Pittsburgh, PA, upstream to the Pennsylvania/New York border, and the Monongahela River from Pittsburgh, PA, to the Pennsylvania/West Virginia border (Fig. 1).

Methods

Sampling methods

In Pennsylvania, we sampled the tailwaters of 11 L/D installations on 4 river systems (Allegheny, Beaver, Monongahela, and Ohio rivers; Fig. 2) using a modified Missouri trawl (2.4 m x 1.2 m, 3.2-mm mesh) electrified with a Smith-Root VI-A

electrofisher (Smith-Root Vancouver, WA) and a 5000-W generator. The unit was powered with an output mode of 6.0 amps, 120 PPS DC, and 6.0 ms pulse width. We established a transect within the tailrace of the L/D (50 m to 150 m below the installation) as the starting point for 7 trawls and placed 1 trawl each within 10 m of the left and right descending bank, 1 trawl at center channel, and the 4 remaining trawls evenly spaced between center-channel and the descending bank. We manually deployed the trawl from the bow of a 6.1-m Sea Ark Jon-type boat with a 115-hp outboard motor moving backwards downstream at a speed slightly faster than river current. We aborted snagged trawls and started a new trawl adjacent to the original location. We used river depth to determine the length of rope deployed with each trawl with the following guidelines: 5.0 m of depth or less = 15.2 m of rope, 5.0 to 10.0 m of depth = 30.5 m of rope. Each trawl consisted of 2 minutes of sampling effort. We identified and enumerated all fish species. In addition to the electrified-benthic trawling, we electrofished 9 tributaries (Smith Root LR-24, backpack electrofisher; single pass) for 100 m starting at the first riffle upstream of the confluence with the main river. We sampled streams with moderate to high flow by electrofishing into a blocking seine (2.4 m x 1.8 m, 3.2-mm mesh). We identified and enumerated all fish species.

In Ohio from 2011 to 2014, we sampled the Ohio River from the Indiana/Ohio border to the Ohio/Pennsylvania border, the entire length of the Muskingum

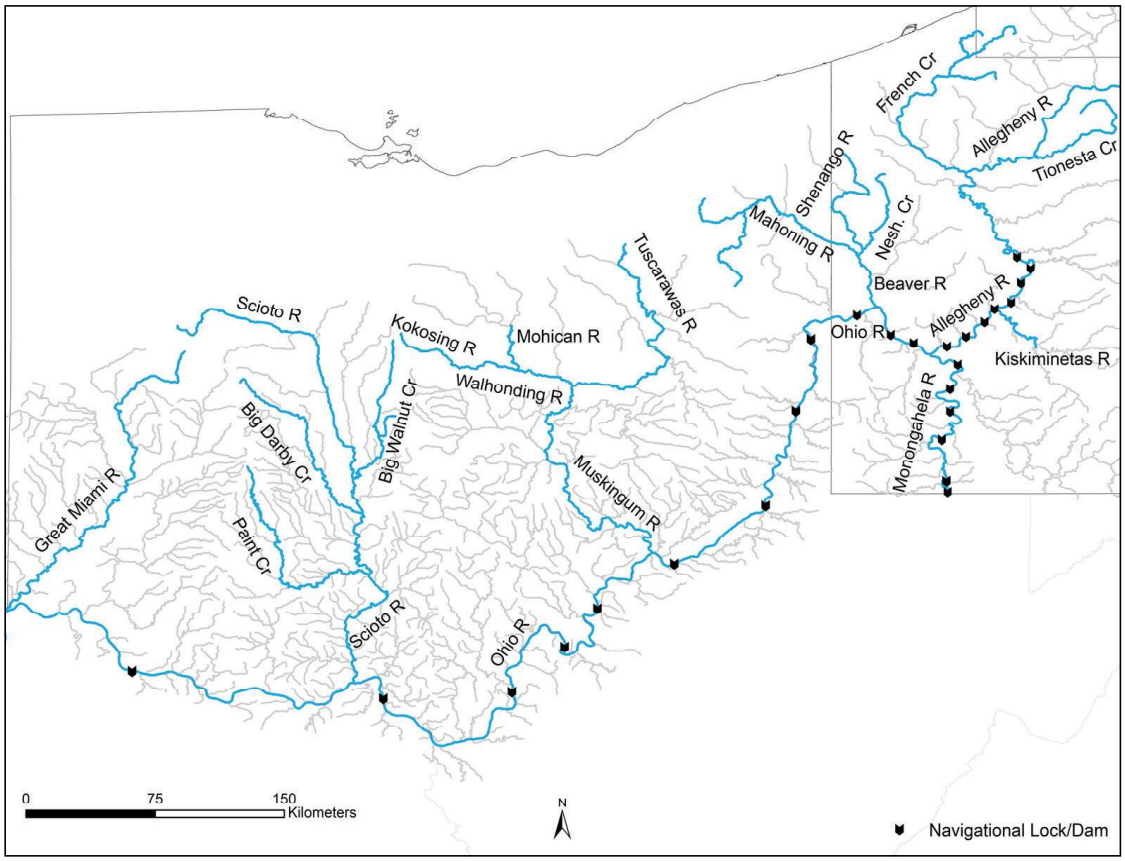

Figure 1. The major rivers and tributaries of the upper Ohio River watershed in Pennsylvania and Ohio.

River, and the Scioto River from Chillicothe to its confluence with the Ohio River at Portsmouth. We sampled all 3 systems approximately every 8 rkm (i.e., about every 5 river miles). The intent of the Ohio portion of this study was to provide presence/absence data for an improved understanding of the Ohio distribution of small benthic fishes in these large rivers. To conduct our samples, we employed a 2.4-m-wide mini-Missouri trawl (Innovative Net Systems, Milton, LA; Herzog et al. 2005) and a 1.4 m x 2.4 m (4.8-mm mesh) seine with a chain added to the lead line for better benthic contact. The trawl was manually deployed from the bow of a 4.8-m flat-bottomed boat (run in reverse) equipped with two 30-hp outboard motors at a target speed of 3.2–6.4 kmph. We classified as suitable habitat in non-wadeable areas all locations with unique features such as depositional zones at tributary mouths, current breaks at upstream and downstream ends of islands, tailwaters, and other areas of significant flow that might support darters and other benthic fishes. We conducted a minimum of 4 trawls varying from 30 s to 60 s in duration at each location. Large areas of suitable habitat were sampled more rigorously (e.g., at least 10 trawls). We sampled wadeable areas in the mainstem Ohio, Muskingum, and Scioto rivers with suitable habitat (as described above) using both kick seining and

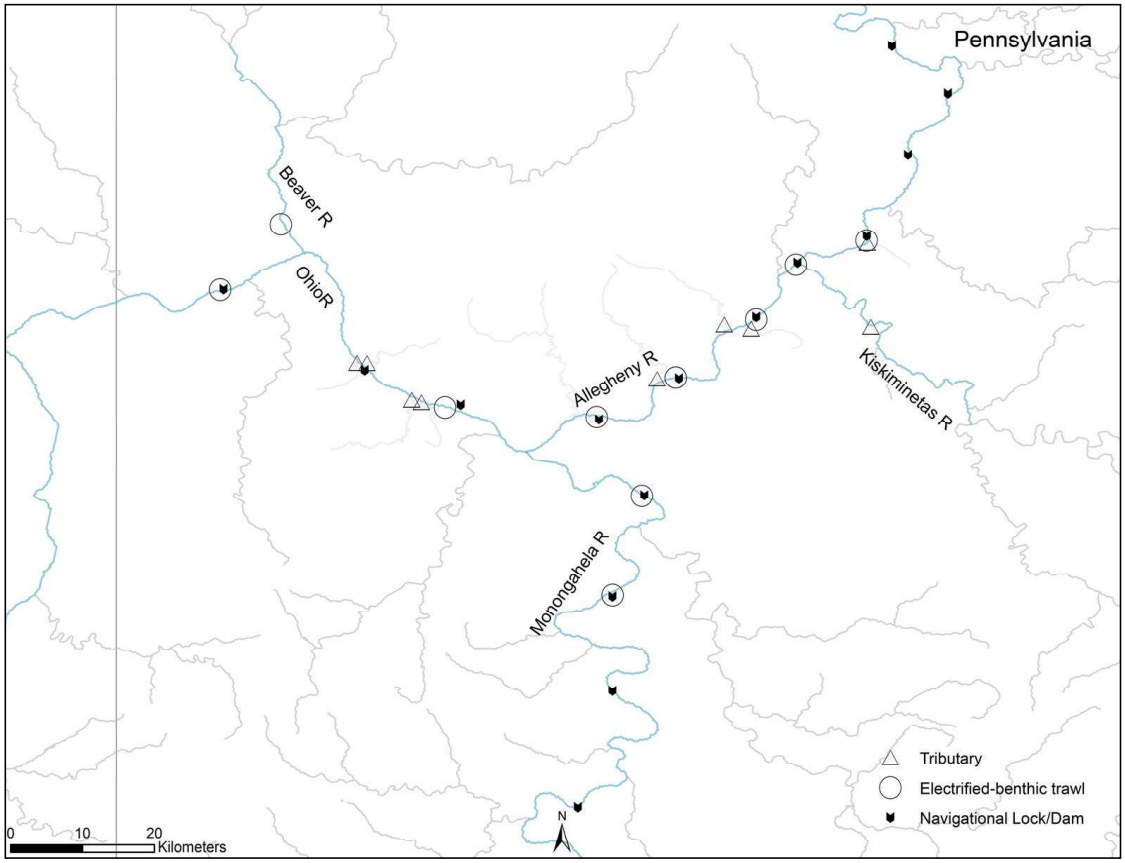

Figure 2. Sample locations of electrified-benthic trawls and the 9 tributaries sampled in this study. Note: Emsworth L/D consists of a main channel and a separate back-channel dam which we counted as 2 separate sample locations. All electrified trawls were conducted in dam tailwaters. The dam on the Beaver River is not a navigational lock and dam; therefore, the sample site is only designated with an open circle.

downstream hauls. When access was possible either by road or boating upstream from the mainstem river, we employed the seining methods described above to sample the first several riffles in the tributaries to these larger rivers. In this study, we use the term "traditional methods" to collectively refer to all fish-survey methods that were used prior to the development and implementation of the mini-Missouri trawl by Herzog et al. (2005). These protocols included backpack electrofishing, boat and tow-barge electrofishing, and all types of seining methods employed to survey for benthic fishes.

Historical data and map construction

We queried historical and contemporary survey data for Ohio from the Ohio State University Museum of Biological Diversity Fish Division database (OSU-MBD 2015; Table 1). OEPA is the largest contributor with over 400,000 records dating back to 1975. We compiled information on relative abundance, specific location, museum record/collection number, and gear type (see Supplemental File 2 available online at http://www.eaglehill.us/NENAonline/suppl-files/n24-2-N1537-Honick-s2, and, for BioOne subscribers, at http://dx.doi.org/10.1656/N1537.s2). Specific sampling protocols can be obtained from each respective agency.

We obtained historic and contemporary data from as many sources as possible for Pennsylvania records (Table 1). Raney (1938) compiled historic records for western Pennsylvania dating back to 1817. We collected information on relative abundance, specific location, museum record/collection number, and gear type (see Supplemental File 2 available online at http://www.eaglehill.us/NENAonline/suppl-files/n24-2-N1537-Honick-s2, and, for BioOne subscribers, at http://dx.doi.org/10.1656/N1537.s2). Specific sampling protocols can be obtained from each respective agency.

We quality-checked historic and contemporary records from Ohio and Pennsylvania for errors (e.g., duplicates, incorrect coordinates). When possible, we assigned coordinates based on site descriptions of the original survey record for historic records that did not have coordinate data and removed ambiguous

Table 1. Sources of historic and contemporary data.

State	Source
Ohio	
	Ohio State University Museum of Biological Diversity, Fish Division Database
	Ohio Environmental Protection Agency (OEPA)
	Ohio Department of Natural Resources - Division of Wildlife
	Ohio River Valley Water Sanitation Commission (ORSANCO) - http://www.orsanco.org/data/fish-population/
Pennsylvania	
	Pennsylvania Fish and Boat Commission (PAFBC)
	Pennsylvania Department of Environmental Protection
	US Army Corps of Engineers - Pittsburgh District
	The Pennsylvania State University Museum - Fish Collection
	California University of Pennsylvania
	Pennsylvania Natural Heritage Program

records. We constructed all distribution maps in ArcMap (v. 10.3.1; ESRI, Redlands, CA). In order to visualize distribution changes, we constructed maps for each species by grouping the data into 5 time-categories: pre-1981, 1981–1990, 1991–2000, 2001–2010, and 2011–2015. We plotted symbols denoting previous survey data (before 2011) on top of the most recent survey data to enhance visualization of distribution changes. For clarification, the terms "record" and "site record" both indicate that the respective species was positively identified during a sampling event at a specific location.

Results

Bluebreast Darter

Pre–1981. Survey records document Bluebreast Darter in a limited number of drainages across Ohio and Pennsylvania (Fig. 3). In Ohio, the Great Miami, Scioto, and Muskingum river watersheds contained extant populations of Bluebreast Darter. We found a total of 106 records, 60 of which were from sites located within Big Darby Creek (Fig. 3; Osburn 1901, OSU-MBD database 2015, Trautman 1981).

In Pennsylvania, populations of Bluebreast Darter were confined to the upper reaches of the Allegheny River, French Creek, and the tributaries that form the Beaver River (Fig. 3). Nine of the 19 records were documented in French Creek

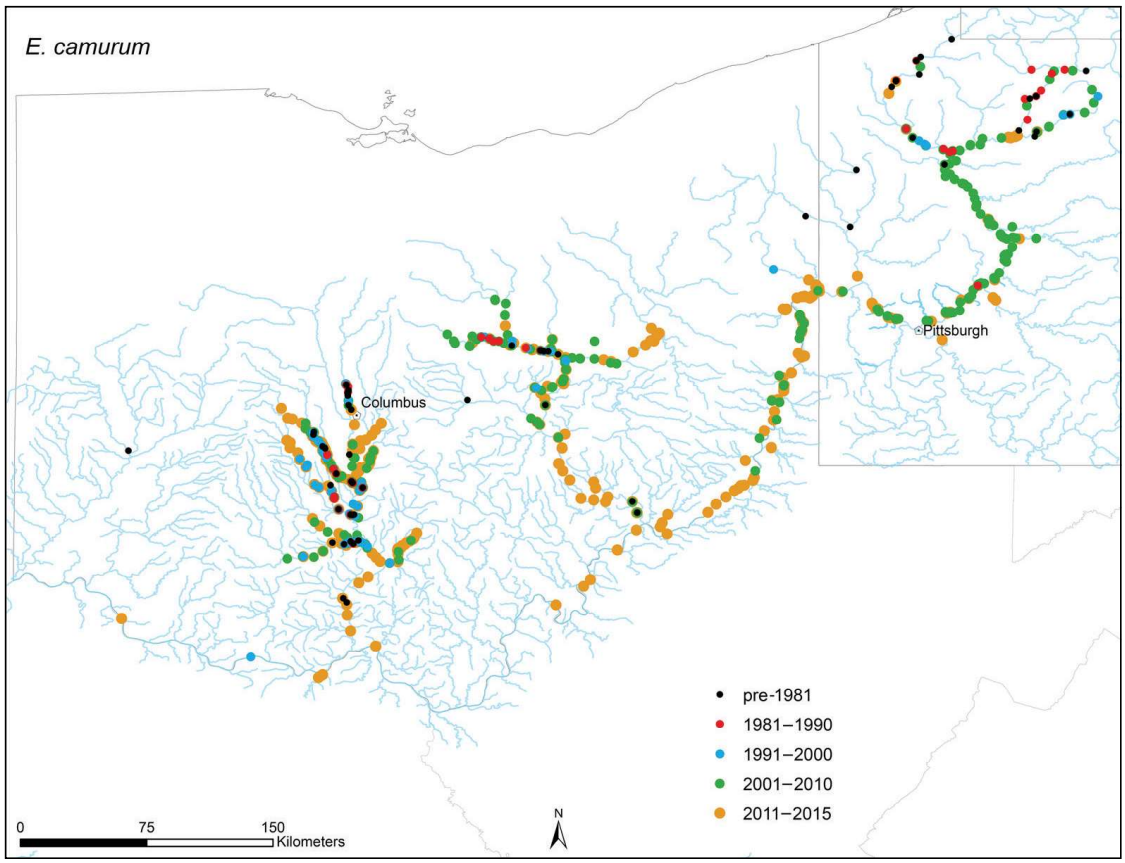

Figure 3. Distribution of Bluebreast Darter in the upper Ohio River system in Ohio and Pennsylvania showing all historic and contemporary data from pre-1981 to 2015.

(Pennsylvania Natural Heritage Program 2015, Raney 1938, Schwartz 1965); the remainder were located in the upper Allegheny River, Tionesta Creek, Little Coon Creek, and Sandy Creek (Pennsylvania Natural Heritage Program 2015, Raney 1938). Two additional locations in the upper Beaver River drainage were documented in the Shenango River and Neshannock Creek (Pennsylvania Natural Heritage Program 2015, Raney 1938).

1981–1990. Additional records for Bluebreast Darter between 1981 and 1990 showed minimal changes in distribution in Ohio and Pennsylvania. Within Ohio, 37 records showed an increased presence in Big Darby Creek, Deer Creek, Paint Creek, Kokosing, Walhonding, and Olentangy rivers (Fig. 3). During this time period, the OEPA greatly increased the amount of fish-sampling effort conducted across Ohio.

In Pennsylvania, records from the Pennsylvania Natural Heritage Program (2015) documented Bluebreast Darter at 14 sites—13 within French Creek and the upper Allegheny River above Tionesta (Fig. 3). Interestingly, 1 verified record in 1986 was collected in the Allegheny River below L/D 5, which was more than 144 rkm downstream from the nearest documented sites in the Allegheny River and French Creek at Franklin, PA (Fig. 3).

1991–2000. By 2000, a total of 93 additional records in Ohio began to show range expansion of Bluebreast Darter within 18 rivers and streams (Fig. 3). Substantial increases were documented within Big Darby Creek (16 sites), Deer Creek (18 sites), Paint Creek (11 sites), and the middle section of the Scioto River from the confluence of Big Darby Creek downstream to Indian Creek, where 26 records documented their presence (OSU-MBD database 2015). Furthermore, Bluebreast Darter was documented in 7 new Ohio tributary locations, including the Middle Fork Little Beaver Creek, Jelloway Creek, Mohican River, Salt Creek, Sugar Run, Tuscarawas River, Wakatamika Creek, and the first mainstem Ohio River site located near Manchester, OH (Fig. 3). In Pennsylvania, only 9 additional records were documented, all of them in the middle to upper reaches of Tionesta Creek, and the previously documented French Creek (Fig. 3).

2001–2010. By 2010, range expansion had become more apparent with 314 Bluebreast Darter records: 182 records in Ohio and 132 in Pennsylvania (Fig. 3).

The first upper Ohio River mainstem record along Ohio's border was documented in 2001 by the Ohio River Valley Water Sanitation Commission (ORSANCO) during boat electrofishing from the Hannibal Pool. Beginning in 2007, use of a modified mini-Missouri benthic trawl (Herzog et al. 2005) by ORSANCO documented 9 sites in the Ohio River from the Pike Island and Hannibal pools (OH) and 1 record in the Scioto River downstream in Chillicothe, OH (OSU-MBD 2015). Additional eastern Ohio range expansion was represented by collections in the Little Muskingum River; Short, Island, and Wheeling creeks; and the Ohio River at the Pike Island tailwater. Elsewhere in Ohio, despite little to no increase in sampling effort, continued surveys by OEPA documented range expansion within Salt Creek, Paint Creek, Walhonding River, and Muskingum River systems (Fig. 3).

Beginning in 2005, benthic trawls were also used to obtain Bluebreast Darter records in Pennsylvania. Of the 132 records, 76 were documented from benthic trawling, and of these, 69 sites were documented in the Allegheny River and 7 records were from the Ohio River below Pittsburgh (Freedman et al. 2009b, ORSANCO 2017). More-efficient sampling methods combined with traditional techniques revealed that Bluebreast Darter in Pennsylvania was present from the upper free-flowing sections of the Allegheny River to below the Montgomery L/D on the Ohio River (Fig. 3).

2011–present. A total of 451 Bluebreast Darter records have been documented in Ohio (367) and Pennsylvania (83) since 2011. In Ohio, 81 trawling records resulting from this study have expanded the known distribution of Bluebreast Darter in the Ohio River from the Pennsylvania state line, downstream to near Indian Creek just southeast of Cincinnati (Fig. 3). We also documented this species in multiple Ohio River tributaries upstream of Marietta, OH, including Yellow, Cross, McMahon, Wegee, Captina, and Sunfish creeks and Croxton Run. Further range documentation included Big Walnut Creek, Tuscarawas River, and progression down the Scioto and Muskingum rivers to near their confluences with the Ohio River mainstem (also documented by this study and continued efforts by OEPA).

Of the 83 records of Bluebreast Darter in Pennsylvania since 2011, 45 records were obtained by trawling with either a modified mini-Missouri trawl or an electrified Missouri trawl. Employing the electrified trawl in Pennsylvania during 2013–2014, we confirmed the presence of Bluebreast Darter in 2 new locations: (1) the lower Beaver River just below the first dam upstream from the confluence with the Ohio River, and (2) expansion of Bluebreast Darter into the lower Monongahela River just below Braddock L/D (Fig. 3). We documented a total of 4 new tributary site records—2 tributaries to the lower Allegheny River (Kiskiminetas River and Bull Creek) and 2 tributaries to the upper Ohio River (Moon Run and Montour Run) (Table 3, Fig. 3).

Tippecanoe Darter

Pre–1981. Prior to 1981, there were 27 records documenting the presence of Tippecanoe Darter in Ohio. All records were within the Scioto and Muskingum River drainages (Fig. 4), 20 of which were in Big Darby Creek (Osburn 1901, OSUMBD 2015, Trautman 1981).

In Pennsylvania, 14 records for Tippecanoe Darter were documented from the upper reaches of the Allegheny River, including 12 records from French Creek and 2 records from the Allegheny River near Tidioute, PA (Cooper 1983, Pennsylvania Natural Heritage Program 2015, Raney 1938).

1981–1990. OEPA sampling effort increased dramatically in this time period, and by 1990, an additional 12 records of Tippecanoe Darter had been collected within the Scioto River drainage in Ohio. Big Darby Creek contained 10 of the 12 records (Fig. 4). Within Pennsylvania, 10 records for Tippecanoe Darter were recorded in the upper Allegheny River and French Creek (Fig. 4).

1991–2000. In Ohio, Tippecanoe Darter started to show signs of distribution changes towards the end of the decade, with a total of 50 site records. New records included Paint Creek (2 sites), and Little Darby Creek (4 sites) with expansion in

the Deer Creek system (6 sites) (Fig. 4). The middle reaches of the Scioto River from Walnut Creek downstream to Paint Creek contained 23 records, and sampling in Big Darby Creek documented the presence of Tippecanoe Darter with 15 records. French Creek was the only location within Pennsylvania, and an additional 9 records of Tippecanoe Darter had been documented by 2000 (Fig. 4).

2001–2010. A total of 100 Tippecanoe Darter records were documented in Ohio between 2001 and 2010. During this period, benthic trawling was added as a new sampling method for both Ohio and Pennsylvania. Of the 100 records for Ohio only 6 were obtained with trawling, but the trawling records documented Tippecanoe Darter in the Ohio River for the first time (Fig. 4, ORSANCO 2017). Other new locations included: Buckskin Creek, Little Beaver Creek (on the Ohio/Pennsylvania border), North Fork Paint Creek, Salt Creek, Walnut Creek, and Wheeling Creek (Table 2, Fig. 4; OEPA 2016, OSU-MBD 2015). Additional records showed an increased presence upstream in Paint Creek (8 records), and 5 records in the lower Muskingum River at the Lowell L/D tailwater (Fig. 4). Further known range expansion of Tippecanoe Darter was documented with 42 records on the Scioto River. A majority of the records from the Scioto River occurred in the reach from the Greenlawn Dam in Columbus, OH, downstream to Big Darby Creek, but 8 more records showed the movement of Tippecanoe Darter downstream to near Candy Run near Lucasville, OH, largely resulting from OEPA standard surveys (Fig. 4).

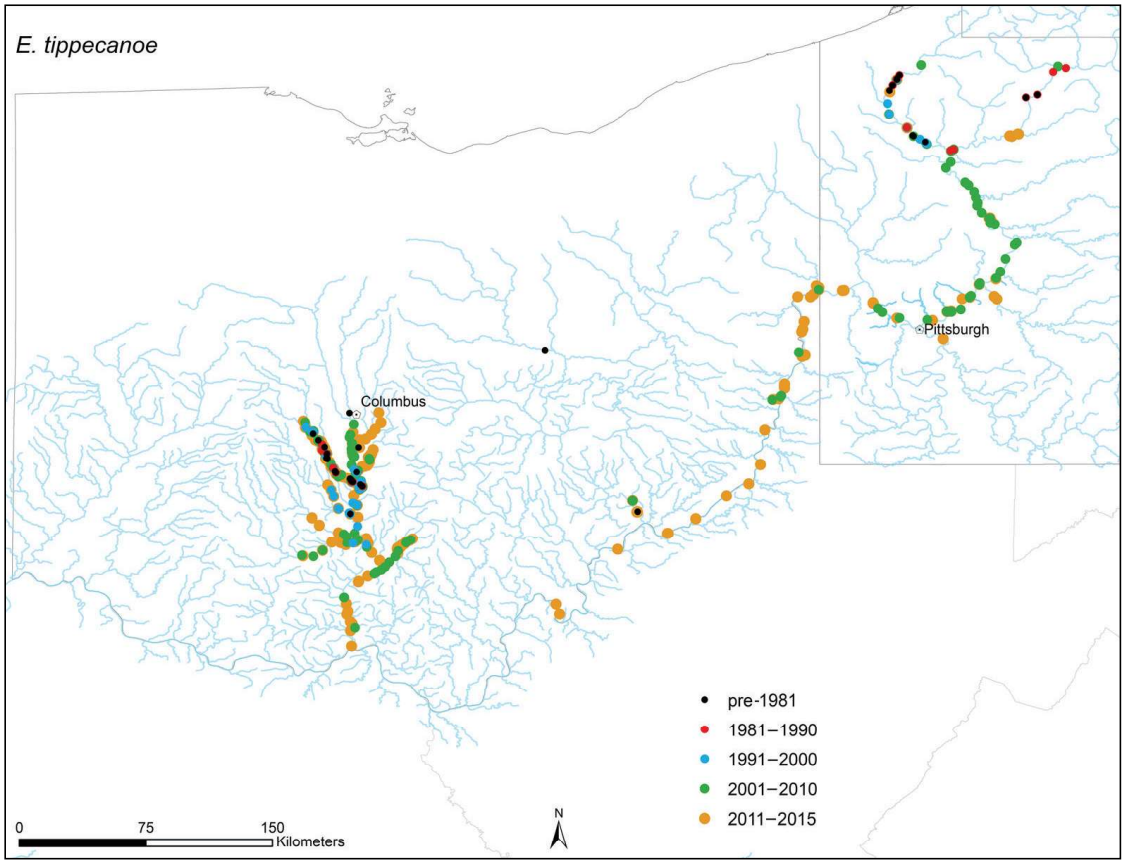

Figure 4. Distribution of Tippecanoe Darter in the upper Ohio River system in Ohio and Pennsylvania showing all historic and contemporary data from pre-1981 to 2015.

In comparison to Ohio, trawling in Pennsylvania accounted for 57% of the records of Tippecanoe Darter (40 out of 70). Trawling data combined with traditional sampling methods helped to document the Tippecanoe Darter's distribution from the free-flowing section of the Allegheny River downstream through the navigable reaches and into the Ohio River below the Dashields L/D. The Dashields record documented the Tippecanoe Darter ~224 rkm downstream from its previously recorded location near Franklin, PA (Fig. 4).

2011–present. Fish surveys in Ohio from 2011 to 2015 resulted in 189 records of Tippecanoe Darter (Fig. 4). Benthic trawling from this study resulted in 40 records, which documented Tippecanoe Darter in the Muskingum, Ohio, and Scioto rivers. Trawls also produced 32 records in the Ohio River, and extended the known range of the Tippecanoe Darter from the Pennsylvania/Ohio border downstream to the Racine L/D tailwater. Our sampling efforts also documented Tippecanoe Darter expansion into the lower portion of Cross Creek, a direct tributary to the Ohio River in eastern Ohio and in the North Fork of Paint Creek and the Scioto River to near its confluence with the Ohio River. Eight new records documented upstream movement past Osburn's 1897 site record on Big Walnut Creek to the confluence with the Rocky Fork (Fig. 4).

In Pennsylvania, there were 56 records for Tippecanoe Darter from 2011 to 2015. Benthic trawling accounted for 61% (34 out of 56) of the records and extended its known range in Pennsylvania with new site records at the tailwaters of the Montgomery L/D on the Ohio River, and up into the lower Monongahela River to the tailwaters of the Braddock L/D (Fig. 4). Additionally, benthic trawling confirmed the presence of Tippecanoe Darter in the tailwaters of the Allegheny River L/Ds 2, 3, 4, 5, and 6. Backpack electrofishing in the Kiskiminetas River and Bull Creek (Tarentum, PA.) produced new site records for Tippecanoe Darter within Pennsylvania (Fig. 4).

Spotted Darter

Pre–1981. Historic records of Spotted Darter in Ohio (total = 38) documented the species in 8 different waterbodies: Big Darby Creek, Big Walnut Creek, Deer Creek, Kokosing River, Mahoning River, Olentangy River, Walhonding River, and Yellow Creek (Mount 1959, Osburn 1901, OSU-MBD database 2015, Trautman 1981; Fig. 5). Twenty-three of these records were from Big Darby Creek.

Prior to 1981, there were 34 records for Spotted Darter in Pennsylvania from 5 streams or rivers: the upper Allegheny River, French Creek, Little Neshannock Creek, Otter Creek, and the Shenango River (Cooper 1983; PAFBC 2015; Pennsylvania Natural Heritage Program 2015; Raney 1938, Raney and Lachner 1939; Fig. 5).Twenty-five of the records were documented in French Creek.

1981–1990. In Ohio, 12 records were documented for Spotted Darter within previously identified locations. Big Darby Creek accounted for 11 of these records, 1 of which occurred upstream near the confluence of Little Darby Creek (Fig. 5). As previously mentioned, sampling effort greatly increased across Ohio in this time period. In Pennsylvania, 8 records for Spotted Darter represented 1 new site

in Sandy Creek (tributary to the upper Allegheny River), 5 in French Creek, and 2 in the Allegheny River near Tidioute, PA (Fig. 5).

1991–2000. There were 7 additional records for Spotted Darter in Ohio between 1991 and 2000. Two new sites were documented in the Scioto River just downstream of Big Darby Creek, and the others were in the Walhonding River and Big Darby Creek (Fig. 5). In Pennsylvania, French Creek contained all 6 Spotted Darter records. No new locations were documented.

2001–2010. Thirty Spotted Darter records were documented in Ohio. Several new site records were documented: near the mouth of Little Darby Creek and Paint Creek, and 4 records were from Walnut Creek just upstream of the confluence with Little Walnut Creek (Table 2, Fig. 5). Additional records documented Spotted Darter presence in the Kokosing River (9 records) and Big Darby Creek (15 records). Trawling did not produce any Spotted Darter records in Ohio.

Of the 42 Pennsylvania records, Spotted Darter was documented at 3 new sites: Woodcock Creek (tributary to French Creek), the mouth of Oil Creek (tributary to the Allegheny River), and the Ohio River just below Pittsburgh. The remainder of the records were within French Creek (11) and the Allegheny River (28). By 2007, the Spotted Darter was documented in the navigable reaches of the Allegheny River below L/D 3 (between the islands that make up Allegheny Islands State Park), and below the Dashields L/D, in the upper Ohio

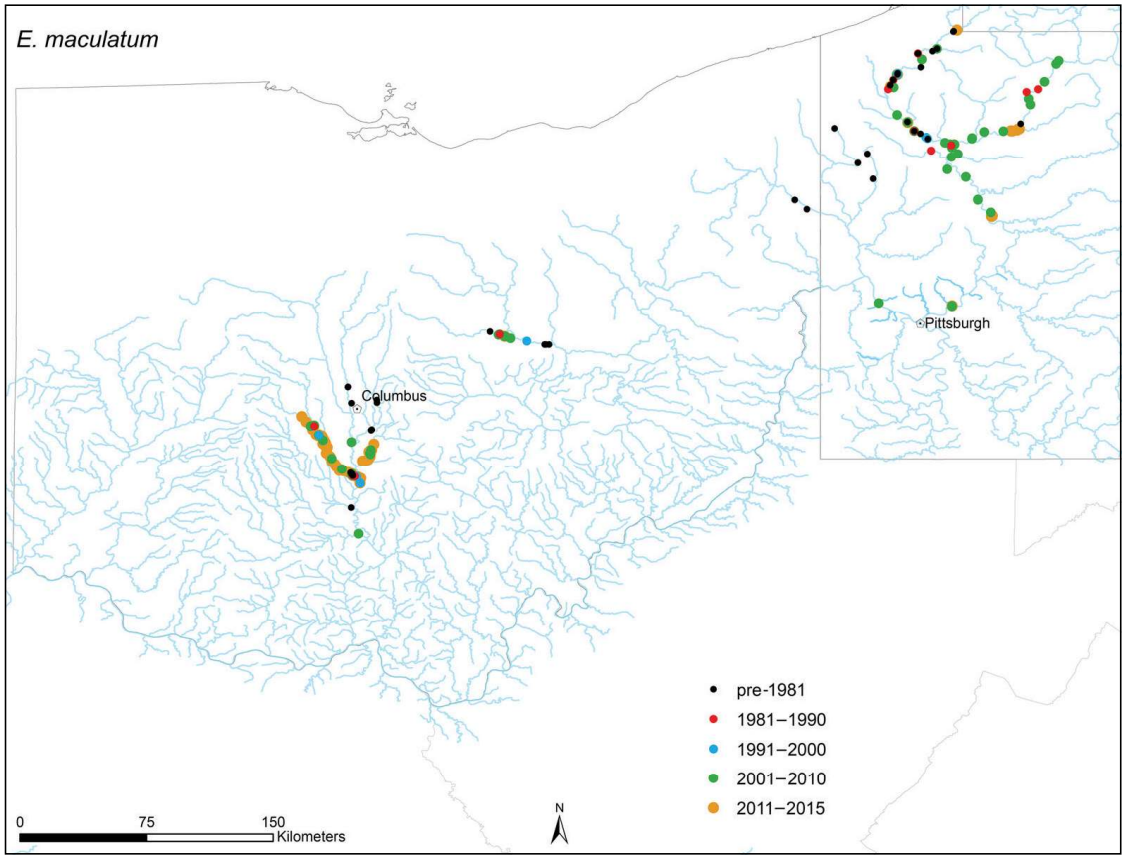

Figure 5. Distribution map for the Spotted Darter in the upper Ohio River system in Ohio and Pennsylvania showing all historic and contemporary data from pre-1981 to 2015.

River below Pittsburgh, PA (Fig. 5). Benthic trawling accounted for 19 of the records (18 in the Allegheny River and 1 in the Ohio River).

2011–present. Our seine sampling efforts helped to document 55 Spotted Darter records in Ohio. Range expansion, however, was minimal (Fig. 5). Additional sites showed slight movement up Little Darby Creek, but the majority of the records documented stable populations in the Kokosing River (3), a slight increase further upstream in Walnut Creek (12), and increased number of records in Big Darby Creek (34). The Spotted Darter has not been documented in the Ohio River within Ohio's borders and was never found during benthic trawling. To date, 21 additional records have documented the Spotted Darter within Pennsylvania. Fifteen of the records were within the Allegheny River and the 6 were in French Creek (Fig. 5). Benthic trawling documented 4 of the records within the Allegheny River.

Extirpations

Bluebreast Darter. There are 3 systems in Ohio where Bluebreast Darter appears to have been extirpated: (1) the Stillwater River (documented 1899), which is a tributary to the Great Miami River in western Ohio; (2) the North Fork of the Licking River (documented in 1899); and (3) Yellow Creek (documented in 1853), a tributary to the Mahoning River on the Ohio/Pennsylvania border (Fig. 3). In Pennsylvania, Bluebreast Darter appears to be extirpated from 2 upper tributaries of the Beaver River system in northwestern PA: (1) Neshannock Creek (documented in 1934), and (2) the Shenango River (documented in 1935) (Fig. 3).

Tippecanoe Darter. The current distribution of Tippecanoe Darter in Ohio illustrates 2 locations where they have been apparently unable to recolonize: (1) the Olentangy River (documented in 1896), which enters the Scioto River near Columbus; and (2) the Walhonding River (documented in 1962) which is in the upper Muskingum River system (Fig. 4). All historical locations for Tippecanoe Darter in Pennsylvania have extant populations.

Spotted Darter. Currently, the Spotted Darter appears to be extirpated from 3 systems in the Scioto River drainage including: (1) the Olentangy River (documented 1958, 1960, and 1963), (2) Big Walnut Creek (documented in 1897, 1959, and 1962), and (3) Deer Creek (documented in 1956) (Fig. 5). No recent surveys on the Ohio/Pennsylvania border have found the Spotted Darter in Yellow Creek (documented 1853) or the adjacent Mahoning River, which is the type locality (Kirtland 1840). The Spotted Darter has apparently not been able to reestablish populations in the Mahoning River since being presumed extirpated in the mid-1850s (Trautman 1981). In Pennsylvania, the Spotted Darter appears to be extirpated from the upper reaches of the Shenango River (documented 1905 and 1934) and Neshannock Creek (documented 1935) (Fig. 5).

Electrified benthic trawling

Electrified benthic trawling surveys below 11 L/D installations yielded varying results. We documented Bluebreast Darter below 8 installations: Allegheny River L/D 2, 3, 4, 5, and 6; Beaver River Dam 1; Monongahela L/D 2; and below the Emsworth back channel L/D on the Ohio River. We also found Tippecanoe Darter below

8 installations: Allegheny River L/D 2, 3, 4, 5, and 6; Monongahela River L/D 2; Ohio River Emsworth back channel; and the Montgomery L/D. We documented the Spotted Darter only below 1 installation: Allegheny River L/D 3.

Discussion

A large proportion of the contemporary survey records illustrate increases in the known ranges of darters of the subgenus *Nothonotus* into the non-wadeable riverine environments of the Allegheny, Ohio, and Monongahela rivers. Regional improvements in water quality that have resulted in improved fish assemblages (Yoder et al. 2005) may have influenced the distribution changes in these focal species. In addition, we suggest that recently developed and improved sampling techniques including the modified Missouri trawl (Herzog et al. 2005, 2009) and the PSU electrified benthic trawl (Freedman et al. 2009a) are responsible for elucidating these new records in the non-wadeable portions of the Allegheny, Monongahela, Muskingum, and upper Ohio rivers in depths >2.0 m where traditional methods can be less effective. In Ohio, the history of increased sampling that coincided with the inception of the OEPA surface-waters sampling program thoroughly documented an increase in distribution of *Nothonotus* darters and many other fish species as water quality improved (OEPA 2016, Yoder et al. 2005). Until recently, non-wadeable stream sampling in the basin was mainly limited to lock-chamber surveys (Thomas et al. 2005), boat electrofishing (Emery et al. 2003, Koryak et al. 2008), hoop/gill netting, beach seining, and various-sized mesh for trawling (Neebling and Quist 2011). Each method has valid applications, but they can also be biased towards certain species, body sizes (Neebling and Quist 2011), and aquatic habitats (e.g., pelagic fish vs. benthic fish). For example, Koryak et al. (2008) surveyed a navigable section of the Allegheny River with both night electrofishing and benthic trawling. Electrofishing resulted in 42 species (834 individuals), while benthic trawling documented 27 species (2903 individuals). Benthic trawling was more effective at collecting species in the family Percidae; electrofishing detected 4 species and trawling documented 12 (Koryak et al. 2008). The use of multiple sample gears to survey for large-river darters was also supported by Neebling and Quist (2011), who compared boat electrofishing, trawling, and shoreline bag-seining in non-wadeable rivers. Those authors surveyed 21 reaches from 3 to 5 km in length and found that 8 species were only detected by trawling and 4 of those species were darters. However, it is important to point out that, in Ohio, the OEPA has shown that boat electrofishing can be effective at detecting the presence of darter species by using an appropriate level of effort and detail within an electrofishing site (Yoder et al. 2005). It should be noted, though, that once depths are consistently > 2 m, effectiveness of this method is diminished. For all data from Ohio and Pennsylvania collected since 2005 and summarized in this study, trawling records accounted for 32% of all records of the 3 focal darter species. The number of trawling records since 2005 also varied by state—20% of the records in Ohio and 57% of the records in Pennsylvania were from trawling. Our surveys and analysis of historical records support previous assessments that concluded it is necessary to utilize benthic and/

or electrified benthic trawling to effectively survey non-wadeable riverine environments for benthic fishes (Freedman et al. 2009a, 2009b; Herzog et al. 2005, 2009; Koryak et al. 2008, 2011). We propose that benthic trawls are an effective

Table 2. Streams in Ohio that were sampled in the same location and the year *Nothonotus* appeared. Full references presented in Table 4. OEPA = Data queried from Ohio State University Museum of Biological Diversity, Fish Division database, analyzed by B.J. Zimmerman.

		Abundance			
Stream/site	Year	*E. camurum*	*E. maculatum*	*E. tippecanoe*	Reference
Middle Fork Salt Creek					
Site 1	1988	-	-	-	OEPA
	1997	-	-	-	OEPA
	2005	1	-	-	OEPA
Salt Creek					
Site 2	1992	-	-	-	OEPA
	2005	10	-	4	OEPA
Site 3	1984	-	-	-	OEPA
	1992	-	-	-	OEPA
	2005	15	-	27	OEPA
Site 4	1992	-	-	-	OEPA
	2005	-	-	3	OEPA
Scioto River					
Site 5	1997	-	-	-	OEPA
	2011	2	-	2	OEPA
Site 12	1979	-	-	-	OEPA
	1988	-	-	-	OEPA
	1992	-	-	-	OEPA
	2002	-	1	-	OEPA
Paint Creek					
Site 6	1992	-	-	-	OEPA
	1997	3	-	-	OEPA
	2006	6	-	2	OEPA
Site 7	1997	-	-	-	OEPA
	2006	6	-	1	OEPA
Site 8	1997	-	-	-	OEPA
	2006	3	-	1	OEPA
North Fork Paint Creek					
Site 9	1985	-	-	-	OEPA
	1997	-	-	-	OEPA
	2006	1	-	-	OEPA
Walnut Creek					
Site 10	1996	-	-	-	OEPA
	2010	72	3	14	OEPA
Site 11	1982	-	-	-	OEPA
	2005	21	2	1	OEPA
Killbuck Creek					
Site 13	1983	-	-	-	OEPA
	2009	1	-	-	OEPA

sampling method for small-bodied benthic species (e.g., darters) in conditions when depths are greater than >2.0 m and/or there is elevated turbidity.

Even though more-efficient sampling techniques may have elucidated the changes documented in the impounded reaches of the non-wadeable rivers, we also show evidence of range expansion of Bluebreast Darter, Tippecanoe Darter, and Spotted Darter into the unimpounded rivers and smaller tributaries. Surveys in multiple streams, with historic and contemporary samples using the same methodology, have recently documented new site records for these 3 darter species. In Ohio, there were at least 13 OEPA survey sites in 7 streams that have newly documented *Nothonotus* records (Table 2, Fig. 6). In Pennsylvania, 6 new records in 6 streams documented the recent expansion of *Nothonotus* species (Table 3, Fig. 7).

Since 2000, outside of the refugia areas of Big Darby Creek, OH, and French Creek, PA, the population sizes of Spotted Darter have been consistently lower than the other 2 focal species (see Supplemental File 2 available online at http://www.eaglehill.us/NENAonline/suppl-files/n24-2-N1537-Honick-s2, and, for BioOne subscribers, at http://dx.doi.org/10.1656/N1537.s2), which justifies continued monitoring. Previously, Lorson (2010) performed benthic trawling surveys of the Allegheny River from its headwaters to Pittsburgh, PA. Within the navigable section of the river, he only documented 1 Spotted Darter below 1 L/D

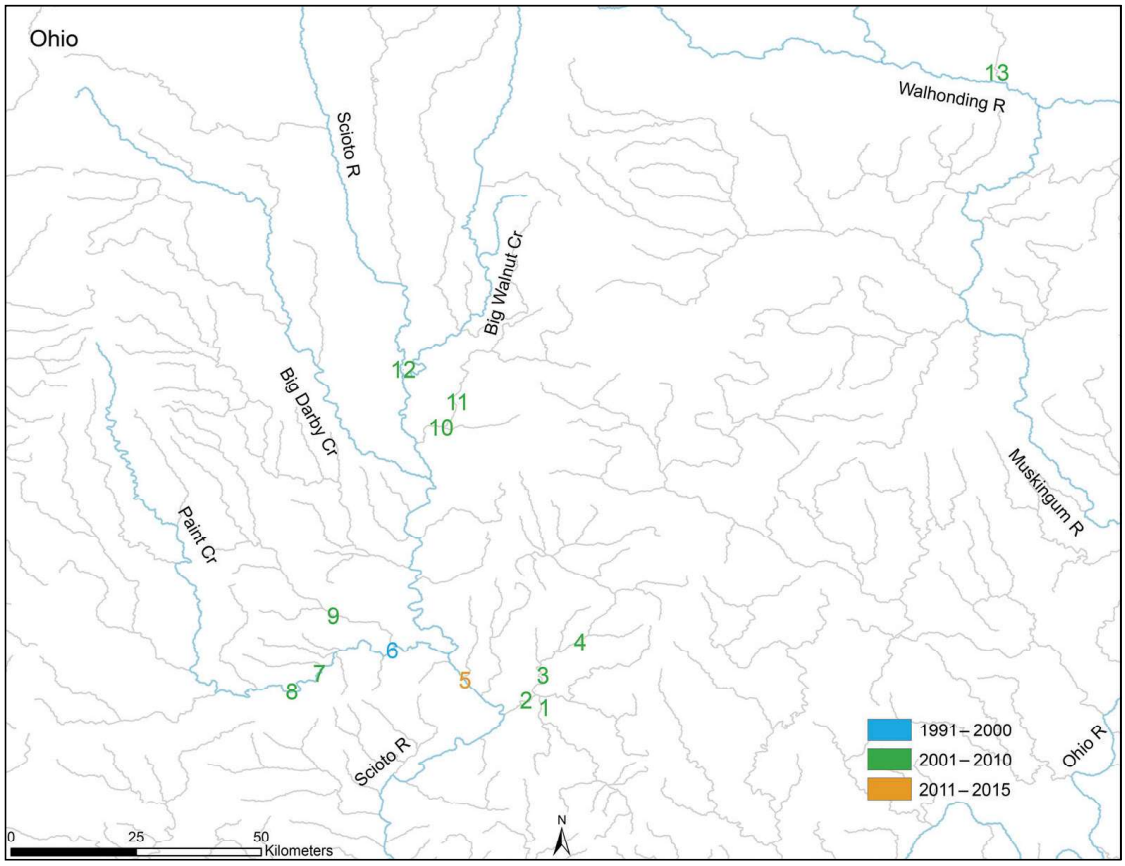

Figure 6. Sites in Ohio where *Nothonotus* species have only recently been documented after years of consistent sampling. Site numbers correspond to Table 2.

installation—the same pool where we documented them (pool 2 below L/D 3). Even within an 81-km reach of the free-flowing section of the upper Allegheny River, Argent and Kimmel (2014) only documented 4 Spotted Darters. In 2015, the Pennsylvania Fish and Boat Commission documented 11 individuals below L/D 3 on the Allegheny River, which suggests that there is a stable population at this location (Fig. 5). However, to our knowledge, documentation of the extension of Spotted Darters downstream into the Montgomery pool of the Ohio River (Freedman et al. 2009b) has not been duplicated, and additional surveys and cautious interpretation of the range extension of stable populations of the Spotted Darter within Pennsylvania are warranted.

Factors effecting Spotted Darter expansion

Habitat considerations and connectivity. The expansion of the known ranges of Bluebreast Darter and Tippecanoe Darter has been robust, but the Spotted Darter has been less successful at utilizing the navigable portions of the Allegheny, Monongahela, Muskingum, and Ohio rivers (Figs. 3, 4, and 5). Reasons for this lack

Table 3. Streams in PA that were sampled in the same location and the year *Nothonotus* appeared. ASH = data collected by A.S. Honick and B.A Porter. BAP = data collected by B.A Porter.

		Abundance			
Stream/site	Year	*E. camurum*	*E. maculatum*	*E. tippecanoe*	Reference
Little Sewickley Creek					
Site 14	2003	-	-	-	Koryak (2003)
	2006	-	-	-	MARIS (2016)
	2012	13	-	-	This study (BAP)
	2013	13	-	-	This study (ASH)
Montour Run					
Site 15	1982	-	-	-	USACE (1997)
	1991	-	-	-	USACE (1997)
	1996	-	-	-	USACE (1997)
	2003	-	-	-	Koryak (2003)
	2014	2	-	-	This study (ASH)
Moon Run					
Site 16	2003	-	-	-	Koryak (2003)
	2014	1	-	-	This study (ASH)
Pine Creek					
Site 17	2002	-	-	-	Hoskin et al. (2003)
	2005	1	-	1	Howell (2007)
Bull Creek					
Site 18	2006	-	-	-	MARIS (2016)
	2014	1	-	15	This study (ASH)
Kiskiminetas River					
Site 19	2009	-	-	-	This study (BAP)
	2010	-	-	-	This study (BAP)
	2011	-	-	-	This study (BAP)
	2013	-	-	4	This study (BAP)
	2013	10	-	25	This study (ASH)

of expansion may be directly related to the availability of optimal habitat. Historically, the Spotted Darter was considered an associate of the Bluebreast Darter and Tippecanoe Darter (Kuehne and Barbour 1983, Raney 1939), but was reported to occupy "deeper parts of riffles" that were often overlooked (Raney and Lachner 1939). In the Ohio River, along the border of Ohio, benthic trawling commonly documented Bluebreast Darters and Tippecanoe Darters in areas of moderate flow, including gravel outwashes near tributaries and the gravel/cobble habitat found up- and downstream of islands. However, benthic and electrified benthic trawling does not support the hypothesis that Spotted Darter is preferentially utilizing similar habitats in the navigable portions of the upper Ohio River watershed. Our electrified benthic trawling surveys below 11 L/D installations documented Bluebreast Darters and Tippecanoe Darters at 8 installations, and have revealed that these 2 species can occupy great depths; ranging from 1.4 m to 4.5 m and 1.4 m to 5.9 m, respectively. In contrast, the Spotted Darter was only found below 1 installation, within a wadeable riffle ~1.0 m deep. Raney and Lachner (1939) described Spotted Darters as occurring in deep, fast riffles and spawning at depths no greater than 0.6 m. Kessler and Thorp (1993) analyzed microhabitat use between the Spotted Darter and *Etheostoma bellum* Zorach (Orangefin Darter) in a tributary of the upper Green River, KY, and documented that Spotted Darters utilized deeper habitats (mean depth = 0.2 m) and were observed mostly under large rocks. Osier

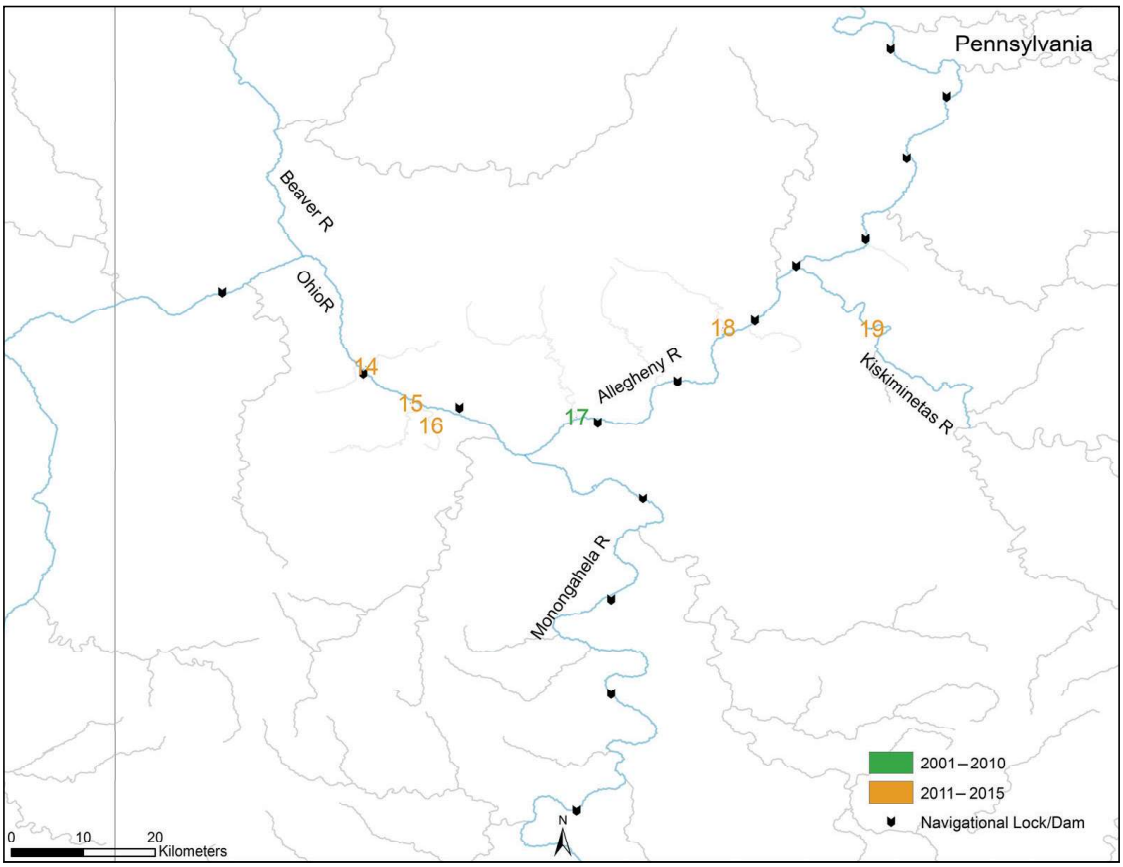

Figure 7. Sites in Pennsylvania where *Nothonotus* species have only recently been documented after years of consistent sampling. Site numbers correspond to Table 3.

and Welsh (2007) studied Spotted Darter habitat in the Elk River, WV, and found that they occurred with large rocks in the glide habitat between the riffles at depths ranging from 0.31 m to 0.49 m. These data suggest that the Spotted Darter is more of a habitat specialist and prefers deeper habitat. However, our data suggest that Spotted Darters may be restricted to shallower habitat within non-wadeable riverine environments, while Bluebreast Darter and Tippecanoe Darter may be benthic generalists that have the ability to utilize the more diverse and deeper habitat found within L/D tailwaters. The restricted expansion of Spotted Darters into the navigable portions of the Allegheny and Ohio rivers could be linked to current and historic dredging of the rivers for commercial aggregates and navigation requirements. Since 2004, the upper Ohio and lower Allegheny rivers have had over 13.6 million metric tons of substrate removed for commercial aggregates, and since the 1800s, it has been estimated that ~0.5 billion metric tons of substrate have been removed (R. Ventorini, Pennsylvania Fish and Boat Commission, Pittsburgh, PA, unpubl. data). Freedman et al. (2013) studied the navigable portion of the Allegheny River and found that dredging significantly changed the benthic fish community, reduced habitat heterogeneity, and shifted the fish assemblage towards habitat generalists. In summary, the Spotted Darter has a limited presence in the navigable portion of the Allegheny River and the upper Ohio River. We were only able to document this species below 1 L/D installation. Freedman et al. (2009b) documented 5 Spotted Darters below Dashields L/D, but that record hasn't been duplicated. Regardless of continued water quality improvements, the historic data compiled by Raney and Lachner (1939), Osburn (1901), and a report from the US Fish and Wildlife Service (2011) suggests that the Spotted Darter was likely never common throughout its range, which may be due to specific habitat requirements (Kessler and Thorp 1993, Osier and Welsh 2007, Raney and Lachner 1939). No surveys have documented Spotted Darters in the Ohio River downstream of the Pennsylvania state line; thus, we contend that the non-wadeable, impounded river environment may not have enough preferred habitat to support robust expansion of the Spotted Darter.

In addition to more-specific habitat requirements, Spotted Darter range extension may be negatively impacted by the Allegheny River's restricted connectivity between the upper free-flowing reaches and the now-lentic habitat of the navigable portion below L/D 9. Recently, Argent and Kimmel (2010) documented that fish community composition immediately above and below L/D installations were markedly different in both the Allegheny and Monongahela rivers. On the Monongahela River, the small-bodied fish assemblages consisted of 12 and 13 species above and below the installations with 2 darter species above and 5 species below. In contrast, in the Allegheny River, the small-bodied fish assemblages consisted of mostly darters, but only 2 darter species were documented above the installations, and 10 darter species were utilizing the tailrace habitat below the L/D installations. Regardless of the fact that the Monongahela River experiences higher lockage-frequency from more boat traffic, neither river indicated a correlation between small-bodied fish diversity and lockage frequency (Argent and Kimmel 2010). Therefore, Argent and Kimmel (2010) indicated that the physical restriction to fish movement posed by L/Ds may result in

isolated populations of darters within navigation pools. The navigational L/D system on the lower Allegheny River may be impeding the movement of Spotted Darters, which suggests that within the upper Ohio River watershed, the species needs to be closely monitored for proper conservation management.

Differences in reproductive strategies, spawning habitat requirements, and larval duration/transport. Field observations directly documenting fecundity and clutch sizes in Bluebreast, Tippecanoe, and Spotted Darters are sparse and are mostly from aquarium studies. Bluebreast Darter and Tippecanoe Darter have been documented as belonging to the egg-burying guild (Kelly et al. 2012, Stiles 1972). Field observations (Stiles 1972, Tiemann 2008) and aquarium experiments in the laboratory (Mount 1959, Page and Simon 1988, Warren et al. 1986) indicated that females of both species bury themselves into the gravel substrate while the males mount them and fertilize the eggs. Tiemann (2008) observed spawning behavior of Bluebreast Darter in the Vermilion River, IL, and documented that males stop defending their territories soon after spawning. Warren et al. (1986) collected Tippecanoe Darters from the Green River, KY, and in aquarium studies, showed that males established territories but quickly abandoned nests after spawning, just like Bluebreast Darters.

In contrast, under field and laboratory conditions, the Spotted Darter uses a different reproductive strategy and has different spawning habitat requirements. Raney and Lachner (1939), Winn (1958), and Stiles (1972) documented the Spotted Darter as belonging to the egg-clumper guild, in which females attach their eggs to the underside of large, flat rocks. They also observed that, in contrast to Bluebreast and Tippecanoe Darters, male Spotted Darters continue to defend their territory after spawning. Additionally, Raney and Lachner (1939) documented that regardless of the amount of suitable spawning habitat, Spotted Darter nests were spaced ≥1.2 m apart. Out of the 14 species of darters that Winn (1958) studied, the Spotted Darter was among the species laying the fewest eggs, and males provided substantial parental care. More recently, Ruble et al. (2016) studied reproductive behaviors of *Etheostoma wapiti* Etnier & J. D. Williams (Boulder Darter), *E. vulneratum* (Cope) (Wounded Darter), and Spotted Darter under laboratory conditions and found that Spotted Darter and Boulder Darter averaged fewer eggs per female and had lower egg-to-juvenile survival rates than Wounded Darter. Therefore, Spotted Darter exhibits characteristics of a K-selected species reproductive strategy, while Bluebreast and Tippecanoe Darters exhibit reproductive strategies more similar to r-selected species. This reproductive strategy and the lack of suitable spawning habitat featuring large unembedded cover stones or large boulders associated with swift currents are likely hindering population expansion by Spotted Darters. In contrast, the impounded portions of the Allegheny and Ohio rivers contain abundant gravel in areas with swift current to prevent siltation where Bluebreast and Tippecanoe Darters can bury their eggs.

Another potential reason for the differences in distribution changes among these 3 species may be linked to temporal variation in pelagic larval duration (PLD) and larval transport. Douglas et al. (2013) studied PLD in 23 darter species and Turner (2001) examined larval transport of 8 darters. Both reports found that PLD and

larval-transport times were highly variable across species of darters. Douglas et al. (2013) documented darter PLDs ranging from 0 to 60 days, with Spotted Darter exhibiting an average PLD of only 18 days. Of the 23 species studied, 12 were listed as imperiled (Douglas et al. 1013) and had PLD averages varying from 9 to 15 days. Short PLDs suggest that the species may have evolved that way to reduce downstream movement in attempts to stay within restricted habitats (Douglas et al. 2013), but reduced dispersal may essentially lead to isolated populations with small ranges (Sorte 2013), which is the pattern observed in Spotted Darter. The shorter PLDs of Spotted Darter relative to Bluebreast and Tippecanoe Darters may also have allowed the latter 2 species to re-establish in the larger rivers after water quality improvements in a shorter amount of time than Spotted Darter.

Summary/Conclusions

Populations of darters classified in the subgenus *Nothonotus* in the upper Ohio River system have historically been described as having disjunct distributions (Cooper 1983, Kuehne and Barbour 1983, Page 1983, Simon and Wallus 2006, Trautman 1981). Our surveys and analysis of ~1700 historic and contemporary survey records revealed major distribution changes for these darters in the upper Ohio River watershed. In Pennsylvania, all 3 species were listed as threatened in 1999 (Pennsylvania Bulletin 1999), and in Ohio, Bluebreast and Tippecanoe Darters were listed as threatened in 1990 and Spotted Darter was listed as endangered in 1974 (15 Ohio Rev. Code § 1531.25 - 2015). In Pennsylvania, assessment of recent survey data led the Pennsylvania Fish and Boat Commission to delist all 3 species in 2014 (Pennsylvania Bulletin 2014). Extensive surveys in Ohio from 2006 to 2012 led to the delisting of Bluebreast Darter in 2012 (ODNR 2012, OSU-MBD 2015) while the Tippecanoe and Spotted Darters maintained their threatened and endangered status, respectively. Our analysis showed that Spotted Darter was less common, had a smaller geographic range, and fewer individuals per sample site compared to Bluebreast and Tippecanoe Darters, which may be related to life-history characteristics, a lack of optimal habitat, and impaired connectivity throughout the navigable portions of the upper Ohio River watershed. Therefore, the stable Spotted Darter source populations should be closely monitored.

Based on previous observations, it is conceivable that the Spotted Darter is not expanding its distribution as effectively because (1) Bluebreast and Tippecanoe Darters employ an r-selected reproductive strategy, while the Spotted Darter displays a K-selected reproductive strategy (Ruble et al. 2016); (2) the Spotted Darter may require larger areas of suitable spawning habitat as a result of maintaining territoriality and nest defense, potentially producing fewer offspring per unit of available habitat; (3) Spotted Darter has been documented as having a short PLD that may be limiting their distance or rate of dispersal; and (4) the navigational L/D system may be restricting movement of Spotted Darters between the free-flowing sections of the upper Allegheny River and the navigable portions of the upper Ohio River watershed.

We were able to collect enough samples of Bluebreast Darter to investigate genetic structure of these populations. The results of our ongoing analysis will provide insight into metapopulation structure and dynamics and reveal if impaired river-connectivity has resulted in many genetically isolated populations within the navigable sections of the rivers. These data will facilitate development of management strategies that emphasize conservation efforts toward maintaining genetically diverse source populations compared to smaller, genetically depauperate, and ephemeral sink populations. In addition, efforts are underway in Ohio to reintroduce all 3 darter species back into historic locations where barriers have prohibited natural recolonization (B. Zimmerman, The Ohio State University, Columbus, OH, unpubl. data).

Acknowledgments

We thank Doug Locy (Aquatic Systems, Inc.) and Mike Koryak for assistance with the electrified benthic surveys, and all the individuals who helped with tributary electrofishing. We are grateful to Elizabeth Dakin for help with sampling and providing comments on the manuscript. Thanks to all the agencies that supplied historic survey data including: Pennsylvania Natural Heritage Program, the Pennsylvania Fish and Boat Commission (Doug Fischer), Pennsylvania Department of Environmental Protection (Rick Spear), the Ohio Environmental Protection Agency, and the Ohio River Valley Water Sanitation Commission. We appreciate the US Geological Survey for providing funds to the Pennsylvania Water Resources Research Center that supported our electrified benthic trawling surveys. Finally, we are grateful to the 2 anonymous reviewers for helping to improve the manuscript.

Literature Cited

Argent, D.G., and W.G. Kimmel. 2010. Influence of navigational lock and dam structures on adjacent fish communities in a major river system. River Research and Applications 27:1325–1333.

Argent, D.G., and W.G. Kimmel. 2014. Patterns of ichthyofaunal distribution and connectivity in navigable and free-flowing reaches of a major river system: The Allegheny River in Pennsylvania. River Research and Applications 30:631–638.

Boschung, H.T., R.L. Mayden, and J.R. Tomelleri. 2004. Fishes of Alabama. Smithsonian Books, Washington, DC. 736 pp.

Brown, T.C., and P. Froemke. 2012. Nationwide assessment of nonpoint-source threats to water quality. Bioscience 62:136–146.

Cooper, E.L. 1983. Fishes of Pennsylvania and the Northeastern United States. The Pennsylvania University Press, University Park, PA. 252 pp.

Cope, E.D. 1870. On Some Etheostomine Perch from Tennessee and North Carolina. Proceedings of the American Philosophical Society 11:261–270.

Douglas, M., B.P. Keck, C. Ruble, M. Petty, J.R. Shute, P. Rakes, and C.D. Hulsey. 2013. Pelagic larval duration predicts extinction risk in a freshwater fish clade. Biology Letters 9(6):20130672.

Emery, E.B., T.P. Simon, F.H. McCormick, P.L. Angermeier, J.E. Deshon, C.O. Yoder, E.E. Sanders, W.D. Pearson, G.D. Hickman, R.J. Reash, and J.A. Thomas. 2003. Development of a multimetric index for assessing the biological condition of the Ohio River. Transactions of the American Fisheries Society 132:791–808.

Etnier, D.A., and W.C. Starnes. 1993. The Fishes of Tennessee. University of Tennessee Press, Knoxville, TN. 689 pp.

Federal Water Pollution Control Act (Clean Water Act). 1972. 33 U.S.C 1251–1376; Chapter 758; P.L. 845, 30 June 1948; amended: 85 stat. 379; P.L. 92-240, 1 March 1972.

Freedman, J.A., T.D. Stecko, B.D. Lorson, and J.R. Stauffer. 2009a. Development and efficacy of an electrified-benthic trawl for sampling large-river fish assemblages. North American Journal of Fisheries Management 29:1001–1005.

Freedman, J.A., T.D. Stecko, R.W. Criswell, and J.R. Stauffer. 2009b. Extensions of the known ranges of *Percina shumardi* Girard and three species of *Etheostoma* (subgenus *Nothonotus*) in Pennsylvania. Journal of the Pennsylvania Academy of Science 83:42–44.

Freedman, J.A., R.F. Carline, and J.R. Stauffer. 2013. Gravel dredging alters diversity and structure of riverine-fish assemblages. Freshwater Biology 58:261–274.

Herzog, D.P., V.A. Barko, J.S. Scheibe, R.A. Hrabik, and D.E. Ostendorf. 2005. Efficacy of a benthic trawl for sampling small-bodied fishes in large river systems. North American Journal of Fisheries Management 25:594–603.

Herzog, D.P., Ostendorf, D.E., Hrabik, R.A. and V.A. Barko. 2009. The mini-Missouri trawl: A useful methodology for sampling small-bodied fishes in small and large river systems. Journal of Freshwater Ecology 24:103–108.

Hoskin, R.H., M. Koryak, and L.J. Stafford. 2003. Fishes of Small Tributaries to the Allegheny and Monongahela Rivers in Urban/Suburban Allegheny County, Pennsylvania. Journal of the Pennsylvania Academy of Science 77:51–58.

Howell, L.A. 2007. Genetic population structure and breeding parameters of three Pennsylvania state threatened darter species: *Etheostoma camurum*, *E. maculatum*, and *E. tippecanoe*. M.Sc. Thesis. Duquesne University, Pittsburgh, PA. 187 pp.

Jelks, H.L., S.J. Walsh, N.M. Burkhead, S. Contreras-Balderas, E. Díaz-Pardo, D.A. Hendrickson, J. Lyons, N.E. Mandrak, F. McCormick, J.S. Nelson,S.P. Platania, B.A. Porter, C.B. Renaud, J.J. Schmitter-Soto, E.B. Taylor, and M.L. Warren Jr. 2008. Conservation status of imperiled North American freshwater and diadromous fishes. Fisheries 33:372–407.

Jordan, D.S,. and B.W. Evermann. 1890. Description of a new species of fish from Tippecanoe River, Indiana. Proceedings of the United States National Museum 13:3–4.

Kelly, N., T. Near, and S. Alonzo. 2012. Diversification of egg-deposition behaviours and the evolution of male parental care in darters (Teleostei: Percidae: Etheostomatinae). Journal of Evolutionary Biology 25:836–846.

Kessler, R.K., and J.H. Thorp. 1993. Microhabitat segregation of the threatened Spotted Darter (*Etheostoma maculatum*) and closely related Orangefin Darter (*E. bellum*). Canadian Journal of Fisheries and Aquatic Sciences 50(5):1084–1091.

Kirtland, J. 1840. Descriptions of four new species of fishes. Boston Journal of Natural History 3:273–277.

Koryak, M. 2003. Fishes of Small Tributaries to the Ohio River in Allegheny County, Pennsylvania. The STUDIO for Creative Inquiry, Carnegie Mellon University, Pittsburgh, PA

Koryak, M., P. Bonislawsky, D. Locy, and B.A. Porter. 2008. Use of benthic trawling to supplement electrofishing in characterizing the fish community of the Allegheny River navigation channel in Pennsylvania, USA. Journal of Freshwater Ecology 23:491–494.

Koryak, M., P.S. Bonislawsky, D.D. Locy, and B.A. Porter. 2009. Typical channel-fish assemblage of the recovering lower Allegheny River navigation system, Pennsylvania, USA. Journal of Freshwater Ecology 24:509–517.

Koryak, M., P.S. Bonislawsky, D.D. Locy, and B.A. Porter. 2011. Gilt Darter (*Percina evides*: Percidae: Etheostomatinae) range expansion, microhabitat selection, and phylogenetics within the Allegheny River navigation system, Pennyslvania, USA. Journal of the Pennsylvania Academy of Science 85:104–108.

Kuehne, R.A., and R.W. Barbour. 1983. The American Darters. University Press of Kentucky, Lexington, KY. 177 pp.

Lorson, B.D. 2010. Distribution and the putative origin of fishes in the Allegheny River, Pennsylvania. M.Sc. Thesis. The Pennsylvania State University, State College, PA. 184 pp.

Mount, D.I. 1959. Spawning behavior of the Bluebreast Darter, *Etheostoma camurum* (Cope). Copeia 3:240–243.

Multistate Aquatic Resources Information System (MARIS). 2016. Data sourced from PA Fish and Boat Commission. Available online at http://www.marisdata.org. Accessed 20 September 2016.

Neebling, T.E., and M.C. Quist. 2011. Comparison of boat electrofishing, trawling, and seining for sampling fish assemblages in Iowa's non-wadeable rivers. North American Journal of Fisheries Management 31:390–402.

Ohio Department of Natural Resources (ODNR). 2012. Ohio's recovery of Bluebreast Darters made possible with improved water quality. Available online at http://wildlife.ohiodnr.gov/stay-informed/news-announcements/post/ohio-s-recovery-of-bluebreast-darters-made-possible-with-improved-water-quality. Accessed 18 May 2016.

ODNR. 2015. Wildlife that are considered to be endangered, threatened, species of concern, special interest, extirpated, or extinct in Ohio. Publication 5356 (R0815), updated October 2015. Columbus, OH.

Ohio Environmental Protection Agency (OEPA). 2016. Ohio 2106 Integrated water-quality monitoring and assessment report, final report. Available online at http://wwwapp.epa.state.oh.us/dsw/2016IR.pdf. Accessed 15 March 2017.

Ohio Revised Code. 2015. Title 15: Conservation of Natural Resources. Chapter 1531.25 - Protection of species threatened with statewide extinction, pursuant to Section 4 of the Endangered Species Act of 1973.

Ohio River Valley Water Sanitation Commission (ORSANCO). 2017. Online fish population database. Available online at http://www.orsanco.org/data/fish-population/. Accessed 15 March 2017.

Ohio State University Museum of Biological Diversity (OSU-MBD). 2015. Fish Division database. Available online at http://osuc.biosci.ohio-state.edu/Fishes/. Accessed in 2015.

Osburn, R.C. 1901. The fishes of Ohio. Ohio State Academy of Sciences Special Paper 4:1–105.

Osier, E.A., and S.A. Welsh. 2007. Habitat use of *Etheostoma maculatum* (Spotted Darter) in Elk River, West Virginia. Northeastern Naturalist 14:447–460.

Page, L.M. 1983. Handbook of Darters. TFH Publications, Neptune City, NJ. 271 pp.

Page, L., and T. Simon. 1988. Observations on the reproductive behavior and eggs of four species of darters, with comments on *Etheostoma tippecanoe* and *E. camurum*. Transactions of the Illinois State Academy of Science 81:205–210.

Pennsylvania Bulletin. 1999. Chapter 75.2, Threatened fish species. 29:4869–4870.

Pennsylvania Bulletin. 2014. Chapter 75.2, Threatened fish species. 44:7876–7878.

Pennsylvania Fish and Boat Commission (PAFBC). 2015. Fish inventory database. No public access; requests for data processed through Douglas Fischer, Nongame Fisheries Biologist Natural Diversity Section, Division of Environmental Services, Bellefonte, PA.

Pennsylvania Natural Heritage Program. 2015. Pennsylvania natural diversity inventory. No public access; data granted upon request from http://www.naturalheritage.state.pa.us. Accessed in 2015.

Raney, E.C. 1938. The distribution of the fishes of the Ohio drainage basin of western Pennsylvania. Ph.D. Dissertation. Cornell University, Ithaca, NY. 102 pp.

Raney, E.C., and E.A. Lachner. 1939. Observations on the Life History of the Spotted Darter, *Poecilichthys maculatus* (Kirtland). Copeia 1939:157–165.

Ruble, C.L., P.L. Rakes, J.R. Shute, and S.A. Welsh. 2016. Captive propagation, reproductive biology, and early life history of *Etheostoma wapiti* (Boulder Darter), *E. vulneratum* (Wounded Darter), and *E. maculatum* (Spotted Darter). Southeastern Naturalist 15:115–126.

Schwartz, F.J. 1965. Densities and ecology of the darters of the upper Allegheny River watershed. Pymatuning Laboratory of Ecology, University of Pittsburgh, Pittsburgh, PA. Special Publication 3:95–103.

Simon, T.P., and R. Wallus. 2006. Reproductive Biology and Early Life History of Fishes in the Ohio River Drainage: Percidae-Perch, Pikeperch, and Darters. CRC Press, Boca Raton, FL. 619 pp.

Sorte, C.J.B. 2013. Predicting persistence in a changing climate: Flow direction and limitations to redistribution. Oikos 122:161–170.

Stauffer, J.R., J.M. Boltz, and L.R. White. 1995. The Fishes of West Virginia. Proceedings of the Academy of Natural Sciences of Philadelphia 146:1–389.

Stauffer, J.R., Jr., R.W. Criswell, and D. P. Fischer. 2016. The Fishes of Pennsylvania. Cichlid Press, El Paso, TX. 556 pp.

Stiles, R.A. 1972. The comparative ecology of three species of *Nothonotus* (Percidae-*Etheostoma*) in Tennessee's Little River. Ph.D. dissertation. University of Tennessee-Knoxville, Knoxville, TN. 96 pp.

Thomas, J.A., E.B. Emery, and F.H. McCormick. 2005. Detection of temporal trends in Ohio River fish assemblages based on lockchamber surveys (1957–2001). American Fisheries Society Symposium 2005:431–449.

Tiemann, J.S. 2008. Distribution and life history characteristics of the state-endangered Bluebreast Darter, *Etheostoma camurum* (Cope), in Illinois. Transactions of the Illinois State Academy of Science 101:235–246.

Trautman, M.B. 1981. The Fishes of Ohio (with illustrated keys). Ohio State University Press, Columbus, OH. 782 pp.

Turner, T.F. 2001. Comparative study of larval transport and gene flow in darters. Copeia 2001:766–774.

US Army Corps of Engineers (USACE). 1997. Montour Run Watershed, Allegheny County, PA: Water quality and aquatic life resources. Pittsburgh, PA.

US Environmental Protection Agency. 2009. National Water Quality Inventory: Report to Congress, 2004 Reporting Cycle. EPA 841-R-08-001.

US Fish and Wildlife Service. 2011. Species assessment and listing priority assignment form: *Etheostoma maculatum* Kirtland. Available online at https://www.fws.gov/midwest/es/soc/fish/SpottedDarterCandidateAssess.html. Accessed 18 May 2016.

Warren, M.L., B.M. Burr, and B.R. Kuhajda. 1986. Aspects of the reproductive biology of *Etheostoma tippecanoe* with comments on egg-burying behavior. American Midland Naturalist 116:215–218.

Winn, H.E. 1958. Comparative reproductive behavior and ecology of fourteen species of darters (Pisces-Percidae). Ecological Monographs 28:155–191.

Yoder, C.O., E.T. Rankin, M.A. Smith, B.C. Alsdorf, D.J. Altfater, C.E. Boucher, R.J. Miltner, D.E. Mishne, R.E. Sanders, and R.F. Thoma. 2005. Changes in fish-assemblage status in Ohio's nonwadeable rivers and streams over two decades. American Fisheries Society Symposium 45:399–429.

Zorach, T., and E.C. Raney. 1967. Systematics of the percid fish, *Etheostoma maculatum* Kirtland, and related species of the subgenus *Nothonotus*. American Midland Naturalist 77:296–322.

CHAPTER TWO

The genetic population structure of *Etheostoma camurum* (Bluebreast Darter) in the upper Ohio River watershed

The Bluebreast Darter, *Etheostoma camurum*, is a small – bodied benthic fish that prefers riffle habitat with silt-free substrates of cobble. It was previously listed as a State – threatened species in Ohio and Pennsylvania. Populations were restricted to areas of refugia after the industrial revolution, but recently they have been documented as expanding their distribution in the upper Ohio River watershed and have since been de-listed. In order to properly manage *E. camurum* recovery, detailed information about the genetic diversity and genetic population structure is needed to identify management units to target conservation efforts towards genetically diverse source populations. Six populations sampled between 2012 to 2014 and two potential source populations (previously collected in 2006) were analyzed for genetic diversity, estimated number of migrants per generation, and genetic population structure using six polymorphic microsatellite loci. Analysis revealed relatively high genetic diversity within the populations and no detectable signs of isolation or inbreeding. Populations showed low levels of divergence between the six centrally located populations with the number of migrants indicating a consistent level of gene flow between the populations. Populations exhibited a lack of structuring consistent with gene flow between populations which suggests minimal impact from the navigational lock and dam system. Finally, a population of *E. camurum* from the Ohio River was consistently identified as a unique population during structure analysis exhibiting unique characteristics that would suggest it might represent an advancing population in the Ohio River mainstem.

2.1 - Introduction

Background

Since the implementation of the Clean Water Act (Clean Water Act, 1972) it has been generally accepted that our nation's waterways have experienced improved water quality. Regardless, recent nationwide studies by the U.S. Environmental Protection Agency (2009) and Brown and Froemke (2012) have documented that the main sources of aquatic ecosystem stress have shifted from point to nonpoint-source pollution (e.g. sedimentation and nutrient runoff). In parallel to this change in stress, Jelks et al. (2008) nationwide assessment of North American fishes showed that the impairment of inland fishes has increased substantially since the American Fisheries Society's last assessment in 1989. These studies on water quality and fish imperilment focus on the national level and may overlook trends documented on regional levels. In the upper Ohio River watershed, severe water quality degradation was the result of years of abandoned mine discharge, industrial effluents and untreated sewage (Argent et al. 2007, Tarr 2002). Additionally, since the early 1800's the Allegheny, Ohio, and Monongahela rivers have essentially been converted from a lotic to a lentic system with the construction of 30 navigational lock and dams (USACE 2003, 2004) which fragments riverine habitat. These navigational waterways are also dredged to meet navigation requirements (minimum channel depth of 2.7 m) and for commercial aggregates. Guenther and Spacie (2006), Santucci Jr et al. (2005), and Argent and Kimmel (2010, 2014) have shown that this interruption in connectivity fragments habitat, and alters and isolates fish communities (Freedman et al. 2013, 2014), which may ultimately result in disjunct and reproductively isolated fish populations. Despite all of these stressors, recent surveys of fish communities in the upper Ohio River watershed (since 2003)

have documented recovery in many areas (Argent and Kimmel 2010, Freedman et al. 2009, Koryak et al. 2009, 2011; Yoder et al. 2005). The focal species, the Bluebreast Darter (*E. camurum*), which was previously listed as a State-threatened species in Ohio (in 1990) and Pennsylvania (in 1999), was historically documented as having a disjunct population in the upper Ohio River watershed (Cooper 1983, Kuehne and Barbour 1983, Page 1983, Stauffer et al. 2016, Trautman 1981). In conjunction with the aforementioned regional recovery, assessments recently completed by Freedman et al. (2009), Honick et al. (2017), Howell (2007), and Ohio EPA (2016) have documented increases in the distribution of *E. camurum*. Therefore, it was hypothesized that *E. camurum* was expanding its distribution downstream from areas of refugia such as French Creek in the upper Allegheny River watershed of Pennsylvania and also using the Ohio River mainstem to move upstream from refugia such as the Kokosing River in Ohio.

Darters, which are small-bodied benthic fish, are considered indicator species of habitat integrity making them ideal species to integrate into aquatic ecosystem monitor protocols. Recent work has shown that small resident fish, such as darters, should be incorporated into models that predict the effects of urbanization (e.g. increased storm run-off and impervious surfaces) (Wenger et al. 2010), and Yeardley (2000) determined that darters are indicator species for bioaccumulation of chemicals that are harmful to humans and biota. In addition, darters have been shown to be important prey items for many sport fish including Smallmouth Bass (*Micropterus dolomieui*) (Rahel and Stein 1988), Largemouth Bass (*Micropterus salmoides*) (Fish and Savitz 1983, Labay and Brandt 1994), and Muskellunge (*Esox masquinongy*) (Kapuscinski and Farrell 2014). Many species of freshwater mollusks are critically imperiled in North America (Campbell et al. 2008, Haag 2012, Haag and Williams 2014) and darters have been documented as larval hosts for distribution of freshwater mussel glochidia (O'Dee and

Watters 2000). Therefore, darters assume many key roles in aquatic ecosystems, and it is crucial to monitor the status of darters species like *E. camurum*.

Microsatellites and Fish Conservation Management

The conservation management of fish species is challenging because of the complexity of aquatic environments resulting from the spatial and temporal variation that is encountered during field surveys (Knouft et al. 2011). Proper conservation management requires collecting data beyond habitat and biotic assessment, and involves incorporating genetic analysis to define management units (Funk et al. 2012, Palsbøll et al. 2007), identify barriers to gene flow (Frankham 2010, Magoulas et al. 2006, Scribner et al. 2016), and to detect genetically diverse source populations from genetically depauperate sink populations (Barson et al. 2009, Gaggiotti 1996, Hänfling and Weetman 2006). According to Scribner et al. (2016), this multi-disciplinary approach is the best management strategy which ultimately helps a manager to understand disturbance, recovery, and the drivers of aquatic biodiversity. There are many molecular markers that can be utilized to assess genetic variation of fish populations - e.g. allozymes, mitochondrial DNA (mtDNA), amplified fragment length polymorphisms (AFLPs), short tandem repeats (STRs, e.g. microsatellites), and single nucleotide polymorphisms (SNPs) (Allendorf 2017, Saura and Faria 2011, Toro et al. 2009). However, microsatellites, which consist of tandemly repeating motifs of nuclear DNA sequences (two to six base pairs in size), have been extensively used in fisheries management since their development in the early 1990s (Allendorf 2017, Saura and Faria 2011, Wan et al. 2004). Microsatellites have been documented as 1) being useful to prioritize conservation units (Avise 2004), 2) providing information to quantify levels of genetic

differentiation within a broad range of fish species such as Sturgeon (*Acipenser spp*.) (Tagliavini et al. 1999), Anchovy (Engraulidae) (Magoulas et al. 2006), Three-Spine and Nine-Spine Stickleback (*Gasterosteus aculeatus*, *Pungitius pungitius*) (Caldera and Bolnick 2008, Shikano et al. 2010), Shoal Bass (*Micropterus cataractae*) (Dakin et al. 2007), and numerous darter species (Davis et al. 2014, Fluker et al. 2010, Ginson et al. 2015, Haponski et al. 2009, Robinson et al. 2013), and 3) able to detect and assess the effects of barriers on gene flow (Beneteau et al. 2009, Magoulas et al. 2006, Roberts et al. 2013).

Microsatellites – Strengths

Because of the extensive use of microsatellites in population genetic analysis, recent assessments have determined the marker's strengths and weaknesses. For example, microsatellites are a useful tool in conservation genetics for determining genetic structuring and genetic differentiation within and between populations. Microsatellites are characterized by high degrees of polymorphism. Polymorphic microsatellite loci exhibit high allelic richness that's derived from having a varying number of repeats within each locus (Putman and Carbone 2014, Wan et al. 2004). Two mechanisms of mutation leading to variably sized microsatellites have been identified as DNA polymerase slippage and unequal crossing-over (Thuillet et al. 2002). Goldstein and Schlotterer (1999) and Sia et al. (1997) indicate that slippage is most likely the major mechanism. In addition, the amount of polymorphism exhibited by microsatellites is linked to the mutation rate. It has been reported that microsatellites evolve 100 – 1000x faster than single copy nuclear DNA (Wan et al. 2004) with mutation rates varying between repeat motif size (Chakraborty et al. 1997), between loci (Chapuis and Estoup 2007), and across species

(Chapuis and Estoup 2007, Putman and Carbone 2014). Ultimately, multi-locus genotyping from microsatellites allows individual to be assigned a genetic identity. Genetic patterns across individuals may provide information about recent evolutionary events among subpopulations of a species and help to determine population structure (Allendorf 2017, Putman and Carbone 2014).

Hypotheses and Objectives

In summary, darters such as *E. camurum* have been documented as indicator species of environmental quality. Due to recent fish surveys indicating expansion of *E. camurum* throughout the region, microsatellite analysis was used to provide insight into the origin(s) of these recently discovered populations and further the understanding of genetic diversity, population structure in regards to recent colonization events, and any potential effects of the navigational lock and dam system on gene flow. I hypothesized that the source population(s) for *E. camurum* living around Pittsburgh, PA were likely to be from refuge populations such as Kokosing R. and French Cr. I also hypothesized that gene flow across *E. camurum* populations would exemplify the Stepping Stone Model (Kimura and Weiss 1964), but the navigational lock and dam system may be impeding gene flow. Microsatellite analysis at six polymorphic loci was utilized to assess the genetic diversity and population structure of *E. camurum* sampled from eight locations in the upper Ohio River watershed (Fig 2.1).

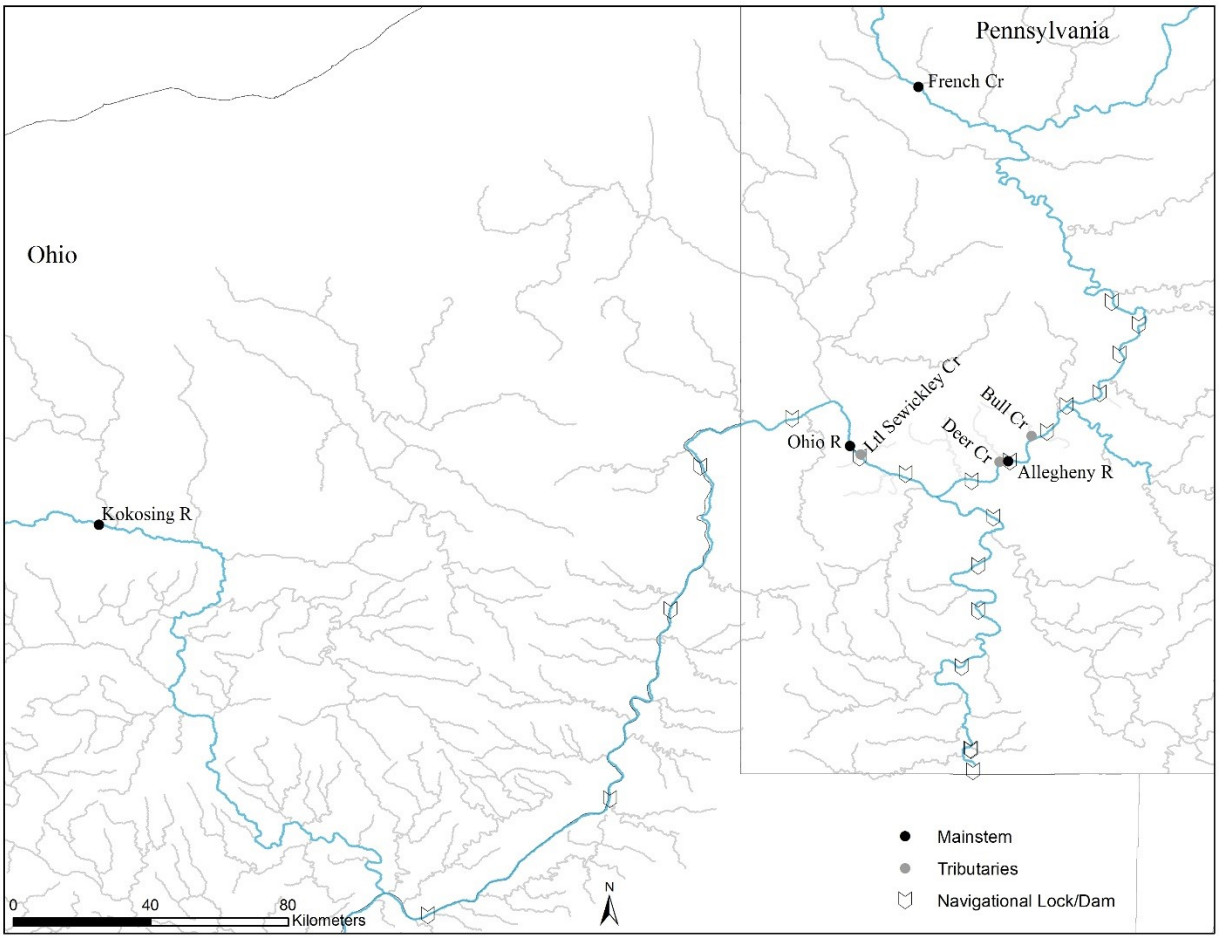

Figure 2.1. Map illustrating the locations of the sample sites. Mainstem river sample sites are indicated with solid black dots and tributary sample sites are indicated by gray dots.

2.2 - Methods

Study Area and Sample Collection

Surveys in the tailwaters of two lock and dam (L/D) installations on the Allegheny and Ohio rivers were completed (2013-2014) with an electrified Missouri trawl, and backpack electrofishing surveys were performed in three tributaries (2013-2014) (Fig. 2.1) (see Chapter One for sampling details). Caudal finclips were collected from *E. camurum* and stored in 95% ethanol at room temperature until DNA extraction. Samples from the Kokosing River in Ohio

(OH) and French Creek in Pennsylvania (PA) were previously collected by Howell (2007). Individuals sampled from Deer Creek in Harmarville, PA were collected at the same location, but two different time periods (Deer Cr. A – 2013, Deer Cr. B – 2012) and thus were kept separate for all analysis. ArcMap (v. 10.3.1) was used to construct the sample location map.

DNA Extraction and Microsatellite Amplification

DNA was extracted from the finclips following a standard phenol:chloroform extraction (Maniatis 1982). DNA from the *E. camurum* populations in the Kokosing River, OH and French Creek, PA was previously extracted by Howell (2007). Twenty-three candidate loci (Table 2.1) that were developed in other darter species (Gabel et al. 2008, Porter et al. 2002, Tonnis 2006) were evaluated for amplification and polymorphism within *E. camurum*. One locus was developed in the Striped Darter, *Etheostoma virgatum*, (Porter 2002), 14 loci were developed in the Rainbow Darter, *Etheostoma caeruleum*, (Tonnis 2006) and 8 loci that were developed in the Cherokee Darter, *Etheostoma scotti* (Gabel 2008). All were tested for amplification in *E. camurum* (Table 2.1). Published PCR and thermocycle conditions were followed initially, but if amplification was unsuccessful conditions were modified in attempts to achieve successful amplification. Four individuals of *E. camurum* from Kokosing R. (2 ea.) and French Cr. (2 ea.) were used for screening. The Kokosing R. and French Cr. individuals (25 from each population) were utilized to confirm informative microsatellite loci that could detect the genetic population structure between the two geographically isolated and potential source populations. Polymerase chain reaction (PCR) and thermocycle protocols were adapted from Gabel et al. (2008) and

optimized with the Kokosing R. and French Cr. individuals (see Appendix B – Table B.1 for detailed protocols). PCR was performed in 12.0 µL reactions with the

Table 2.1. List of 23 microsatellites that were screened and the species the microsatellite was developed in. Bold indicates the loci selected for this study.

Locus	Species	Repeat Motif	Reference
Cv24	***E. virgatum***	**-**	**Porter, 2002**
Eca6	*E. caeruleum*	(GATA)GAAA(GATA)	Tonnis, 2006
Eca10	*E. caeruleum*	(GATA)	Tonnis, 2006
Eca11	***E. caeruleum***	**(GATA)N(GATA)N**	**Tonnis, 2006**
Eca13	*E. caeruleum*	(TAGA)	Tonnis, 2006
Eca14	*E. caeruleum*	(TAGA)	Tonnis, 2006
Eca22	*E. caeruleum*	TAGA…(TAGA)…(TAGA)	Tonnis, 2006
Eca24	*E. caeruleum*	(GATA)N(GATA)	Tonnis, 2006
Eca36	*E. caeruleum*	(TAGA)…(TAGA)…(TAGA)	Tonnis, 2006
Eca37	*E. caeruleum*	(GATA)	Tonnis, 2006
Eca44	*E. caeruleum*	(TAGA)	Tonnis, 2006
Eca46	***E. caeruleum***	**(TAGA)**	**Tonnis, 2006**
Eca48	***E. caeruleum***	**(TAGA)CTTA(TAGA)**	**Tonnis, 2006**
Eca49	*E. caeruleum*	(GATA)	Tonnis, 2006
Eca70	***E. caeruleum***	**(GATA)**	**Tonnis, 2006**
Esc18	*E. scotti*	(GATA)	Gabel, 2008
Esc26b	*E. scotti*	(TAGA)	Gabel, 2008
Esc57	*E. scotti*	(GATA)	Gabel, 2008
Esc68	*E. scotti*	(AGAT)	Gabel, 2008
Esc96	*E. scotti*	(CTAT)	Gabel, 2008
Esc120	*E. scotti*	(AGAT)	Gabel, 2008
Esc132b	***E. scotti***	**(CTAT)**	**Gabel, 2008**
Esc187	*E. scotti*	(GTCT)	Gabel, 2008

following reagents (final concentrations): Fisher buffer B (1x), $MgCl_2$ (2.5 mM), dNTPs (0.8 mM), forward and reverse primers (0.25 µM each), *Taq* polymerase (0.8 units), and either 24 or 36 ng of DNA (see Appendix B – Table B.1 for locus specific details). Each forward primer was labeled with a specific fluorophore color (6-FAM, NED, PET, or VIC) for fragment analysis on an ABI 3130 genetic analyzer. Individuals were genotyped by allele size in base pairs using Peak

Scanner (v 2.0) utilizing the parameters for GeneScan™ - 500 Liz™ size standard. Genotyping data was scored independently by three individuals to minimize scoring errors. Samples with low quality fragment analysis data were re-amplified and re-run on the genetic analyzer. After all samples were genotyped, allele sizes were assigned by binning individuals to a size range (e.g. an allele of 200 base pairs could range from 199.85 to 200.46 base pairs on Peak Scanner). When the cutoff size between bins was ambiguous, individuals were completely re-processed (from PCR to fragment analysis) to confirm allele sizes.

Marker Validation

Genotyping data was tested for departure from Hardy-Weinberg equilibrium (HWE) in all locus-by-site combinations using Arlequin (v. 3.5) (Excoffier 2005) using the Markov Chain Monte Carlo (MCMC) method with 100,000 dememorization and 1,000,000 Markov chain steps. With this test, the observed genotype frequencies are compared to expected genotype frequencies under the conditions of no mutation, no migration, no selection, infinite population size, and no non-random mating. GenePop v. 4.2 (Rousset 2008) was used to assess the alternative hypotheses of heterozygote excess and deficiency for each locus in each population following Rousset and Raymond (1995). Markov Chain parameters were 10,000 dememorizations, 1,000 batches, and 10,000 iterations. Linkage disequilibrium between all pairs of loci was analyzed in GenePop (v. 4.2) (Rousset 2008) using the log likelihood ratio statistic test at default parameters. Linkage disequilibrium tests confirm that the alleles at each locus are randomly associated, and thus, are not inherited together because of being in close proximity to each other on a chromosome. The B-Y correction method for multiple simultaneous tests was applied to

determine the level of significance in departure from HWE and linkage disequilibrium (Benjamini et al. 2001, Narum 2006). Micro-Checker (v. 2.2.3) (Van Oosterhout et al. 2004) was used to detect the presence of null alleles, large allele dropout, and genotyping errors (related to stutter) by calculating null allele frequencies (based on expected Hardy-Weinberg proportions) following Chakraborty et al. (1992) and Brookfield (1996). A Monte Carlo bootstrap simulation method generates allele size class differences of expected homozygote and heterozygote frequencies. Chapuis and Estoup (2007) demonstrated that the Expectation Maximization algorithm of Dempster et al. (1977) more accurately estimated null allele frequencies when compared to Chakraborty et al. (1992) and Brookfield (1996). Of note, the presence of null alleles may bias downstream assessments of genetic differentiation (i.e. an artificial increase in F_{ST}) and calculations of genetic distance (Chapuis and Estoup 2007, Putman and Carbone 2014). Therefore, F_{ST} Refined Estimation by Excluding Null Alleles (FreeNA) analysis program (Chapuis and Estoup 2007) was used to estimate null allele frequencies and adjust genotype frequencies to remove bias from the presence of null alleles in F_{ST} estimates.

Genetic Diversity and Population Structure

Statistics summarizing the genetic variation within each population was calculated using GenAlEx (v. 6.5) (Peakall and Smouse 2012, Peakall and Smouse 2006) and included: number of alleles (N_A), number of effective alleles (N_E), number of private alleles (N_P), observed heterozygosity (H_O), expected heterozygosity (H_E), and the fixation index (F). Allelic richness (A_R) and the inbreeding coefficient (F_{IS}) for each population were calculated with FSTAT v. 2.9.3 (Goudet 2001).

Multiple indices of genetic differentiation between populations were calculated. Pairwise F_{ST} values, which describe the amount of genetic differentiation among populations (Nei and Chesser 1983), were calculated in Arlequin following Weir and Cockerham (1984). Analysis of molecular variance (AMOVA) (Excoffier et al. 1992) was used to determine the significance of pairwise comparisons with estimators based on 10,000 permutations and a null hypothesis of no differentiation (i.e. $F_{ST} = 0$). The B-Y correction method for multiple simultaneous tests was applied. Additional indices were calculated in GenAlEx and included: G_{ST}, G'_{ST}, G''_{ST}, and Jost's D. G_{ST} is said to be equivalent to F_{ST} (Whitlock 2011), but G_{ST} corrects for multiple alleles within a locus (Nei 1973). Because high mutation rates (and thus high levels of genetic variation) decrease values for G_{ST} (Whitlock 2011), G_{ST}'s calculated value can never reach 1 even with absolute differentiation between populations. To standardize comparisons and account for high levels of variation among individuals, Hedrick (2005) developed G'_{ST}. Again, even under conditions of absolute differentiation between populations, calculated values of G'_{ST} never reach 1. Meirmans and Hedrick (2011) indicate that G'_{ST} tends to underestimate differentiation under a small number of populations, therefore, they developed G''_{ST} to correct for this bias. Lastly, Jost (2008) indicates that problems arise in interpreting G-statistics because they rely upon mean subpopulation heterozygosity. Therefore, Jost (2008) argues that the D statistic provides a better estimation of population differentiation because it is based off of the effective number of alleles. GenAlEx was also used to perform a Mantel test of matrix correspondence (Mantel 1967) to evaluate the validity and significance of comparing different indices of differentiation (e.g. F_{ST} and G''_{ST}). A random permutation test (9999 permutations) was used to establish significance with a null hypothesis of no significant relationship between the two matrices (Smouse et al. 1986). To determine the spatial structure or source of variation defining the genetic diversity, a

nested analysis of molecular variance (AMOVA) was completed in Arlequin following Excoffier et al. (1992). Analysis was performed using pairwise F_{ST} values with significance based on 1 x 10^4 permutations with the B-Y correction applied. Individuals were grouped by sample location and river. To understand dynamics related to gene flow, the number of migrants per generation (N_M) was calculated. Whitlock and McCauley (1999) showed that using F_{ST} to estimate gene flow (as N_M) would rarely generate an accurate estimate of N_M from real populations because of the violations of assumptions needed for the model to work. For example, estimates of the number of migrants (m) from F_{ST} follow the original, simplified equation from Wright (1943) where :

$$F_{ST} \approx \frac{1}{1 + 4Nm}$$

However, this equation was derived from another equation (Wright 1943) that showed that the mutation rate (u) influences F_{ST} estimation:

$$F_{ST} \approx \frac{1}{1 + 4Nu + 4Nm}$$

As a result, F_{ST} estimation (and the subsequent re-arrangement of the equation for the number of migrants) makes the assumption that the mutation rate is much lower than the migration rate, which is not likely when using microsatellites (Meirmans and Hedrick 2011). Therefore, N_M was also calculated in GenAlEx using the Shannon Diversity Partition multiple hierarchical level algorithm of Smouse et al. (2015). The Shannon index which is a common measure of community similarity in ecology has been shown to provide robust information when applied at the genetic level (Sherwin et al. 2006, Smouse et al. 2015).

Several methods were implemented to visualize population structure. Using a Mantel test of the statistical relationship between elements of matrices, isolation by distance (IBD) was calculated to look for a correlation between genetic differentiation of the populations and geographic distance. GenAlEx was used to generate a matrix of genetic distance from linearized pairwise F_{ST} values (Slatkin 1995) and a linearized geographic distance matrix followed by the Mantel test for significance (9999 permutations) (Peakall and Smouse 2006). Because of a potential for bias from the source populations being extremely geographically isolated (Koizumi et al. 2006), IBD was also calculated by removing Kokosing R. and French Cr. The program STRUCTURE (v. 2.3) (Pritchard et al. 2000) was used to assign individuals to clusters (i.e. populations – denoted as K) based on Bayesian analysis that utilizes allele frequencies to calculate the probability of an individual belonging to a particular cluster. Analysis was conducted under two conditions for comparison 1) the admixture model which assigns probabilities with the condition that individuals likely have mixed ancestry and can belong to any potential population, and 2) the admixture model modified by Hubisz et al. (2009) named "locprior" which adds sample location data to the genetic data. While the admixture model assumes that all individuals have an equally likely chance of belonging to any particular cluster, the locprior model may help to detect weak structure among individuals by adding the condition that individuals from the same sample location "may" have similar ancestry. All STRUCTURE analyses were run with a burn-in of 1×10^5 simulations to minimize effects from the starting configuration and 1×10^6 MCMC simulations. The number of genetic clusters was run from K = 1 to the total number of sample locations plus two (K = 10) with 15 iterations completed per K. In order to select the most appropriate number of clusters from the analysis, STRUCTURE HARVESTER (v. 0.6.94) (Earl and vonHoldt 2012) was used to implement the ΔK analysis of

Evanno et al. (2005) which selects the most likely K (i.e. number of clusters or populations) by calculating the rate of change of the log probability data between K simulations. After selection of the most probable K, CLUMPP (v. 1.1.2) (Jakobsson and Rosenberg 2007) was used to permute the 15 iterations into a mean cluster to overcome label switching and multimodality that results from repeated runs of STRUCTURE. To view the output from CLUMPP, DISTRUCT (v. 1.1) (Rosenberg 2004) was used to control the final graphics of the figure.

Additionally, population structuring was analyzed using principal coordinate analysis (PCoA). PCoA finds patterns by scaling multidimensional data into separate axes (with Eigen vectors) which proportionately explain the total variation of the dataset (Borcard and Legendre 2002). A genetic distance matrix calculated in GenAlEx was used to calculate Eigen values following the algorithm for data standardization (Orlóci 1975).

2.3 - Results

A total of 156 individuals were collected from eight sample sites (Fig 2.1) and genotyped for six loci (Table 2.1 and Appendix A - Table A.2, see Table A.3 for raw scoring data). Only six loci were selected from the initial 23 because either *E. camurum* failed to amplify or the Kokosing R. and French Cr. populations exhibited major deviations from Hardy-Weinberg Equilibrium. A total of five individuals could not be genotyped across all six loci: two individuals from the Ohio River (missing locus Eca 70), one individual from the Allegheny River (missing locus Eca 46), one individual from Bull Creek (missing locus Eca 46), and one individual from Little Sewickley Creek (missing locus Eca 70) (Appendix A - Table A.2).

Marker Validation

Significant deviations from HWE were documented in two loci in four separate populations following B-Y correction (adjusted $\alpha = 0.02041$), (Table 2.2). Out of the two potential source populations, only French Cr. exhibited one locus out of HWE – Eca70. For all loci out of HWE, observed heterozygosity (H_O) was lower than expected heterozygosity (H_E) (Appendix A - Table A.4). Exact tests of heterozygote deficit calculated in GenePop showed that locus Eca70 in DeerA Cr. was approaching significance ($p = 0.0665$) and Eca46 in Bull Cr. was significant before B-Y correction ($p = 0.0202$). One allelic combination from the Allegheny River showed significant linkage disequilibrium ($p = 0.00069$) between locus Eca48 and Esc132b after B-Y correction (adjusted $\alpha = 0.01507$). In Micro-Checker, the increased frequency of homozygotes predicted by the four estimators in locus CV24, DeerA Cr. and in Eca46 in Bull Cr. signaled the potential for null alleles. There was no evidence of large allelic dropout or scoring errors due to stutter.

Genetic Diversity

Genetic diversity across the sample sites was high with the average number of alleles ranging from 5.333 to 9.166 and the average number of effective alleles ranging from 3.206 to 4.963 (Table 2.3). Allele frequency histograms were created for each locus to visually represent diversity by grouping the populations by river (Appendix A - Fig. A.2 – A.4). The number of private alleles was highly variable and ranged from 0 to 14 (Table 2.3). Average allelic richness ranged from 4.293 to 7.487 across all locations. The average observed heterozygosity exceeded

Table 2.2. Hardy-Weinberg Equilibrium (HWE) analysis of each microsatellite locus reported as the p-value after B-Y correction. Locus Eca 46 was significantly out of HWE in Allegheny R. and Bull Cr. populations while locus Eca 70 was significantly out of HWE in Deer Cr. in 2013 and French Cr.

Locus	Kokosing	Ohio	Ltl Sewickly	DeerA (2013)	DeerB (2012)	Allegheny	Bull	French
Eca46	0.184	0.364	0.641	0.224	0.922	**0.000**	**0.001**	0.419
Eca70	0.197	0.325	0.097	**0.000**	0.385	0.094	0.866	**0.012**
CV24	0.533	0.249	0.154	0.040	0.916	0.780	0.431	0.325
Eca48	0.308	0.503	0.264	0.333	0.716	0.570	0.454	0.598
Eca11	0.556	0.153	0.075	0.117	0.731	0.920	0.204	0.900
Esc132b	0.393	0.518	0.827	0.385	0.197	0.950	1.000	0.466

Bold values indicate significance after B-Y correction (adjusted $\alpha = 0.02041$)

Table 2.3. Genetic variation patterns within each sample location. *N* the number of individuals, N_A average number of alleles per locus, N_E average number of effective alleles, A_R average allelic richness, N_P number of private alleles across all loci, H_O average observed heterozygosity, H_E average expected heterozygosity, F fixation index

Pop	*N*	N_A (SE)	N_E (SE)	A_R	N_P	H_O (SE)	H_E (SE)	F (SE)
Kokosing	25	5.333 (0.843)	3.206 (0.458)	4.293	2	0.653 (0.043)	0.674 (0.039)	0.011 (0.036)
Ohio	11	8.166 (1.166)	4.963 (0.769)	7.487	14	0.728 (0.051)	0.806 (0.042)	0.054 (0.027)
Ltl Sewickley	13	6.000 (0.447)	4.021 (0.362)	5.416	0	0.791 (0.038)	0.773 (0.019)	-0.065 (0.047)
DeerA (2013)	27	7.833 (1.249)	4.848 (0.67)	5.943	6	0.741 (0.051)	0.792 (0.023)	0.045 (0.068)
DeerB (2012)	10	5.833 (0.477)	3.932 (0.37)	5.643	1	0.800 (0.036)	0.771 (0.029)	-0.100 (0.067)
Allegheny	30	9.166 (2.023)	4.751 (1.055)	6.148	2	0.799 (0.064)	0.762 (0.038)	-0.062 (0.059)
Bull	15	6.333 (1.201)	3.943 (0.718)	5.398	0	0.718 (0.072)	0.731 (0.049)	-0.024 (0.088)
French	25	7.000 (0.816)	3.673 (0.454)	5.249	2	0.726 (0.065)	0.719 (0.037)	-0.023 (0.059)

SE = Standard Error

the average expected heterozygosity in Little Sewickley Cr., DeerB Cr., Allegheny R., Bull Cr., and French Cr. The Kokosing R., Ohio R., and DeerA Cr. exhibited average observed heterozygosity that was below their average expected heterozygosity (Table 2.3).

Estimates of genetic differentiation using pairwise F_{ST} comparisons showed low levels of differentiation (values < 0.05) for the majority of the comparisons, but eight comparison involving Kokosing R., Ohio R., Deer(A & B) Cr., and Bull Cr. showed moderate levels of differentiation (values 0.05 – 0.15) (Table 2.4, see Table 2.7 for the guidelines of interpreting differentiation from F-statistics). All comparisons with the Kokosing R. were significant with three of the comparisons indicating moderate levels of differentiation. The highest level of differentiation was detected between the potential source populations Kokosing R. and French Cr. (F_{ST} = 0.097) (Table 2.4, see Appendix A - Fig. A.5 for a graphic illustrating F_{ST} comparisons in a linear manner). F_{IS} values for each population were low indicating that inbreeding is not likely occurring. When using F_{ST} to estimate the number of migrants per generation (N_M), the lowest value was between the Kokosing R. and French Cr. (N_M = 2.325) and the largest N_M was between DeerB Cr. and Little Sewickley Cr. (N_M = 1249.750) (Table 2.4). The number of migrants calculated via Shannon's index showed a similar but slightly different trend. The lowest N_M was between Little Sewickley Cr. and the Ohio R. (N_M = 0.250), and the highest was between the Allegheny R. and DeerB Cr. (N_M = 2.069) (Table 2.5). Meirmans and Hedrick (2011) suggests that G''_{ST} (which corrects for multiple alleles and bias from a small number of populations) provides the most relevant comparison with F_{ST} when interested in migration. The G''_{ST} pairwise comparisons follow a very similar pattern as compared to F_{ST}. All comparisons with Kokosing R. and French Cr. were significant (i.e. $F_{ST} \neq 0$). In contrast to F_{ST}, only three G''_{ST} comparisons showed values consistent with low levels of

Table 2.4. Genetic differentiation between populations using pairwise F_{ST} values (below the diagonal). The diagonal (in italics) is the inbreeding coefficient (F_{IS}) for each population. Above the diagonal is the number of migrants (N_M) estimated using F_{ST}.

	Kokosing	**Ohio**	**Ltl Sewickley**	**DeerA (2013)**	**DeerB (2012)**	**Allegheny**	**Bull**	**French**
Kokosing	*0.031*	3.616	4.983	2.978	7.280	4.772	6.024	2.325
Ohio	**0.065**	*0.097*	7.528	4.363	15.838	26.865	10.000	3.175
Ltl Sewickley	**0.048**	**0.032**	*-0.023*	14.975	1249.750	18.519	8.425	4.917
DeerA (2013)	**0.077**	**0.054**	0.016	*0.066*	7.358	5.597	4.634	4.619
DeerB (2012)	**0.033**	0.016	0.000	**0.033**	*-0.038*	37.860	167.535	3.734
Allegheny	**0.050**	0.009	0.013	**0.043**	0.007	*-0.049*	13.898	4.983
Bull	**0.040**	0.024	**0.029**	**0.051**	0.001	**0.018**	*0.016*	5.391
French	**0.097**	**0.073**	**0.048**	**0.051**	**0.063**	**0.048**	**0.044**	*-0.010*

Bold values indicate differentiation was significant after B-Y correction (adjusted $\alpha = 0.01928$)

Table 2.5. Genetic differentiation between populations using pairwise G″ST values (below the diagonal). The diagonal (in italics) is the inbreeding coefficient (F_{IS}) for each population. Above the diagonal is the number of migrants estimated using Shannon's Diversity index

	Kokosing	**Ohio**	**Ltl Sewickley**	**DeerA (2013)**	**DeerB (2012)**	**Allegheny**	**Bull**	**French**
Kokosing	*0.031*	0.263	0.626	0.401	0.867	0.861	0.718	0.386
Ohio	**0.232**	*0.097*	0.250	0.278	0.287	0.630	0.315	0.264
Ltl Sewickley	**0.170**	**0.171**	*-0.023*	1.069	1.350	1.652	0.729	1.098
DeerA (2013)	**0.289**	**0.310**	0.081	*0.066*	0.994	1.000	0.718	0.763
DeerB (2012)	**0.114**	0.077	0.010	**0.150**	*-0.038*	2.069	1.229	0.637
Allegheny	**0.177**	0.050	0.062	**0.190**	0.032	*-0.049*	1.888	1.112
Bull	**0.130**	0.097	**0.116**	0.214	0.007	**0.072**	*0.016*	0.846
French	**0.320**	**0.307**	**0.187**	**0.209**	**0.244**	**0.185**	**0.158**	*-0.010*

Bold values indicate differentiation was significant after B-Y correction (adjusted $\alpha = 0.01928$)

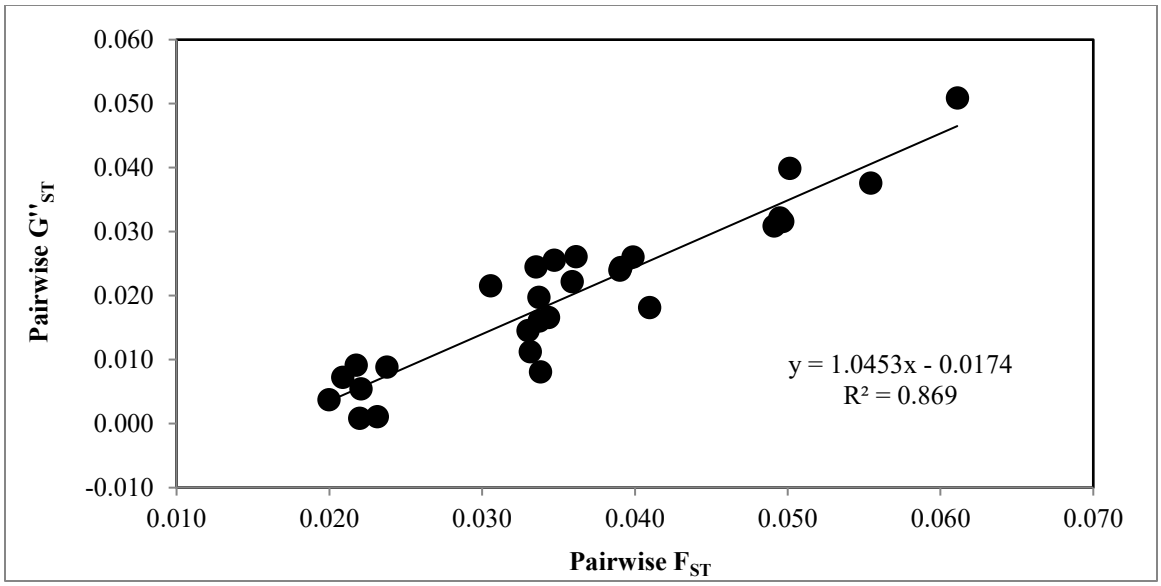

Figure 2.2. Mantel test for matric correspondence showing the validity of the relationship between F_{ST} and G''_{ST} (p = 0.0003). The strong (R^2 = 0.86904), positive relationship between F_{ST} and G''_{ST} indicates that these indices are showing very similar trends and can be used to draw similar inferences.

differentiation: Little Sewickley/DeerB Cr. (0.010), Allegheny R./DeerB Cr. (0.032), and Bull Cr./DeerB Cr. (0.007) (Table 2.5). The highest level of differentiation was between Kokosing R. and French Cr. (0.320). The Mantel test of matrix correspondence between F_{ST} and G''_{ST} showed a significant (p = 0.0003), strong (R^2 = 0.869), positive relationship indicating that G''_{ST} is an appropriate diversity index for comparison with F_{ST} (Fig. 2.2). The nested analysis of molecular variance (AMOVA) across all populations indicated that the majority of the differentiation was due to variation within individuals of the populations (94.53%) (Table 2.6). The fixation indices across all loci were generally low indicating little differentiation overall.

Genetic population structure was visualized with several methods. Isolation by distance analysis among all populations showed a significant (p = 0.032), positive, but weak (R^2 = 0.4372) relationship (Fig. 2.3, Panel A). However, when Kokosing R. (Fig. 2.3, Panel B) and French Cr. (Fig. 2.3, Panel C) were removed, the relationship between genetic differentiation and distance dissolved and was no longer significant. The Bayesian analysis by STRUCTURE was

Table 2.6. Analysis of molecular variance (AMOVA) showing that the major source of variation could be explained by differences within individuals

Source of Variation	df	Sum of Squares	Variance Components	% Variation	Fixation Indices
Among Pop	7	43.98	0.10628	4.52	F_{IS} = 0.00996
Within Pop	148	333.49	0.02235	0.95	F_{ST} = 0.04521
Within Indiv	156	345.00	2.22217	94.53	F_{IT} = 0.05472

Table 2.7. General guidelines for interpretation of levels of genetic differentiation from pairwise F – statistics (Hartl 1997).

F_{ST}/G_{ST} Value	Level of Population Differentiation
0.0	No differentiation
< 0.05	Little differentiation
0.05 - 0.15	Moderate differentiation
> 0.15	Great differentiation
1	Absolute differentiation

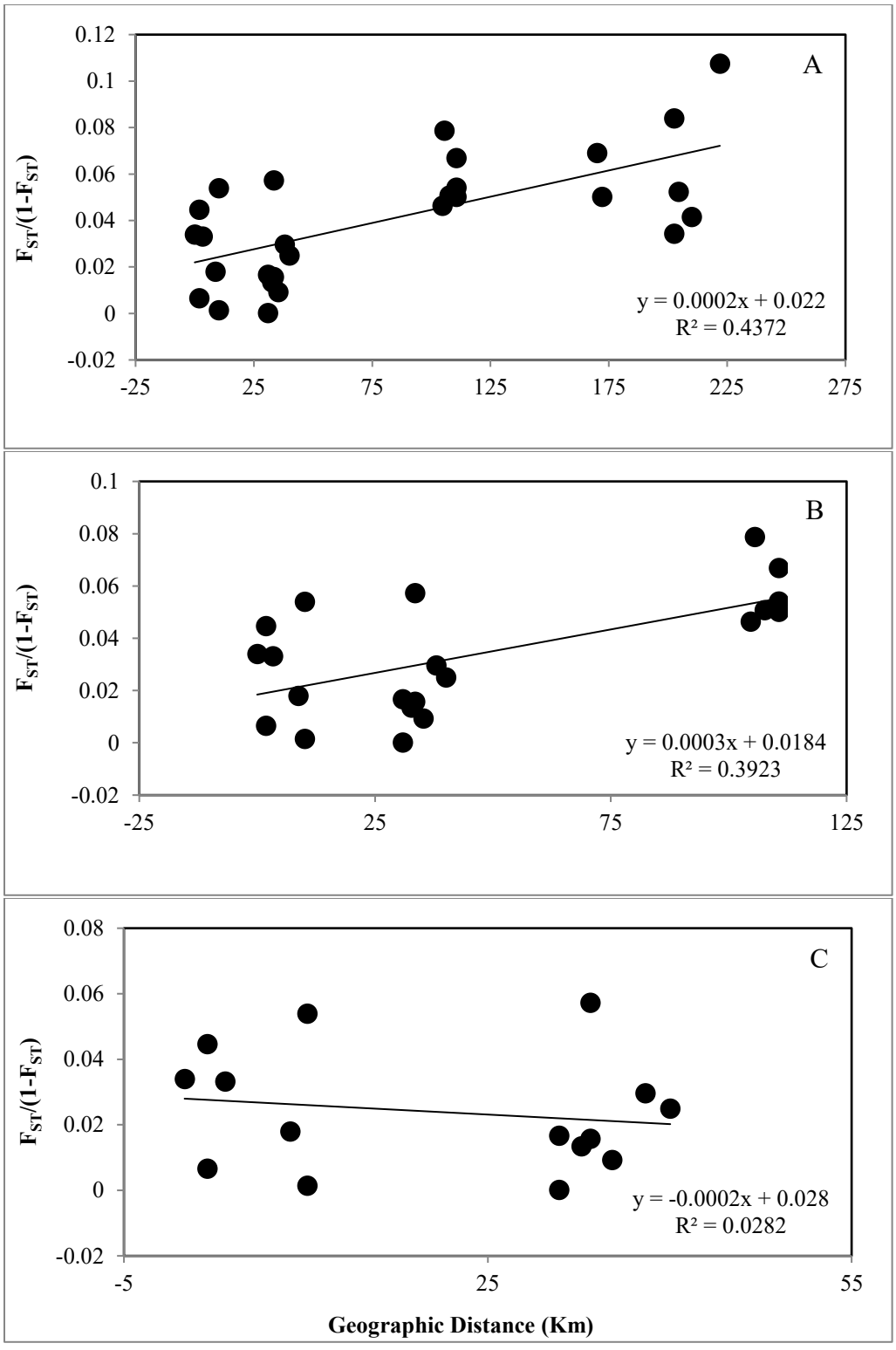

Figure 2.3. Mantel test for isolation by distance (IBD). Panel A - IBD across all populations shows a weak but significant (p = 0.032) correlation between population differentiation and distance. Panel B - IBD calculated after removal of Kokosing R. (p = 0.130) and Panel C - IBD after removing Kokosing R. and French Cr. (p = 0.321).

completed under several scenarios to look for differences in assignment probabilities in the presence of weak genetic structure. The first analysis was to confirm if the microsatellite loci were informative enough to distinguish the two geographically isolated populations which served as the potential genetic source populations for this study. Next, because the diversity indices indicated the potential for weak population structure, the samples were analyzed 1) with all sample localities run under the admixture model, 2) with the admixture model but with Kokosing R. samples removed from the analysis - because of its geographic isolation and initial STRUCTURE tests clearly indicating its assignment, and 3) the locprior model (described above) with the Kokosing R. samples removed. The microsatellite loci were informative enough to clearly distinguish the two geographically isolated populations with ΔK analysis clearly selecting K = 2 clusters (Fig. 2.4). STRUCTURE and ΔK analysis of all locations under the admixture model selected the most appropriate number of clusters at K = 3 (Fig. 2.5, see Appendix A, - Table A.5 and Fig. A.6 for examples of ΔK analysis and Structure Harvester output). Kokosing R. and French Cr. clearly formed two of the clusters. Ohio R. was assigned mainly to a third cluster while the remaining individuals could not be clearly assigned to Kokosing R., French Cr., or the Ohio R. population (Fig. 2.5). Removing the Kokosing R. from the data set showed varying results in defining population structure with the remaining samples. Delta K analysis selected K = 4 clusters, but the data supporting the selection was ambiguous (Appendix A – Fig. A.7). As a result, clusters K = 2 – 4 were plotted for comparison (Fig. 2.6). Plots K = 2 and K = 3 both indicate that the individuals from the Ohio R. are genetically distinct from the other individuals. The locprior model of analysis in STRUCTURE and subsequent ΔK analysis indicated a likely number of clusters at K = 3 (Fig. 2.7). Again, Ohio R. clearly showed

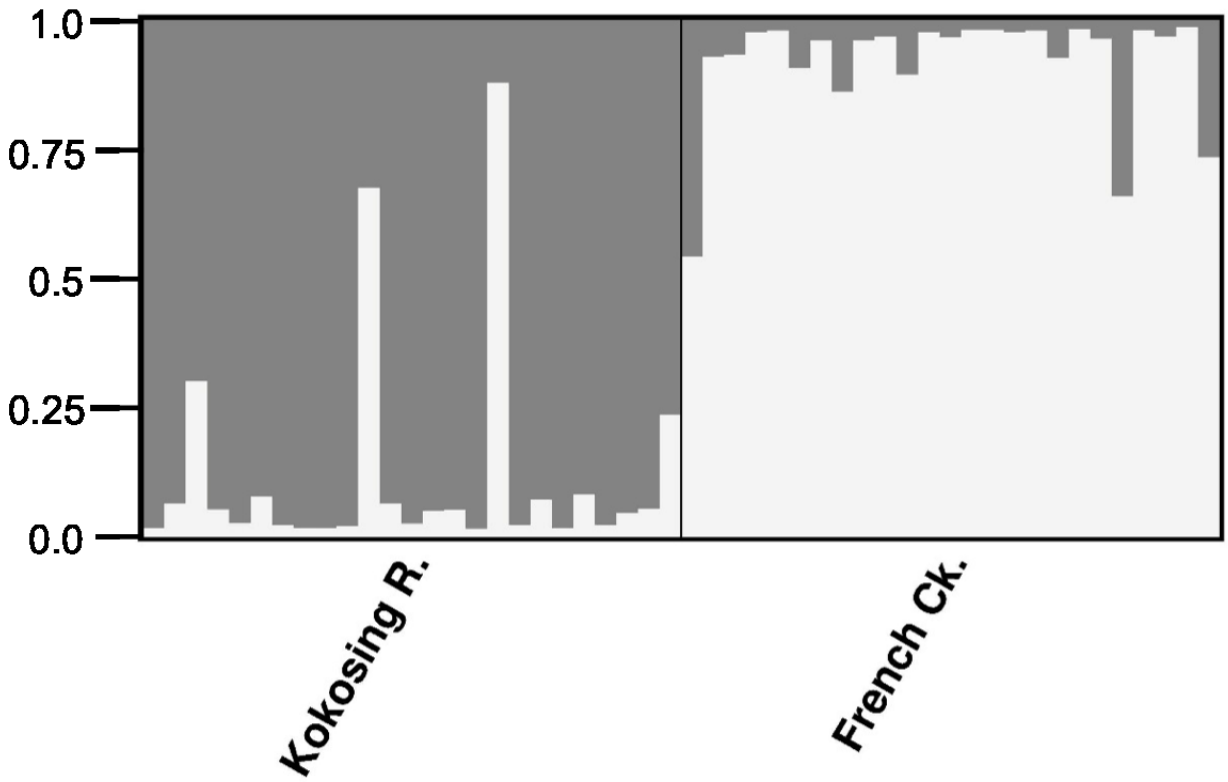

Figure 2.4. STRUCTURE analysis of Kokosing R. and French Cr. to validate that the microsatellites were informative. ΔK analysis clearly indicates the number of clusters at K = 2. The y-axis indicates the probability of belonging to a cluster, 0 to 1.

a distinct population. DeerA Cr. appeared to form its own cluster, and French Cr. clustered with individuals from Little Sewickley Cr., DeerB Cr., Allegheny R., and Bull Cr. (Fig. 2.7). Because some of the sample sizes were < 25 individuals which may be considered inadequate for genetic analysis, STRUCTURE was run again with the populations that had ≥ 25 individuals: Kokosing R., DeerA Cr., Allegheny R., and French Cr. Delta K analysis selected K = 2 as the most likely number of clusters with the Kokosing R. being distinguished from the other three locations (Fig. 2.8). Principal coordinate analysis (PCoA) was also used to delineate population structure with axis one (x – axis) able to explain 11.11 % of the variation and axis two (y – axis) explaining 8.84 %. Similar to the patterns described by STRUCTURE, PCoA delineated the Ohio R. from French Cr., but Allegheny R. had no definite delineation or clustering (Fig. 2.9).

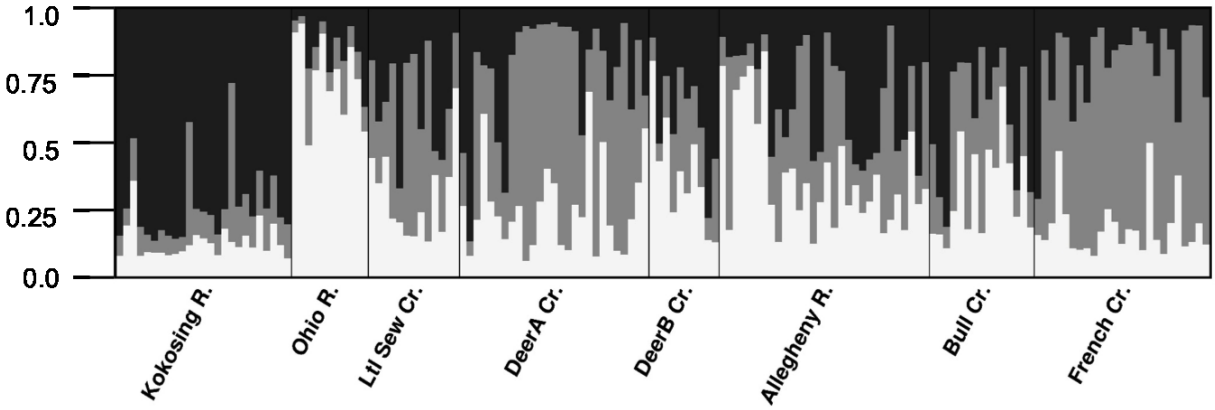

Figure 2.5. STRUCTURE analysis of all sample locations with ΔK analysis indicating the number of clusters at K = 3: Kokosing R. – K1, French Cr. – K2, and Ohio R. – K3. All other locations were not clearly assigned to a particular population. The y-axis indicates the probability of belonging to a cluster.

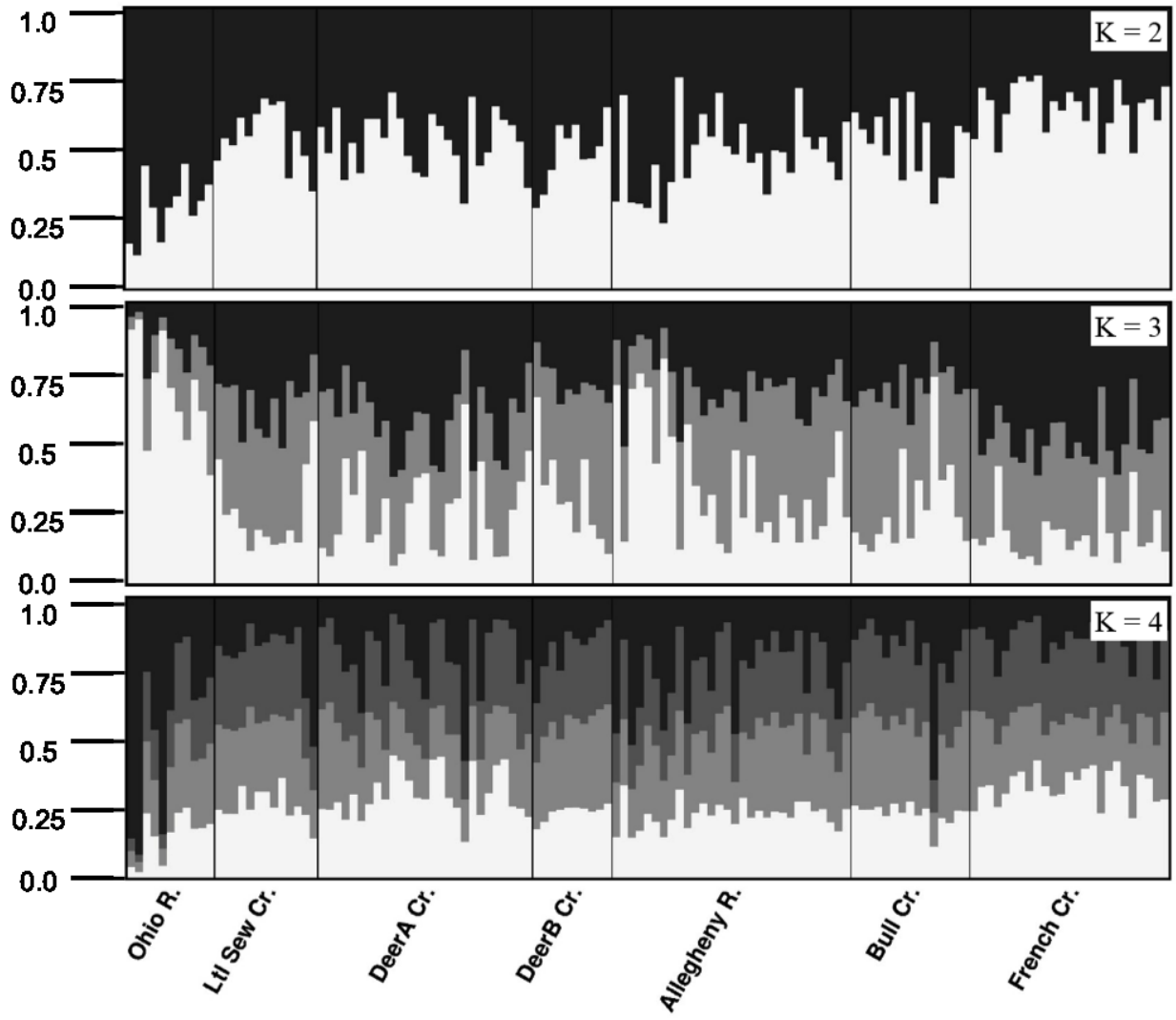

Figure 2.6. STRUCTURE plots showing K = 2 – 4 with Kokosing R. removed from analysis. K = 3 indicates that the Ohio R. and French Cr. are likely separate populations. All remaining locations and individuals could not be clearly assigned to either the Ohio R. or French Cr. with STRUCTURE indicating the presence of a third cluster. This supports the hypothesis that there may be an additional source population that I did not sample. The y-axis indicates the probability of belonging to a cluster.

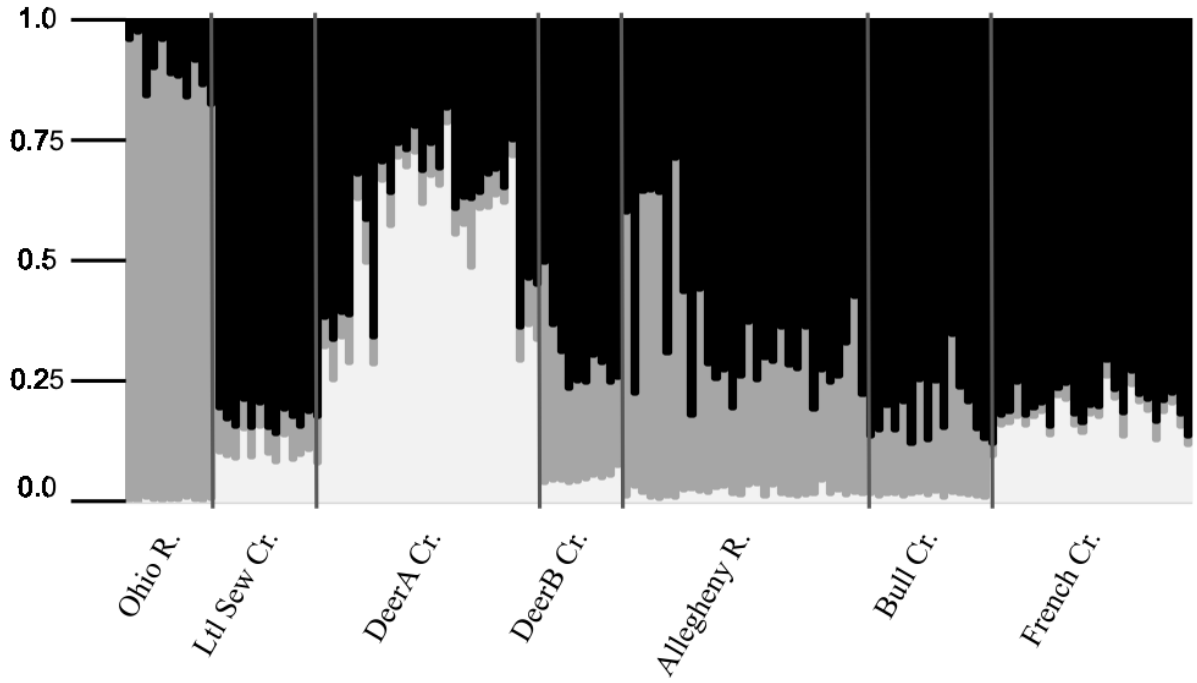

Figure 2.7. STRUCTURE plot showing the results from the 'a prior location' model. Delta K analysis indicated K = 3. Ohio R. and DeerA Cr. were assigned as unique populations while the remaining populations were largely assigned to a third cluster that was shared by French Cr. The y-axis indicates the probability of belonging to a cluster.

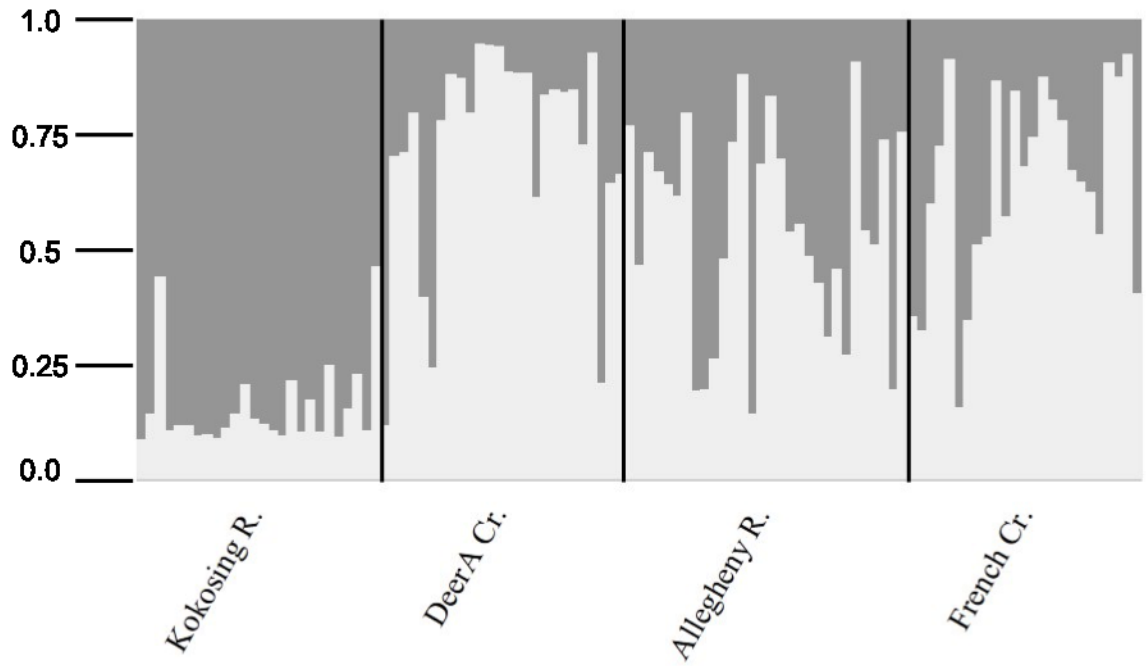

Figure 2.8. STRUCTURE analysis of the populations containing ≥ 25 individuals. Delta K analysis selected K = 2. The Kokosing R. forms a cluster while the other three locations were largely assigned to a separate cluster. This reinforces the concept that Kokosing R. is likely not a source population for the local *E. camurum*. The y-axis indicates the probability of belonging to a cluster.

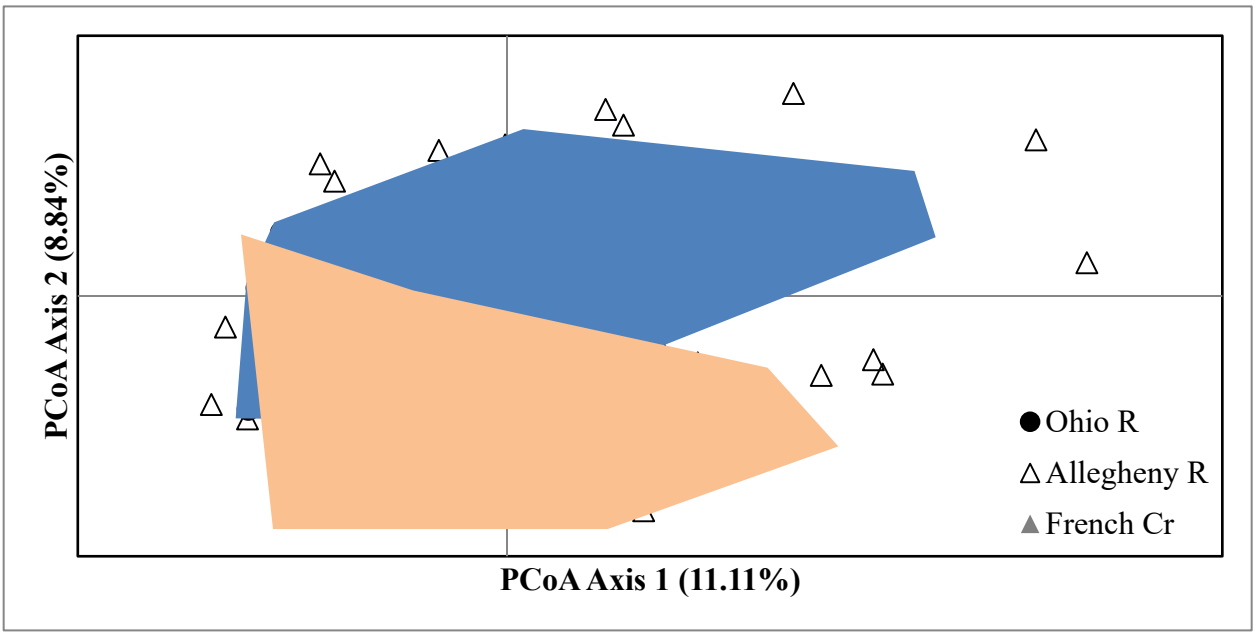

Figure 2.9. Principal coordinate analysis (PCoA) of genetic distance with the populations grouped by river. The Ohio R. and French Cr. show some differentiation, but the Allegheny R. doesn't cluster with either. Kokosing R. was removed from the analysis. This supports the concept that there is likely another source population that I did not sample for this analysis. Also, regardless of the navigational lock and dam system there appears to be sufficient gene flow between populations.

2.4 - Discussion

Objectives and Rationale for Hypotheses

Because recent studies have revealed distribution changes in *E. camurum*, it was hypothesized that these new occurrences in the upper Ohio River watershed reflect the movement of *E. camurum* from areas of refugia by moving downstream from French Creek (PA) or moving upstream via the mainstem Ohio River from areas such as the Kokosing River (OH), or a combination of dispersal with upstream and downstream movement. My goals were to 1) document genetic population structure and genetic diversity, 2) provide information regarding the direction of gene flow and estimate a potential source population(s) for this expansion, and 3) to see if the navigational L/D system was having an impact on genetic structuring or diversity. In general, the dendritic and linearized flow patterns in streams and rivers tend to establish genetic patterns in fish of low genetic diversity within populations but high genetic differentiation among populations (Barson et al. 2009, Sasaki et al. 2016, Shikano et al. 2010). Darters (Austin et al. 2011, Beneteau et al. 2009) and Sculpin (Hänfling and Weetman 2006) have demonstrated high levels of genetic structuring and differentiation due to dispersal restrictions from anthropogenic barriers. It was also hypothesized that *E. camurum* might display genetic patterns similar to the Stepping Stone Model (Kimura and Weiss 1964) of genetic differentiation from isolation by distance – where gene flow happens in a step-wise fashion between neighboring populations with genetic divergence increasing as geographical distance increases. It was surmised that French Cr. and/or the Kokosing R. were potential source populations and that local populations of *E. camurum* would exhibit a genetic signature overlapping with the source population (e.g. similar allele frequencies). French Cr. was thought to be the most likely source

population due to unidirectional stream flow which strongly influences larval and adult dispersal (Douglas et al. 2013, Shikano et al. 2010). Lastly, because Argent and Kimmel (2010) showed that local darter communities varied above and below L/D installations, I hypothesized that the navigational L/D system, as a barrier to dispersal, had the potential to create a series of isolated populations which would be signified by genetic structuring among the populations and inbreeding within each population.

Microsatellites – Weaknesses

According to Putman and Carbone (2014) and Chapuis and Estoup (2007), the same fundamentals that make microsatellites useful in conservation genetics also lead to their weaknesses. For example, many of the fundamental theories behind population genetics and mutation processes were developed under the assumptions of the infinite allele model (Tajima 1996) where mutations give rise to unique alleles which only happen once (Putman and Carbone 2014). Microsatellites more appropriately follow the stepwise mutation model (Slatkin 1995) in which new alleles are derived from either a loss or gain of the repeat motif due to strand slippage. The main issue associated with the stepwise mutation model is that microsatellites are most often genotyped by size of the allele in base pairs. Therefore, homoplasy may be likely because two individuals from different populations may experience a mutation in a locus resulting in an allele of the same size. These individuals at this particular locus appear to be identical in size, but do not have the same descent (Putman and Carbone 2014, Saura and Faria 2011). As aforementioned, the characteristics of each microsatellite locus and thus how it evolves can vary with length of the repeat, the total size of the amplicon, and by species

(Chakraborty et al. 1997, Chapuis and Estoup 2007) which according to Piry et al. (1999) occasionally places microsatellites between the infinite allele and the stepwise mutation models that may complicate data interpretation. Another complication due to genotyping based on amplicon size is the inability to identify diagnostic SNPs which may occur within the amplicon and the flanking regions where the polymerase chain reaction (PCR) primers attach. Barthe et al. (2012) showed that sequencing the flanking regions and the entire allele revealed polymorphism due to SNPs which significantly contributed to the source of differentiation documented both among individuals and populations. Several studies have been conducted on microsatellite null alleles and the potential effects on estimation of population differentiation and data interpretation. The effective definition of a null allele is an allele that consistently fails to amplify from PCR (Dakin and Avise 2004). Common causes of null alleles include sequence variation in the flanking region that prohibits primer annealing, preferential amplification of short amplicons because of DNA quality or quantity, and polymerase slippage during PCR causing a size variant that is not a true allele (Chapuis and Estoup 2007, Dakin and Avise 2004, Putman and Carbone 2014). The most notable effects of null alleles are the appearance of reduced heterozygosity (i.e. increased homozygosity and thus deviation from Hardy-Weinberg equilibrium) within a population which may result in the over-estimation of genetic differentiation among populations (e.g. increased F_{ST} values) (Chapuis and Estoup 2007, Dakin and Avise 2004, Putman and Carbone 2014). Lastly, genotyping errors from microsatellites may also arise from allelic dropout. Allelic dropout is the stochastic failure of PCR to amplify a particular allele or preferential amplification of a certain allele in heterozygotes (Pompanon et al. 2005, Soulsbury et al. 2007). Dropout is thought to be linked to low quality and quantity of DNA, such as in an environmental sample (e.g. DNA from hair or feces) (Pompanon et al. 2005,

Taberlet et al. 1996). However, Soulsbury et al. (2007), using fox tissue, demonstrated that high allelic dropout could occur with high quality DNA. Genotyping errors from allelic dropout have consequences similar to those of null alleles (Pompanon et al. 2005). Regardless of these deficiencies, microsatellites are a well-established and useful conservation genetic tool that when cautiously interpreted and used with multiple lines of evidence promote proper conservation management (Allendorf 2017). Their strength lies in the power of highly polymorphic loci and mutation rates that can identify recent changes in genetic population structure.

Hardy Weinberg Equilibrium, Null Alleles, and Linkage Disequilibrium

Two loci within four populations showed deviation from HWE and were identified as having the potential for null alleles. DeerA Cr. and Bull Cr. showed heterozygote deficiency consistent with null alleles and was likely the source of deviation from HWE. Evidence for linkage disequilibrium (LD) was only found between two loci in one population. None of these loci were removed from further analysis because Carlsson (2008) documented that null alleles had a minimal impact on F_{ST} and genetic structure analysis (using the software STRUCTURE). Chapuis and Estoup (2007) documented that null allele frequencies as high as 20% had minimal effects on genetic differentiation estimates. In this study, the highest estimated null allele frequency, 12%, occurred in locus Eca46 with the most conservative estimator from Chakraborty et al. (1992). Carlsson (2008) suggested that multi-locus analysis combined with careful interpretation of the data should still yield dependable results. The two potentially linked loci were kept in the analysis because neither exhibited deviation from HWE (Johnson et al. 2006). Additionally, Davis et al. (2014) and Ginson et al. (2015) argued that inconsistent patterns of

HWE deviation, null alleles, and LD across the affected loci and populations indicate that the impact on results should be minimal. The inbreeding coefficient (F_{IS}) was low in all populations suggesting that inbreeding depression was not a concern (Table 4.2). The genetic diversity indices indicated high levels of diversity within the populations which parallels F_{IS}, and rejects the hypothesis of *E. camurum* being isolated to small, fragmented populations. Isolated populations would be characterized by increased levels of homozygosity, high values of F_{IS}, and a low number of alleles (from inbreeding) (Sasaki et al. 2016). The average allelic richness (A_R) and the average number of alleles per population also indicated high genetic diversity and reinforces the concept that *E. camurum* dispersal is not very restricted. The Ohio R. had the highest A_R and number of private alleles (N_P). The elevated N_P suggests that the individuals in the Ohio R. may be from a genetically distinct population and have recently colonized the region. However, these statements are made with caution because genetic diversity based on allele frequencies is sensitive to sample size (Kalinowski 2004).

Genetic Differentiation

Pairwise F_{ST} indicated consistently low levels of genetic differentiation between populations even in regards to the geographically isolated populations of the Kokosing R. and French Cr. (Table 2.4, Appendix A - Fig. A.5). Both populations are more than 200 river kilometers away from the nearest populations that were sampled in the Allegheny and Ohio rivers. Therefore, based on the Stepping Stone Model, it was expected that comparison between local populations (i.e. those centrally located to the study area in the Ohio and Allegheny river) and the geographically isolated populations would consistently indicate structure with at least

moderate levels of differentiation. Because analysis was performed with multiple loci and a relatively small number of populations, G''_{ST} was calculated to correct for those biases (Meirmans and Hedrick 2011). Pairwise G''_{ST} comparisons showed moderate to great differentiation between Kokosing R., French Cr., and the local populations indicating this index may be more informative under the conditions of this study. A dendritic stream network under the assumptions of the Stepping Stone Model would be expected to reveal a pattern of linearized increases in differentiation with increased geographic separation from a source population – also referred to isolation by distance (IBD). Even though there was a general pattern of increased divergence with increased distance, there were differences depending on directionality of gene flow: upstream from the Kokosing R. vs. downstream from French Cr. For example, the Allegheny R. (G''_{ST} = 0.177) and Bull Cr. (G''_{ST} = 0.130) populations indicated less differentiation from the Kokosing R. than the Ohio R. (G''_{ST} = 0.232) population which is geographically closer. In contrast, the mainstem populations that are downstream of French Cr. did follow a linearized progression of increased differentiation, thus supporting the Stepping Stone Model. The Allegheny R. (G''_{ST} = 0.048) showed less differentiation from French Cr. than the Ohio R. (G''_{ST} = 0.073) population. Regardless of index, patterns of differentiation between the tributaries and the mainstem populations were variable and for the most part did not follow expectations. For example, Little Sewickley Cr. joins the mainstem Ohio River in the same pool (the Montgomery Pool) as the Ohio R. population and is approximately 3.0 river kilometers apart with no apparent barriers. However, pairwise G''_{ST} comparisons indicated that Little Sewickley Cr. was more differentiated with the Ohio R. (G''_{ST} = 0.171) population than either of the more distant Allegheny R. and Bull Cr. ($G''ST$ = 0.50 and 0.097, respectively) populations. In comparison with the Allegheny R. population, Deer Cr. is located within the same pool and is

approximately 2.5 river kilometers downstream while Bull Cr. is located upstream 12.0 river kilometers and separated from the Allegheny R. population by a navigational L/D. The Little Sewickley Cr. population is approximately 28 river kilometers downstream and separated by three navigational L/Ds from the Allegheny R. population. Counter intuitive to expectations of isolation by distance, Little Sewickley Cr. showed less differentiation ($G''_{ST} = 0.062$) than DeerA Cr. ($G''_{ST} = 0.190$, sampled 2013). Even being separated from the Allegheny R. population by a navigational L/D, Bull Cr. indicated lower differentiation ($G''_{ST} = 0.072$) than DeerA Cr. which is nearly adjacent to the Allegheny R. population. In summary, even though there is variation in the patterns of genetic differentiation, relatively low to moderate levels of genetic differentiation between the populations of *E. camurum* is indicative of gene flow in the absence of barriers.

Because interpretation of the data using F-statistics doesn't present a model pattern of differentiation, it's prudent to briefly discuss some of the assumptions that go along with utilizing F-statistics and microsatellites. F_{ST} and related G-statistics were modeled under several assumptions: the populations are in a state of migration – drift equilibrium, the genetic marker(s) mutate under the infinite allele model, and the populations follow structure patterns of the island population model (Meirmans and Hedrick 2011, Palsbøll et al. 2007, Putman and Carbone 2014). Migration – drift equilibrium describes the exchange of migrants and the effects of genetic drift between two populations: one population with a large effective population size and smaller population. Essentially, alleles that are gained due to migration are lost via drift (and vice versa) which over many generations establishes an equilibrium (Roderick and Navajas 2003). These conditions are likely to be violated when using markers that demonstrate high mutation rates that can vary across loci and species (e.g. microsatellites) (Putman and Carbone 2014), and with populations that have recently been founded (Roderick and Navajas 2003). Next, the assumption

that new alleles are derived strictly by the mutational processes of the infinite allele model (IAM) is not entirely compatible with microsatellites. As previously discussed, the IAM assumes that mutations arise once at a given locus (Tajima 1996), but microsatellites more closely follow the Stepwise Mutation Model (SMM) (Slatkin 1995) where new alleles are the result of the gain or loss of the repeat unit and in theory can happen repeatedly in either direction. Even though microsatellites are generally considered to follow the SMM (Barthe et al. 2012), studies have shown that variability between microsatellite loci and across species does not support strict adherence to one type of mutational model (Palsbøll et al. 2007, Putman and Carbone 2014). Lastly, F-statistics best describe differentiation between populations that follow the island model of migration where gene flow into equally sized sub-populations is equally likely among all the populations (Lowe and Allendorf 2010). Natural populations and specifically, fish, in linearized dendritic networks are less likely to satisfy this assumption. Regardless of these drawbacks, F – statistics have been extensively used to evaluate natural populations with alternative indices developed to overcome some of the aforementioned limitations. Details providing the reasoning behind utilizing G''_{ST} for this study were previously discussed, but there is one more index that deserve attention when discussing issues related to microsatellites. In particular, R_{ST} (Slatkin 1995) was developed to deal with the high mutation rates of microsatellites and its effects on inferring gene flow with F_{ST}. Unfortunately, R_{ST} estimates are only accurate when the genetic marker strictly follows the assumptions of the SMM (Hardy et al. 2003). Despite R_{ST}'s ability to model differentiation independent of the mutation rate, R_{ST} was not utilized in this study because 1) the mutation process in microsatellites varies (Meirmans and Hedrick 2011), 2) research suggests that R_{ST} is very sensitive to deviations from the SMM (Balloux et al. 2000), and 3) R_{ST} has been shown to be less accurate when populations exhibit weak structure (which is likely in

dendritic aquatic networks) or low diversity (Balloux and Lugon-Moulin 2002, Putman and Carbone 2014).

Population Structure Analysis

Particular patterns of population structure were exhibited with IBD analysis. Isolation by distance analysis is useful for detecting disruption to migration – drift equilibrium which underpins the original concept of the Stepping Stone Model and indicates that even though populations may be continuously connected, divergence occurs with increasing distance because gene flow is more likely to occur between the adjacent populations and not across the entire geographic range (Kimura and Weiss 1964). Roberts et al. (2013) showed that barriers such as hydro-electric dams can disrupt migration – drift equilibrium in small benthic fish like the Roanoke Logperch (*Percina rex*). The restricted movement of *P. rex* led to strong population structure patterns with distinct boundaries from isolation and restricted movement/gene flow. Koizumi et al. (2006) and Raeymaekers et al. (2008) showed that IBD analysis can be easily biased by including outlier populations (i.e. populations that are extremely geographically isolated). This could lead to misinterpretation of the data in two ways. First, statistically significant IBD (with outlier populations) may show that the populations are in migration – drift equilibrium, or there could be other causes to the divergence such as physical barriers like dams (Koizumi et al. 2006). Next, true patterns of IBD may be concealed for centralized populations when the sampling range occurs over a large geographic range (Koizumi et al. 2006). For these reasons, IBD analysis in this study was conducted for all populations and then compared with the removal of Kokosing R. and French Cr. Including both geographically isolated populations, IBD

showed a significant positive relationship (Fig. 2.3). Removal of Kokosing R. resulted in a weaker correlation (R^2 decreased from 0.43 to 0.39) and was no longer significant, but still indicated a positive relationship with French Cr (Fig. 2.4). Removing Kokosing R. and French Cr. revealed that the local populations (i.e. centralized) no longer exhibited IBD and showed a slightly negative and very weak relationship ($p = 0.321$, $R^2 = 0.02$) (Fig. 2.5). The pairwise comparisons from F_{ST} and G''_{ST} parallel the findings of the IBD analysis which denote that the magnitude of the geographic isolation of Kokosing R. and French Cr. was contributing to the degree of pairwise differentiation and the significant, positive relationship of IBD. The lack of IBD, low levels of differentiation, and the number of migrants (N_M) between the local populations implies that gene flow is occurring which is homogenizing the populations and limiting differentiation regardless of the navigational L/D system. Darters have been documented as having variable dispersal abilities (Douglas et al. 2013, Turner 2001), but have not been documented dispersing over long distances (Davis et al. 2014, Page 1983). However, Bronnenhuber et al. (2011) suggests that short, repeated dispersals are sufficient to prevent inbreeding and loss of genetic diversity in small fish populations that rely on areas of patchy habitat. The data in this study support the conclusion that *E. camurum* may be dispersing in short distances regardless of the navigational L/Ds and keeping the local populations from suffering the consequences of inbreeding depression and subsequent loss of genetic diversity.

The Bayesian assignment program, STRUCTURE, was used to visualize population structure under several different scenarios to assign an individual's probability of belonging to a particular cluster or group (see Methods and Results sections for details). Initial tests (data not shown) and analysis of all populations under the admixture model clearly showed that Kokosing R. individuals were a separate population (Fig. 2.5). In conjunction with IBD confirming the bias

from the Kokosing R. being extremely geographically isolated, it was decided to remove those individuals from further population structure analysis. Analysis of the remaining seven populations showed a lack of strong population structure (Fig. 2.6). Delta K analysis defined K = 4 as the most probable number of clusters, but Figure 2.6 shows that almost all individuals had approximately a 25 % chance of belong to any of the four suggested clusters. Assignment data for K = 3 possibly shows a more accurate assignment because the Ohio R. population and French Cr. (marginally) form separate clusters from the other individuals which parallels data from the pairwise F_{ST} and G''_{ST} comparisons. The lack of differentiation between the centralized populations, which is evident with F-statistics and IBD analysis, supports the findings of STRUCTURE that the centralized populations exhibit weak population structure due to gene flow. However, careful interpretation of this data extends to understanding a few of the limitations of population assignment using STRUCTURE. STRUCTURE utilizes allele frequency data which is sensitive to sample size. Therefore, analysis was limited to including only populations that had at least ≥ 10 individuals. Puechmaille (2016) used simulated and empirical datasets to assess STRUCTURE's ability to correctly assign individuals using uneven sample sizes. They found that uneven sample size and subsequent ΔK analysis led to improper hierarchical structuring and downward bias in the selection of the true number of subpopulations. This data set contains uneven sample sizes, so STRUCTURE analysis was also completed using only the populations that had ≥ 25 individuals. Regardless, K = 2 was the most likely number of clusters between the Kokosing R., DeerA Cr., Allegheny R., and French Cr (Fig. 2.8). The lack of structure from the populations in the Allegheny River drainage support the previous concept that there is sufficient gene flow occurring between those populations to homogenize allele frequencies regardless of the navigational L/D system. Even though Falush et al. (2003)

documented that weak population structure could accurately be detected with STRUCTURE when F_{ST} values were low (ranging from 0.02 to 0.10), there is concern with this dataset because 28% of the F_{ST} values fall below that criterion. To compensate for some of the shortcomings of the dataset, the locprior model was utilized to detect population structure by adding location data. Analysis indicated K = 3 which is consistent with the admixture model, but this time structure was more clearly defined (Fig. 2.7). Again, the Ohio R. population was clearly assigned as a unique cluster with Little Sewickley Cr., DeerB Cr., Allegheny R., Bull Cr., and French Cr. all having a large probability of belonging to the same cluster. Interestingly, DeerA was assigned to its own unique cluster. The locprior model reinforces the idea of gene flow between the Allegheny River populations. Though it's hard to explain, Little Sewickley Cr. consistently shows divergence from the adjacent Ohio R. population, and shows less divergence from the populations in the Allegheny River drainage with the exception of French Cr. Little Sewickley Cr. shows a large number of migrants with those populations which are all upstream and doesn't have any private alleles which may suggest that the Little Sewickley Cr. population was originally founded by migrants from the Allegheny River. The Ohio R. population was consistently identified from structure analysis and shows some unique characteristics that may be signs of a newly advancing population of *E. camurum*. For example, Dlugosch and Hays (2008), Dlugosch and Parker (2008), and Lowe and Allendorf (2010) indicate that populations in the process of successful invasion usually exhibit unique genetic qualities such as increased genetic variation and a novel combination of unique alleles. This is counter to traditional signals of decreased genetic diversity and bottlenecks of founding/invading populations, but Dlugosch and Parker (2008) reviewed 80 studies focused on species invasions and found varying degrees of loss in genetic variation and found that several species didn't show signs of genetic bottleneck or

founder effect. They also contend that the major factor contributing to successful invasion is repeated attempts and the changes that occur to genetic diversity over that time span are often not thoroughly documented. However, depending on the species, they also suggest that decades could pass before successful invasion and if the population happens to be discovered at the correct time, unique characteristics such as private alleles will still be detectable because they remain unique as a function of time, migration, and mutation rate. Recently, Bronnenhuber et al. (2011) documented that invasions of the Round Goby (*N. melanostomus*) from the Great Lakes into 20 rivers exemplified a stratified dispersal mechanism of long and short distances that averaged 500 m of advancement per year. Within four years, founder effects were mitigated in the river populations with the lake and river populations exhibiting similar levels of genetic diversity. The Ohio R. population exhibits a large number of private alleles (14), the highest allelic richness (7.487), and is significantly diverged from three out of the other six populations (excluding Kokosing R.) which at the very least may indicate that there is a source population contributing to the region that was not identified in this study. The sample size of the Ohio R. population was small (N = 11), so these data should be cautiously interpreted. When grouped by river, the PCoA, which may be biased because of the stark differences between Little Sewickley Cr. and Ohio R. populations, doesn't completely delineate the Ohio R. population as being unique. Nonetheless, PCoA does corroborate the lack of population structuring and confirms that the Ohio R. and French Cr. show slight genetic divergence, but the remaining sites in the Allegheny River drainage indicate no apparent pattern.

Hypotheses for the Differences Documented at Deer Creek

Lastly, sample collection took place over a period of three years (2012 – 2014). The geographically isolated and potential source populations (Kokosing R. and French Cr.) were collected in 2006. Deer Cr. was sampled in September of 2012 (DeerB) and in June of 2013 (DeerA), and was kept separate throughout the analysis because of potential effects from temporal variation. Remarkably, there was a drastic difference between the two sample years regardless of identical sample location. The source of the drastic change has not been identified, but there is a hypothesis that may explain the variation. First, the DeerA Cr. population showed evidence for a null allele and was heterozygote deficient resulting in deviation from HWE. The average number of alleles, average number of effective alleles, and allelic richness were not substantially different between years. However, variation was evident in the sample size (N = 10 in 2012, N = 27 in 2013), the number of private alleles (1 in 2012 vs. 6 in 2013), and levels of observed heterozygosity (2012 showed an excess, 2013 showed a deficit). Howell (2007) documented that Deer Cr. was being utilized by *E. camurum* for spawning between the months of May through August. Cooper (1983) and Trautman (1981) also documented that *E. camurum* make seasonal migrations from deeper waters to shallow riffles to spawn and then return to their deeper habitat. It is hypothesized that the sampling in June of 2013 (DeerA) likely collected a subpopulation of *E. camurum* that had migrated into Deer Cr. for spawning. In contrast, the 2012 (DeerB) sampling that occurred in September likely collected individuals that may have represented a resident population that was originally established from the Allegheny R. population and is separate from the seasonal spawning migrants. The genetic divergence data supports this hypothesis. G''_{ST} shows little differentiation between DeerB and the Allegheny R.

population, but there is moderate differentiation between DeerA and the Allegheny R. population. This indicates that the individuals from 2013 may have migrated into the stream for spawning (from an undocumented source) and were not represented at any other sample location. Another scenario is that DeerA may represent a combination of resident and seasonally spawning individuals that represent two populations but were sampled as one. This would result in cryptic subpopulation structure (the Wahlund effect) (Dharmarajan et al. 2013) in which the two subpopulations were sufficiently diverged with independent allele frequencies; which may explain the deficit in heterozygosity and the fact that DeerA had six private alleles. The concept that a new population of individuals may have migrated into Deer Cr. for spawning is also supported by the fact that the Allegheny R. population exhibited excess heterozygosity, which is a signature of admixture where previously isolated populations have recently interbred (Roderick and Navajas 2003).

2.5 - Conclusions

Genetic analysis of *E. camurum* populations in the upper Ohio River watershed shows relatively high levels of genetic diversity within their populations and no detectable evidence of inbreeding. Population structure analysis confirms that the Kokosing R. is not a likely source population and that French Cr. has/is minimally influencing the genetic structure of the local populations. The low levels of genetic differentiation, lack of IBD and population structuring indicate that the navigational L/D system has not greatly impacted gene flow in these populations. The tributaries show low levels of differentiation from the mainstem Allegheny R. population which suggests that these populations were likely established with migrants from the

Allegheny River drainage. Finally, structural analysis indicates that the Ohio R. population is likely from an undocumented source population and may represent the progression of *E. camurum* from a source that is moving upstream via the Ohio River mainstem.

Future directions should include gathering data from sample sites between these local populations and French Cr. and searching for potential source populations from downstream locations that are less geographically isolated than the Kokosing R. For example, the lower Scioto and Muskingum rivers in Ohio which have been documented with large, stable populations of *E. camurum*. Future analysis should include more rigorous estimates of gene flow (e.g. Bayesian inference of recent migration using multi-locus genotypes), incorporating more microsatellite loci, or utilizing SNPs, both of which should increase power of the analysis.

Literature Cited

Allendorf, F.W. 2017. Genetics and the conservation of natural populations: allozymes to genomes. Molecular Ecology 26:420-430.

Argent, D.G. and Kimmel, W.G. 2010. Influence of navigational lock and dam structures on adjacent fish communities in a major river system. River Research and Applications 27:1325-1333.

Argent, D.G. and Kimmel, W.G. 2014. Patterns of ichthyofaunal distribution and connectivity in navigable and free-flowing reaches of a major river system: The Allegheny river in Pennsylvania. River Research and Applications 30:631-638.

Argent, D.G., Kimmel, W.G., Lorson, R. and Emery, E. 2007. Ichthyofauna of the monongahela river basin in Pennsylvania: A contemporary evaluation. Journal of Freshwater Ecology 22:617-628.

Austin, J.D., Jelks, H.L., Tate, B., Johnson, A.R. and Jordan, F. 2011. Population genetic structure and conservation genetics of threatened Okaloosa darters (Etheostoma okaloosae). Conservation Genetics 12:981-989.

Avise, J.C. 2004. Molecular markers, natural history, and evolution. Sinauer Associates, Sunderland. 684 pp.

Balloux, F., Brünner, H., Lugon-Moulin, N., Hausser, J. and Goudet, J. 2000. Microsatellites can be misleading: an empirical and simulation study. Evolution 54:1414-1422.

Balloux, F. and Lugon-Moulin, N. 2002. The estimation of population differentiation with microsatellite markers. Molecular Ecology 11:155-165.

Barson, N.J., Cable, J. and van Oosterhout, C. 2009. Population genetic analysis of microsatellite variation of guppies (Poecilia reticulata) in Trinidad and Tobago: evidence for a dynamic source-sink metapopulation structure, founder events and population bottlenecks. Journal of Evolutionary Biology 22:485-497.

Barthe, S., Gugerli, F., Barkley, N.A., Maggia, L., Cardi, C. and Scotti, I. 2012. Always Look on Both Sides: Phylogenetic Information Conveyed by Simple Sequence Repeat Allele Sequences. Plos One 7:e40699.

Beneteau, C.L., Mandrak, N.E. and Heath, D.D. 2009. The effects of river barriers and range expansion of the population genetic structure and stability in Greenside Darter (Etheostoma blennioides) populations. Conservation Genetics 10:477-487.

Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N. and Golani, I. 2001. Controlling the false discovery rate in behavior genetics research. Behavioural Brain Research 125:279-284.

Borcard, D. and Legendre, P. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecological Modelling 153:51-68.

Boschung, H.T., Mayden, R.L., and Tomelleri, J.R. 2004. Fishes of Alabama. Smithsonian Books. 736 pp.

Bronnenhuber, J.E., Dufour, B.A., Higgs, D.M. and Heath, D.D. 2011. Dispersal strategies, secondary range expansion and invasion genetics of the nonindigenous round goby, Neogobius melanostomus, in Great Lakes tributaries. Molecular Ecology 20:1845-1859.

Brookfield, J.F.Y. 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. Molecular Ecology 5:453-455.

Brown, T.C. and Froemke, P. 2012. Nationwide assessment of nonpoint source threats to water quality. Bioscience 62:136-146.

Caldera, E.J. and Bolnick, D.I. 2008. Effects of colonization history and landscape structure on genetic variation within and among threespine stickleback (Gasterosteus aculeatus) populations in a single watershed. Evolutionary Ecology Research 10:575-598.
Campbell, D.C., Johnson, P.D., Williams, J.D., Rindsberg, A.K., Serb, J.M., Small, K.K. and Lydeard, C. 2008. Identification of 'extinct' freshwater mussel species using DNA barcoding. Molecular Ecology Resources 8:711-724.
Carlsson, J. 2008. Effects of Microsatellite Null Alleles on Assignment Testing. Journal of Heredity 99:616-623.
Chakraborty, R., Andrade, M.D., Daiger, S.P. and Budowle, B. 1992. Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. Annals of Human Genetics 56:45-57.
Chakraborty, R., Kimmel, M., Stivers, D.N., Davison, L.J. and Deka, R. 1997. Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. Proceedings of the National Academy of Sciences of the United States of America 94:1041-1046.
Chapuis, M.-P. and Estoup, A. 2007. Microsatellite Null Alleles and Estimation of Population Differentiation. Molecular Biology and Evolution 24:621-631.
Cooper, E.L. 1983. Fishes of Pennsylvania and the North-Eastern United States. The Pennsylvania University Press, University Park, PA. 252 pp.
Dakin, E. and Avise, J. 2004. Microsatellite null alleles in parentage analysis. Heredity 93:504-509.
Dakin, E.E., B.A. Porter, B.J. Freeman, and J.M. Long. 2007. Genetic integrity of an isolated population of shoal bass (*Micropterus cataractae*) in the upper Chattahoochee River basin. Report NPS/NRWRD/NRTR - 2007/366. National Park Service. Fort Collins, Colorado.
Davis, D.J., Wieman, A.C. and Berendzen, P.B. 2014. The influence of historical and contemporary landscape variables on the spatial genetic structure of the rainbow darter (Etheostoma caeruleum) in tributaries of the upper Mississippi River. Conservation Genetics 16:167-179.
Dempster, A.P., Laird, N.M. and Rubin, D.B. 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the royal statistical society. Series B (methodological):1-38.
Dharmarajan, G., Beatty, W.S. and Rhodes, O.E. 2013. Heterozygote deficiencies caused by a Wahlund effect: Dispelling unfounded expectations. The Journal of Wildlife Management 77:226-234.
Dlugosch, K.M. and Hays, C.G. 2008. Genotypes on the move: Some things old and some things new shape the genetics of colonization during species invasions. Molecular Ecology 17:4583-4585.
Dlugosch, K.M. and Parker, I.M. 2008. Founding events in species invasions: Genetic variation, adaptive evolution, and the role of multiple introductions. Molecular Ecology 17:431-449.
Douglas, M., Keck, B.P., Ruble, C., Petty, M., Shute, J.R., Rakes, P. and Hulsey, C.D. 2013. Pelagic larval duration predicts extinction risk in a freshwater fish clade. Biology Letters 9.
Earl, D.A. and vonHoldt, B.M. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4:359-361.

Environmental Protection Agency, U.S. 2009. National Water Quality Inventory: Report to Congress, 2004 Reporting Cycle. EPA 841-R-08-001.
Etnier, D.A.a.S., W.C. 1993. The Fishes of Tennessee. University of Tennessee Press - Knoxville. 689 pp.
Evanno, G., Regnaut, S. and Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611-2620.
Excoffier, L., Laval, G., and Schneider, S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1:47-50.
Excoffier, L., Smouse, P.E. and Quattro, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics 131:479-491.
Falush, D., Stephens, M. and Pritchard, J.K. 2003. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. Genetics 164:1567-1587.
Clean Water Act - Federal Water Pollution Control Act of 1972, 33 U.S.C 1251-1376; Chapter 758; P.L. 845, June 30, 1948; amended: 85 stat. 379; P.L. 92-240, March 1, 1972.
Fish, P.A. and Savitz, J. 1983. Variations in Home Ranges of Largemouth Bass, Yellow Perch, Bluegills, and Pumpkinseeds in an Illinois Lake. Transactions of the American Fisheries Society 112:147-153.
Fluker, B.L., Kuhajda, B.R., Lang, N.J. and Harris, P.M. 2010. Low genetic diversity and small long-term population sizes in the spring endemic watercress darter, Etheostoma nuchale. Conservation Genetics 11:2267-2279.
Frankham, R. 2010. Challenges and opportunities of genetic approaches to biological conservation. Biological Conservation 143:1919-1927.
Freedman, J., Lorson, B., Taylor, R., Carline, R. and Stauffer, J., Jr. 2014. River of the dammed: longitudinal changes in fish assemblages in response to dams. Hydrobiologia 727:19-33.
Freedman, J.A., Carline, R.F. and Stauffer, J.R. 2013. Gravel dredging alters diversity and structure of riverine fish assemblages. Freshwater Biology 58:261-274.
Freedman, J.A., Stecko, T.D., Criswell, R.W., and Stauffer, J.R. 2009. Extensions of the known ranges of *Percina shumardi* Girard and three species of *Etheostoma* (subgenus *Nothonotus*) in Pennsylvania. Journal of the Pennsylvania Academy of Science 83:42-44.
Funk, W.C., McKay, J.K., Hohenlohe, P.A. and Allendorf, F.W. 2012. Harnessing genomics for delineating conservation units. Trends in Ecology and Evolution 27:489-496.
Gabel, J.M., Dakin, E.E., Freeman, B.J. and Porter, B.A. 2008. Isolation and identification of eight microsatellite loci in the Cherokee darter (Etheostoma scotti) and their variability in other members of the genera Etheostoma, Ammocrypta, and Percina. Molecular Ecology Resources 8:149-151.
Gaggiotti, O.E. 1996. Population genetic models of source-sink metapopulations. Theoretical Population Biology 50:178-208.
Ginson, R., Walter, R.P., Mandrak, N.E., Beneteau, C.L. and Heath, D.D. 2015. Hierarchical analysis of genetic structure in the habitat-specialist Eastern Sand Darter (Ammocrypta pellucida). Ecology and Evolution 5:695-708.
Goldstein, D.B. and Schlotterer, C. 1999. Microsatellites: evolution and applications. Oxford university press. 352 pp.

Goudet, J. 2001. FSTAT: A program to estimate and test gene diversities and fixation indices (version 2.9.3). Available online: http://www2.unil.ch/popgen/softwares/fstat.htm Lausanne, Switzerland, Institute of Ecology.

Guenther, C.B. and Spacie, A. 2006. Changes in fish assemblage structure upstream of impoundments within the Upper Wabash River Basin, Indiana. Transactions of the American Fisheries Society 135:570-583.

Haag, W.R. 2012. North American freshwater mussels: Natural history, ecology, and conservation. 521 pp.

Haag, W.R. and Williams, J.D. 2014. Biodiversity on the brink: An assessment of conservation strategies for North American freshwater mussels. Hydrobiologia 735:45-60.

Hänfling, B. and Weetman, D. 2006. Concordant genetic estimators of migration reveal anthropogenically enhanced source-sink population structure in the river sculpin, Cottus gobio. Genetics 173:1487-1501.

Haponski, A.E., Bollin, T.L., Jedlicka, M.A. and Stepien, C.A. 2009. Landscape genetic patterns of the rainbow darter Etheostoma caeruleum: a catchment analysis of mitochondrial DNA sequences and nuclear microsatellites. Journal of Fish Biology 75:2244-2268.

Hardy, O.J., Charbonnel, N., Fréville, H. and Heuertz, M. 2003. Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. Genetics 163:1467-1482.

Hartl, D.L.a.C., A.G. 1997. Principles of population genetics. Sinauer associates Sunderland. 635 pp.

Hedrick, P.W. 2005. A standardized genetic differentiation measure. Evolution 59:1633-1638.

Honick, A.S., Zimmerman, B.J., Stauffer, J.R. Jr., Argent, D.G., and Porter, B.A. 2017. Expanded Distributions of Three *Etheostoma* Darters (subgenus *Nothonotus*) within the Upper Ohio River Watershed. Northeastern Naturalist. In Press.

Howell, L.A. 2007. Genetic population structure and breeding parameters of three Pennsylvania state threatened darter species: Etheostoma camurum, E. maculatum, and E. tippecanoe. Duquesne University, Pittsburgh, PA. 187 pp.

Hubisz, M.J., Falush, D., Stephens, M. and Pritchard, J.K. 2009. Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources 9:1322-1332.

Jakobsson, M. and Rosenberg, N.A. 2007. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801-1806.

Jelks, H.L., Walsh, S.J., Burkhead, N.M., Contreras-Balderas, S., Díaz-Pardo, E., Hendrickson, D.A., Lyons, J., Mandrak, N.E., McCormick, F., Nelson, J.S., Platania, S.P., Porter, B.A., Renaud, C.B., Schmitter-Soto, J.J., Taylor, E.B. and Warren Jr, M.L. 2008. Conservation status of imperiled north American freshwater and diadromous fishes. Fisheries 33:372-407.

Johnson, R.L., Mitchell, R.M. and Harp, G.L. 2006. Genetic variation and genetic structuring of a numerically declining species of darter, Etheostoma moorei Raney & Suttkus, endemic to the upper Little Red River, Arkansas. American Midland Naturalist 156:37-44.

Jost, L. 2008. G_{ST} and its relatives do not measure differentiation. Molecular Ecology 17:4015-4026.

Kalinowski, S.T. 2004. Counting alleles with rarefaction: private alleles and hierarchical sampling designs. Conservation Genetics 5:539-543.
Kapuscinski, K.L. and Farrell, J.M. 2014. Habitat factors influencing fish assemblages at muskellunge nursery sites. Journal of Great Lakes Research 40:135-147.
Kimura, M. and Weiss, G.H. 1964. The Stepping Stone Model of Population Structure and the Decrease of Genetic Correlation with Distance. Genetics 49:561-576.
Knouft, J.H., Caruso, N.M., Dupre, P.J., Anderson, K.R., Trumbo, D.R. and Puccinelli, J. 2011. Using fine-scale GIS data to assess the relationship between intra-annual environmental niche variability and population density in a local stream fish assemblage. Methods in Ecology and Evolution 2:303-311.
Koizumi, I., Yamamoto, S. and Maekawa, K. 2006. Decomposed pairwise regression analysis of genetic and geographic distances reveals a metapopulation structure of stream-dwelling Dolly Varden charr. Molecular Ecology 15:3175-3189.
Koryak, M., Bonislawsky, P.S., Locy, D.D. and Porter, B.A. 2009. Typical Channel Fish Assemblage of the Recovering Lower Allegheny River Navigation System, Pennsylvania, USA. Journal of Freshwater Ecology 24:509-517.
Koryak, M., Bonislawsky, P.S., Locy, D.D., and Porter, B.A. 2011. Gilt Darter (*Percina evides*: Percidae: Etheostomatinae) Range Expansion, Microhabitat Selection, and Phylogenetics Within The Allegheny River Navigation System, Pennyslvania, USA. Journal of the Pennsylvania Academy of Science 85:104-108.
Kuehne, R.A. and Barbour, R.W. 1983. The american darters. University Press of Kentucky.
Labay, A.A. and Brandt, T.M. 1994. Predation by Cyclops vernalis on Florida Largemouth Bass and Fountain Darter Larvae. The Progressive Fish-Culturist 56:37-39.
Lowe, W.H. and Allendorf, F.W. 2010. What can genetics tell us about population connectivity? Molecular Ecology 19:3038-3051.
Magoulas, A., Castilho, R., Caetano, S., Marcato, S. and Patarnello, T. 2006. Mitochondrial DNA reveals a mosaic pattern of phylogeographical structure in Atlantic and Mediterranean populations of anchovy (Engraulis encrasicolus). Molecular Phylogenetics and Evolution 39:734-746.
Maniatis, T., Fristch, E.F., and Sambrook, J. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. 545 pp.
Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. Cancer research 27:209-220.
Meirmans, P.G. and Hedrick, P.W. 2011. Assessing population structure: FST and related measures. Molecular Ecology Resources 11:5-18.
Narum, S.R. 2006. Beyond Bonferroni: Less conservative analyses for conservation genetics. Conservation Genetics 7:783-787.
Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences of the United States of America 70.
Nei, M. and Chesser, R.K. 1983. Estimation of fixation indices and gene diversities. Annals of Human Genetics 47:253-259.
O'Dee, S.H. and Watters, G.T. 2000. New or confirmed host identifications for ten freshwater mussels. Proceedings of the Conservation, Captive Care, and Propagation of Freshwater Musslels Symposium. Ohio Biological Survey. Columbus, Ohio. 77-82 pp.

Ohio Environmental Protection Agency. 2016. Ohio 2016 Integrated water quality monitoring and assessment report – Final Report. Available online at http://wwwapp.epa.state.oh.us/dsw/2016IR.pdf. Accessed 15 March 2017.
Orlóci, L. 1975. Multivariate analysis in vegetation research. The Hague. 274 pp.
Page, L.M. 1983. Handbook of Darters. TFH Publications, Neptune City, NJ.
Palsbøll, P.J., Bérubé, M. and Allendorf, F.W. 2007. Identification of management units using population genetic data. Trends in Ecology and Evolution 22:11-16.
Peakall, R. and Smouse, P.E. 2012. GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537-2539.
Peakall, R. and Smouse, P.E. 2006. GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6:288-295.
Piry, S., Luikart, G. and Cornuet, J.M. 1999. BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. Journal of Heredity 90:502-503.
Pompanon, F., Bonin, A., Bellemain, E. and Taberlet, P. 2005. Genotyping errors: causes, consequences and solutions. Nat Rev Genet 6:847-846.
Porter, B.A., Cavender, T.M. and Fuerst, P.A. 2002. Molecular phylogeny of the snubnose darters, subgenus Ulocentra (Genus Etheostoma, family Percidae). Molecular Phylogenetics and Evolution 22:364-374.
Pritchard, J.K., Stephens, M. and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.
Puechmaille, S.J. 2016. The program STRUCTURE does not reliably recover the correct population structure when sampling is uneven: sub-sampling and new estimators alleviate the problem. Molecular Ecology Resources:n/a-n/a.
Putman, A.I. and Carbone, I. 2014. Challenges in analysis and interpretation of microsatellite data for population genetic studies. Ecology and Evolution 4:4399-4428.
Raeymaekers, J.A.M., Maes, G.E., Geldof, S., Hontis, I., Nackaerts, K. and Volckaert, F.A.M. 2008. Modeling genetic connectivity in sticklebacks as a guideline for river restoration. Evolutionary Applications 1:475-488.
Rahel, F.J. and Stein, R.A. 1988. Complex predator-prey interactions and predator intimidation among crayfish, piscivorous fish, and small benthic fish. Oecologia 75:94-98.
Roberts, J.H., Angermeier, P.L. and Hallerman, E.M. 2013. Distance, dams and drift: What structures populations of an endangered, benthic stream fish? Freshwater Biology 58:2050-2064.
Robinson, J.D., Simmons, J.W., Williams, A.S. and Moyer, G.R. 2013. Population structure and genetic diversity in the endangered bluemask darter (Etheostoma akatulo). Conservation Genetics 14:79-92.
Roderick, G.K. and Navajas, M. 2003. Genes in new environments: genetics and evolution in biological control. Nat Rev Genet 4:889-899.
Rosenberg, N.A. 2004. DISTRUCT: A program for the graphical display of population structure. Molecular Ecology Notes 4:137-138.
Rousset, F. 2008. GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources 8:103-106.
Rousset, F. and Raymond, M. 1995. Testing heterozygote excess and deficiency. Genetics 140:1413-1419.

Santucci Jr, V.J., Gephard, S.R. and Pescitelli, S.M. 2005. Effects of multiple low-head dams on fish, macroinvertebrates, habitat, and water quality in the Fox River, Illinois. North American Journal of Fisheries Management 25:975-992.

Sasaki, M., Hammer, M.P., Unmack, P.J., Adams, M. and Beheregaray, L.B. 2016. Population genetics of a widely distributed small freshwater fish with varying conservation concerns: the southern purple-spotted gudgeon, Mogurnda adspersa. Conservation Genetics 17:875-889.

Saura, M. and Faria, R. 2011. Genetic tools for restoration of fish populations. Journal of Applied Ichthyology 27:5-15.

Scribner, K.T., Lowe, W.H., Landguth, E., Luikart, G., Infante, D.M., Whelan, G.E. and Muhlfeld, C.C. 2016. Applications of Genetic Data to Improve Management and Conservation of River Fishes and Their Habitats. Fisheries 41:174-188.

Sherwin, W.B., Jabot, F., Rush, R. and Rossetto, M. 2006. Measurement of biological information with applications from genes to landscapes. Molecular Ecology 15:2857-2869.

Shikano, T., Shimada, Y., Herczeg, G. and Merilä, J. 2010. History vs. habitat type: Explaining the genetic structure of European nine-spined stickleback (Pungitius pungitius) populations. Molecular Ecology 19:1147-1161.

Sia, E.A., Jinks-Robertson, S. and Petes, T.D. 1997. Genetic control of microsatellite stability. Mutation Research/DNA Repair 383:61-70.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457-462.

Smouse, P.E., Long, J.C. and Sokal, R.R. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. Systematic zoology 35:627-632.

Smouse, P.E., Whitehead, M.R. and Peakall, R. 2015. An informational diversity framework, illustrated with sexually deceptive orchids in early stages of speciation. Molecular Ecology Resources 15:1375-1384.

Soulsbury, C.D., Iossa, G., Edwards, K.J., Baker, P.J. and Harris, S. 2007. Allelic dropout from a high-quality DNA source. Conservation Genetics 8:733-738.

Stauffer, J.R., Boltz, J.M. and White, L.R. 1995. The Fishes of West Virginia. Proceedings of the Academy of Natural Sciences of Philadelphia 146:1-389.

Stauffer, J.R., Jr., R. W. Criswell, and D. P. Fischer. 2016. The Fishes of Pennsylvania. Cichlid Press, El Paso Texas. 556 pp.

Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., Waits, L.P. and Bouvet, J. 1996. Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic Acids Research 24:3189-3194.

Tagliavini, J., Conterio, F., Gandolf, G. and Fontana, F. 1999. Mitochondrial DNA sequences of six sturgeon species and phylogenetic relationships within Acipenseridae. Journal of Applied Ichthyology 15:17-22.

Tajima, F. 1996. Infinite-allele model and infinite-site model in population genetics. Journal of Genetics 75:27-31.

Tarr, J.A. 2002. The metabolism of the industrial city - The case of Pittsburgh. Journal of Urban History 28:511-545.

Thuillet, A.-C., Bru, D., David, J., Roumet, P., Santoni, S., Sourdille, P. and Bataillon, T. 2002. Direct Estimation of Mutation Rate for 10 Microsatellite Loci in Durum Wheat, Triticum turgidum (L.) Thell. ssp durum desf. Molecular Biology and Evolution 19:122-125.

Tonnis, B.D. 2006. Microsatellite DNA markers for the rainbow darter, Etheostoma caeruleum (Percidae), and their potential utility for other darter species. Molecular Ecology Notes 6:230-232.
Toro, M.A., Fernández, J. and Caballero, A. 2009. Molecular characterization of breeds and its use in conservation. Livestock Science 120:174-195.
Trautman, M.B. 1981. The fishes of Ohio: with illustrated keys. Ohio State University Press. 782 pp.
Turner, T.F. 2001. Comparative Study of Larval Transport and Gene Flow in Darters. Copeia 2001:766-774.
USACE - Army Corps of Engineers, U.S. – Pittsburgh District. 2003. Ohio River Navigation Charts – Pittsburgh, Pennsylvania to New Martinsville, West Virginia. Available online: http://www.lrp.usace.army.mil/Portals/72/docs/navigation/OhioRiverNavigationChart.pdf
USACE - Army Corps of Engineers, U.S. – Pittsburgh District. 2004. Allegheny River Navigation Charts – Pittsburgh, Pennsylvania to East Brady, Pennsylvania. Available online:http://www.lrp.usace.army.mil/Portals/72/docs/navigation/AlleghenyRiverNavigationChart.pdf
Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M. and Shipley, P. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4:535-538.
Wan, Q.H., Wu, H., Fujihara, T. and Fang, S.G. 2004. Which genetic marker for which conservation genetics issue? Electrophoresis 25:2165-2176.
Weir, B.S. and Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38:1358-1370.
Wenger, S.J., Freeman, M.C., Fowler, L.A., Freeman, B.J. and Peterson, J.T. 2010. Conservation planning for imperiled aquatic species in an urbanizing environment. Landscape and Urban Planning 97:11-21.
Whitlock, M.C. 2011. G′ST and D do not replace FST. Molecular Ecology 20:1083-1091.
Whitlock, M.C. and McCauley, D.E. 1999. Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm + 1)$. Heredity 82:117-125.
Wright, S. 1943. Isolation by distance. Genetics 28:114.
Yeardley, R.B. 2000. Use of small forage fish for regional streams wildlife risk assessment: Relative bioaccumulation of contaminants. Environmental Monitoring and Assessment 65:559-585.
Yoder, C.O., Rankin, E.T., Smith, M.A., Alsdorf, B.C., Altfater, D.J., Boucher, C.E., Miltner, R.J., Mishne, D.E., Sanders, R.E. and Thoma, R.F. 2005. Changes in fish assemblage status in Ohio's non-wadeable rivers and streams over two decades. Pp. 399, In Editor (Ed.)^(Eds.). Book Changes in fish assemblage status in Ohio's non-wadeable rivers and streams over two decades. Vol. 45. American Fisheries Society

CHAPTER THREE

Development of an aquatic environmental DNA (eDNA) method for detection of darters (subgenus *Nothonotus*) in the upper Ohio River watershed

Conservation management of imperiled fishes is difficult due to temporal and spatial variation in aquatic environments and the challenges inherently associated with monitoring species low in abundance. For effective management, data is gathered on species abundance and distribution using traditional techniques such as seining, boat and backpack electrofishing, and benthic or electrified – benthic trawling in non-wadeable riverine environments. In Pennsylvania, three species of darter (subgenus *Nothonotus*), Bluebreast Darter (*Etheostoma camurum*), Spotted darter (*Etheostoma maculatum*), and Tippecanoe darter (*Etheostoma tippecanoe*) were imperiled and listed as either threatened or endangered. Recent data showed that these species had expanded distributions throughout the non-wadeable portions the Allegheny and Ohio rivers and were de-listed in 2014. As a result, it's crucial that their recovery is monitored for future assessment of imperilment. In Pennsylvania, the current State – mandated method for sampling non-wadeable rivers is electrified – benthic trawling which effectively surveys deeper aquatic habitats, but it is costly, time and labor intensive, and often harms small fish like darters because of abrasion that occurs from sediment and debris entering the trawl. To address this problem, a protocol was developed using environmental DNA (eDNA) from water samples to non-invasively sample for the three focal species. Water samples were collected prior to performing traditional sampling below eleven navigational lock and dams and nine adjacent tributaries. eDNA extraction protocols were evaluated and polymerase chain reaction (PCR) primers were developed and tested on known tissue samples for species specificity. A novel detection mechanism utilizing fragment analysis was tested and compared to agarose gel – electrophoresis with ethidium bromide visual detection. One primer set, under laboratory conditions, could accurately detect all three focal species. A second primer set was specific to *E. tippecanoe* detection, and was tested with eDNA water samples from four locations where traditional sampling verified their presence. Using fragment analysis to visualize eDNA results, *E. tippecanoe* was detected at all four locations, but gel – electrophoresis visualization could not confirm *E. tippecanoe* at any location. In this study, eDNA methodologies were developed for monitoring *E. tippecanoe* with DNA extracted from water samples, and a novel detection mechanism using fragment analysis was developed for visualization of eDNA results. These newly developed tools will benefit the conservation management of imperiled *Nothonotus* darters in the upper Ohio River watershed. Additionally, this general approach may be modified for simultaneous detection of multiple species from eDNA with multiplex PCR that may have the ability to provide data at the community level.

3.1 - Introduction

Background

In the upper Ohio River watershed three species of small – bodied benthic fish, *Etheostoma camurum* (Bluebreast Darter), *Etheostoma maculatum* (Spotted Darter), and *Etheostoma tippecanoe* (Tippecanoe Darter), were recently classified as either threatened or endangered by the Ohio Division of Wildlife (Ohio Department of Natural Resources – Division of Wildlife 2015; 15 Ohio Rev. Code § 1531.25 - 2015) and the Pennsylvania Fish and Boat Commission (Pennsylvania Bulletin 1999). Extensive fish surveys in Ohio and Pennsylvania have documented increases in these species' distributions (Honick et al. 2017) which have led to changes in their conservation status. In 2012, Ohio delisted *E. camurum*, but kept *E. tippecanoe* and *E. maculatum* listed as threatened and endangered, respectively (Ohio Department of Natural Resources – Division of Wildlife 2015; 15 Ohio Rev. Code § 1531.25 - 2015). In Pennsylvania, all three species have been delisted since 2014 (Pennsylvania Bulletin 2014). However, in order to properly manage their return, it's important to continually track distribution changes for future assessments of imperilment. In general, the conservation management of fish species can be challenging due to the spatial and temporal variation of aquatic ecosystems (Knouft et al. 2011). These three focal species present an additional challenge because surveys by Yoder et al. (2005), Freedman (2009b), Argent and Kimmel (2010, 2014), and Honick et al. (2017) show that these species are utilizing deeper, non-wadeable habitat within the mainstem rivers. Traditional sampling techniques (defined as backpack electrofishing, boat electrofishing, and seining) lose effectiveness when depths increase to greater than 2.0 m. As an alternative, benthic trawling and

electrified-benthic trawling utilizing a modified mini-Missouri trawl (Freedman et al. 2009a, Herzog et al. 2005, Herzog et al. 2009) have been implemented to survey benthic fish in these deeper habitats. In Pennsylvania, electrified-benthic trawling is currently the State-mandated method for surveying non-wadeable rivers, and even though this method has shown to be effective (Freedman et al. 2009a,b; Lorson 2010) there are drawbacks especially when surveying for imperiled species. Trawling is costly and labor intensive. A demonstration of the financial burden comes from Honick et al. (2017) where it cost approximately $5,000 and labor from three volunteers to perform electrified - benthic trawling surveys at only ten locations in the upper Ohio River watershed. Another drawback, which can be particularly problematic when surveying for imperiled species in large river surveys, is that benthic trawling and electrified-benthic trawling often harms the fish by abrasion from sediment and debris entering the trawl.

Previous eDNA Methodologies

Recently, a new, non-invasive molecular approach utilizing environmental DNA (eDNA) has been developed for species detection from water samples. eDNA can be defined as DNA that is extracted from environmental samples (e.g. water, air, soil, sediment) and is isolated before physical/visual detection of the target species (Rees et al. 2014, Taberlet et al. 2012). There has been a large expansion in the application of species detection with eDNA techniques since 2011 (scientific articles published with eDNA as the keyword: 2011 – 64, 2016 – 187). Conservation management of rare and invasive species is challenging because population assessments rely on visual/physical detection which is inherently problematic for species in low densities (Jerde et al. 2011, Rees et al. 2014). With the aforementioned challenges of surveying aquatic ecosystems,

eDNA has advantages because 1) aquatic environments keep DNA suspended in the water column and therefore, rare species can be detected without ever being seen (Jane et al. 2014, Jerde et al. 2011), 2) target species don't need to be removed from their environment to take a sample, 3) DNA can be used for accurate species-level identification, 4) enhanced species detection sensitivity, and 5) the potential to decrease sampling costs (Evans et al. 2017, Jane et al. 2014, Turner et al. 2014b). eDNA methodologies have been applied to a wide range of environments and species (this is only a partial list for demonstrating the numerous applications) including the invasive Asian Carp (Bighead Carp - *Hypophthalmichthys nobilis*, Silver Carp - *Hypophthalmichthys molitrix*) (Jerde et al. 2011, Jerde et al. 2013), Brook and Bull Trout (*Salvelinus fontinalis* and *Salvelinus confluentus*, respectively) (Baldigo et al. 2017, Wilcox et al. 2013), the Slackwater Darter (*Etheostoma boschungi*) (Janosik and Johnston 2015), Rocky Mountain Tailed Frogs (*Ascaphus montanus*) (Goldberg et al. 2011), the Trinidad Golden Tree Frog (*Phyllodytes auratus*) (Brozio et al. 2017), the Idaho Giant Salamander (*Dicamptodon aterrimus*) (Goldberg et al. 2011), Eastern Hellbenders (*Cryptobranchus alleganiensis*) (Olson et al. 2012), Unionid mussels (Cho et al. 2016), and even an invasive freshwater diatom (Keller et al. 2017). The predominant sources of fish eDNA are derived from their urine, feces, and sloughing epidermal cells (Rees et al. 2014, Thomsen et al. 2012b).

When fisheries managers first started using eDNA for fish species detection the most common approach was to design species-specific primers targeting one species at a time, amplify the eDNA, and then confirm detection with electrophoresis and agarose gel-based visualization (Janosik and Johnston 2015, Jerde et al. 2011). Due to eDNA methodologies being new to conservation management, published protocols vary quite drastically. Water sample collection methods and volumes vary by application and have ranged from collecting 15.0 mL grab

samples (Ficetola et al. 2008) to taking peristaltic pumps on site and collecting 6.0 to 12.0 L of water or until the filter clogs (Turner et al. 2014a, Wilcox et al. 2013). The generally accepted protocol has been to collect multiple 2.0 L grab samples (before any traditional sampling takes place) resulting in total water volumes of approximately 10.0 to 14.0 L of water per sample site (Jerde et al. 2011, Mahon et al. 2013, Minamoto et al. 2012). Another area that has been debated is filter pore size and the filter material which provides the highest capture efficiency of eDNA from water. Early experiments used 0.45 µm cellulose-nitrate filters (Goldberg et al. 2011, Pilliod et al. 2014) or 1.5 µm glass fiber filters (Jerde et al. 2013, Jerde et al. 2011, Mahon et al. 2013). However, Minamoto et al. (2012) used 3.0 µm isopure polycarbonate filters and Kelly et al. (2014) used a durapore membrane filter with a pore size of 0.22 µm. Turner et al. (2014a) performed an extensive examination on the particle size distribution of eDNA for the Common Carp (*Cyprinus carpio*) and used sequential pore size filtration to determine which pore size captured the most eDNA. Interestingly, they selected a combination of filter materials with the large pore sizes being nylon net filters and the smaller pores sized filters (20.0, 10.0, 1.0 and 0.22 µm) made from polycarbonate. They showed that 0.22 µm pore sized filters captured the largest amount of aqueous eDNA, but it was suggested that to overcome the logistics of filter clogging that layered sequential filtering was recommended. However, it has been implicated that capture efficiency will also vary depending on the species being sought and characteristics of the water quality (e.g. low turbidity vs. high turbidity) (Rees et al. 2014, Turner et al. 2014a, Turner et al. 2014b). There are applications where eDNA was not captured via filtration, but by precipitating all DNA from the water sample (Ficetola et al. 2008, Goldberg et al. 2011). This method is cited less often probably due to the logistics of precipitating DNA from large volumes of water.

Another variable attribute of eDNA methodology has been extraction of eDNA from the

filter. Once eDNA is captured on the filter media, various methods have been used to extract the DNA. Methods have ranged from extraction with commercial kits such as the DNeasy Blood and Tissue or MoBio PowerWater kits (Jerde et al. 2011, Minamoto et al. 2012, Takahara et al. 2012) to a modified phenol:chloroform extraction (Deiner and Altermatt 2014). Comparison tests done by Deiner et al. (2015) between precipitation methods, filtration methods, and three commercial extraction kits showed that combinations of the methods yielded different results and that final protocols need to be carefully selected based on the species being targeted. Eichmiller et al. (2015) used Common Carp eDNA to compare the affects of filter media, pore size, and six types of commercially available DNA extraction kits on the optimal conditions for detection and quantification of aquatic eDNA. They found that certain kits extracted higher amounts of eDNA (DNeasy Blood and Tissue Kit), but polymerase chain reaction (PCR) was negatively impacted by high levels of inhibitors. In contrast, the kits that yielded lower amounts of DNA (MoBio PowerWater) resulted in more consistent PCR that would be ideal for making comparison across a wide range of aquatic environments. Eichmiller et al. (2015) showed that a 1.5 μm glass fiber filter with the MP Biomedicals FastDNA Spin Kit extracted a relatively large amount of eDNA with fewer instances of PCR inhibition and thus, consistent amplification. Therefore, this kit was recommended for presence/absence detection objectives because of the balance between DNA yield and consistent levels of successful PCR.

Quantifying eDNA Detection Limits

Another hurdle that eDNA methodologies encounter is the determination and reporting the limits of detection (LOD). Currently, the most common techniques for eDNA detection are

quantitative PCR (Hunter 2017, Rees 2014) and newly developed next generation sequencing (Jerde 2013, Rees 2014). These methods may be robust for eDNA detection, but problems exist with reporting their limits of detection. The limits of detection for eDNA assays have been shown to vary across species which may be influenced by the specificity of the PCR primers (which influences false positives and false negatives), the method of detection (e.g. qPCR vs. digital PCR vs. PCR and gel electrophoresis), and how the researcher chooses to determine and calculate the LOD (Hunter et al. 2017). According to Hunter et al. (2017) a common method to determine eDNA LOD using qPCR assays has been adaptation of the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE). The problem with MIQE guided detection levels is that interpretation of the guidelines is subject to bias because the definition only specifies that the LOD is derived from the lowest concentration in which 95% of the samples have positive detection (Bustin et al. 2009). Across eDNA experiments, Hunter et al. (2017) claims that the guidelines are too stringent because whether you perform three or six replicates the power of detection likely won't increase because target eDNA is expected to be in such low quantities that stochastic variation from sub-sampling an extraction may result in replicates of non-detection. Lacking a standard for reporting LODs from qPCR has resulted in variable reporting levels such as Turner et al. (2014b) who reported an LOD of 10 copies with three replicates and Takahara et al. (2012) who defined their LOD as detecting one copy in two out of three replicates. LOD are also poorly defined with other detection methods. For example, detection limits from traditional PCR detection with agarose gel-base visualization was reported by Jerde et al. (2011) as one positive PCR reaction per sample site that was standardized to catch-per-unit effort while Janosik and Johnston (2015) simply reported the percentage of

positive detections from all three PCR replicates vs. the number of sites where only one or two of the replicates showed positive detection.

Objectives and Hypothesis

Regardless of these challenges, eDNA detection methods have been successfully demonstrated across multiple species and habitats. In this study, my original goals were to develop an eDNA assay for the detection of all three focal species. Specifically, my goals were to develop an eDNA assay with a multi-plex PCR to simultaneously amplify eDNA from all three species and to develop a detection assay using fragment analysis which I hypothesized should be more sensitive than electrophoresis and agarose gel-based visualization assays. It was also surmised that because genetic analyzers such as the ABI 3130 have been widely used in genetic research that many research and academic institutions already have the technology in place which should reduce the cost and implementation time of employing a new technique utilizing eDNA.

3.2 - Methods

Study Area and Sample Collection

Over the summers of 2013 and 2014, non-wadeable habitats in the tailwaters of eleven L/D installations on four river systems (Allegheny, Beaver, Monongahela, and Ohio) were surveyed using a modified Missouri trawl (2.4 m x 1.2 m, 3.2 mm mesh) electrified with a

Smith-Root VI-A electrofisher and a 5000 W generator (see Chapter 1 for details). A transect within the tailrace of the L/D (50 to 150 m below the installation) was established as the starting point for seven trawls – one trawl each within 10 m of the left and right descending bank, one trawl at center channel, and the four remaining trawls were evenly spaced between center channel and the descending bank. Each trawl consisted of two minutes of sampling effort. All fish species were identified and enumerated (Appendix A – Table A.6, A.7). Before trawling began, seven eDNA water samples were collected along the transect following the same spacing pattern. At each location, a 2.0 L water sample was collected by lowering a horizontal Kemmerer bottle to the bottom of the river and upon detecting river bottom the messenger was released forcing the endcaps to close. Samples were transferred to 1.0 L leak-proof polypropylene bottles and stored on ice (a total of 14.0 L of water collected per site).

In addition to the electrified-benthic trawling, nine tributaries were electrofished (Smith Root LR-24, backpack electrofisher – single pass) for 100 m starting at the first riffle upstream of the confluence with the main river. Streams with moderate to high flow were sampled by electrofishing into a blocking seine (2.4 m x 1.8 m, 3.2 mm mesh). All fish species were identified and enumerated (Appendix A – Table A.8, A.9). Before electrofishing the stream, six 1.0 L grab samples were collected (1.0 L polypropylene bottles) at the electrofishing starting location and stored on ice (a total of 6.0 L of water per tributary sample site). For all sample sites, two 1.0 L bottles were filled with Millipore water at the laboratory. One bottle was taken into the field as a field control and the other remained at the lab for a lab/filtering control. In addition, at all sample locations, water quality parameters were recorded with a calibrated YSI multi-parameter sonde (pH, dissolved oxygen, conductivity) and turbidity was taken with a Hach

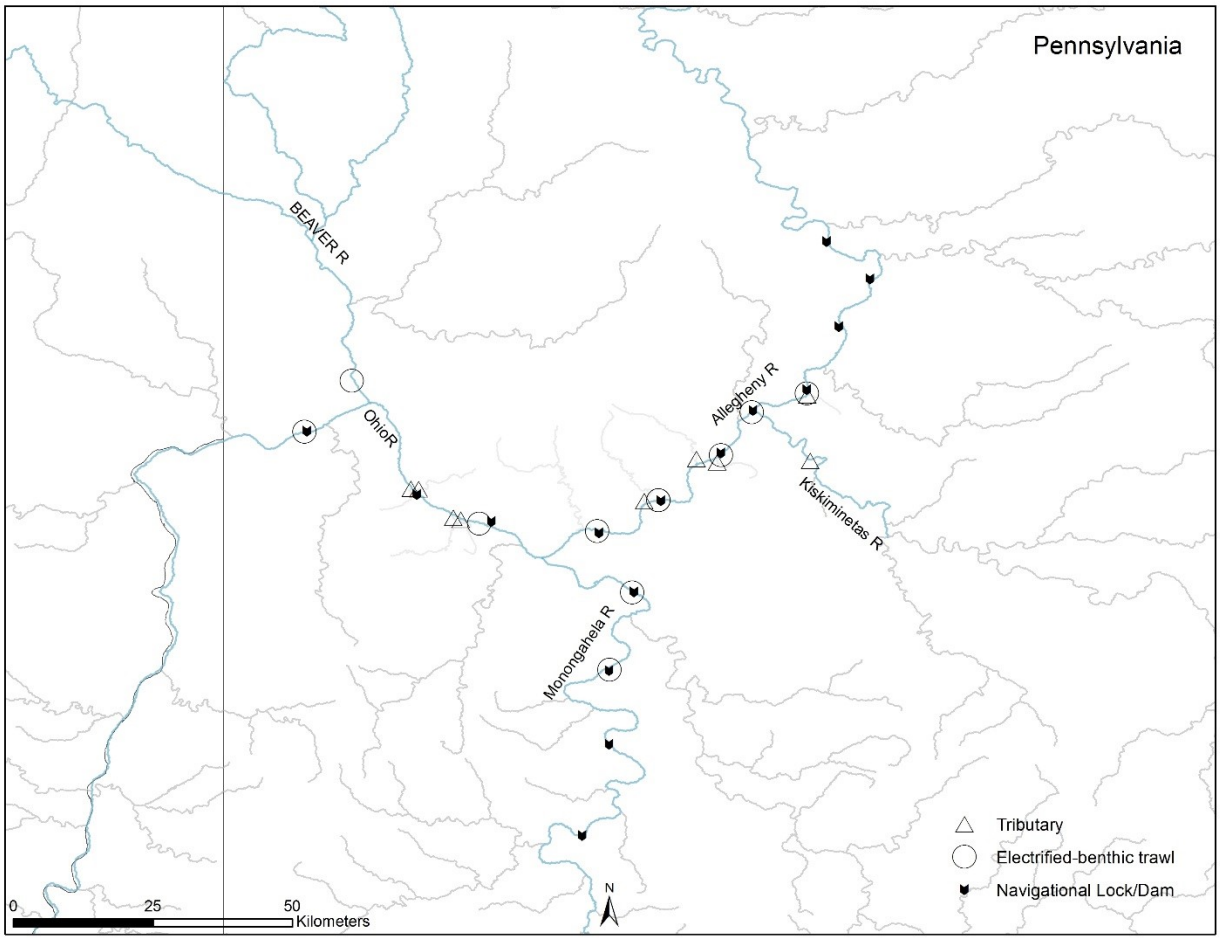

Figure 3.1. Map illustrating the locations of the sample sites. Water samples were collected from nine tributaries (triangles) that were adjacent to the lock and dam collection sites (open circles). Six liters of water were collected from each tributary and 14.0 L were collected at each lock and dam. All water samples were collected prior to performing a traditional fish survey at each location.

2100P Turbidimeter and recorded in Nephelometric Turbidity Units (NTU – a standard measure of the degree to which incident light is scattered and then measured at a 90 ° angle, (U.S.EPA Method 180.1)). See Figure 3.1 for sample locations. All equipment was sterilized with 20 % bleach between sample sites.

Water Filtration

eDNA water samples were processed within 24 h of collection. All samples were filtered at Duquesne University's forensics laboratory inside of a UV sterilized biological hood within a dedicated pre-PCR room. Water filtration pre-testing indicated that filters with 0.2 µm and 0.45 µm pores clogged after filtering less than 500 mL of water, which logistically, would increase sample processing costs and increase sample processing times substantially. The eDNA scientific literature was consulted to find the best filter pore size that would accommodate these conditions but still retain eDNA. The water samples (2.0 L per filter) were vacuum filtered using 0.7 µm sterile glass-fiber filters. Turbid samples (> 8.0 NTU) were filtered by stacking a 1.5 µm glass-fiber filter on top off the 0.7 µm filter to decrease pore clogging. Filters were placed into individual storage bags and stored at -80 ºC until DNA extraction.

DNA Extraction

Two different DNA extraction kits were selected to compare eDNA extraction efficiency – MoBio PowerWater DNA Isolation Kit and Qiagen DNeasy Blood and Tissue Kit. MoBio PowerWater extractions followed the manufacturer's instructions but with the following modifications to accommodate extraction from glass fiber filters: 1) extra filter material around the outside edge was removed to reduce bulk, 2) vortexing time during the lysis step was reduced to 2.5 minutes, and 3) after lysis - contents were placed into a 5.0 mL syringe which was placed within a 15.0 mL conical tube for centrifugation – this allowed for better recovery of the lysate. DNeasy Blood and Tissue extractions were completed following a protocol developed by the

U.S. Fish and Wildlife Service (Appendix B – Figure B.1), but with the following changes: 1) after lysis - contents were transferred to a 5.0 mL syringe which was placed within a 15.0 mL conical tube for centrifugation and lysate recovery, 2) buffer AE was warmed to 55 °C, 3) each spin column was eluted 2x with 100 µL of buffer AE, and 4) the contents from the separate elutions were combined by drying off the contents of one tube at 65 °C and then using the 100 µL from the other elution for resuspension.

DNA extraction kit yields were compared by using the replicate water samples that were collected at Allegheny River L/D #6 and Dashields L/D. From each location, six 2.0 L water samples were filtered with 1.5 µm filters stacked on top of 0.7 µm filters to prevent filter clogging. The filters were extracted and processed separately and evenly between the kits. Extracted DNA was quantified using a Qubit Fluorometer (2.0) and a dsDNA broad range assay kit. Samples were prepared per manufacturers' specifications. DNA concentrations were averaged and standard deviations calculated.

Primer Design

Primer design was completed following two different protocols because the first round was screened and found to be non-specific and insufficient for eDNA application. Regardless, methods will be given for each approach. eDNA in aquatic systems consists of a mixture of DNA from multiple organisms and due to environmental degradation likely persists as small fragments (Taberlet et al. 2012, Wilcox et al. 2013). Therefore, to increase the chances of finding target species eDNA, primers were designed for amplification of a DNA fragment (≤150 base pair (bp)) from the multi-copy mitochondrial gene *NADH dehydrogenase 2* (*ND2*). This

selection was based on previous work by Ritchea (2006) who showed that in nine species of darters, *ND2* was the most variable gene across the entire mitochondrial genome. *ND2* sequences for the three focal species and two positive control species (Rainbow Darter – *E. caeruleum* and Channel Darter – *P. copelandi*) were downloaded from GenBank (Clark et al. 2016) and aligned using DNASTAR- MEGALIGN software (v. 8.1.3). In order to design unique forward primers and a common reverse primer for a multi-plex PCR, SPecies IDentity and Evolution in R (SPIDER, v. 1.1-0) was used to perform a sliding window analysis of *ND2* to locate the most diagnostic portion of the gene. Analysis was completed with a 100 bp sliding window in steps of 1 bp. The most diagnostic region was identified between 550 – 750 bp. Primer locations were manually plotted within the identified region. Integrated DNA Technologies (IDT) web-based analysis software, OligoAnalyzer (v. 3.1) (https://www.idtdna.com/calc/analyzer), was used to test the primer pairs for thermodynamic compatibility: melting temperature mismatch between forward and reverse (Tm), hairpin formation, and self-dimerization. CLUSTAL Omega (v. 1.2.0), was used to make a visual representation of primer locations (Fig. 2.3). All forward primers were initially ordered without fluorescent tags to develop PCR and thermocycle protocols.

A different approach was taken regarding eDNA primer re-design. The search for species – specific primers was extended to two additional mitochondrial genes: *cytochrome c oxidase subunit 1* (*COI*), and *cytochrome b* (*Cytb*). Primers were developed using a web-based design tool: Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). Primer-BLAST software was developed to combine the ability to design species-specific primers and to perform thermodynamic analysis, but with the added function of utilizing the Basic Local Alignment Sequence Tool (BLAST) to check for non-target amplification (Ye et al. 2012). Default software

settings were followed except for the following conditions: 1) target amplicon size – 100 to 250 bp and 2) created a customized database for cross-amplification detection. The customized cross-amplification database was constructed by downloading available darter sequences for *ND2*, *COI* and *Cytb* genes from the National Center for Biotechnology Information (NCBI) Nucleotide database. Sequences were then restricted to all darter species that were documented within the upper Ohio, Allegheny, and Monongahela Rivers according to the Pennsylvania Fish and Boat Commission (Steiner 2000) (Table 3.1, see Appendix A – Table A.10 for accession numbers and details of the customized cross-amplification database). Primer-BLAST generated five primer pairs which were ranked according to the likelihood of cross-amplification and meeting the pre-determined requirements (e.g. amplicon size, # of mismatches tolerated, likely hairpins, etc.). Any primer pairs identified as having the potential to cross-amplify a non-target species were removed. The top two ranking primer pairs for each species were selected for specificity screening (Table 3.2). Because it is crucial to confirm that non – detection from eDNA is not due to PCR failure, primers were also designed for two common darter species that are likely to be at any sampling location: *E. caeruleum* (for stream sample sites) and *P. copelandi* (for non-wadeable river sample sites). For future reference in primer design, nucleotide diversity tests were calculated between each species for *ND2* and *COI* using the online software DnaSP (v. 5.10.01) (Librado and Rozas 2009). Calculations determined the average number of nucleotide differences per site between two sequences.

DNA Amplification and Specificity Testing

A high-fidelity polymerase, Phusion DNA polymerase, was selected for PCR because it has an error rate 50-fold lower than *Taq* polymerase and 3′ to 5′ exonuclease activity. Specificity testing was completed with DNA extracted from fin clips or muscle tissue using a standard phenol:chloroform extraction (Maniatis 1982). Primers were tested individually with PCR following Phusion guidelines in a 20.0 µL reaction (final concentration): Phusion HF buffer (1X), dNTPS (0.4 mM), forward and reverse primer (0.5 µM each), Phusion polymerase (0.2 units/µL), and 25 ng of DNA. A gradient thermocycle (60 °C to 72 °C, Bio-Rad C-1000 Touch Thermocycler) was used to determine optimal annealing temperatures with a magnesium chloride concentration of 1.5mM at the final volume of 20.0 µL. See Appendix B – Table B.2-B.4 for detailed PCR and thermocycle conditions. Each primer set was screened for cross amplification in 13 species of darters (Table 3.1) and two very common species likely to be found within the region, the Smallmouth Bass (*Micropterus dolomieu*) and the Mottled Sculpin (*Cottus bairdii*). The final stage of specificity testing was completed with PCR reactions that included the target species and any other species that showed possible amplification in fragment analysis. A competitive screening test was conducted where target DNA was reduced incrementally in separate reactions down to 0.025 ng per 20.0 µL reaction. A total of 25 ng of DNA per reaction was maintained with the remaining amount split evenly amongst the species being screened against. Results of the PCR were visualized with gel electrophoresis and fragment analysis methods (see below). The relative fluorescence units (RFUs) from fragment analysis (read in Peak Scanner) were documented for the three focal species using 25 ng tissue-extracted DNA per reaction down to 0.025 ng DNA per reaction. The idea was to develop a ratio

Table 3.1. List of species used to screen eDNA primers for specificity *in silico* and for experiments utilizing tissue-derived DNA extracted from positively identified specimens. All of these species have been documented in the upper Ohio River watershed by the Pennsylvania Fish and Boat Commission.

Genus species	Common Name
Etheostoma blennioides	Greenside Darter
Etheostoma caeruleum	Rainbow Darter
Etheostoma camurum	Bluebreast Darter
Etheostoma flabellare	Fantail Darter
Etheostoma maculatum	Spotted Darter
Etheostoma nigrum	Johnny Darter
Etheostoma tippecanoe	Tippecanoe Darter
Etheostoma variatum	Variegate Darter
Etheostoma zonale	Banded Darter
Percina caprodes	Log Perch
Percina copelandi	Channel Darter
Percina evides	Gilt Darter
Percina macrocephala	Longhead Darter
Percina shumardi	River Darter
Micropterus dolomieu	Smallmouth Bass
Cottus bairdii	Mottled Sculpin

of peak intensity (at the target amplicon size) in the 25 ng DNA positive control to look for a trend under decreasing target DNA in the presence of non – target species.

PCR of eDNA Samples

PCR amplification of eDNA samples was performed following the optimal PCR conditions but with different total DNA concentrations per reaction: 50 ng, 20 ng, 10 ng, 5 ng, and two dilutions, 1:5 and 1:50, of the original extract regardless of DNA concentration. eDNA amplification pre-testing and Eichmiller et al. (2015) showed that extracted eDNA samples may contain variable levels of potential PCR inhibitors. Therefore, if a sample contained a high level

of inhibitors, having less of the DNA extract per reaction may dilute the inhibitors below a threshold allowing PCR to work.

eDNA Detection

One of my main goals of the eDNA method development was to test eDNA detection using agarose gel - based visualization compared to fragment analysis. Gel electrophoresis of the PCR product was run with 5.0 µL of product mixed with 2.0 µL of loading dye on a 2 % agarose gel infused with 1 % ethidium bromide and sized against a 100 bp DNA ladder. Products were visualized under a UV light. Each forward primer was labeled with a species - specific fluorophore color (*E. camurum*: PET, *E. maculatum*: 6-FAM, and *E. tippecanoe*: VIC) for fragment analysis on an ABI 3130 genetic analyzer. PCR product was prepared according to the GeneScan™ - 500 Liz™ size standard protocol: 0.5 µL PCR product + 0.25 µL size standard + 9.25 µL of Hi-Di™ formamide. In order to increase the sensitivity of detection from a potentially weak PCR signal, a customized analysis module was created for the ABI 3130 genetic analyzer. Attempting to increase the fluorescent signal by adding more PCR product during fragment analysis sample preparation resulted in sizing abnormalities from disruption of the sample : size standard : formamide ratio (Applied Biosystems 2000). Therefore, the sample injection time of the new module was increased from 12 sec to 24 sec. Fragment analysis results were visualized using Peak Scanner (v 2.0).

3.3 - Results

A total of 1,974 fish were collected from electrified – benthic trawling below eleven L/D installations and 2,183 fish were collected from backpack electrofishing nine tributaries (Appendix A - Tables A.6 – A.9). A total of 44 species were represented which included 13 darter species.

The comparison of the DNA extraction kits showed a substantial difference between the quantities of total DNA extracted. Using the MoBio kit, the average amount of DNA extracted from the 0.7 µm and 1.5 µm pore – sized filters was 3.9 ng/µL and 3.1 ng/µL, respectively. Using the Qiagen DNeasy kit, the average amount of DNA extracted from the 0.7 µm and 1.5 µm pore – sized filters was 12.4 ng/µL and 23.4 ng/µL, respectively (Fig. 3.2 and 3.3).
The primers developed in the first phase of the project were not species – specific and were found to cross - amplify with non-target species and among the three focal species. The primers did not perform robustly under ideal laboratory conditions. Primer re-design using Primer-BLAST resulted in ten primer pairs that were screened for species – specificity (Table 3.2). After screening, one primer set for each species was selected for eDNA field testing. Primers for *E. camurum* (Ecam-ND2-P2) and *E. maculatum* (Emac-ND2-P2) were located within *ND2* with target amplicon sizes of 159 bp and 140 bp, respectively (Table 3.2, 3.3). The primer set for *E. tippecanoe* (Etip-COI-P2) was located within *COI* with a target amplicon of 231 bp (Table 3.2, 3.4). Primers for *E. caeruleum* (Ecaer-COI-P1) were in *COI* with an amplicon size of 234 bp, and primers for *P. copelandi* (Pcope-ND2-P1) were in *ND2* with an amplicon size of 241bp. Nucleotide diversity tests, θ (π), showed very low sequence diversity between the three focal species for *ND2* and *COI*. Comparisons in *ND2*

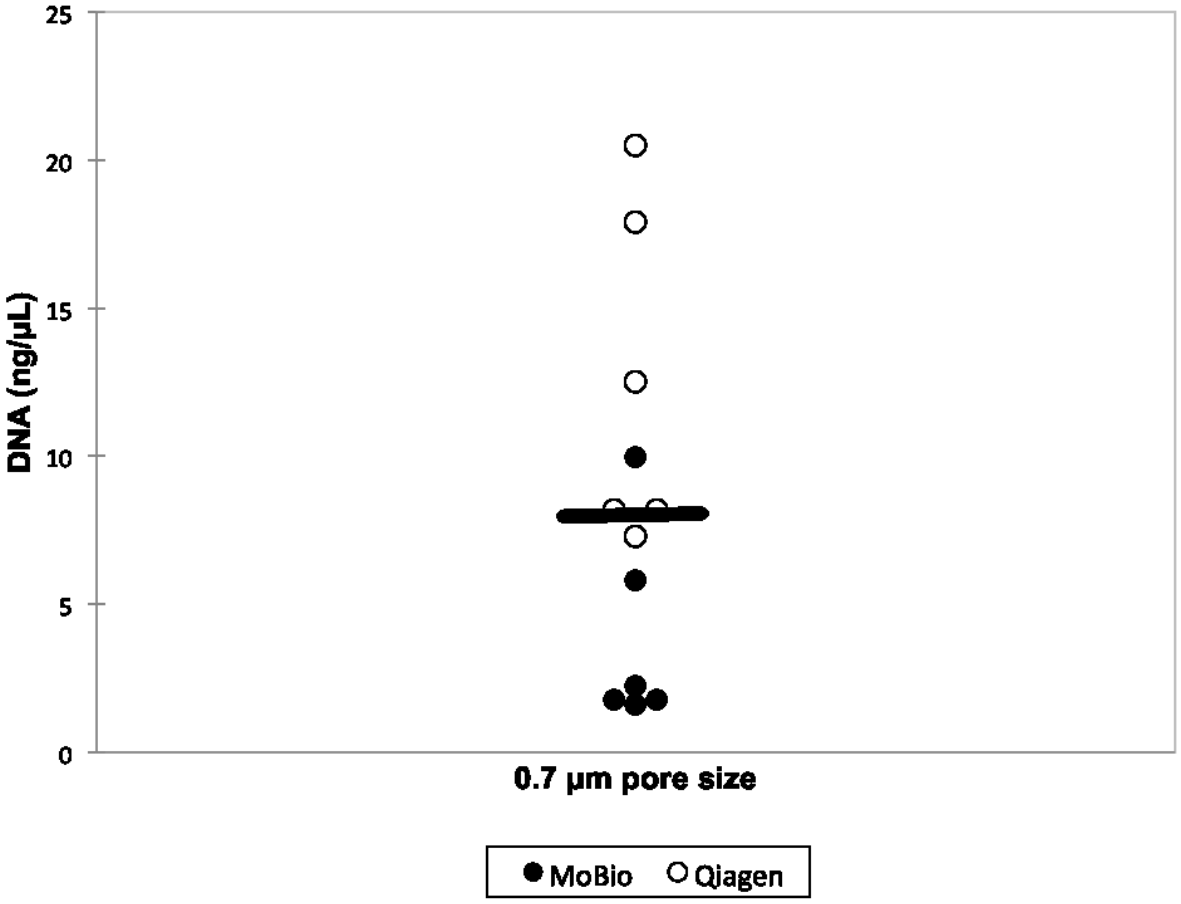

Figure 3.2. Beeswarm plot of the quantity of eDNA extracted from the 0.7 µm glass-fiber filters using the different extraction kits. The black line represents the mean across all extractions. The MoBio kit consistently extracted less eDNA than the Qiagen kit.

between *E. camurum-E. maculatum*, *E. camurum-E. tippecanoe*, and *E. maculatum-E. tippecanoe* were $\theta = 0.035$, 0.071, and 0.069, respectively. Comparisons in *COI* between *E. camurum-E. maculatum*, *E. camurum-E. tippecanoe*, and *E. maculatum-E. tippecanoe* were $\theta =$ 0.019, 0.043, and 0.054, respectively.

Primers for *E. caeruleum* were species – specific, but primers for *P. copelandi* weakly amplified *E. flabellare*, *E. nigrum*, *E. zonale*, and *P. caprodes*. However, multiple PCR screenings using different *P. copelandi* individuals confirmed that *P. copelandi* could be distinguished by the presence of a secondary diagnostic band at 219 bp. *E. maculatum* primers very weakly amplified *E. variatum* and *P. evides*, and fragment analysis showed consistent but

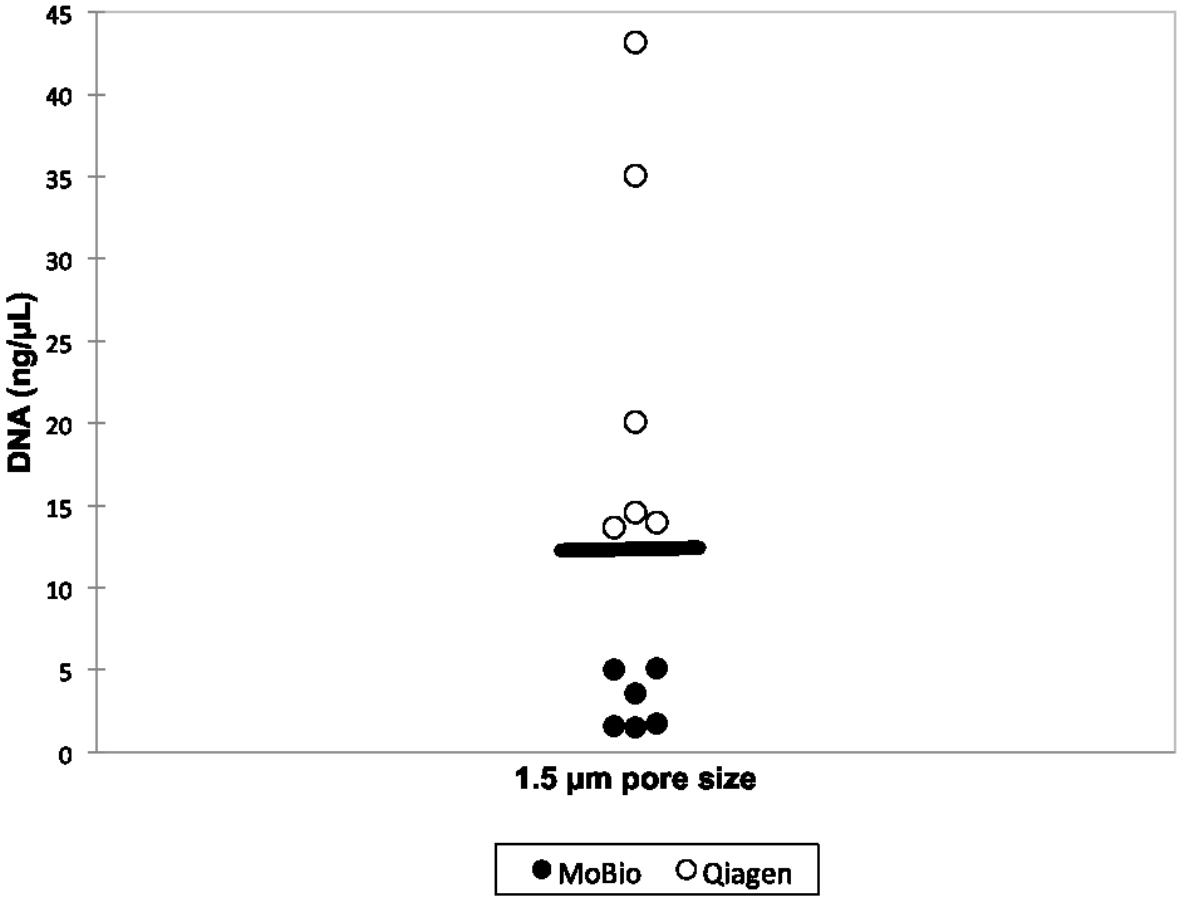

Figure 3.3. Beeswarm plot of the quantity of eDNA extracted from the 1.5 µm glass-fiber filters using the different extraction kits. The black line represents the mean across all extractions. The MoBio kit consistently extracted less eDNA than the Qiagen kit.

weak amplification of *E. camurum.* In comparison, *E. camurum* primers weakly amplified *E. nigrum*, but consistently amplified *E. maculatum* and *E. tippecanoe*. Screening of *E. tippecanoe* showed a potential for weak cross – amplification with *E. camurum*, *E. caeruleum*, *E. nigrum*, and *P. copelandi*. However, the competitive screening test showed that under decreasing amounts of target DNA, a non – specific peak showed up at 246/247 bp. Single PCR specificity screening confirmed that the 246/247 bp amplicon was very weak amplification from *E. caeruleum* or *P. copelandi*. Regardless, *E. tippecanoe* also exhibited a diagnostic non – specific amplicon at 213 bp which was detectable in all fragment analysis screenings and only disappeared when the PCR reaction did not contain any target DNA. The presence of the non –

Table 3.2. Summary of primers developed in Primer-BLAST and screened for specie-specificity. Primers in bold were determined to be species-specific and selected for eDNA fragment analysis. Tm = melting temperature in degrees Celsius. GC% = the percentage of the primer sequence comprised of G/C nucleotides.

Primer Name	Sequence (5' ➔ 3')	Product Size (bp)	Length (bp)	Start	Stop	Tm (°C)	GC%
Ecam-ND2-P2							
Forward primer	CTCGCCTACTCATCCATCGC	159	20	532	551	60.39	60
Reverse primer	GGCGAGCATGTTGAGAGTGG		20	690	671	61.36	60
Ecam-ND2-P3							
Forward primer	ACTCATCCATCGCCCATCTT	151	20	539	558	58.86	50
Reverse primer	GCGAGCATGTTGAGAGTGG		19	689	671	58.92	58
Emac-ND2-P1							
Forward primer	GCCTATTCATCCATCGCCCA	145	20	535	554	59.96	55
Reverse primer	TGAGCGTGGTCGACTTACTG		20	679	660	59.76	55
Emac-ND2-P2							
Forward primer	TGCCTATTCATCCATCGCCC	145	20	534	553	59.96	55
Reverse primer	GAGCGTGGTCGACTTACTGA		20	678	659	59.48	55
Etip-COI-P1							
Forward primer	CTCCTCGGGTGTAGAGGCT	234	19	292	310	60.08	63
Reverse primer	GCGGTAATCAGGACAGCCC		19	525	507	60.52	63
Etip-COI-P2							
Forward primer	CTCGGGTGTAGAGGCTGGA	231	19	295	313	60.38	63
Reverse primer	GCGGTAATCAGGACAGCCCA		20	525	506	62.25	60
Ecaer-COI-P1							
Forward primer	ACTACTTGCYTCTTCCGGGG	234	20	283	302	59.39	55
Reverse primer	AGTACGGCCCACACGAATAG		20	516	497	59.54	55
Ecaer-ND2-P4							
Forward primer	CCCYCTCCCTACAACCCTTA	107	20	273	292	59.95	60
Reverse primer	CGGTAGTAAGRTCCAACCCC		20	379	360	59.46	60
Pcope-ND2-P1							
Forward primer	GGGGCCTTCCCCCATTAAC	241	19	752	770	60.08	63.16
Reverse primer	GTGGCTGTCGTTGTCGCTA		19	992	974	60.37	57.89
Pcope-ND2-P2							
Forward primer	CAAGAACTTGCCAAGCAGGA	198	20	799	818	58.68	50
Reverse primer	GGTGGTGGCTGTCGTTGTC		19	996	978	61.25	63.16

Table 3.3. Locations of the primers for *E. camurum* (Ecam-ND2-P2-F and Ecam-ND2-P2-R) and *E. maculatum* (Emac-ND2-P1-F and Emac-ND2-P1-R) within *ND2*. The light grey box shows the locations of the forward primers and the dark grey box shows the location of the reverse primers. Sequence variation within the gene is shown at each polymorphic site and *E. tippecanoe* is shown for comparison purposes. Numbers above the sequences indicate location (in bp) in the gene.

	530		532			535															551			554					559
E.cam	-	C/T	-	-	C/T	-	-	C/A	-	-	C	-	-	A/G	-	-	-	-	-	C	-	-	-	-	-	C/T	-	-	-	-
E.mac	-	T	-	-	T	-	-	C	-	-	T	-	-	A	-	-	-	-	-	C	-	-	-	-	-	T	-	-	-	-
E.tip	-	C	-	-	T	-	-	A	-	-	C	-	-	G	-	-	-	-	-	C/A	-	-	-	-	-	C	-	-	-	-

	660										671								679										690
E.cam	-	-	A/G	-	-	-	A	-	-	A/T	-	-	-	-	-	T	C/A	-	C/T	-	-	C/T	-	-	G	-	-	C/A	-	-
E.mac	-	-	G	-	-	-	G	-	-	G	-	-	-	-	-	G	C	-	C	-	-	T	-	-	A	-	-	C	-	-
E.tip	-	-	G	-	-	-	A	-	-	T	-	-	-	-	-	T	A	-	T	-	-	T	-	-	G	-	-	A	-	-

Table 3.4. Locations of the primers for *E. tippecanoe* within *COI* (Etip-COI-P2-F and Etip-COI-P2-R). The light grey box shows the locations of the forward primer and the dark grey box shows the location of the reverse primer. Sequence variation within the gene is shown at each polymorphic site and *E. camurum* and *E. maculatum* are shown for comparison purposes. Numbers above the sequences indicate location (in bp) in the gene.

	290					295																	313								320
E.cam	-	-	-	-	-	-	-	-	A	-	-	A	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	A	-
E.mac	-	-	-	-	-	-	-	-	A	-	-	A	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	T	-
E.tip	-	-	-	-	-	-	-	-	G	-	-	T	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	C	-
	506																			525											536
E.cam	-	-	A	-	-	C	-	-	A	-	-	T	-	-	-	-	-	T	-	-	A	-	-	-	-	-	T	-	-	-	-
E.mac	-	-	A	-	-	T	-	-	A	-	-	T	-	-	-	-	-	T	-	-	A	-	-	-	-	-	C	-	-	-	-
E.tip	-	-	G	-	-	T	-	-	C	-	-	G	-	-	-	-	-	C	-	-	C	-	-	-	-	-	T	-	-	-	-

specific amplicon at 213 bp was diagnostic to *E. tippecanoe*, but the intensity of the peak was 55x less than the 231 bp peak and would likely be lost to background noise with concentrations of target DNA less than 0.025 ng per reaction. The ratio between the RFUs of the 231 bp peak at 0.025 ng DNA and the positive control peak was 6.7 : 1, but when no target DNA was present a weak signal (RFU = 74) of a slightly smaller peak less than 231 bp (230.9) remained in a ratio of 111.5 : 1 to the positive control. In all tests with target DNA present, the peak was always between 231 - 232 bp, so the shift to a peak < 231 bp in the absence of target DNA combined with a visible peak at 247 bp is diagnostic to confirming the presence of either *E. caeruleum* or *P. copelandi*. Under these conditions, the confidence in detection of *E. tippecanoe* would be reduced.

Because *E. tippecanoe* primers were confirmed species – specific, eDNA extracted from four locations with confirmed *E. tippecanoe* presence were tested: Allegheny River L/D #6, two sites in the Kiskiminetas River (with three separate extractions), and Deer Creek. Not all eDNA samples were tested from Allegheny L/D #6 because all of the water samples were filtered separately and extracted separately resulting in 18 extractions. Selection was based on the extractions with the highest DNA yields. Because of the variation between eDNA samples (i.e. DNA extraction yields) and the potential for PCR inhibition, multiple PCRs were run for each eDNA extraction with the following amounts of total DNA (ng) per PCR reaction: 50, 20, 10, 5, and two dilutions, 1:5 and 1:10, of the original extraction (regardless of the DNA concentration). *E. tippecanoe* was detected in nine out of 48 PCR tests (Table 3.5), but was detected in at least one water sample at all four sites. Field and laboratory controls were negative and all positive control reactions worked. The positive control PCR consisted of 25 ng of tissue-derived *E.*

Table 3.5. Table summarizing detection of *E. tippecanoe* eDNA from water samples collected from Allegheny River L/D 6, Kiskiminetas River, and Deer Creek. (+) indicates positive detection using fragment analysis. The number adjacent to (+) indicates the lane of the agarose gel in figures 3.4 and 3.5. * indicates the DNA concentration of the extraction was low and the maximum volume of sample was added to the PCR to get as close to 50 ng per reaction as possible.

			DNA (ng)						No. of *E.tip*
Location	Pore Size (µm)	Extraction Kit	50	20	10	5	1:5	1:50	Collected
AR6-2*	0.7	MoBio	+[1]	+[7]	—	—	—	—	43
AR6-2	1.5	MoBio	—	—	—	—	—	—	43
AR6-5	0.7	Qiagen	—[2]	—	—	—	—	—	43
AR6-5a	1.5	Qiagen	—	—	—	—	—	—	43
Kiski-1a	0.7	Qiagen	+[3]	+[8]	+[9]	—	—	—	25
Kiski-1b	0.7	MoBio/Qiagen	+[4]	—	—	—	—	—	25
Kiski-2*	0.7	MoBio	+[5]	—	—	—	—	—	6
DeerCk*	0.45	MoBio	+[6]	+	—	—	—	—	2

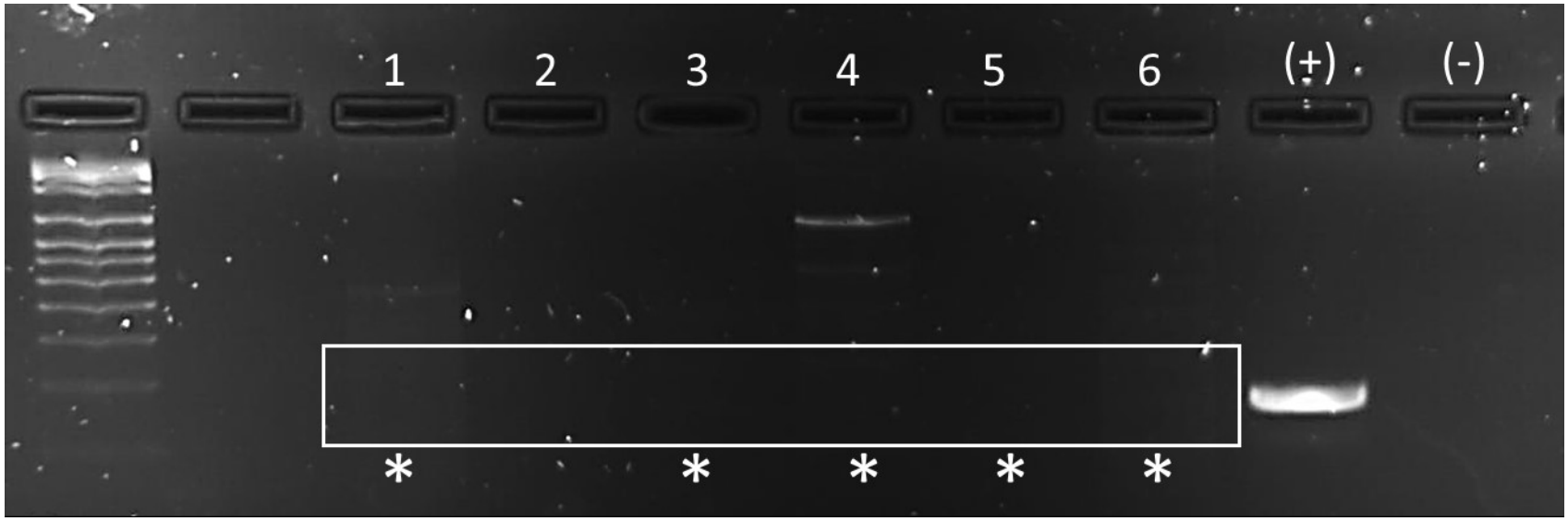

Figure 3.4. Agarose gel visualization of PCR amplification with 50 ng eDNA per reaction. The white box indicates the area where the target amplicon (231bp) would be seen with positive detection. * indicates samples that were confirmed with positive detection via fragment analysis: (1) Allegheny River L/D 6 – trawl #2, (3) Kiski1a, (4) Kiski1b, (5) Kiski2, and (6) Deer Creek. (+) and (-) indicate the positive and negative controls, respectively. Numbers of each lane correspond to Table 3.5.

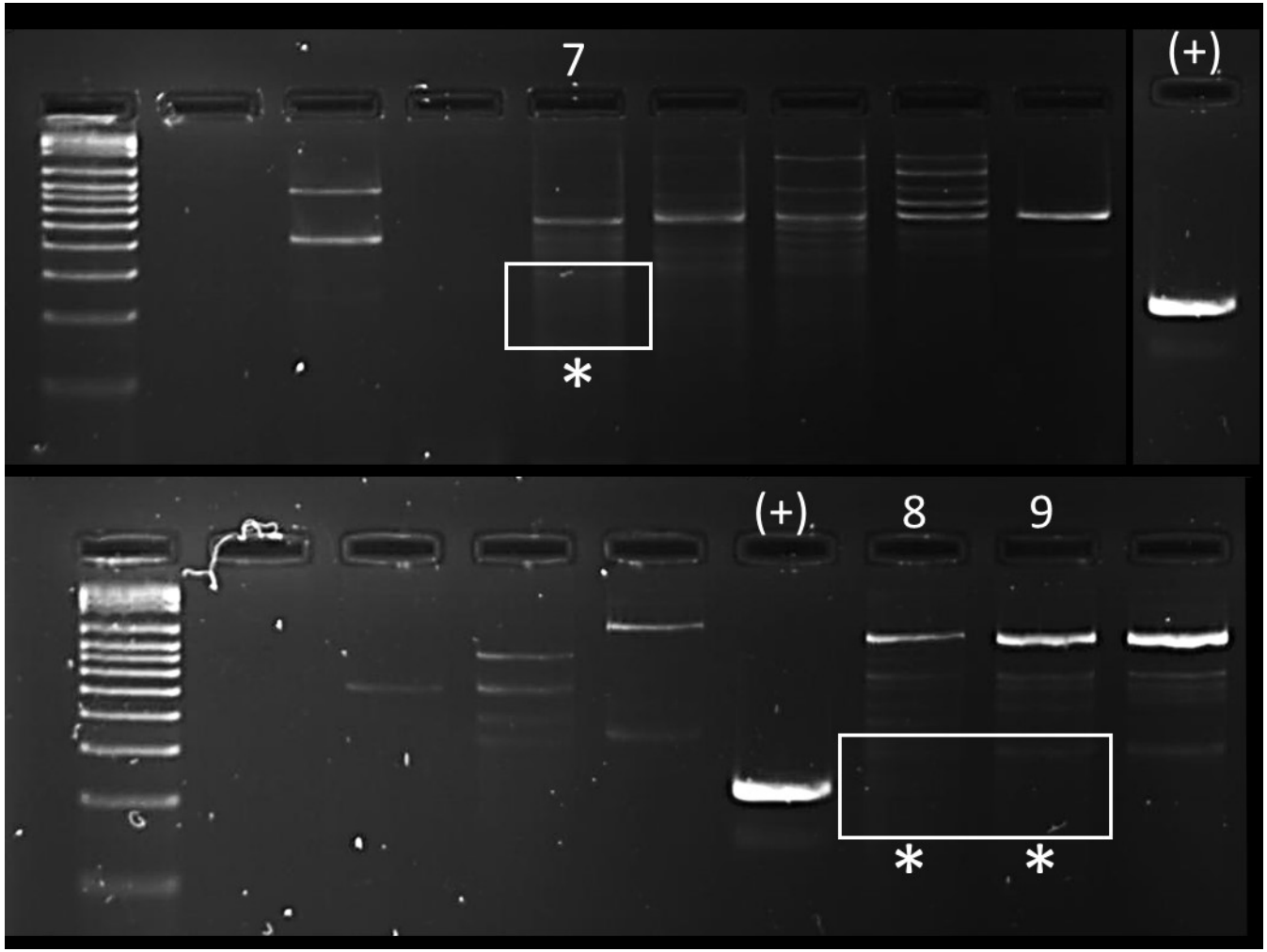

Figure 3.5. Agarose gel visualization of PCR amplification with 20 ng (1 and 2) and 10 ng (3) eDNA per reaction. The white boxes indicate the area where the target amplicon (231bp) would be seen with positive detection. * indicates samples that were confirmed with positive detection via fragment analysis: (7) Allegheny River L/D 6 – trawl #2, (8) Kiski1a – 20 ng, and (9) Kiski1a – 10 ng. (+) indicates the positive control. Numbers of each lane correspond to Table 3.5.

tippecanoe DNA per reaction. PCR with 50 ng of DNA per reaction showed positive detection in five out of eight reactions. Three out of eight tests with 20 ng of DNA per reaction were positive for *E. tippecanoe* while only one out of eight were positive with the 10 ng DNA reactions, (Table 3.5). Electrophoresis and agarose gel-based visualization failed to confirm the detection of *E. tippecanoe* in all cases (data for Deer Creek not shown) (Fig. 3.4 and 3.5). Modified injection time for fragment analysis increased the intensity of the target peaks (RFUs) for all of the samples. The increase ranged from 25 % for Deer Creek (20 ng PCR – normal injection = 28 RFU, increased injection = 35 RFU) to over a 900 % increase for Kisk1a (50 ng PCR – normal injection = 54 RFU, increased injection = 590 RFU). Kiski 2 was a non – detect under the normal injection protocol. See Figures 3.6 – 3.8 for electropherograms of fragment analysis that show *E. tippecanoe* detection and an example of what an eDNA negative control looks like on Peak Scanner.

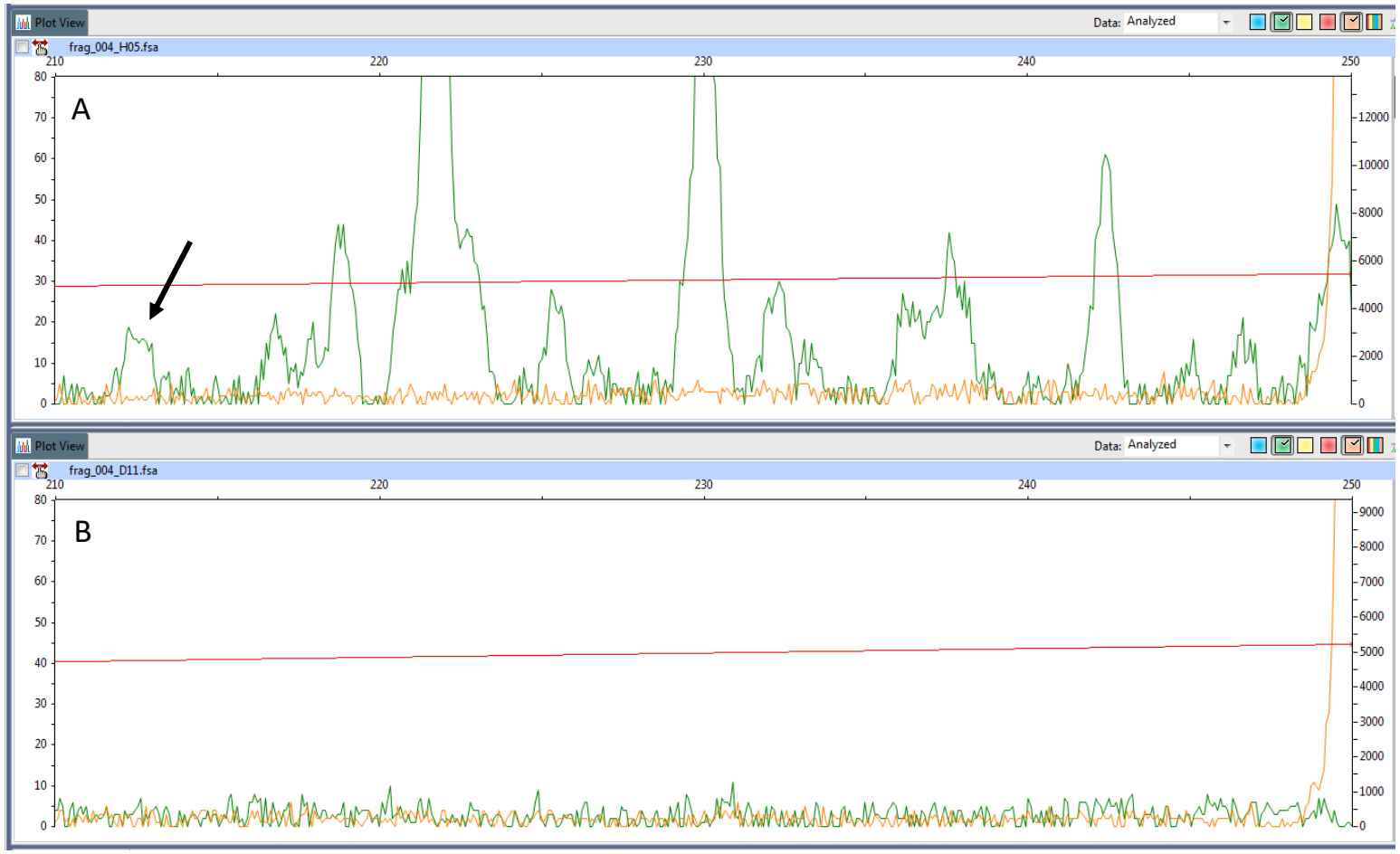

Figure 3.6. Electropherograms from Peak Scanner showing positive detection of *E. tippecanoe* from Allegheny River L/D 6 with a peak at 231 bp (Panel A – 20 ng total eDNA). The electropherogram also shows the diagnostic 213 bp peak (arrow). Panel B shows what an eDNA negative control looks like in Peak Scanner.

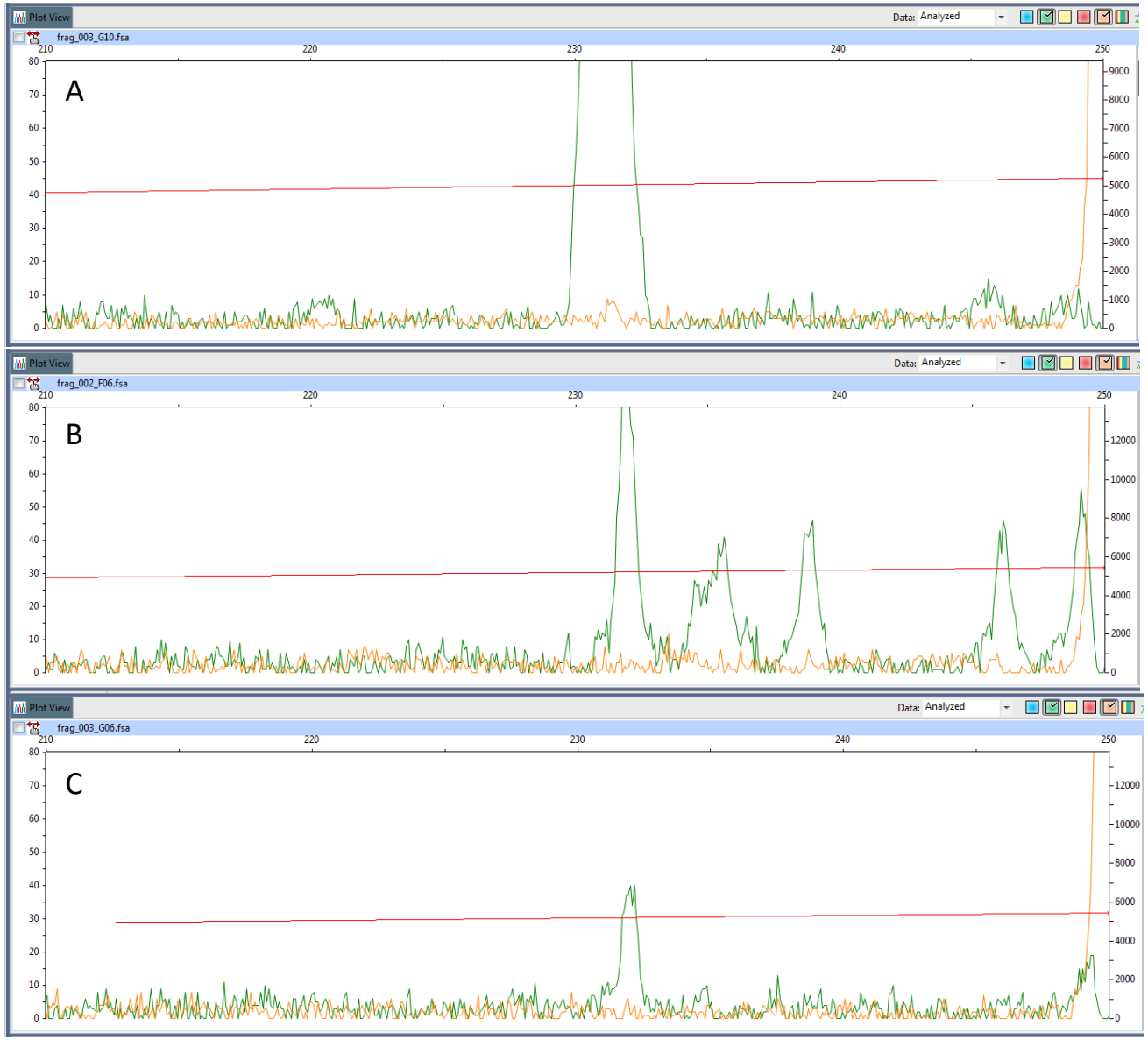

Figure 3.7. Electropherograms from Peak Scanner showing positive detection of *E. tippecanoe* from site Kiski1a with a peak at 231 bp. Panel A – 50 ng total eDNA, Panel B – 20 ng total eDNA, and Panel C – 10 ng total eDNA

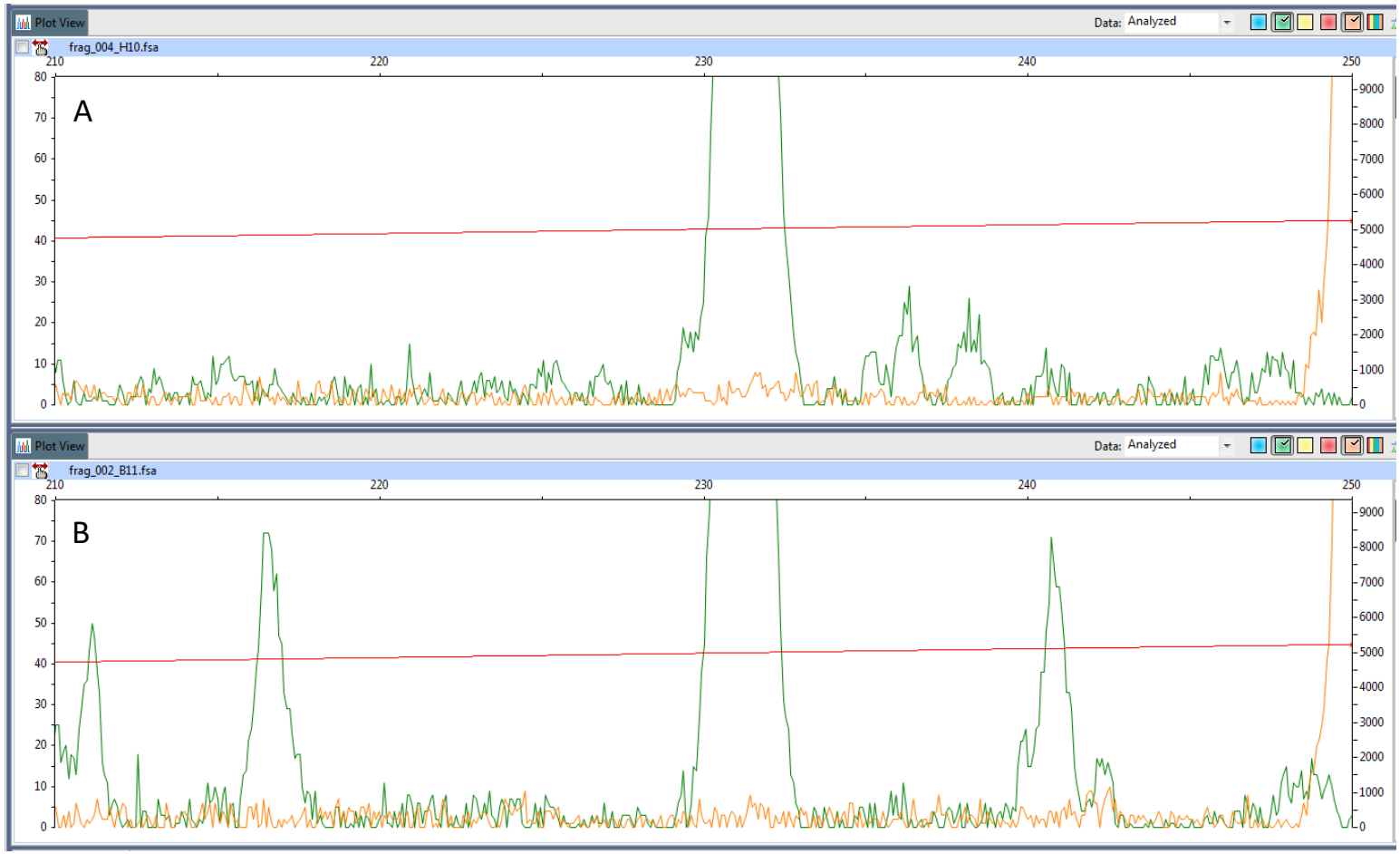

Figure 3.8. Electropherograms from Peak Scanner showing positive detection of *E. tippecanoe* from site Kiski1b (Panel A – 50 ng) and Deer Creek (Panel B – 50ng) with a peak at 231 bp.

3.4 - Discussion

Summary of Findings

Even though the initial attempts to develop a multi-plex primer set for eDNA applications failed, the second round of primer design resulted in two primer sets for eDNA surveys that can aide traditional sampling efforts for darters in the subgenus *Nothonotus*. The primer set for *E. camurum* was not field tested with eDNA samples, but all laboratory tests suggest that this primer set will be able to detect all three focal species (*E. camurum*, *E. maculatum*, and *E. tippecanoe*) from eDNA extracted from water samples. Until 2014 in Pennsylvania, all three species were State – listed as either threatened or endangered. Therefore, the *E. camurum* primer set has the potential to become a screening tool for surveying locations were these species have not been previously documented. If detected using this eDNA assay, traditional fish surveying methods could then be utilized to conduct a thorough survey to verify species presence. The primer set that was developed for *E. tippecanoe* was used to screen eDNA extracted from water samples collected at locations where traditional sampling methods verified their presence. This newly developed assay detected *E. tippecanoe* at all four locations (a 100% detection rate). The abundance of *E. tippecanoe* was: 43 individuals at Allegheny River L/D #6, 25 individuals at Kiski 1a/b, six individuals at Kiski 2, and only two individuals were documented at Deer Creek. In all cases, visualizing the products from eDNA PCR with electrophoresis on an agarose gel failed to identify any location where *E. tippecanoe* was present. Fragment analysis documented detection at all sites, and using fragment analysis to visualize eDNA detection is a novel approach that does not exist in the eDNA literature.

Limitations of eDNA Data

Even though eDNA methodologies have many benefits for species detection, the data have limitations to any interpretation beyond simple detection. For example, this study confirmed detection of *E. tippecanoe* below L/D #6 on the Allegheny River, but where was the source of the eDNA? *E. tippecanoe* has been documented at many sites further upstream of L/D 6 meaning that my analysis can't say for certain where the eDNA came from. Therefore, one limitation of eDNA analysis is related to the dynamics of eDNA persistence in aquatic systems. eDNA in aquatic systems is subject to environmental degrading and is influenced by factors such as UV light, mechanical forces, water temperature, water chemistry (e.g. pH), and microbial degradation (Barnes et al. 2014, Rees et al. 2014, Takahara et al. 2012). Studies of eDNA persistence have mainly been conducted under laboratory conditions in which the species is removed from a mesocosm and then degradation is assessed daily with measures of total eDNA and the point until which successful PCR no longer occurs (i.e. no detectable eDNA) (Barnes et al. 2014). Persistence has been shown to vary by species and habitat type ranging from 0.9 days for marine fish in aquaria (Thomsen et al. 2012a) to 21 days for a New Zealand Mud Snail (*Potamopyrgus antipodarum*) (Brown and Froemke 2012). Recently, Jane et al. (2014) placed caged Brook Trout (*Salvelinus fontinalis*) (five per cage) in the headwaters of several trout-less streams and took eDNA water samples at regular intervals downstream of the trout for six months. Results were correlated to flow conditions and the biomass of the caged trout. Interestingly, Brook Trout eDNA could be detected up to 239.5 m downstream regardless of flow conditions, but the amount of eDNA was inversely proportional to flow. Under low flow conditions, eDNA was more concentrated near the cages and quickly dissipated downstream, but

under high flow conditions the concentration of eDNA was low regardless of proximity to the cages (Jane et al. 2014) which suggests that larger volumes of water more thoroughly mix the eDNA causing it to be less concentrated. Lastly, Jane et al. (2014) sampled eDNA late into the fall in the presence of high levels of leaf deposition (with the waters likely having increased dissolved organic carbon and tannic acids) and documented complete inhibition of PCR even when eDNA exceeded 2,000 copies. This confirms that temporal variation and dissolved organic matter can influence eDNA detection and that care needs to be taken when designing a study. Next, conservation management of fishes is often concerned with the biomass and/or abundance of a species and a few eDNA studies have examined the correlation that exists between the amount of target eDNA and biomass. For example, Jane et al. (2014) showed that the copy number of eDNA was positively correlated to the biomass of caged brook trout. Using aquaria and experimental ponds, Takahara et al. (2012) documented that biomass of the Common Carp (*Cyprinus carpio*) was positively correlated to carp biomass, and interestingly, water temperature – warmer waters contained more eDNA. Recently, Lacoursière-Roussel et al. (2016) placed various numbers of Brook Charr (*Salvelinus fontinalis*) into aquaria at two different temperatures. There was a positive correlation between biomass and eDNA concentration, but the relationship was impacted by water temperature and the type of filter used to capture eDNA. Again, warmer water temperatures contained more eDNA which may have been the consequences of increased fish metabolism and movement; potentially leading to increased production of feces, urine and elevated sloughing of mucosal cells. Lacoursière-Roussel et al. (2016) also demonstrated that higher amounts of eDNA were captured using 0.7 and 1.2 µm glass – fiber (GF) filters. In this study, I stacked 1.5 µm GF filters on top of 0.07 µm GF filters. Even though these studies confirm a positive correlation between concentration of eDNA and

fish biomass/abundance, methods have not been developed for using eDNA to accurately estimate population size. Also, these studies confirm that factors like filter type and water temperature can significantly affect the amount of eDNA that is captured from the environment. Careful experimental design can help reduce uncertainty, but eDNA dynamics have been shown to vary temporally, by species, and across environments (e.g. lotic vs. lentic systems).

In this study, eDNA was collected at different times of the year and in systems that varied substantially in water discharge, all of which may have influenced eDNA concentrations. For example, when Allegheny River L/D #6 was sampled the water temperature was 16.2 °C and discharge was 25,000 ft^3/sec. In contrast, when eDNA water samples were collected from the Kiskiminetas River, water temperature was 23.4 °C and discharge was 561 ft^3/sec. The highest amount of eDNA extracted from any of the Kiski River sites vs. Allegheny River L/D #6 was 370.3 ng/µL and 16.8 ng/ µL, respectively. The discharge at Allegheny River L/D #6 was 44.5x higher than the Kiski, and the extracted eDNA concentrations reflect those conditions. However, it cannot be certain that discharge is the only factor influencing eDNA concentration because water temperature was different along with a host of other factors that weren't documented. In the early stages of method development, it was decided to keep the multiple extractions per site separated. Despite of Allegheny River L/D #6 having the largest number of documented *E. tippecanoe*, the eDNA yields and the fact that only one out of four filters showed positive detection for *E. tippecanoe* suggests that pooling eDNA extractions at sites with large discharge may improve detection rates. Cross – amplification tests and the competitive PCR assay were used to provide guidelines for determining the level of confidence in detection from the fragment analysis method. In the absence of *E. tippecanoe* DNA the electropherogram may show the following features: the target peak shifts to < 231 bp, a peak shows up at 246/247 bp, and the

ratio between the RFUs of the 230/231 bp peak and the positive control exceeds 111 : 1. Positive detection of *E. tippecanoe* may be assisted by the presence of a non – specific amplicon at 213 bp, but it was 55x less intense than the 231 bp peak at the lowest tested concentration and may likely be undetectable with low eDNA concentrations. See Fig. 3.9 for a graphical representation of the aforementioned detection guidelines.

3.5 - Conclusions

This study confirms that eDNA detection of darters via fragment analysis is more robust than electrophoresis and agarose gel – based methods. Early eDNA studies with fish (Jerde et al. 2011) and some as recently as 2015 with the Slackwater Darter (*Etheostoma boschungi*) (Janosik and Johnston 2015) implement regular PCR and gel detections methods. The first attempt to design species-specific primers for multiple closely related species failed. Cross-amplification occurred across the closely related species, which indicates the importance of primer specificity, and that challenges exist when developing aquatic eDNA primers in closely related fish species such as darters. The nucleotide diversity tests showed very low amounts of nucleotide diversity between the species with the highest diversity between *E. tippecanoe* and *E. camurum*, 7.1 % across the entire *ND2* gene. *E. camurum* and *E. maculatum* primers, regardless of the gene, always cross amplified each other. This was reflected in the nucleotide diversity tests which indicated that very low levels of diversity between the two species: $\theta = 0.035$ (*ND2*) and $\theta = 0.019$ (*COI*). These low levels of diversity complicate primer design and should be strongly considered when developing eDNA assays. Here, two primer sets were developed in which one, under laboratory conditions, can accurately detect all three focal species, and the other primer set

successfully detected *E. tippecanoe* with eDNA derived from water samples. A protocol was developed for water sample collection, eDNA extraction, eDNA amplification, and guidelines for determining the confidence in detection. Both primer sets should be evaluated with all of the eDNA samples that were collected during this dissertation. However, there are a few suggestions for changes that should be strongly considered: 1) another type of positive control should be added for more accurate comparison of eDNA extraction efficiency and RFU ratios - a deionized water sample should be spiked with a known quantity of target DNA and then processed following the same protocol as the eDNA water samples, 2) extend the aforementioned to create a test for determining detection limits, 3) attempt to increase eDNA extraction yields by following Eichmiller et al. (2015) whose comparison included the kits used in this study and determined that the Biomedicals FastDNA Spin Kit outperformed the MoBio and DNeasy kits, and 4) should consider pooling water samples taken from the mainstem river sites (which exhibit large discharges) to increase eDNA concentrations. However, pooling eDNA extractions could be problematic because this may concentrate PCR inhibitors into one sample. Depending on the application, pooling eDNA extractions may also result in diminished resolution when using eDNA for targeting species in very specific habitats or locations within a river. These methods may need to be further refined, but the primers developed in this study and the novel mechanism of visualizing detection will be useful tools in monitoring the recovery of the three focal species.

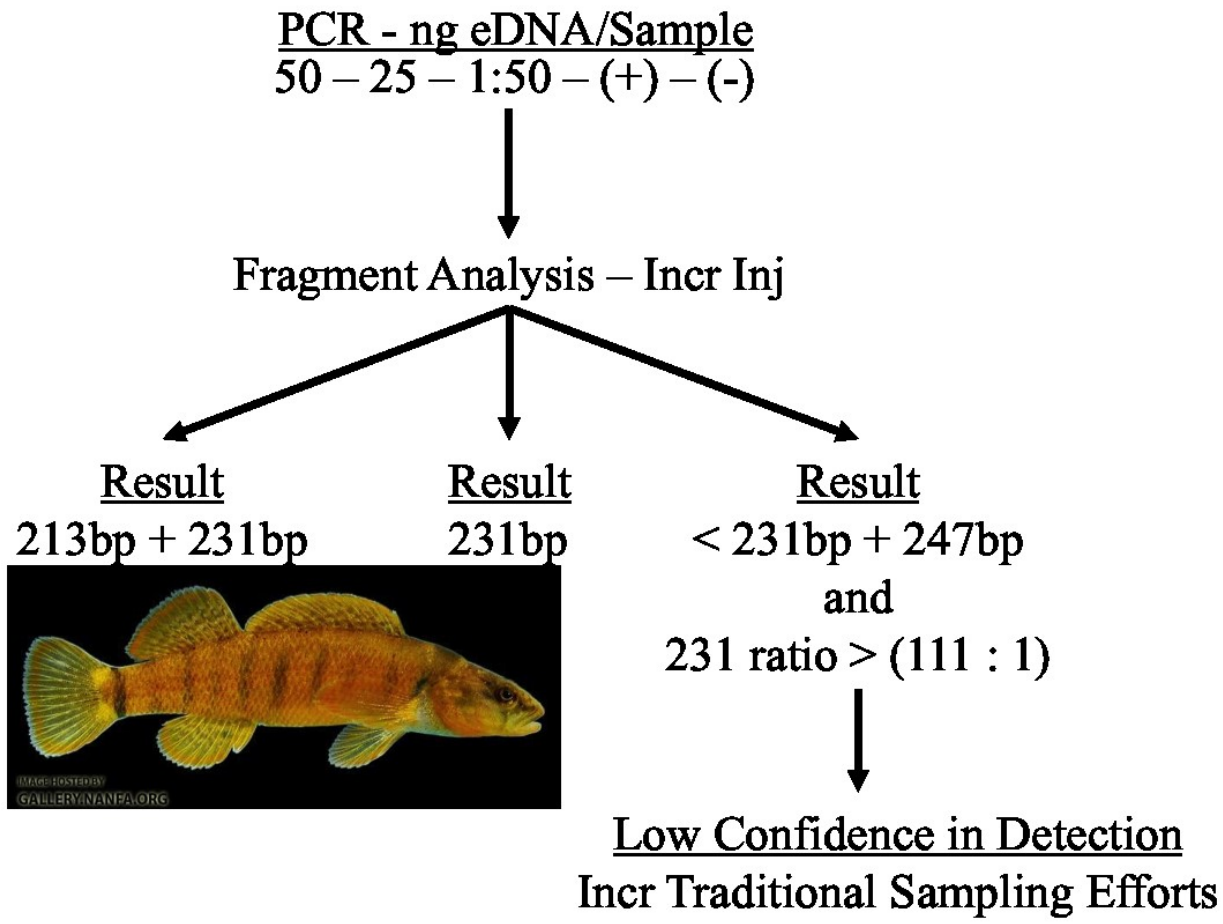

Figure 3.9. Flow chart and guidelines for targeting *E. tippecanoe* eDNA detection from water samples. Analysis begins by performing PCR with varying quantities of eDNA, positive controls (using tissue-derived DNA from the target species and an additional PCR using the primers developed for the common darter found at the sample site), and a negative control (deionized water). After PCR, fragment analysis should be completed with an increased injection time protocol on the genetic analyzer. A positive detection results from the presence of both the 213 bp and 231 bp peak or the presence of only a 231 bp peak. A slight shift in peak size to 229-230 bp, the presence of a 247 bp peak, or a 231 bp peak RFU ratio > 111:1 (relative to the positive control) would signal low confidence in detection with eDNA. If *E. tippecanoe* was not detected initially by traditional sampling efforts, an increase in traditional sampling effort should be performed to verify presence.

Literature Cited

Argent, D.G. and Kimmel, W.G. 2010. Influence of navigational lock and dam structures on adjacent fish communities in a major river system. River Research and Applications 27:1325-1333.

Argent, D.G. and Kimmel, W.G. 2014. Patterns of ichthyofaunal distribution and connectivity in navigable and free-flowing reaches of a major river system: The Allegheny river in Pennsylvania. River Research and Applications 30:631-638.

Baldigo, B.P., Sporn, L.A., George, S.D. and Ball, J.A. 2017. Efficacy of Environmental DNA to Detect and Quantify Brook Trout Populations in Headwater Streams of the Adirondack Mountains, New York. Transactions of the American Fisheries Society 146:99-111.

Barnes, M.A., Turner, C.R., Jerde, C.L., Renshaw, M.A., Chadderton, W.L. and Lodge, D.M. 2014. Environmental conditions influence eDNA persistence in aquatic systems. Environmental Science and Technology 48:1819-1827.

Brown, T.C. and Froemke, P. 2012. Nationwide assessment of nonpoint source threats to water quality. Bioscience 62:136-146.

Brozio, S., Manson, C., Gourevitch, E., Burns, T.J., Greener, M.S., Downie, J.R. and Hoskisson, P.A. 2017. Development and application of an eDNA method to detect the critically endangered Trinidad golden tree frog (Phytotriades auratus) in bromeliad phytotelmata. Plos One 12.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W. and Shipley, G.L. 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clinical chemistry 55:611-622.

Cho, A., Morris, T., Wilson, C. and Freeland, J. 2016. Development of species-specific primers with potential for amplifying eDNA from imperilled freshwater unionid mussels. Genome 59:1141-1149.

Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J. and Sayers, E.W. 2016. GenBank. Nucleic Acids Research 44:D67-D72.

Deiner, K. and Altermatt, F. 2014. Transport distance of invertebrate environmental DNA in a natural river. Plos One 9.

Deiner, K., Walser, J.C., Mächler, E. and Altermatt, F. 2015. Choice of capture and extraction methods affect detection of freshwater biodiversity from environmental DNA. Biological Conservation 183:53-63.

Eichmiller, J.J., Miller, L.M. and Sorensen, P.W. 2015. Optimizing techniques to capture and extract environmental DNA for detection and quantification of fish. Molecular Ecology Resources 16:56-68.

Evans, N.T., Shirey, P.D., Wieringa, J.G., Mahon, A.R. and Lamberti, G.A. 2017. Comparative Cost and Effort of Fish Distribution Detection via Environmental DNA Analysis and Electrofishing. Fisheries 42:90-99.

Ficetola, G.F., Miaud, C., Pompanon, F. and Taberlet, P. 2008. Species detection using environmental DNA from water samples. Biology Letters 4:423-425.

Freedman, J.A., Stecko, T.D., Lorson, B.D. and Stauffer, J.R. 2009a. Development and Efficacy of an Electrified Benthic Trawl for Sampling Large-River Fish Assemblages. North American Journal of Fisheries Management 29:1001-1005.

Freedman, J.A., Stecko, T.D., Criswell, R.W., and Stauffer, J.R. 2009b. Extensions of the known ranges of *Percina shumardi* Girard and three species of *Etheostoma* (subgenus *Nothonotus*) in Pennsylvania. Journal of the Pennsylvania Academy of Science 83:42-44.

Goldberg, C.S., Pilliod, D.S., Arkle, R.S. and Waits, L.P. 2011. Molecular Detection of Vertebrates in Stream Water: A Demonstration Using Rocky Mountain Tailed Frogs and Idaho Giant Salamanders. Plos One 6.

Herzog, D.P., Barko, V.A., Scheibe, J.S., Hrabik, R.A. and Ostendorf, D.E. 2005. Efficacy of a benthic trawl for sampling small-bodied fishes in large river systems. North American Journal of Fisheries Management 25:594-603.

Herzog, D.P., Ostendorf, D.E., Hrabik, R.A. and Barko, V.A. 2009. The mini-Missouri trawl: a useful methodology for sampling small-bodied fishes in small and large river systems. Journal of Freshwater Ecology 24:103-108.

Honick, A.S., Zimmerman, B.J., Stauffer, J.R. Jr., Argent, D.G., and Porter, B.A. 2017. Expanded distributions of three *Etheostoma* Darters (subgenus *Nothonotus*) within the upper Ohio River watershed. Northeastern Naturalist In Press.

Hunter, M.E., Dorazio, R.M., Butterfield, J.S.S., Meigs-Friend, G., Nico, L.G. and Ferrante, J.A. 2017. Detection limits of quantitative and digital PCR assays and their influence in presence–absence surveys of environmental DNA. Molecular Ecology Resources 17:221-229.

Jane, S.F., Wilcox, T.M., McKelvey, K.S., Young, M.K., Schwartz, M.K., Lowe, W.H., Letcher, B.H. and Whiteley, A.R. 2014. Distance, flow and PCR inhibition: EDNA dynamics in two headwater streams. Molecular Ecology Resources.

Janosik, A.M. and Johnston, C.E. 2015. Environmental DNA as an effective tool for detection of imperiled fishes. Environmental Biology of Fishes 98:1889-1893.

Jerde, C.L., Chadderton, W.L., Mahon, A.R., Renshaw, M.A., Corush, J., Budny, M.L., Mysorekar, S. and Lodge, D.M. 2013. Detection of Asian carp DNA as part of a Great Lakes basin-wide surveillance program. Canadian Journal of Fisheries and Aquatic Sciences 70:522-526.

Jerde, C.L., Mahon, A.R., Chadderton, W.L. and Lodge, D.M. 2011. "Sight-unseen" detection of rare aquatic species using environmental DNA. Conservation Letters 4:150-157.

Keller, S.R., Hilderbrand, R.H., Shank, M.K. and Potapova, M. 2017. Environmental DNA genetic monitoring of the nuisance freshwater diatom, Didymosphenia geminata, in eastern North American streams. Diversity and Distributions 23:381-393.

Kelly, R.P., Port, J.A., Yamahara, K.M. and Crowder, L.B. 2014. Using environmental DNA to census marine fishes in a large mesocosm. Plos One 9.

Knouft, J.H., Caruso, N.M., Dupre, P.J., Anderson, K.R., Trumbo, D.R. and Puccinelli, J. 2011. Using fine-scale GIS data to assess the relationship between intra-annual environmental niche variability and population density in a local stream fish assemblage. Methods in Ecology and Evolution 2:303-311.

Lacoursière-Roussel, A., Rosabal, M. and Bernatchez, L. 2016. Estimating fish abundance and biomass from eDNA concentrations: variability among capture methods and environmental conditions. Molecular Ecology Resources 16:1401-1414.

Librado, P. and Rozas, J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451-1452.

Lorson, B.D. 2010. Distribution and the putative origin of fishes in the Allegheny River, Pennsylvania. Master of Science. The Pennsylvania State University, State College. 184 pp.
Mahon, A.R., Jerde, C.L., Galaska, M., Bergner, J.L., Chadderton, W.L., Lodge, D.M., Hunter, M.E. and Nico, L.G. 2013. Validation of eDNA Surveillance Sensitivity for Detection of Asian Carps in Controlled and Field Experiments. Plos One 8.
Maniatis, T., E.F. Fristch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
Minamoto, T., Yamanaka, H., Takahara, T., Honjo, M.N. and Kawabata, Z. 2012. Surveillance of fish species composition using environmental DNA. Limnology 13:193-197.
Ohio Revised Code. 2015. Title 15: Conservation of Natural Resources. Chapter 1531.25 - Protection of species threatened with statewide extinction, pursuant to Section 4 of the Endangered Species Act of 1973.
Olson, Z.H., Briggler, J.T. and Williams, R.N. 2012. An eDNA approach to detect eastern hellbenders (Cryptobranchus a. alleganiensis) using samples of water. Wildlife Research 39:629-636.
Pennsylvania Bulletin. 2014. Chapter 75.2, Threatened Fish Species. 44:7876-7878.
Pennsylvania Bulletin. 1999. Chapter 75.2, Threatened Fish Species. 29:4869-4870.
Pilliod, D.S., Goldberg, C.S., Arkle, R.S. and Waits, L.P. 2014. Factors influencing detection of eDNA from a stream-dwelling amphibian. Molecular Ecology Resources 14:109-116.
Rees, H.C., Maddison, B.C., Middleditch, D.J., Patmore, J.R. and 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.
Ritchea, S.B. 2006. Genetic population structure of the federally endangered Etowah darter, *Etheostoma etowahae*. Master of Science. Duquesne University, Pittsburgh, PA. 131pp.
Steiner, L. 2000. Pennsylvania Fishes. Pennsylvania Fish and Boat Commission.
Taberlet, P., Coissac, E., Hajibabaei, M. and Rieseberg, L.H. 2012. Environmental DNA. Molecular Ecology 21:1789-1793.
Takahara, T., Minamoto, T., Yamanaka, H., Doi, H. and Kawabata, Z. 2012. Estimation of Fish Biomass Using Environmental DNA. Plos One 7.
Thomsen, P.F., Kielgast, J., Iversen, L.L., Møller, P.R., Rasmussen, M. and Willerslev, E. 2012a. Detection of a Diverse Marine Fish Fauna Using Environmental DNA from Seawater Samples. Plos One 7.
Thomsen, P.F., Kielgast, J., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T. P., Orlando, L., Willerslev, E. 2012b. Monitoring endangered freshwater biodiversity using environmental DNA. Molecular Ecology 21:2565-2573.
Turner, C.R., Barnes, M.A., Xu, C.C.Y., Jones, S.E., Jerde, C.L. and Lodge, D.M. 2014a. Particle size distribution and optimal capture of aqueous macrobial eDNA. Methods in Ecology and Evolution 5:676-684.
Turner, C.R., Miller, D.J., Coyne, K.J. and Corush, J. 2014b. Improved methods for capture, extraction, and quantitative assay of environmental DNA from Asian bigheaded carp (hypophthalmichthys spp.). Plos One 9.
Wilcox, T.M., McKelvey, K.S., Young, M.K., Jane, S.F., Lowe, W.H., Whiteley, A.R. and Schwartz, M.K. 2013. Robust Detection of Rare Species Using Environmental DNA: The Importance of Primer Specificity. Plos One 8.

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. 2012. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics, 13, 134

Yoder, C.O., Rankin, E.T., Smith, M.A., Alsdorf, B.C., Altfater, D.J., Boucher, C.E., Miltner, R.J., Mishne, D.E., Sanders, R.E. and Thoma, R.F. 2005. Changes in fish assemblage status in Ohio's non-wadeable rivers and streams over two decades. American Fisheries Society Symposium 45:399-429.

CHAPTER FOUR

Major Conclusions and Future Directions

4.1 - Expanded Distributions of Three Etheostoma Darters (subgenus Nothonotus) within the Upper Ohio River Watershed

A total of 1,974 fish were collected from electrified – benthic trawling below eleven L/D installations and 2,183 fish were collected from backpack electrofishing nine tributaries. A total of 44 species were represented which included 13 darter species. Analysis of contemporary and historic distribution records of *E. camurum, E. tippecanoe,* and *E. maculatum* showed that *E. maculatum* had a less continuous distribution, smaller geographic range, and feweonvention set earlier in the document.onvention set earlier in the document.r individuals per sample site compared to *E. camurum* and *E. tippecanoe*, which may be related to life history characteristics, a lack of optimal habitat, and impaired connectivity throughout the navigable portions of the upper Ohio River watershed. Historically, the three focal species were reported to occupy similar habitats, but with *E. maculatum* occupying deeper riffle habitat. I documented that only *E. camurum* and *E. tippecanoe* are effectively utilizing deeper habitat below navigational lock and dam installations in the upper Ohio River watershed.

There are several potential reasons why *E. maculatum* is not expanding its distribution as effectively. First, *E. camurum* and *E. tippecanoe* illustrate characteristics of an r-selected type of reproductive strategy while *E. maculatum* employs characteristics of a K-selected reproductive strategy. Additionally, *E. maculatum* may require larger areas of suitable spawning habitat as a result of maintaining territoriality and nest defense - potentially producing fewer offspring per

unit of available habitat. *E. maculatum* has also been documented as having a short pelagic larval duration that may be limiting their distance (or rate) of dispersal, and lastly, the navigational L/D system may be restricting movement of *E. maculatum* between the free-flowing sections of the upper Allegheny River and the navigable portions of the upper Ohio River watershed. Therefore, the stable, source populations of *E. maculatum* should be closely monitored.

Future analysis should focus on more extensive surveys throughout the navigable reaches of the upper Ohio River watershed and its tributaries to document spawning grounds and monitor distribution changes in all three species. Also, efforts should include obtaining sample sizes large enough to complete genetic analysis to determine source/sink dynamics for each species.

4.2 - The genetic population structure of Etheostoma camurum (Bluebreast Darter) in the upper Ohio River watershed

Sample sizes large enough for genetic analysis were only obtained for *E. camurum*. Fragment analysis of six loci in *E. camurum* populations in the upper Ohio River watershed showed relatively high levels of genetic diversity within their populations and no detectable evidence of inbreeding. Population structure analysis confirmed that the Kokosing R. is not a likely source population and that French Cr. has/is minimally influencing the genetic structure of the populations in the study area. The low levels of genetic differentiation, lack of IBD (isolation by distance) and population structuring indicated that the navigational L/D system has not greatly impacted gene flow in these populations. The tributaries showed low levels of differentiation from the mainstem Allegheny R. population which suggests that these populations were likely established with migrants from the Allegheny River drainage. Finally, structural analysis indicates that the Ohio R. population is likely from an undocumented source population

and may represent the progression of *E. camurum* from a source that is moving upstream via the Ohio River mainstem or a population in the Allegheny River that was not documented in this study.

Future directions should include gathering data from sample sites between these local populations and French Cr. and searching for potential source populations from downstream locations that are less geographically isolated than the Kokosing R. For example, the lower Scioto and Muskingum rivers in Ohio, which have been documented with large, stable populations of *E. camurum*. Future analysis should include more rigorous estimates of gene flow (e.g. Bayesian inference of recent migration using multi-locus genotypes), incorporating more microsatellite loci, or utilizing SNPs, both of which should increase power of the analysis.

4.3 – Development of an aquatic environmental DNA (eDNA) method for detection of darters (subgenus Nothonotus) in the upper Ohio River watershed

Developing eDNA primers and fragment detection assays for closely related darter species presented several challenges. I showed that when nucleotide diversity between closely related fish species falls below 7 %, eDNA primer development will be challenging. PCR cross-amplification occurred between species when nucleotide diversity across the gene was between 1.9 – 4.3 %. *E. tippecanoe* eDNA extracted from water samples taken from the Kiskiminetas and Allegheny rivers was successfully PCR amplified with *E. tippecanoe* specific primers and detected using fragment analysis. I also showed that fragment analysis was more sensitive for eDNA detection compared to electrophoresis and agarose gel-based visualization.

Future directions should include verifying the *E. tippecanoe* eDNA primers with water samples extracted from the additional sample sites. Also, positive controls to verify PCR is

working in addition to the target species were developed for common darters likely to be at any sample site (e.g. Rainbow Darter – *E. caeruleum*, and Channel Darter – *P. copelandi*) and should be included in future eDNA analysis. Lastly, these methods may be expanded for eDNA assays that have the ability to detect community-level species data by designing primers that amplify a diagnostic region of a gene across multiple species such as those applied in DNA metabarcoding studies. These methods could be developed to detect fish, mussel, or insect communities.

Appendix A:
Supplementary Results and Raw Data

Table A.1 References for Tables 1.2 and 1.3.

Reference
Hoskin, R.H., Michael Koryak, and Linda J. Stafford. 2003. Fishes of Small Tributaries to the Allegheny and Monongahela Rivers in Urban/Suburban Allegheny County, Pennsylvania. Journal of the Pennsylvania Academy of Science 77:51-58.
Howell, L.A. 2007. Genetic population structure and breeding parameters of three PA state threatened darter species: E. camurum, E. maculatum, and E. tippecanoe. Master's Thesis. Duquesne University, Pittsburgh, PA. 187 pp.
Koryak, M. 2003. Fishes of Small Tributaries to the Ohio River in Allegheny County, Pennsylvania. The STUDIO for Creative Inquiry, Carnegie Mellon University, Pittsburgh, PA.
Multistate Aquatic Resources Information System (MARIS). 2016. Data sourced from PA Fish and Boat Commission. Available online at http://www.marisdata.org. Accessed 20 Sept. 2016.
U.S. Army Corps of Engineers. 1997. Montour Run Watershed, Allegheny County, PA: Water Quality and Aquatic Life Resources. Pittsburgh, PA: U.S. Army Corps of Engineers
Ohio Environmental Protection Agency (OhioEPA). Data queried from Ohio State University Museum of Biological Diversity Fish Division database, analyzed by Zimmerman, B. J.
This study (ASH) - data collected by Honick, A. S. and Porter, B. A.
This study (BAP) - data collected by Porter, B. A.

Table A.2. Genotypes of all 156 *E. camurum* individuals with site names – final binning.

SITE	Eca46		Eca70		CV24		Eca48		Eca11		Esc132b	
Kok	228	240	202	210	130	130	160	160	166	170	182	182
Kok	280	282	197	202	130	144	160	160	166	170	186	186
Kok	228	271	202	210	128	154	164	167	149	166	178	182
Kok	228	250	197	197	130	130	160	171	166	166	178	182
Kok	250	282	197	197	130	144	160	171	166	170	182	182
Kok	238	250	210	210	130	130	160	167	166	170	170	182
Kok	250	250	197	202	128	130	160	171	166	170	170	182
Kok	228	282	197	197	130	154	160	160	149	149	182	186
Kok	224	282	197	197	130	130	171	160	149	170	182	186
Kok	238	250	202	197	130	154	167	167	166	166	170	170
Kok	240	240	197	197	128	130	142	160	170	166	178	182
Kok	238	250	202	202	126	128	171	160	149	170	170	182
Kok	228	240	197	210	130	154	175	160	149	170	182	186
Kok	282	282	202	197	128	130	160	175	166	166	178	182
Kok	224	228	202	197	130	130	167	167	166	166	178	182
Kok	240	282	202	202	130	154	171	171	166	170	170	170
Kok	232	240	197	197	128	130	160	167	170	166	178	178
Kok	228	240	197	197	130	154	160	171	149	166	170	186
Kok	228	240	202	210	130	130	164	171	149	166	178	182
Kok	228	228	197	202	130	130	160	171	170	149	182	182
Kok	228	228	197	197	130	154	160	160	149	166	153	186
Kok	228	240	197	197	130	130	160	171	166	166	170	186
Kok	228	240	197	202	130	154	171	175	166	166	170	178
Kok	228	282	192	197	130	130	160	167	149	149	182	182
Kok	224	250	197	197	128	128	160	160	149	166	182	186
Ecam	287	301	197	197	130	154	160	164	166	170	174	178
Ecam	224	228	197	197	128	128	142	171	166	174	178	178
Ecam	224	250	119	197	128	154	164	171	166	166	178	182
Ecam	228	240	192	197	128	130	142	171	149	166	170	178
Ecam	290	297	197	197	144	156	167	167	149	178	182	182
Ecam	228	232	197	202	130	154	160	164	149	166	178	182
Ecam	232	254	197	202	130	154	167	171	149	166	178	182
Ohio	297	301	000	000	175	183	179	191	170	170	170	174

SITE	Eca46		Eca70		CV24		Eca48		Eca11		Esc132b	
Ohio	262	294	144	144	122	181	179	191	125	170	200	200
Ohio	287	297	197	197	130	130	167	171	149	149	170	178
Ohio	204	230	197	197	130	130	160	175	166	174	174	178
Ohio	297	297	000	000	178	181	183	187	170	170	170	170
Ohio	230	235	197	202	130	154	155	167	166	170	178	182
Ohio	230	244	197	202	130	144	142	171	166	174	178	178
Ohio	230	244	197	197	128	130	160	171	149	166	178	178
Ohio	230	230	170	197	130	144	164	171	149	174	178	194
Ohio	244	316	197	202	126	130	160	164	149	174	178	186
Ohio	244	256	192	197	154	154	160	160	166	174	182	186
Ecam	287	287	144	144	150	150	183	183	162	162	200	244
Ecam	235	297	197	197	128	130	146	167	149	170	178	182
Ecam	230	256	192	202	130	130	160	171	149	170	178	182
AGR	256	297	197	197	130	144	142	171	162	170	178	186
AGR	228	240	197	197	154	156	142	164	149	166	170	178
AGR	235	235	197	197	154	156	155	171	170	170	182	182
AGR	230	316	197	197	130	130	146	160	149	170	178	182
AGR	230	244	202	202	130	144	142	160	149	170	178	178
AGR	254	297	197	197	130	154	142	164	166	170	170	178
AGR	230	235	197	202	130	144	150	167	149	174	170	178
AGR	238	240	202	202	130	156	160	171	166	170	170	178
AGR	228	228	197	197	128	130	142	160	149	166	178	186
AGR	224	240	197	202	128	130	160	171	145	170	170	178
AGR	228	294	197	197	144	154	160	171	149	170	170	178
AGR	228	232	197	200	130	130	142	171	149	153	178	182
AGR	228	232	170	200	130	154	164	171	166	178	166	178
AGR	228	232	197	197	130	154	160	171	166	170	182	186
AGR	224	250	192	192	130	154	160	164	166	174	178	178
AGR	228	228	197	200	130	152	142	155	149	170	153	178
AGR	000	000	200	200	126	130	142	167	170	170	182	186
AGR	232	297	197	202	126	154	160	167	170	170	178	178
AGR	224	250	170	200	130	130	142	164	149	149	170	182
AGR	224	250	170	188	130	130	167	171	170	174	170	186
AGR	238	250	188	197	130	130	142	164	149	166	178	182
AGR	240	301	197	202	154	154	160	167	149	166	178	178

SITE	Eca46		Eca70		CV24		Eca48		Eca11		Esc132b	
AGR	250	314	170	197	154	154	160	171	149	166	178	186
AGR	228	232	197	197	130	154	142	171	166	170	178	182
AGR	228	232	197	200	152	152	164	175	170	178	178	178
AGR	224	250	170	197	126	144	142	171	166	170	178	182
AGR	238	250	197	200	130	152	142	160	149	166	178	186
AGR	228	297	192	197	130	154	164	167	170	174	170	178
AGR	280	282	197	197	130	154	160	164	170	170	178	186
AGR	228	306	197	200	126	154	160	164	149	170	178	186
Deer(A)	228	250	170	197	130	144	164	171	149	166	178	182
Deer(A)	224	250	197	202	154	154	167	167	149	166	178	182
Deer(A)	232	232	170	197	154	156	160	164	149	166	178	186
Deer(A)	294	294	170	197	130	130	142	167	149	170	178	178
Deer(A)	228	280	200	204	130	144	160	160	149	174	178	182
Deer(A)	280	282	200	200	130	130	167	171	170	170	182	186
Deer(A)	224	250	197	197	128	144	160	164	166	170	178	182
Deer(A)	228	228	204	204	130	154	164	171	166	174	170	182
Deer(A)	228	232	200	208	130	152	164	171	166	174	170	182
Deer(A)	228	232	200	200	132	156	160	171	149	166	178	182
Deer(A)	228	228	200	204	156	156	160	160	170	170	170	178
Deer(A)	228	294	200	200	132	132	142	167	174	174	162	182
Deer(A)	290	301	200	200	156	156	160	171	149	174	174	178
Deer(A)	232	301	204	204	132	152	160	164	166	183	170	178
Deer(A)	228	240	200	200	128	132	160	164	162	174	178	182
Deer(A)	228	287	200	204	132	156	171	171	166	170	170	182
Deer(A)	232	297	200	204	130	154	142	171	166	174	170	178
Deer(A)	250	282	200	204	154	154	142	160	162	166	182	182
Deer(A)	250	276	204	204	130	130	142	155	149	166	174	186
Deer(A)	232	287	200	204	128	130	142	160	149	166	178	182
Deer(A)	228	294	204	204	130	154	150	171	149	166	178	182
Deer(A)	238	250	200	200	132	144	167	171	174	174	182	182
Deer(A)	224	240	200	204	128	132	171	171	166	166	178	182
Deer(A)	232	287	200	204	132	132	164	171	149	166	178	186
Deer(A)	232	232	197	202	130	130	160	171	166	170	178	186
Deer(A)	232	240	197	197	130	156	142	171	149	149	166	170
Deer(A)	290	297	170	197	130	144	160	160	149	166	170	182

SITE	Eca46		Eca70		CV24		Eca48		Eca11		Esc132b	
Bull	000	000	197	197	126	152	160	171	166	166	182	186
Bull	224	228	197	202	126	144	160	164	149	166	182	186
Bull	238	250	197	197	154	154	160	167	166	174	178	182
Bull	232	232	188	197	130	130	160	171	166	174	178	186
Bull	228	306	188	197	130	154	164	171	174	174	178	186
Bull	232	232	192	197	130	144	164	167	166	166	178	182
Bull	290	297	197	202	130	130	142	160	166	166	170	182
Bull	232	232	197	197	130	130	142	171	149	166	178	178
Bull	228	301	197	202	154	154	160	171	174	174	170	178
Bull	228	306	197	197	130	144	142	160	149	166	178	178
Bull	228	282	202	202	126	156	142	179	149	149	178	190
Bull	301	301	197	202	130	154	164	167	149	166	178	182
Bull	282	294	197	202	126	130	160	167	149	166	178	186
Bull	228	228	188	197	144	154	142	171	166	170	174	178
Bull	224	254	197	197	126	130	160	171	166	166	178	186
Ecam	224	250	197	202	130	130	142	160	166	174	174	182
Deer(B)	228	244	197	202	130	144	164	171	170	174	174	276
Deer(B)	224	250	197	202	130	130	164	171	149	166	174	276
Deer(B)	228	297	197	202	130	144	171	171	149	166	178	174
Deer(B)	228	240	197	197	154	154	160	167	170	178	170	178
Deer(B)	232	290	197	202	130	154	142	164	166	166	178	178
Deer(B)	228	228	188	202	130	130	142	164	166	174	178	182
Deer(B)	228	290	197	202	128	130	160	171	149	149	174	178
Deer(B)	250	250	197	200	126	154	167	171	153	174	178	186
Deer(B)	224	240	182	197	130	154	160	167	149	174	182	182
Deer(B)	228	250	197	200	130	154	160	164	149	174	182	182
Ltl Sew	232	297	000	000	130	130	160	171	149	170	170	178
Ltl Sew	228	294	192	197	130	154	171	171	149	170	178	182
Ltl Sew	228	228	182	182	126	130	160	171	149	149	174	178
Ltl Sew	228	240	197	200	154	156	160	171	166	170	170	174
Ltl Sew	228	250	182	182	130	154	142	167	166	166	178	182
Ltl Sew	232	294	200	200	130	130	160	167	166	170	178	182
Ltl Sew	228	232	197	200	130	154	160	164	149	170	170	178
Ltl Sew	228	228	192	197	130	154	164	171	166	166	174	182
Ltl Sew	240	287	197	200	126	130	164	164	149	149	178	182

SITE	Eca46		Eca70		CV24		Eca48		Eca11		Esc132b	
Ltl Sew	224	294	182	188	130	154	160	171	149	149	178	186
Ltl Sew	224	250	197	200	126	144	160	171	166	178	182	182
Ltl Sew	240	294	182	197	128	130	160	175	166	170	178	178
Ltl Sew	232	297	182	197	126	144	164	167	149	174	170	190
French	224	254	197	197	152	154	160	167	166	166	182	186
French	232	240	197	197	128	130	142	171	166	170	178	178
French	228	240	197	197	126	128	160	171	166	174	174	182
French	232	301	197	197	128	128	146	171	170	170	178	178
French	228	254	192	200	126	156	164	171	149	166	178	178
French	228	240	197	197	128	128	142	167	166	170	182	186
French	228	228	197	197	128	156	160	167	166	166	178	178
French	228	228	197	200	126	154	160	167	149	166	178	182
French	228	228	197	200	128	128	142	160	149	166	170	178
French	232	232	197	200	128	152	171	175	170	174	170	178
French	228	228	197	197	142	144	142	171	166	166	178	182
French	228	240	192	200	126	144	160	167	166	178	178	178
French	228	240	197	200	126	128	142	160	170	174	178	186
French	228	240	188	200	128	144	164	167	166	170	178	178
French	232	310	200	200	128	128	160	167	166	174	170	178
French	228	287	197	200	128	144	142	164	149	166	178	182
French	232	297	170	170	128	144	142	160	166	170	178	186
French	240	282	200	200	128	144	160	160	166	174	170	178
French	228	240	197	200	128	128	142	164	149	170	178	178
French	228	232	197	197	128	156	155	167	149	166	178	182
French	240	254	197	197	130	130	155	160	166	174	178	182
French	232	287	197	200	126	126	160	164	149	174	178	182
French	228	232	200	200	130	142	142	171	149	166	178	178
French	232	232	192	200	126	152	142	160	166	166	178	190
French	228	232	197	200	130	154	160	164	149	166	186	186

Table A.3. Table of raw scoring data from PeakScanner. CAM = Kokosing R., Ecam = Honick dissertation samples, FCCAM = French Cr.

Sample	Eca46		Eca70		CV24		Eca48		Eca11		Esc132b	
CAM2	227.54	240.48	201.52	210.13	130.34	130.34	159.56	159.56	166.61	170.13	182.41	182.41
CAM3	279.02	281.99	197.38	201.44	130.35	144.60	159.63	159.63	166.70	170.28	185.65	185.65
CAM4	227.97	271.19	201.51	210.23	128.41	155.37	163.67	166.42	149.61	166.42	177.41	181.38
CAM5	229.99	250.00	197.48	197.48	130.35	130.35	159.70	171.56	166.04	166.04	178.19	181.65
CAM6	250.00	281.07	197.31	197.31	130.35	144.56	159.54	171.45	166.35	170.37	182.04	182.04
CAM7	237.80	250.06	210.16	210.16	130.28	130.28	159.47	166.29	166.29	170.26	170.05	182.08
CAM8	250.00	250.00	197.29	201.34	128.65	130.54	160.53	171.57	166.39	170.30	170.17	181.91
CAM9	227.66	282.02	196.27	196.27	130.10	154.85	159.64	159.64	149.93	149.93	181.85	186.20
CAM10	223.27	282.04	197.32	197.32	130.28	130.28	171.32	159.57	149.68	170.12	181.42	185.40
CAM11	237.75	250.00	201.35	197.26	130.26	154.91	166.36	166.36	166.93	166.93	169.55	169.55
CAM12	240.52	240.52	197.57	197.57	128.38	130.24	142.08	159.55	170.59	166.82	177.50	181.54
CAM13	237.86	250.00	201.35	201.35	127.10	128.25	171.37	159.64	150.30	170.30	170.01	181.97
CAM14	227.63	240.64	197.26	210.07	130.16	154.90	175.38	159.57	150.08	170.23	181.36	185.37
CAM15	282.84	282.84	201.44	197.32	128.42	130.26	159.63	175.43	166.52	166.52	177.42	181.37
CAM16	223.35	227.62	201.33	197.31	130.30	130.30	166.35	166.35	166.35	166.35	177.43	181.43
CAM17	240.60	282.06	201.51	201.51	130.26	154.92	171.63	171.63	166.43	170.49	169.47	169.47
CAM18	231.98	240.59	197.35	197.35	128.41	130.26	159.55	166.37	170.31	166.23	177.82	177.82
CAM19	227.72	240.70	197.34	197.34	130.29	154.95	159.63	171.53	149.34	166.41	169.42	185.43
CAM20	227.64	240.58	201.26	209.96	130.24	130.24	163.55	171.36	150.38	166.03	177.49	181.52
CAM21	227.71	227.71	197.29	201.41	130.22	130.22	159.64	171.45	169.78	150.61	182.47	182.47
CAM22	227.61	227.61	197.49	197.49	130.29	155.08	159.47	159.47	149.68	166.24	152.77	185.47
CAM23	227.68	240.67	196.35	196.35	130.33	130.33	159.46	171.53	166.03	166.03	170.41	186.21
CAM24	227.63	240.60	197.30	201.34	130.35	154.81	171.43	175.43	166.85	166.85	170.16	178.03
CAM25	227.63	282.04	192.37	196.29	130.15	130.15	159.59	167.46	150.58	150.58	182.55	182.55
CAM26	225.66	250.06	197.14	197.14	128.37	128.37	159.46	159.46	149.52	166.27	182.42	186.30
Ecam1	286.87	302.17	197.49	197.49	130.27	155.21	159.87	163.93	166.61	170.65	174.43	178.51
Ecam2	224.60	228.95	193.69	197.62	128.35	128.35	142.47	171.59	166.19	174.18	178.40	178.40
Ecam3	225.91	250.00	119.27	197.97	128.32	155.28	163.90	171.81	166.60	166.60	178.39	182.41
Ecam4	228.96	242.16	193.94	197.93	130.54	130.54	142.23	171.88	149.86	166.75	170.27	178.45
Ecam5	290.21	297.68	197.79	197.79	144.72	157.34	167.51	167.51	150.00	178.31	182.11	182.11
Ecam6	228.24	232.77	197.76	201.94	130.16	155.25	159.72	163.61	149.64	166.49	178.25	182.27
Ecam7	232.80	253.85	197.43	201.64	130.25	155.14	167.71	171.62	149.63	166.45	177.69	181.65
Ecam8	296.02	301.86	0.00	0.00	175.18	182.87	179.76	191.65	170.51	170.51	169.89	173.96

Sample	Eca46		Eca70		CV24		Eca48		Eca11		Esc132b	
Ecam9	262.38	293.69	144.27	144.27	122.48	180.80	179.60	191.50	125.04	170.32	200.35	200.35
Ecam10	288.50	295.90	197.73	197.73	130.13	130.13	167.60	171.49	149.63	149.63	170.13	178.26
Ecam11	204.73	230.64	197.79	197.79	130.15	130.15	159.72	175.57	166.50	174.51	174.30	178.36
Ecam12	297.36	297.36	0.00	0.00	178.94	180.81	183.41	187.36	170.43	170.43	169.80	169.80
Ecam13	230.40	235.21	197.70	201.87	130.13	155.00	155.36	167.45	166.30	170.28	178.57	182.63
Ecam14	230.43	244.07	197.72	201.86	130.18	144.53	142.14	171.54	166.31	174.35	178.52	178.52
Ecam15	230.43	244.01	197.69	197.69	128.21	130.05	159.64	171.49	149.54	166.33	178.62	178.62
Ecam16	227.93	293.62	170.19	196.82	129.64	144.02	163.53	171.41	149.20	174.01	178.20	194.50
Ecam17	244.34	317.08	197.75	201.94	128.63	130.52	159.55	163.58	149.36	174.07	178.43	186.54
Ecam18	244.39	256.71	193.69	197.71	155.11	155.11	159.57	159.57	166.31	174.35	182.36	186.41
Ecam19	286.44	286.44	144.29	144.29	151.37	151.37	183.49	183.49	163.19	163.19	200.35	243.88
Ecam20	235.57	295.64	197.76	197.76	128.30	130.22	146.57	167.59	149.07	169.92	178.45	182.47
Ecam21	230.48	256.48	193.70	201.86	130.08	130.08	159.64	171.49	149.55	170.35	177.95	181.92
Ecam22	250.00	293.17	197.66	197.66	130.14	144.41	142.03	171.44	162.43	169.88	178.00	186.02
Ecam23	229.53	242.96	197.79	197.79	155.15	157.11	142.08	163.61	149.62	166.47	169.87	177.90
Ecam24	235.10	235.10	197.71	197.71	155.11	157.16	155.40	171.59	170.43	170.43	181.96	181.96
Ecam25	230.52	316.97	197.76	197.76	130.13	130.13	146.53	159.58	149.55	170.33	178.28	182.26
Ecam26	230.45	244.11	201.85	201.85	130.12	144.41	142.03	159.64	149.48	170.43	178.23	178.23
Ecam27	254.07	297.97	197.70	197.70	130.15	155.04	142.06	163.42	166.28	170.30	170.04	178.22
Ecam28	230.67	235.48	197.73	201.97	130.15	144.41	150.97	167.54	149.47	174.37	170.14	178.14
Ecam29	237.41	241.72	201.86	201.86	130.13	157.23	159.65	171.55	166.39	170.41	170.29	178.40
Ecam30	228.54	228.54	197.66	197.66	128.24	130.18	142.09	159.72	149.55	166.44	178.48	186.63
Ecam31	224.25	241.73	197.70	201.82	128.25	130.09	159.58	171.45	145.18	170.30	170.38	178.50
Ecam32	227.89	293.57	197.72	197.72	144.57	155.38	159.85	171.71	149.92	170.65	170.26	178.40
Ecam 33	227.80	232.16	195.58	199.49	129.57	129.57	141.87	171.52	149.09	153.66	178.18	182.17
Ecam 34	227.81	232.22	170.03	199.63	129.53	154.76	163.59	171.54	166.12	178.11	166.08	178.13
Ecam 35	227.88	232.32	197.39	197.39	129.60	154.69	159.61	171.57	166.19	170.19	181.18	185.13
Ecam 36	225.79	250.09	191.79	191.79	129.58	154.69	159.61	163.68	166.27	174.33	177.49	177.49
Ecam 37	227.67	227.67	195.66	199.59	129.47	152.40	141.76	155.24	149.22	170.05	153.60	178.11
Ecam 38	0.00	0.00	199.59	199.59	127.63	129.54	141.85	167.35	170.30	170.30	182.12	186.09
Ecam 39	232.05	297.61	197.30	201.46	127.71	154.56	159.45	167.38	170.20	170.20	177.18	177.18
Ecam 40	225.72	250.05	170.00	199.59	129.50	129.50	141.85	163.52	149.68	149.68	169.36	181.19
Ecam 41	225.65	249.95	169.97	187.75	129.52	129.52	167.44	171.50	170.09	174.13	170.03	186.10
Ecam 42	237.82	250.05	187.76	195.66	129.58	129.58	141.85	163.61	149.27	166.11	178.19	182.11
Ecam 43	240.55	301.66	197.39	201.56	154.59	154.59	159.41	167.31	149.20	165.99	177.41	177.41

Sample	Eca46		Eca70		CV24		Eca48		Eca11		Esc132b	
Ecam 44	250.00	314.72	170.44	197.21	154.53	154.53	159.40	171.50	149.21	166.09	178.00	185.22
Ecam 45	227.61	231.97	197.42	197.42	129.51	154.58	141.75	171.42	166.01	170.01	178.47	182.07
Ecam 46	228.57	233.08	196.35	200.36	152.43	152.43	163.47	175.34	170.09	177.89	178.19	178.19
Ecam 47	225.74	250.09	169.95	195.74	127.64	143.79	141.79	171.43	166.00	170.05	178.19	182.21
Ecam 48	237.73	249.95	196.36	200.36	129.61	152.39	141.89	159.49	149.15	166.10	177.20	185.22
Ecam 49	228.65	298.00	192.82	196.75	129.49	154.45	163.49	167.39	170.02	173.99	169.83	178.76
Ecam 50	278.56	282.37	197.37	197.37	130.39	155.30	159.85	163.95	169.86	169.86	178.00	186.01
Ecam 51	227.60	306.06	196.40	200.36	127.66	154.43	159.35	163.41	149.11	169.93	178.09	186.06
Ecam 52	229.77	250.00	170.09	195.73	129.48	143.83	163.57	171.51	149.31	166.12	177.27	181.25
Ecam 53	225.71	250.09	197.44	201.59	154.51	154.51	167.43	167.43	149.26	166.05	177.80	181.96
Ecam 54	231.91	231.91	169.98	195.72	154.49	156.54	159.36	163.43	149.16	165.92	178.15	186.17
Ecam 55	293.82	293.82	170.04	195.70	129.52	129.52	141.78	167.44	149.10	170.05	178.33	178.33
E cam 56	228.74	279.04	200.46	205.11	130.23	144.65	160.29	160.29	150.16	174.86	178.24	183.21
Ecam 57	279.04	282.75	200.43	200.43	130.26	130.26	167.44	171.56	170.38	170.38	182.27	186.26
E cam 58	225.79	250.05	197.54	197.54	128.35	144.57	159.62	163.65	166.43	170.44	178.25	182.22
Ecam 59	227.96	227.96	205.28	205.28	130.21	155.19	163.85	171.68	166.59	174.63	170.28	182.34
Ecam60	228.48	232.84	199.75	208.80	130.20	153.11	163.71	171.55	166.52	174.51	170.35	182.27
Ecam61	228.44	232.75	199.76	199.76	130.53	155.31	160.10	172.09	150.42	167.50	178.37	182.34
Ecam62	228.39	228.39	199.60	204.15	156.16	156.16	160.10	160.10	171.74	171.74	170.38	178.25
Ecam63	228.51	293.91	199.75	199.75	131.08	131.08	142.69	167.89	175.51	175.51	162.13	182.29
Ecam64	291.81	300.08	199.41	199.41	156.08	156.08	160.00	171.94	150.50	175.14	174.24	178.31
Ecam65	232.70	301.48	204.02	204.02	131.10	153.89	160.09	164.10	167.44	183.44	170.11	178.15
Ecam66	228.30	241.16	199.63	199.63	128.65	130.54	160.09	164.07	162.84	175.34	178.24	182.26
Ecam67	228.41	286.31	199.41	203.84	131.06	156.08	171.85	171.85	167.43	170.81	170.44	182.25
Ecam68	233.17	297.67	199.14	203.56	130.13	155.12	141.72	171.41	166.03	173.98	170.22	178.24
Ecam69	250.00	281.10	199.03	203.40	155.24	155.24	141.95	159.61	162.37	166.38	182.22	182.22
Ecam70	250.00	276.98	203.86	203.86	130.25	130.25	142.14	155.44	149.83	166.44	174.19	186.18
Ecam71	233.36	286.93	199.24	203.57	128.28	130.14	141.91	159.61	149.67	166.44	178.25	182.26
Ecam72	228.55	293.91	203.35	203.35	130.06	155.30	151.28	171.30	149.66	166.09	178.33	182.28
Ecam73	237.94	250.00	197.66	197.66	130.91	145.17	167.78	171.81	175.39	175.39	182.52	182.52
Ecam74	224.20	241.49	199.02	203.41	128.50	130.40	171.72	171.72	167.35	167.35	178.24	182.22
Ecam75	233.34	286.93	199.04	203.47	131.01	131.01	164.01	171.92	150.49	166.83	178.31	186.30
Ecam76	232.03	232.03	197.55	201.75	130.18	130.18	159.60	171.55	166.62	170.33	178.39	186.33
Ecam77	232.72	241.32	197.65	197.65	130.07	157.25	141.99	171.34	150.40	150.40	166.14	170.17
Ecam78	290.20	297.67	170.61	197.55	130.12	144.68	159.77	159.77	149.66	166.22	169.66	181.74

Sample	Eca46		Eca70		CV24		Eca48		Eca11		Esc132b	
Ecam79	000.00	000.00	197.57	197.57	127.77	152.49	159.45	171.43	165.93	165.93	182.34	186.35
Ecam80	224.13	228.44	197.52	201.71	127.80	144.03	159.56	163.59	149.54	165.97	182.25	186.13
Ecam81	238.01	250.07	197.67	197.67	154.76	154.76	159.50	167.43	166.07	174.10	178.34	182.33
Ecam82	233.42	233.42	189.70	197.67	129.72	129.72	159.50	171.47	166.38	174.35	178.12	186.12
Ecam83	228.96	306.74	189.49	197.39	129.67	154.59	163.45	171.39	174.44	174.44	177.54	185.60
Ecam84	233.45	233.45	193.47	197.45	129.69	143.97	163.49	166.01	166.49	166.49	178.09	182.11
Ecam85	290.16	297.63	197.44	201.56	129.66	129.66	141.83	159.43	166.41	166.41	170.12	182.13
Ecam86	233.56	233.56	197.48	197.48	129.71	129.71	141.85	171.36	149.58	166.17	177.61	177.61
Ecam87	229.02	302.38	197.36	201.49	154.64	154.64	159.44	171.36	174.41	174.41	170.21	178.24
Ecam88	229.02	306.82	197.56	197.56	129.57	143.97	141.78	159.50	149.11	166.01	178.04	178.04
Ecam89	228.49	282.74	201.57	201.57	127.72	156.72	141.77	179.36	149.70	149.70	178.19	190.11
Ecam90	300.00	300.00	197.47	201.64	129.69	154.75	163.54	167.40	149.58	165.90	178.12	181.33
Ecam91	282.77	293.87	197.42	201.56	127.77	129.64	159.44	167.41	149.53	166.30	178.07	186.09
Ecam92	227.97	227.97	189.38	197.36	144.00	154.65	141.85	171.32	166.25	169.86	173.39	177.81
Ecam93	224.76	254.74	197.41	197.41	127.75	129.69	159.44	171.42	166.39	166.39	178.23	186.21
Ecam94	226.00	249.96	197.41	201.50	129.71	129.71	141.83	159.49	166.26	174.24	174.38	182.46
Ecam95	227.94	245.10	197.33	201.43	129.68	143.91	163.45	171.37	169.81	174.22	174.08	276.04
Ecam96	225.97	250.07	197.47	201.56	129.62	129.62	163.55	171.49	149.48	166.36	174.04	276.88
Ecam97	227.38	297.00	197.40	201.50	129.63	143.94	171.40	171.40	149.29	166.08	177.74	173.37
Ecam98	227.82	240.55	197.70	197.70	154.70	154.70	159.50	167.56	170.13	178.09	169.88	177.97
Ecam99	232.88	290.20	197.38	201.51	129.74	154.74	141.94	163.52	166.47	166.47	178.25	178.25
Ecam100	228.04	228.04	189.37	201.50	129.72	129.72	141.92	163.51	165.99	174.01	177.80	181.95
Ecam101	228.05	290.23	197.42	201.59	128.63	129.75	159.56	171.45	149.54	149.54	174.07	177.76
Ecam102	250.00	250.00	196.67	200.72	127.68	154.27	167.35	171.33	153.46	173.96	177.61	185.81
Ecam103	223.68	241.00	182.61	196.71	129.63	154.11	159.41	167.33	149.16	173.96	182.14	182.14
Ecam104	229.81	250.10	196.38	200.44	129.53	154.29	159.41	163.47	149.20	173.96	181.48	181.48
Ecam105	232.42	297.72	000.00	000.00	129.54	129.54	159.36	171.31	149.11	169.97	169.79	177.80
Ecam106	227.77	293.90	192.38	196.35	129.64	154.59	171.33	171.33	149.18	170.01	177.84	181.78
Ecam107	227.97	227.97	182.73	182.73	127.70	129.59	159.37	171.31	149.55	149.55	173.93	177.85
Ecam108	227.76	240.68	196.35	200.36	154.15	156.58	159.41	171.34	166.04	170.00	169.58	173.71
Ecam109	229.84	250.00	182.62	182.62	129.59	154.53	141.80	167.40	166.31	166.31	178.14	181.99
Ecam110	232.07	293.73	200.71	200.71	129.51	129.51	159.43	167.38	166.02	170.01	177.80	181.78
Ecam111	227.66	232.06	197.56	201.66	129.60	154.32	159.47	163.49	149.33	170.04	170.01	177.97
Ecam112	227.76	227.76	192.64	196.60	129.61	154.45	163.55	171.42	166.30	166.30	174.08	181.97
Ecam113	240.67	286.32	196.40	200.36	127.69	129.59	163.46	163.46	149.56	149.56	177.85	181.82

Sample	Eca46		Eca70		CV24		Eca48		Eca11		Esc132b	
Ecam114	223.38	293.82	182.78	186.83	129.57	154.56	159.39	171.40	149.51	149.51	177.83	185.77
Ecam115	225.67	249.95	196.41	200.43	127.78	143.97	159.42	171.42	166.05	178.04	181.84	181.84
Ecam116	240.54	293.85	182.77	196.71	128.21	130.22	159.47	175.46	166.07	170.08	177.87	177.87
Ecam117	232.06	297.62	182.59	196.75	127.68	143.84	163.43	167.38	149.13	173.95	169.97	189.76
FCCAM1	224.03	253.91	197.49	197.49	153.76	155.00	159.80	167.58	166.49	166.49	182.05	186.19
FCCAM2	232.63	241.26	197.47	197.47	128.34	130.22	141.87	171.48	166.12	170.14	178.37	178.37
FCCAM3	228.33	241.36	197.53	197.53	127.13	129.06	159.73	171.58	166.44	174.46	174.01	182.12
FCCAM4	232.70	301.74	197.30	197.30	128.96	128.96	146.51	171.58	170.42	170.42	178.33	178.33
FCCAM5	228.23	253.67	193.45	201.53	127.02	156.16	163.76	171.57	149.64	166.47	178.22	178.22
FCCAM6	228.24	241.12	201.53	201.53	128.88	128.88	142.06	167.60	166.51	170.49	182.06	186.19
FCCAM7	228.37	228.37	197.41	197.41	128.95	156.21	159.73	167.64	166.48	166.48	178.13	178.13
FCCAM8	227.73	227.73	197.41	201.47	128.58	155.39	159.66	167.49	149.64	166.47	177.91	181.99
FCCAM9	228.21	228.21	197.45	201.52	128.40	128.40	141.92	159.58	149.63	166.28	170.03	178.11
FCCAM10	232.61	232.61	197.41	201.50	130.38	155.39	171.51	175.52	170.40	174.42	170.04	178.14
FCCAM11	228.27	228.27	197.34	197.34	142.21	144.52	141.83	171.43	166.68	166.68	178.14	182.15
FCCAM12	228.22	241.15	192.62	200.56	128.58	144.82	159.66	167.60	166.44	178.38	178.22	178.22
FCCAM13	228.25	241.18	196.35	200.35	126.98	128.85	142.06	159.59	170.37	174.44	178.10	186.06
FCCAM16	228.23	241.15	188.60	200.56	128.90	143.09	163.70	167.60	166.43	170.45	178.18	178.18
FCCAM17	232.58	310.39	201.54	201.54	129.14	129.14	159.66	167.58	166.48	174.48	170.15	178.13
FCCAM18	228.23	286.47	197.37	201.40	128.92	143.06	142.06	163.69	149.57	166.40	178.50	182.02
FCCAM19	232.61	297.63	193.44	201.48	128.88	143.09	142.08	159.66	166.43	170.45	177.94	186.04
FCCAM20	241.19	282.68	201.51	201.51	128.90	143.16	159.59	159.59	166.41	174.42	170.11	178.09
FCCAM21	228.20	241.17	197.38	201.51	128.88	128.88	142.01	163.65	149.57	170.34	178.15	178.15
FCCAM22	228.18	232.63	197.44	197.44	129.07	156.14	155.25	167.53	149.57	166.43	178.01	182.11
FCCAM23	241.19	253.68	197.36	197.36	130.16	130.16	155.50	159.79	166.44	174.48	177.99	182.08
FCCAM24	232.63	286.41	197.40	201.50	126.99	126.99	159.59	163.61	149.56	174.33	178.09	182.13
FCCAM27	228.24	232.65	200.57	200.57	130.50	130.50	141.90	171.52	150.28	166.02	178.22	178.22
FCCAM31	232.47	232.47	192.59	200.57	127.02	153.74	142.01	159.59	166.40	166.40	178.01	190.15
FCCAM35	228.21	232.61	196.51	200.49	130.54	155.50	159.78	163.86	149.92	166.66	186.04	186.04

Table A.4. Observed (H_O) and expected heterozygosity (H_E) of the loci that were out of HWE

	Deer (A)		Allegheny		Bull		French	
Locus	**Ho**	**He**	**Ho**	**He**	**Ho**	**He**	**Ho**	**He**
Eca46	0.81481	0.89099	0.89655	0.91228	0.64286	0.88889	0.7600	0.77388
Eca70	0.55556	0.73305	0.53333	0.65763	0.6000	0.55172	0.4800	0.60653
CV24	0.62963	0.8225	0.73333	0.7339	0.66667	0.76322	0.7200	0.79347
Eca48	0.77778	0.79804	1.000	0.83277	1.000	0.81839	0.9600	0.82449
Eca11	0.77778	0.76520	0.83333	0.74972	0.53333	0.64138	0.800	0.69878
Esc132b	0.88889	0.74633	0.800	0.68814	0.86667	0.71954	0.6400	0.62204

Table A.5. Summary table of ΔK analysis from Evanno (2005). The shaded run indicates the most probable number of clusters at K = 3. Selection of the most probable K is seen in the large rate of change that occurs between ΔK calculations of K = 3 and K = 4.

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	15	-3470.927	0.260	NA	NA	NA
2	15	-3617.353	40.357	-146.427	290.767	7.205
3	15	-3473.013	8.018	144.340	174.913	21.814
4	15	-3503.587	9.889	-30.573	20.467	2.070
5	15	-3554.627	12.133	-51.040	52.773	4.349
6	15	-3658.440	93.540	-103.813	62.793	0.671
7	15	-3825.047	102.756	-166.607	9.173	0.089
8	15	-3982.480	67.750	-157.433	35.547	0.525
9	15	-4104.367	68.610	-121.887	24.120	0.352
10	15	-4202.133	101.948	-97.767	NA	NA

Table A.6. Fish species and abundance from electrified-benthic trawling in 2013. Bold indicates target species. LD = lock and dam, BC = back channel.

Species	River	Allegheny	Beaver	Monongahela	Ohio		
	Lock/Dam	LD2	Dam1	LD2/Braddock	Emsworth	Emsworth BC	Montgomery
A. grunniens			2			1	1
E. blennioides				1		1	
E. camurum		**4**	**3**	**2**		**3**	
E. flabellare		2		3		1	
E. nigrum		1		1	3	2	
E. tippecanoe		**2**		**2**		**16**	**2**
E. zonale			6	11		1	
H. nigricans			1				
I. bubalus							
I. niger					1		
I. punctatus			1	1			1
M. anisurum			2				
M. dolomieu		1				1	
N. vol-wickliffi					6	2	
N. volucellus		9		113	4	12	
P. caprodes		4	6	2	2	4	
P. copelandi		53		204	56	37	3
P. evides		2				9	
P. shumardi				1		1	19
S. canadensis		1			1		
Total		79	21	341	73	91	26

Table A.7. Fish species and abundance from electrified-benthic trawling in 2014. Bold indicates target species. LD = lock and dam.

Species	River	Allegheny				Monongahela
	Lock/Dam	LD3	LD4	LD5	LD6	LD3/Elizabeth
A. grunniens			4		3	3
E. blennioides			2			2
E. camurum		**37**	**56**	**67**	**23**	
E. flabellare						5
E. maculatum		**1**				
E. nigrum		1	6			16
E. tippecanoe		**37**	**41**	**13**	**43**	
E. variatum		3	1	2		
E. zonale		2	6	2	2	3
H. amblops			1			
H. tergisus			1		3	
I. punctatus			2	2	9	1
I. niger			1	1		
M. dolomieu			1			
M. erythrurum			2		1	
M. macrolepidotum			5	1	2	
N. flavus		1				
N. volucellus			7			2
N. wickliffi			1			7
P. caprodes		3	19	5	11	5
P. copelandi		113	505	103	84	39
P. evides		7	20	10	15	
P. flavescens		1				
P. macrocephala			1			
P. notatus		4	1			
P. omiscomaycus		14				
S. canadensis						2
Total		224	683	206	196	85

Table A.8. Fish species and abundance from tributary backpack electrofishing in 2013. Bold indicates target species.

Species	**Stream** **Tributary to:**	DeerA(2013) AlleghenyR	DeerB(2012) AlleghenyR	Flaugherty Run OhioR	Kiski(A) AlleghenyR	Kiski(B) AlleghenyR	Ltl Sewickley OhioR	Taylor Run AlleghenyR
C. anomalum		1		10			6	1
C. commersonii		6		9			3	3
C. bairdii		2	2	74			124	
C. spiloptera			15			6		2
E. blennioides		10	43		3	4		
E. caeruleum		32	38	27			36	8
E. camurum		**27**	**11**		**10**	**17**	**13**	
E. flabellare		3						
E. nigrum			6					
E. tippecanoe		**2**			**25**	**31**		
E. variatum					7	11		
E. zonale		8	33	28	48	64		
H. amblops								1
H. nigricans		6					1	
M. dolomieu				2				1
N. micropogon							1	
N. atherinoides			55					
N. rubellus					2	33		19
N. volucellus			9		16	17	9	15
P. caprodes		1						
P. notatus		1	11	1			1	35
R. atratulus		1	1	109			24	40
R. cataractae		1		9			35	
S. canadensis			1					
S. atromaculatus				47			16	8
Total		101	225	316	111	183	269	133

Table A.9. Fish species and abundance from tributary backpack electrofishing in 2014. Bold indicates target species.

Species	**Stream Tributary to:**	Bull Cr AlleghenyR	Chartier's Run AlleghenyR	Montour Run OhioR	Moon Run OhioR	Raccoon Cr OhioR
C. anomalum		2	1	4	2	13
C. commersonii			1			
C. bairdii			40			
C. spiloptera		1		9		
D. cepedianum						11
E. blennioides		1	2			55
E. caeruleum		3	69		10	93
E. camurum		**15**		**2**	**1**	
E. flabellare		2	1	1	4	
E. tippecanoe		**1**				
E. variatum						7
E. zonale						175
F. cingulatus					1	
H. nigricans		1		8	3	3
L. chrysocephalus						2
M. dolomieu				1		
N. atherinoides		5				
N. rubellus		201				
N. stramineus						12
N. volucellus		44		3		
N. flavus						9
P. caprodes		1				1
P. notatus		5		4		3
R. atratulus			2			
R. cataractae			2			1
S. atromaculatus					7	
Total		282	118	32	28	385

Table A.10. Species and GenBank accession numbers used to create the customized cross-amplification database in PrimerBLAST.

Species	Gene	Accession #s		
		ND2	***COI***	***Cytb***
Ammocrypta pellucida		JQ088502	JN024788	FJ381008
Etheostoma blennioides		FJ381263	HQ579050	HQ128093
		JQ088546	JN025638	HQ128092
			JN025622	AF288426
			JN025621	EU296665
			EU524017	AF386539
Etheostoma caeruleum		FJ381267	JN025719	EU046707
		FJ381265	JN025716	DQ465226
		FJ381268	JN025715	DQ465225
		FJ381266	JN025710	DQ465224
		EF027187	EU524023	DQ465200
			EU524022	DQ465199
Etheostoma camurum		EU814337	JN025723	AF045348
		EU814335	JN025724	GU015083
		EU814333	JN025722	GU015082
		EU814331	JN025721	EU094672
		EU814329		EU094671
		EU814327		EU094670
		EU814366		
		EU814364		
		EU814362		
		EU814360		
		EU814358		
		EU814356		
		EU814354		
		EU814352		
		EU814350		
		EU814348		
		EU814346		
		EU814344		
		EU814342		
		EU814340		
		EU814338		
		EU814336		
		EU814334		
		EU814332		
		EU814330		
		EU814328		
		EU814326		
		JQ088518		
		FJ381262		
Etheostoma exile		EF027194	JN025831	AF386541
Etheostoma flabellare		JQ088540	KF929867	HQ128131
		AF412540	HQ557469	AF045342
			HQ557272	AF412526
			JN025849	AF386544
			JN025848	
			EU524038	

Species	Gene	Accession #s		
		ND2	***COI***	***Cytb***
Etheostoma maculatum		EU810789	HQ557544	HQ005525
		EU810791	HQ557543	HQ005524
		EU810793	HQ557542	HQ005523
		EU810795		GU015095
		EU810797		GU015094
		EU810778		AY742663
		EU810780		
		EU810782		
		EU810784		
		EU810786		
		EU810788		
		EU810790		
		EU810792		
		EU810794		
		EU810796		
Etheostoma nigrum		JQ088561	JX516785	GQ183677
				AF183945
				GQ183676
				GQ183675
				GQ183674
				GQ183673
Etheostoma tippecanoe		EU814368	JN026471	AF274471
		EU814369	JN026470	AF274470
		EU814377	JN026469	GU015280
		EU814375		EU094715
		EU814373		EU094714
		EU814371		EU094713
		EU814406		
		EU814404		
		EU814402		
		EU814400		
		EU814398		
		EU814396		
		EU814394		
		EU814392		
		EU814390		
		EU814388		
		EU814386		
		EU814384		
		EU814382		
		EU814380		
		EU814378		
		EU814376		
		EU814374		
		EU814372		
		EU814370		
		EU814397		
		EU814395		

Species	Gene	Accession #s		
		ND2	***COI***	***Cytb***
Etheostoma variatum		EF027230	JN026501	AY964688
			JN026500	EU296693
			JN026499	AY964687
			JN026498	HQ128251
			JN026497	AF289266
			JN026496	
Etheostoma zonale		EF027233	KF929879	KF592447
			HQ579051	HQ128252
			JN026596	AY964706
			JN026595	AY964705
			JN026578	KF592394
			JN026576	KF592390
Perca flavescens		EF027169	JX517165	AF045357
			JX517139	AF386600
			JX517095	AY374280
			KC819884	AF546115
			EU524245	
			EU524244	
Percina caprodes		EF027178	JN027956	AY770841
		AY770850	JN027955	KC211182
		AY770849	JN027954	EU379095
		EU379081	KC819890	EU379094
		EU379080	EU524249	DQ493490
		EU379079	EU524248	DQ493489
		DQ493531		
Percina copelandi		AY770860	JN027975	AY374283
			JN027976	AF386568
			JN027974	
			JN027973	
			JN027972	
			EU524252	
Percina evides		JQ088622	JN027999	AF375955
		DQ493545	KF930249	AF375939
			JN027998	AF375938
			JN027997	AF375942
			JN027993	AF375952
			JN027984	
Percina macrocephala		JQ088628	JN028028	DQ493501
		DQ493546		
Percina maculata		AY517725	JN028051	AF045353
Percina oxyrhynchus		JQ088632	JN028079	KM209982
Percina shumardi		JQ088635	JN028137	AF386572
			JN028136	AF386571
			JN028135	
			JN028133	
			JN028132	
			EU524260	

Species	Gene	Accession #s		
		ND2	***COI***	***Cytb***
Sander canadensis		JQ088642	KC819869	KC819818
			KC819868	KC819817
			KC819867	KC819816
			EU524373	KC819815
			EU524372	KC819814
			EU524371	
Sander vitreus		JQ088645	KF930366	KC819821
			JN028405	KC819820
			JN028404	KC819819
			KC819874	KC819822
			KC819873	
			KC819872	

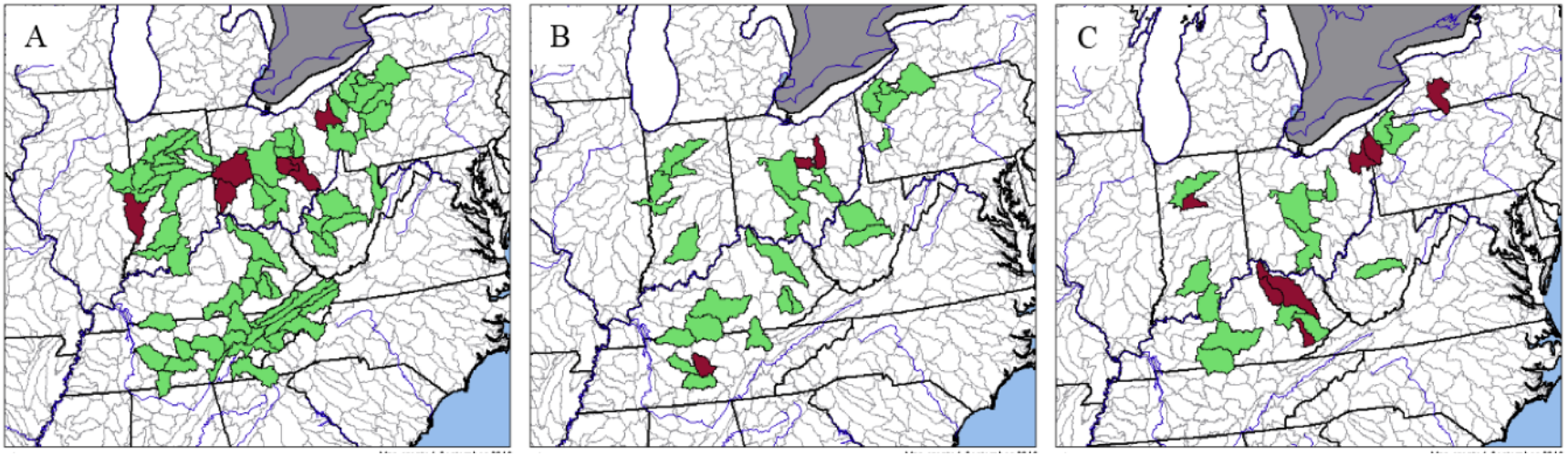

Figure A.1. NatureServe maps indicating the watersheds where Bluebreast Darter (Panel A), Tippecanoe Darter (Panel B), and Spotted Darter (Panel C) have been documented. Green watersheds are currently known distributions and red watersheds represent areas of extirpation.

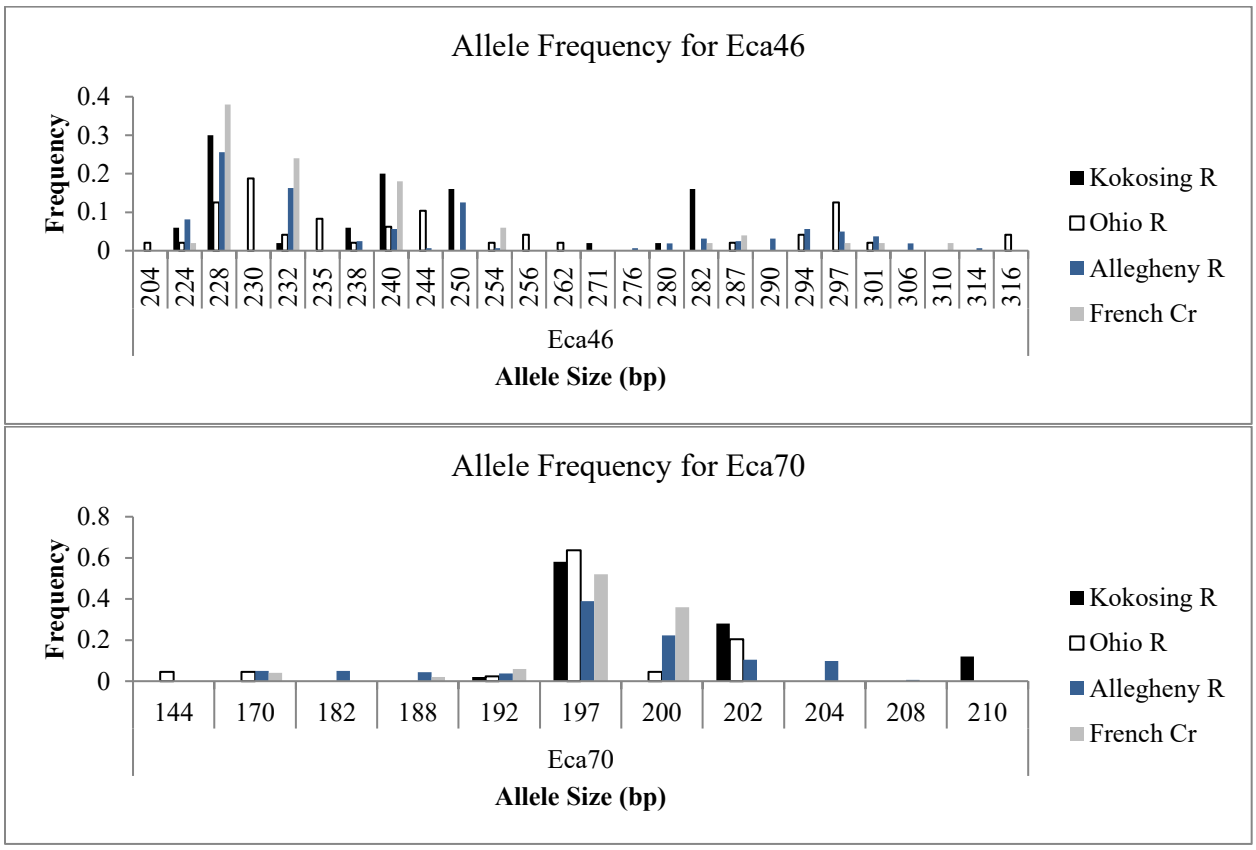

Figure A.2. Allele frequency histograms for loci Eca46 (top panel) and Eca7- (bottom panel).

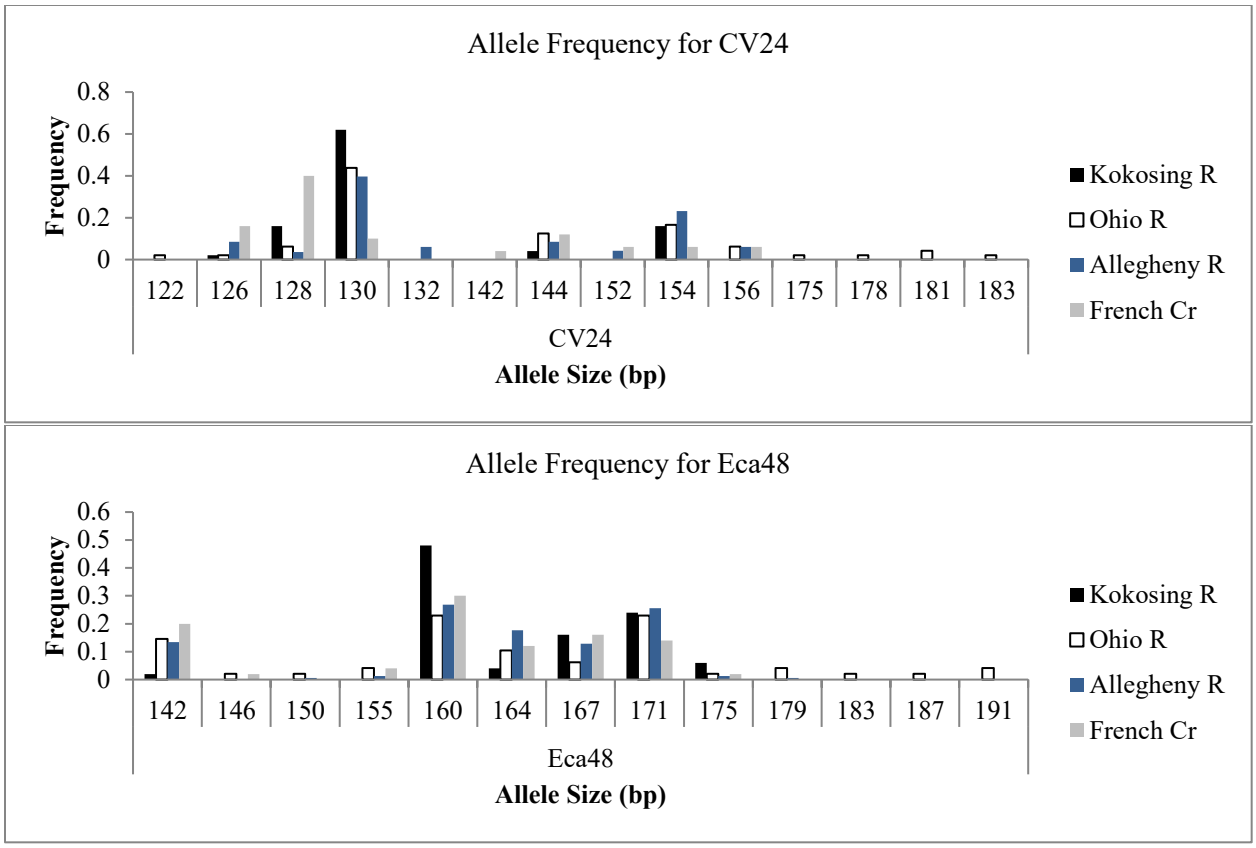

Figure A.3. Allele frequency histograms for loci CV24 (top panel) and Eca48 (bottom panel).

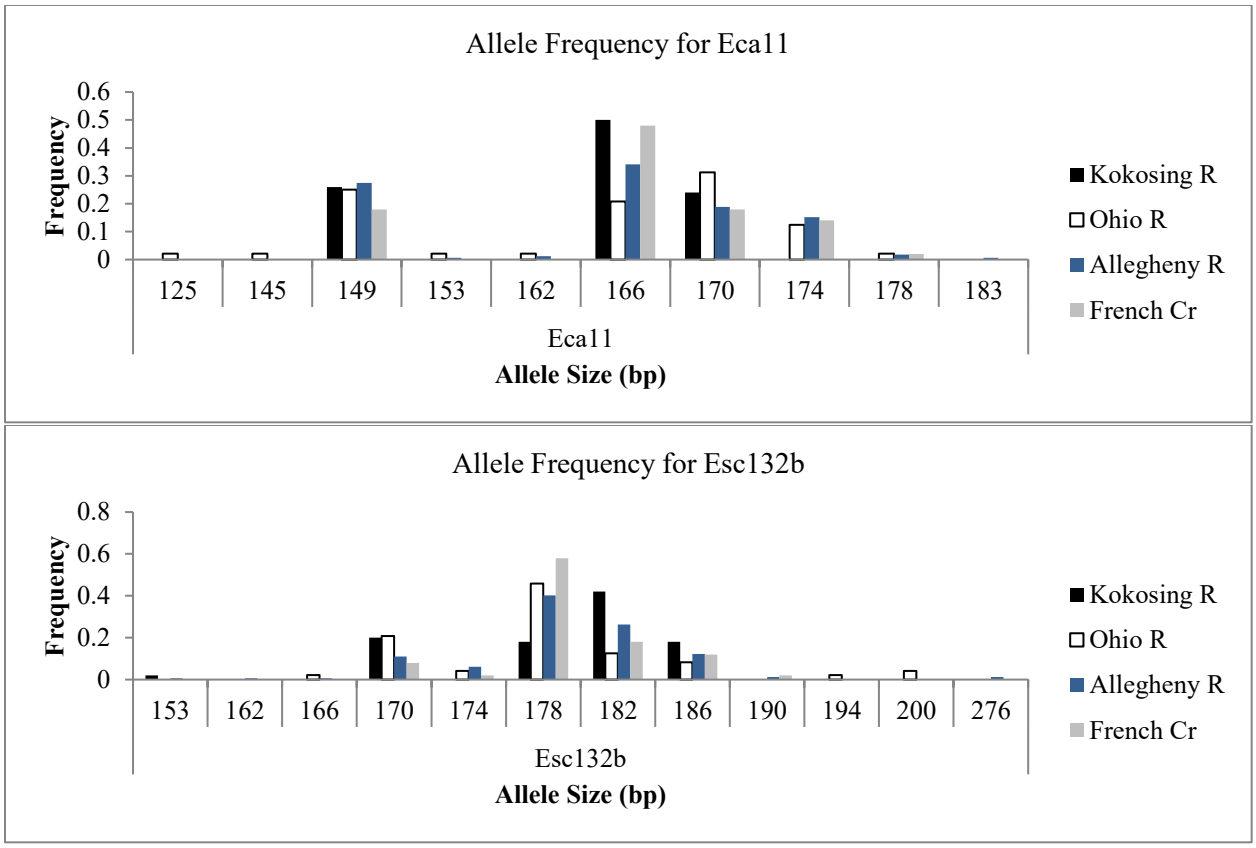

Figure A.4. Allele frequency histograms for loci Eca11 (top panel) and Esc132b (bottom panel).

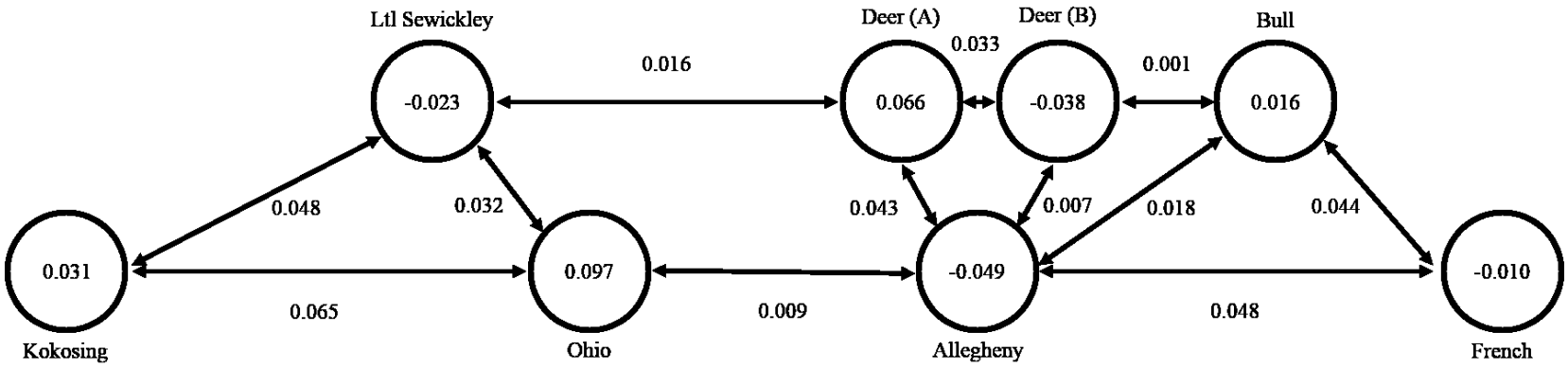

Figure A.5. Graphic representing the pairwise F_{ST} values (along the arrows) in a linearized manner that aids visualization of the differentiation of adjacent populations within the stream network (note: not all comparisons shown). The inbreeding coefficient (F_{IS}) is inside the circles.

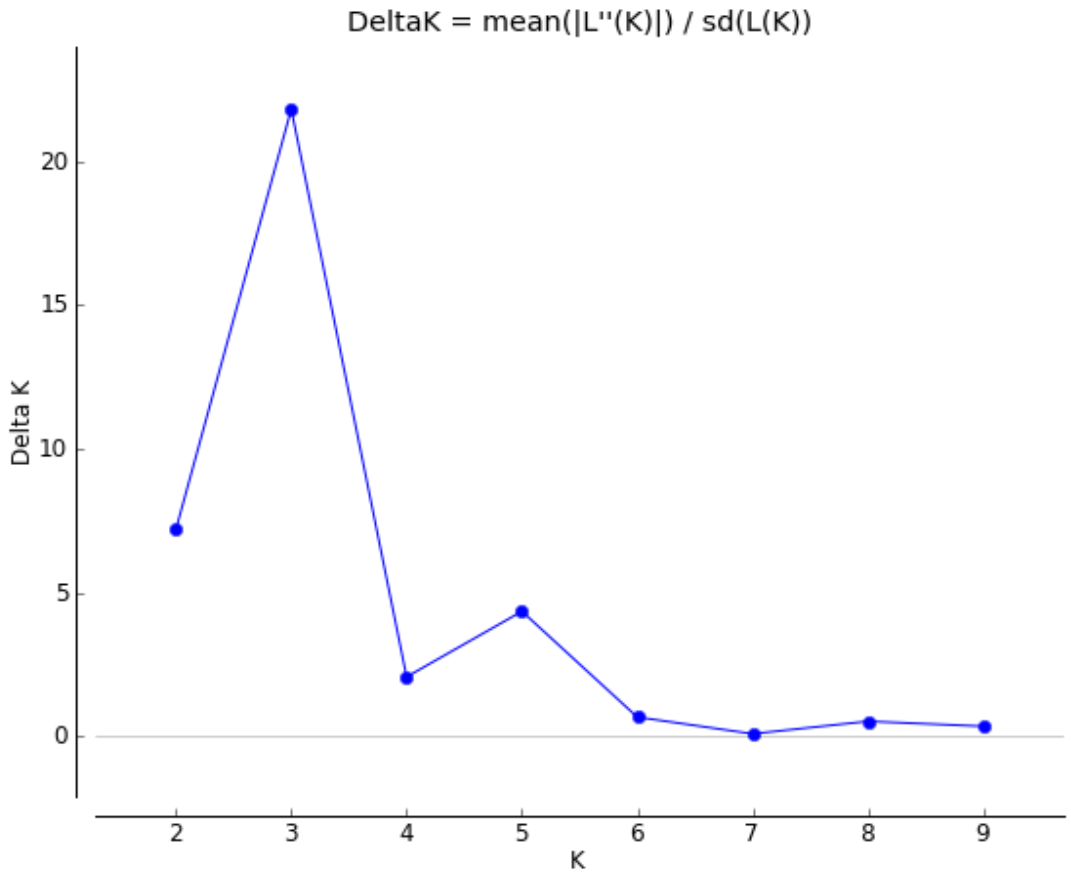

Figure A.6. Figure from Structure Harvester and the ΔK analysis of Evanno (2005) representing the rate of change between ΔK calculations of simulations K = 1 – 10 for all populations. The most likely number of clusters is signaled by the large drop in the rate of change between K = 3 and K = 4.

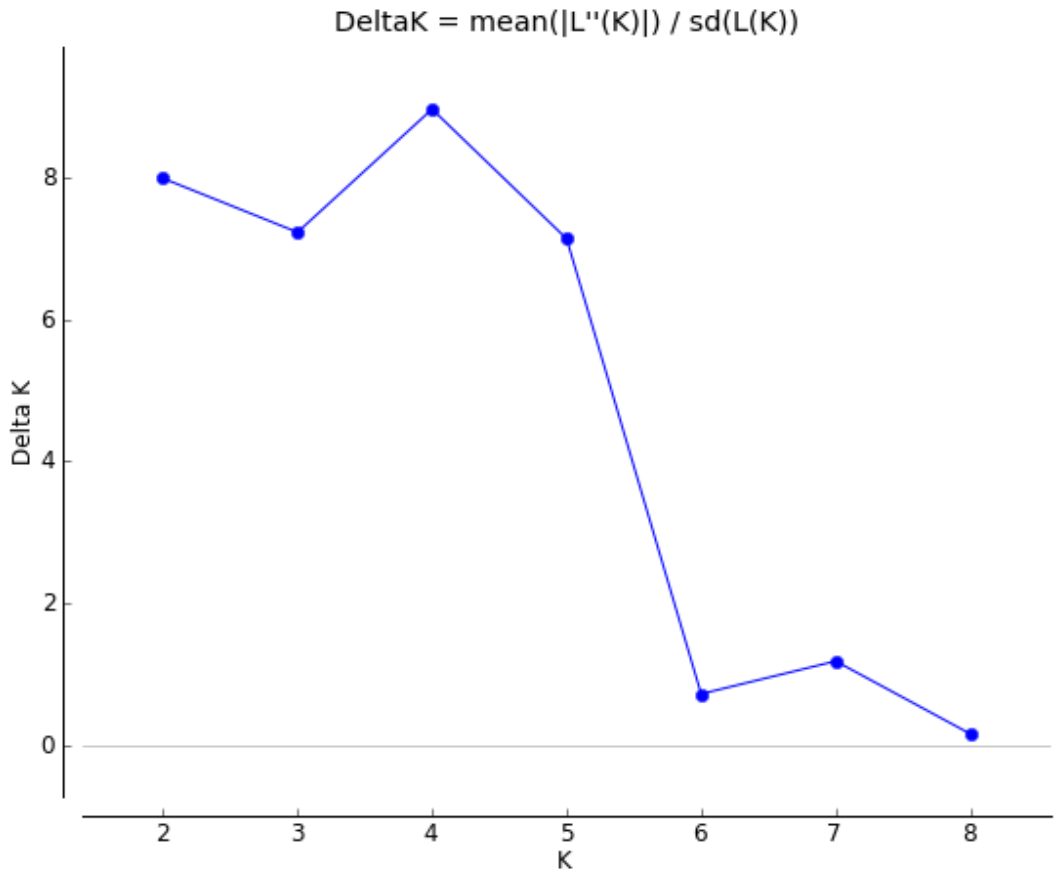

Figure A.7. Figure representing the ambiguous selection of K = 4 when Kokosing R. was removed from STRUCTURE analysis.

Appendix B:
Expanded Methods

Table B.1. Summary of the PCR master mix, thermocycle, and annealing temperature for each locus used for fragment analysis. Each reaction was in a final volume of 12.0 µL. All reagents (except the amount of DNA) are given in µL.

Reagent	Eca46	Eca70	CV24	Eca48	Eca11	Esc132b
Buffer B	1.2	1.2	1.2	1.2	1.2	1.2
dNTPs	2.0	1.92	2.0	2.0	2.0	1.92
$MgCl_2$	1.2	0.96	1.2	1.2	1.2	0.96
For	0.3	0.3	0.3	0.3	0.3	0.3
Rev	0.3	0.3	0.3	0.3	0.3	0.3
Taq	0.16	0.1	0.16	0.16	0.16	0.16
H_2O	4.84	4.22	4.84	4.84	4.84	5.16
DNA (ng)	24	36	24	24	24	24
Annealing T (°C)	53	45	53	53	53	57
Thermocycle	Esc153_53	Esc153_45	Esc153_53	Esc153_53	Esc153_53	Esc132

Table B.2. Phusion polymerase PCR master mix protocol for a 20.0 µL final reaction volume used in all eDNA analysis.

Reagent	1 RXN (µL)
Nuclease-free H_2O + template DNA	12.2 - amt of template
5X Phusion HF Buffer	4.0
5 mM dNTPs	1.6
Forward primer	1.0
Reverse primer	1.0
Phusion Poymerase	0.2
Total Volume	7.8

*Add reagents in exact order for Phusion Polymerase and keep on ice

Table B.3. Primer set and thermocycle used for eDNA analysis.

Primer Set	Annealling T (°C)	Target Amplicon(bp)	Thermocycle
Ecam-ND2-P2-PET	67.5	159	PhueDNA_67.5
Etip-COI-P2-VIC	68.0	231	PhueDNA_68
Emac-ND2-P1-6FAM	67.5	145	PhueDNA_67.5
Ecaer-COI-P1-NED	66.0	234	PhueDNA_66
Pcope-ND2-P1-NED	67.5	241	PhueDNA_67.5

Table B.4. Thermocycle conditions for eDNA analysis.

Step	Temperature (°C)	Time
1	98	0:30
2	98	0:10
3	Annealing	0:30
4	72	0:30
5	GoTo step 2	49x
6	72	10:00
7	12	Hold

DNA Extraction from Filters or centrifuged tubes using DNEasy Blood and Tissue DNA Isolation Kit

1. Obtain samples to be processed:
    a. Remove filter samples from freezer.
    b. Remove centrifuged samples from the drying hood.
    c. Remove one positive and one negative extraction control from freezer.
2. Label 1.7-ml MCT extraction tubes and add 370 µL ATL to each tube.
3. Add 30 µL *proteinase K* to each tube. Vortex for 5 - 10 seconds.
4. Add DNA to prepared extraction tubes.
    a. Filters: Remove each filter and carefully tear off filter ring, fold and roll the filter so that it fits into the tube. Only one filter can be extracted per centrifuge tube. Be sure to push filter down into buffer (use a clean pipette tip).
    b. Centrifuge: Place a sterile swab into each tube of extraction buffer to moisten swab.
        i. Use one swab to wipe the pellet from each field replicate. Visually inspect the field sample tube and swab any particulates from the walls. Rinse swab as needed.
        ii. After swabbing all replicate tubes for a single field sample, break off the swab into the extraction tube, close the tube. Change gloves between each sample.
5. Vortex tubes with filters or swabs, Incubate at 55°C for 1 hour.
6. Centrifuge at max speed for 5 minutes.
7. Transfer supernatant to a new 1.7 mL centrifuge tube. Archive tube and filter or swab at -80°C.
8. Add 400 µL Buffer AL. Mix thoroughly by vortexing.
9. Add 400 µL ethanol (96 – 100%). Mix thoroughly by vortexing.
10. Transfer about half of mixture by pipet into a DNeasy Mini spin column placed in a 2 mL collection tube. Centrifuge at ≥ 6000 x g for 1 minute. Discard flow-through and collection tube, save spin column.
11. Transfer remaining mixture by pipet into the same Mini spin column, place in a new 2 mL collection tube. Centrifuge at ≥ 6000 x g for 1 minute. Discard flow-through and collection tube.
12. Place spin column in a new 2 mL collection tube. Add 500 µL Buffer AW1. Centrifuge at ≥ 6000 x g for 1 minute. Discard flow-through and collection tube.
13. Place spin column in a new 2 mL collection tube. Add 500 µL Buffer AW2. Centrifuge at 18,000 x g for 3 minutes. Discard flow-through and collection tube.
14. Transfer the spin column to a new 1.7 mL MCT.
15. If there are 8 or fewer filters for a single sample, elute the DNA by adding 200 µL Buffer AE to the center of the spin column membrane. If there are more than 8 filters, elute with only 100 µL Buffer AE. Incubate for 1 minute at room temperature (15 - 25°C). Centrifuge at ≥ 6000 x g for 1 minute.
16. Discard the spin column, combine multiple extractions for a single sample if necessary. Store in extraction room freezer if not proceeding immediately to amplification.

Figure B.1 A copy of the U.S. Fish and Wildlife eDNA extraction protocol that was modified for eDNA extraction.