Sturgeon (Acipenseridae) phylogeny, biogeography, & ontogeny.

Katie May Laumann

College of William and Mary - Virginia Institute of Marine Science

Follow this and additional works at: https://scholarworks.wm.edu/etd

Recommended Citation

This Dissertation is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Dissertations, Theses, and Masters Projects by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.
Sturgeon (Acipenseridae)

Phylogeny, Biogeography, & Ontogeny

A Dissertation

Presented to

The Faculty of the School of Marine Science
The College of William & Mary in Virginia

In Partial Fulfillment

Of the requirements for the Degree of

Doctor of Philosophy

By

Katie May Laumann

2016
APPROVAL SHEET

This dissertation is submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Katie May Laumann

Approved by the committee, April 2016

Eric Hilton, Ph.D.
Committee Chairman / Advisor

Casey Dillman, Ph.D.
Smithsonian Institution
Washington, D.C.

J. Emmett Duffy, Ph.D.
Smithsonian Institution
Washington, D.C.

John Graves, Ph.D.

Lynne Parenti, Ph.D.
Smithsonian Institution
Washington, D.C.
DEDICATION

For my family, my wonderful parents Bonnie May Kersta and Rupert Laumann and my sister Jessie May Laumann, and most of all for my husband Jon Lefcheck, and Berlioz May Fett Laumann-Lefcheck, who have provided unequaled love and support.
# TABLE OF CONTENTS

- **ACKNOWLEDGEMENTS** ......................................................................................................................... iii
- **LIST OF TABLES** ....................................................................................................................................... 4
- **LIST OF FIGURES** ..................................................................................................................................... 5
- **ABSTRACT** ............................................................................................................................................... 7
  - Sturgeon (Acipenseridae): .................................................................................................................... 9
  - Phylogeny, Biogeography, & Ontogeny ............................................................................................... 9
- **INTRODUCTION** .................................................................................................................................... 10
  - References ....................................................................................................................................... 13
- **Figures** .................................................................................................................................................. 14
- **CHAPTER 1. Phylageny of Sturgeons (Actinopterygii: Acipenseriformes: Acipenseridae) Based on Full Mitogenomic Data** .................................................................................................................. 16
  - **Abstract** ............................................................................................................................................ 17
    - 1. Introduction ................................................................................................................................. 19
    - 2. Methods ....................................................................................................................................... 29
      - 2.1 Taxon Sampling and Summary............................................................................................. 29
      - 2.2 DNA Purification and Sequencing of *A. fulvescens* and *H. dauricus* ....................... 30
      - 2.3 DNA Purification and Sequencing for Other Species ......................................................... 31
      - 2.4 Alignments and Model Tests ................................................................................................ 32
      - 2.5 Phylogenetic Analyses ........................................................................................................... 33
      - 2.6 Tree Comparisons .................................................................................................................. 34
      - 2.7 Comparisons to Previous Studies ......................................................................................... 35
    - 3. Results .......................................................................................................................................... 36
      - 3.1 Sequences and Alignment.................................................................................................... 36
      - 3.2 Full mitogenome results ...................................................................................................... 37
        - 3.2.1 Phylogenetic Hypothesis Support Values ........................................................................ 37
        - 3.2.2 Full Mitogenome Hypotheses ........................................................................................... 37
      - 3.3 Gene Tree Analysis and Comparison to the Full Mitogenome ........................................... 39
        - 3.3.1 tRNA Gene Family Phylogeny.......................................................................................... 39
1.2.4 Recent Geologic History: East China, Yellow, and Japan Seas ...................................... 101
1.2.5 Recent Geologic History: Interior North America, The Mississippi and Missouri Rivers and their tributaries .................................................................................................................. 101

2.0 Methods ................................................................................................................................... 102
2.1 Phylogenies for Comparison .............................................................................................. 103
2.2 Species Ranges and Phylogenies........................................................................................ 104

3.0 Results ...................................................................................................................................... 105
3.1 Areas of Endemism ............................................................................................................. 105
3.2 Intersection Tree ................................................................................................................. 108

4.0 Discussion and Conclusions ................................................................................................... 108
4.1 Anadromy ............................................................................................................................ 108
4.2 Area Relationships .............................................................................................................. 109
4.2.1 Atlantic Clade ................................................................................................................... 109
4.2.1a Atlantic Clade: North American and Baltic/North Sea/Ponto-Caspian Regions ...... 109
4.2.1b Atlantic Clade: North American Regions ...................................................................... 110
4.2.1c Atlantic Clade: Baltic/North Sea and Ponto-Caspian Regions ..................................... 110
4.3 Pacific Clade ........................................................................................................................111
4.3.1 Pacific Clade: Asia and Pacific North America................................................................. 111
4.3.2 Pacific Clade: Pacific and Arctic ....................................................................................... 112
4.4 Bering Sea ........................................................................................................................... 114
4.5 Conclusions ......................................................................................................................... 115

Acknowledgements ....................................................................................................................... 116
References ..................................................................................................................................... 117
Tables ............................................................................................................................................. 130
Figures............................................................................................................................................ 133

CHAPTER 3. Utility of Ontogenetic Characters in Phylogenetic Studies ...................................... 154
Abstract .......................................................................................................................................... 155
1. Introduction .......................................................................................................................... 156
2. Methods ..................................................................................................................................... 160
2.1 Taxon Sampling and Data Collection .............................................................................. 160
ACKNOWLEDGEMENTS

Many thanks to my committee, Drs. Eric Hilton, John Graves, Emmett Duffy, Casey Dillman, and Lynne Parenti, for their assistance, advice, and support throughout this work. Thanks also to Jan McDowell for support, mentorship, and friendship, and to Dr. Masaki Miya for mentorship and the amazing opportunity to work with his lab at the Natural History Museum and Institute in Chiba, Japan. Thanks to my family my good friends Derek and Rebekah Loftis and Ashley Morton for encouragement and support. Finally, thank you Jon and Berlioz.
LIST OF TABLES

CHAPTER 1. Phylogeny of Sturgeons (Actinopterygii: Acipenseriformes: Acipenseridae) Based on Full Mitogenomic Data .................................................................................................................................................................................. 16

Table 1. GenBank Accession Numbers .................................................................................................................................................................................. 75

Table 2. Gene Families, Lengths, and Best Models .................................................................................................................................................. 76

Table 3. Rzhetsky-Nei Distances: Bayesian Mitogenome Topology and Gene-Family Topologies ................................................................................................................................................. 77

Table 4. Full Mitogenome Hypothesis Clade Constituents, and Presence in Gene-Family Trees ........................................................................................................................................................................... 77

Table 5. Rzhetsky-Nei Distances between Bayesian Mitogenome Topology and Gene Trees ............................................................................................................................................................................. 78

Table 6. Birstein et al. 2002 Parsimony Clade Constituents ................................................................................................................................................. 79

Table 7. Birstein et al. 2002 Maximum Likelihood Clade Constituents ................................................................................................................................................. 80

Table 8. Birstein et al. 2002 Combined Evidence Clade Constituents ................................................................................................................................................. 80

CHAPTER 2. The Biogeography of Extant Sturgeons .......................................................................................................................................................................................... 90

Table 1: Sturgeon species ranges .......................................................................................................................................................................................... 130

Table 2: Lamprey species ranges .......................................................................................................................................................................................... 131

Table 3: Pinniped species ranges .......................................................................................................................................................................................... 132

CHAPTER 3. Utility of Ontogenetic Characters in Phylogenetic Studies .................................................................................................................................................................................. 154

Table 1: Character matrix of three behavioral traits .............................................................................................................................................................................. 185

Table 2: Earliest occurrences of ontogenetic changes in four species .............................................................................................................................................................................. 186

Table 3: Calculations of phylogenetic signal in morphological characters based on age at which milestones first occur. *s indicate significance .............................................................................................................................................................................. 187

Table 4. Calculations of phylogenetic signal in morphological characters based on total length (mm) at which milestones first occur .............................................................................................................................................................................. 188
INTRODUCTION.......................................................................................................................... 10

Figures ........................................................................................................................................ 14

Figure 1. *Acipenser gueldenstaedtii* (Russian sturgeon) with protruding jaw ....................... 14

Figure 2. *Acipenser gueldenstaedtii* (Russian sturgeon) with distinctive scutes and barbels .. 15

CHAPTER 1. Phylogeny of Sturgeons (Actinopterygii: Acipenseriformes: Acipenseridae) Based
on Full Mitogenomic Data ........................................................................................................ 16

Figure 1: Parsimony tree based on 12s, 16s, cyt b, D-loop, and NADH 5 .......................... 81

Figure 2: Maximum Likelihood tree based on 12s, 16s, cyt b, D-loop, and NADH 5 ......... 82

Figure 3: Combined Evidence tree based on 12s, 16s, cyt b, D-loop, NADH 5, and morphological data ................................................................. 83

Figure 4: Bayesian full mitogenome phylogeny ................................................................. 84

Figure 5: Maximum Likelihood Full mitogenome phylogeny ............................................ 85

Figure 6: Parsimony full mitogenome phylogeny ............................................................ 86

Figure 7: Bayesian phylogenetic hypothesis based on concatenated tRNA sequences ....... 87

Figure 8: Bayesian phylogenetic hypothesis based on concatenated rRNA sequences ...... 88

Figure 9: Bayesian phylogenetic hypothesis based on concatenated protein-coding sequences. .................................................................................. 89

CHAPTER 2. The Biogeography of Extant Sturgeons ............................................................ 90

Table 1: Sturgeon species ranges ..................................................................................... 130

Table 2: Lamprey species ranges .................................................................................... 131

Table 3: Pinniped species ranges ................................................................................... 132

Figure 1: Bayesian full mitogenome phylogeny ............................................................ 133

Figure 2: Croizat’s global biogeography map ............................................................... 134

Figure 3: Areagram of Croizat’s (1958) biogeographic regions ...................................... 135

Figure 4: Map of sturgeon Areas of Endemism; ............................................................. 136

Figure 5: Morphological Lamprey Phylogeny ............................................................... 137

Figure 6: Molecular Pinniped Phylogeny .................................................................... 138
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Range maps of sturgeon species included in biogeographic analyses</td>
<td>139</td>
</tr>
<tr>
<td>8</td>
<td>Sturgeon Areagram</td>
<td>150</td>
</tr>
<tr>
<td>9</td>
<td>Lamprey Areagram</td>
<td>151</td>
</tr>
<tr>
<td>10</td>
<td>Pinniped Areagram</td>
<td>152</td>
</tr>
<tr>
<td>11</td>
<td>Sturgeon-Lamprey-Pinniped Intersection Tree</td>
<td>153</td>
</tr>
<tr>
<td>1</td>
<td>Bayesian full mitogenome phylogeny</td>
<td>189</td>
</tr>
<tr>
<td>2</td>
<td>Example of cleared and stained specimens</td>
<td>190</td>
</tr>
<tr>
<td>3</td>
<td>Full mitogenome Bayesian phylogeny, pruned</td>
<td>191</td>
</tr>
<tr>
<td>4</td>
<td>Full mitogenome Bayesian phylogeny</td>
<td>192</td>
</tr>
<tr>
<td>5</td>
<td>Topology shown by single-character (yolk sac and teeth) trees</td>
<td>193</td>
</tr>
<tr>
<td>6</td>
<td>Topology shown by single-character (jaws elements) tree</td>
<td>194</td>
</tr>
<tr>
<td>7</td>
<td>Topology shown by single-character (migration) tree</td>
<td>195</td>
</tr>
<tr>
<td>8</td>
<td>Topology shown by single-character (nocturnality) trees</td>
<td>196</td>
</tr>
<tr>
<td>9</td>
<td>Topology shown by single-character (phototaxis) trees</td>
<td>197</td>
</tr>
<tr>
<td>10</td>
<td>Phylogenetic Independent Contrast values</td>
<td>198</td>
</tr>
<tr>
<td></td>
<td>CONCLUSIONS</td>
<td>199</td>
</tr>
</tbody>
</table>
ABSTRACT

Thought of as “ancient” fishes, 25 broadly recognized extant sturgeon species are classified in four genera (Acipenser, Huso, Pseudoscaphirhynchus, and Scaphirhynchus). Molecular and morphological analyses have led to broad but conflicting changes to sturgeon phylogeny. For example, the position of Scaphirhynchus among other sturgeons had been contentious, and various sets of sturgeon species have been proposed to make up the subfamily Husinae. Here, a molecular phylogeny of sturgeons, based on the full mitogenome, is presented. In this phylogeny, Scaphirhynchus is recovered with strong support as basal to the other sturgeons. Huso huso is recovered as basal within a clade containing P. kaufmanni and several species of Acipenser, and is proposed as a new, monotypic subfamily Husinae. This phylogeny is used to examine phylogenetic signal in individual genes and in gene families. The protein coding genes as a unit, and individually, along with 16s rRNA, show phylogenetic signal most similar to that of the full mitogenome. The phylogeny, along with evolutionary relationships of pinnipeds and lampreys, provides the basis for the exploration of sturgeon biogeography. Relationships among geographic areas inhabited by sturgeons are found, finding two sets of related areas- a Pacific area group and an Atlantic group. Relationships of areas within and between these groups reflect area relationships proposed by previous biogeographic and geologic studies. Phylogenetic signal is tested amongst ontogenetic characters, and is recovered in the timing at which larval sturgeon teeth are completely resorbed, indicating that the timing of ontogenetic milestones can carry signal. The phylogeny is used to remove confounding signal from, and
investigate correlations among, behavioral and morphological ontogenetic characters.

Correlation is found between one pair of characters.
Sturgeon (Acipenseridae):

Phylogeny, Biogeography, & Ontogeny
INTRODUCTION
The 25 broadly recognized extant sturgeon species are classified in four genera (Acipenser, Huso, Pseudoscaphirhynchus, and Scaphirhynchus) and are widely acknowledged to form the monophyletic family Acipenseridae. Thought of as “ancient” fishes, sturgeons occupy a basal position relative to other ray-finned fishes (Actinopterygii). Together with their sister group, the paddlefishes (Polyodontidae), they form the extant sister-group to Neopterygii (the group including gars, bowfins, and teleostean fishes; Grande and Bemis 1996). Sturgeons range throughout the freshwaters and coastal regions of Europe, Asia, and North America. Most sturgeons are diadromous, moving between fresh and saltwater for feeding (primarily downstream movement) and reproductive (upstream movement) purposes (Bemis and Kynard 1997, Peng 2007). Sturgeons have protrusible jaws that enable suction feeding (Figure 1), chemosensory barbels (Figure 2), and bony plates called scutes along their bodies in rows (Figure 2). These charismatic fishes are culturally and economically important (e.g., as the source of true caviar). The oldest known sturgeon fossils are from the Late Cretaceous (85 million years; Hilton and Grande 2006), but the sister-group relationship between Acipenseridae and Polyodontidae, combined with the existence of paddlefish fossils from the Early Cretaceous, indicates that sturgeons are much older (Peng et al. 2007). Populations of many sturgeon species are currently in sharp decline due in part to overfishing and habitat alterations (Holzkamm and Waisberg 2004), and all species of sturgeons are provided conservation protection (IUCN 2014). Even with the imperiled status of these fishes, the study of their evolutionary history through examination of character states and cladistic methods began only
relatively recently (Findeis 1997), and a well-resolved, broadly accepted phylogeny for the family remains elusive.

The first explicitly cladistic study of the family Acipenseridae, based on morphological data (Findeis 1997), recovered the genus *Huso* as the sister group of the other three genera and recovered *Scaphirhynchus* and *Pseudoscaphirhynchus* as sister taxa (forming the tribe Scaphirhynchini, recognized as the subfamily Scaphirhynchinae by some authors). Molecular data were introduced into the study of sturgeon phylogenetics at about the same time (e.g., Birstein and DeSalle 1998) and recent analyses using both molecular and morphological data (independent of one another and together) have resulted in markedly different phylogenetic hypotheses than those of Findeis (1997) (Hilton et al. 2011). Several broad changes to Findeis’ phylogeny of Acipenseridae have been proposed across morphological and molecular studies, for example the dissolution of Scaphirhynchini. The details of these differences, however, vary from study to study; for example, the position of *Scaphirhynchus* is has been debated. Some differences in hypothesized relationships of sturgeons may be due to incomplete taxon sampling and the use of different molecular markers and morphological characters across studies.

Variation in the details of recent sturgeon phylogenetic hypotheses illustrates the need for further examination. The goal of this dissertation research was to use modern phylogenetic methods and various types of data to work toward a well-resolved sturgeon phylogeny. I constructed a molecular phylogeny, sequencing the full mitogenome and including near-complete taxon sampling, explored the effects of using different segments of mitochondrial DNA on phylogeny, and different optimality criteria for phylogenetic hypotheses. I used the resulting
full mitogenome phylogenetic hypothesis to investigate area relationships of different areas inhabited by sturgeons. Finally, I investigated jaw development in five North American species in the genus *Acipenser*, and identified and tested potential morphological and behavioral ontogenetic characters for use in phylogenetic studies.

References


Figures

Figure 1. *Acipenser gueldenstaedtii* (Russian sturgeon) with protruding jaw, enabling suction feeding. Photo taken at Volga River Exhibit at the Tennesse Aquarium in Chattanooga.
Figure 2. *Acipenser gueldenstaedtii* (Russian sturgeon) with distinctive scutes and barbels indicated. Photo taken at Volga River Exhibit at the Tennesse Aquarium in Chattanooga.
CHAPTER 1.

Phylogeny of Sturgeons (Actinopterygii: Acipenseriformes: Acipenseridae)
Based on Full Mitogenomic Data
Abstract

The 25 commonly recognized species of sturgeons, family Acipenseridae (Actinopterygii, Acipenseriformes), are traditionally classified in four genera (Acipenser, Huso, Scaphirhynchus, and Pseudoscaphirhynchus). Recent phylogenetic studies of the family, using both morphological and molecular data, generally recover Acipenser as paraphyletic, Huso nested within Acipenser, and Pseudoscaphirhynchus as sister to A. stellatus. The details of phylogenetic relationships among species within the family, however, vary from study to study. Previous molecular-based studies of the family have been limited to the use of fragments of the mitochondrial gene to estimate the phylogenetic relationships among sturgeons. In this study, I present a phylogenetic hypothesis based on analysis of the complete mitogenomic sequences of 23 sturgeon taxa (22 species and both subspecies of A. oxyrinchus) and discuss the implications of the proposed relationships. This analysis supports the inclusion of Huso spp. in “Acipenser” and proposes that “Scaphirhynchini” is nonmonophyletic. Huso huso is recovered as sister to a clade that includes A. ruthenus, A. fulvescens, A. baerii, A. naccarii, A. gueldenstaedtii, A. brevirostrum, and A. stellatus and P. kaufmanni as sister species. Huso dauricus is recovered in a clade with A. medirostris and A. mikadoi, and this clade is sister to A. dabryanus, A. sinensis, A. transmontanus, and A. schrenckii. Scaphirhynchus is recovered as sister to all other sturgeons. Acipenser sturio and both subspecies of A. oxyrinchus form another monophyletic group. I explore the influence of each mitochondrial gene region on full mitogenome topologies. Although they are linked within the same molecule, individual rRNA, tRNA, and protein-coding mitochondrial genes do not all exhibit the same phylogenetic signal as one another or as the full mitogenome as a whole. Several genes, including all of the protein-coding and rRNA genes, as well as D-loop, recover significant phylogenetic structure among sturgeons. Of those genes that
recover structure, the protein coding genes and 16s recover relationships similar to those found by the full mitogenome, whereas 12s and D-loop recover structure different from that shown by the full mitogenome. Other genes, including the tRNAs, recover very little phylogenetic structure, showing either complete polytomies or several groups, within which there is little or no resolution.
1. Introduction

The 25 commonly recognized extant species of sturgeon, traditionally classified in four genera (*Acipenser*, *Huso*, *Pseudoscaphirhynchus*, and *Scaphirhynchus*), form the monophyletic family *Acipenseridae* Bonaparte 1831 (Berg 1904, Bemis et al. 1997, Birstein et al. 1997, Findeis 1997, Hilton et al. 2011). Two monotypic genera, †*Priscosturion* (Grande and Hilton 2006) and †*Protoscapkirhynchus* (Wilimovsky 1956), are known only as fossils. The paddlefishes (Polyodontidae), two entirely fossil families (†*Chondrostidae*, †*Peipiaosteidae*), and *Acipenseridae* together form the Acipenseriformes, which is broadly regarded as the extant sister-group to Neopterygii (gars, bowfins, and teleosts; Grande and Bemis 1996, Bemis et al. 1997, Nelson 2006, Hilton and Forey 2009, Hilton et al. 2011). The earliest fossil evidence of sturgeons dates to the Late Cretaceous, 85 million years ago ago (Hilton and Grande 2006), but these fishes are likely older, as polyodontid fossils date back to the Early Cretaceous (Bemis et al. 1997, Grande et al. 2002). Due in part to overfishing and habitat alterations (Holzkamm and Waisberg 2004), nineteen species have been listed as endangered or critically endangered by the IUCN (2014) and at least one (*P. fedtschenkoi*) is likely extinct (IUCN 2014). International trade of all sturgeon species is closely monitored (ICES 12.7 CoP16).

A broad understanding of sturgeon evolutionary history may have implications for management and conservation efforts. These efforts focus on regions, species, and the evolutionary process itself (Forest et al. 2007). Conservation managers recognize that species represent evolutionary history, and treat the loss of “old” taxa, taxa undergoing rapid diversification, evolutionarily distinct taxa, and taxa at high risk of extinction as priorities for conservation (Crozier 1997, Isaac et al. 2007, Forest et al. 2007, Rodrigues and Gaston 2002, Moritz 1994, Rodrigues and Gaston 2002). Determining whether a species or group of species fits into these categories requires phylogenetic information, and although incomplete
phylogenies may be useful, hypotheses based on comprehensive datasets are the most informative for management and conservation planning (Moritz 1994, Rodrigues et al. 2011). As conservation and management strategies aim to conserve phylogenetic diversity (evolutionary differences across all biodiversity, Lefcheck et al. 2014), phylogenetic hypotheses are playing a larger role in conservation efforts (Faith 1992, Avise 2010). Understanding phylogeny also allows conservation managers and planners to estimate historic population size, identify trends in gene flow, and determine how bottlenecks have historically impacted currently imperiled species (Moritz 1994). These data can help conservation managers determine the best methods for conservation of various taxa (Moritz 1994), helping maintain gene flow in extant populations and preventing or mitigating the effects of bottlenecks. All of the criteria listed above apply to sturgeons— they are considered “ancient” fishes, and are declining across their ranges. A well-supported sturgeon phylogeny will help resource managers to determine the conservation priority of these fishes as well as helping identify avenues for the protection of these species (for example by helping to determine when hybridization with escaped or released hatchery-raised individuals might pose a risk to wild populations, and mitigate this threat). In addition to conservation and management value, a better understanding of sturgeon phylogeny may be broadly informative for the evolution of all fishes, as sturgeons generally represent the most species-rich clade of basal ray-finned fishes (Actinopterygii).

There have been numerous attempts to resolve the evolutionary history of Acipenseridae from many perspectives; however, a consensus on sturgeon phylogeny has not been reached. Based on morphological data, Findeis (1997) found Huso to be sister to all other sturgeons and recovered *Pseudoscaphirhynchus* and *Scaphirhynchus* as a monophyletic group, Scaphirhynchini (also sometimes classified as a subfamily, Scaphirhynchinae; e.g., see Mayden and Kuhajda 1996). Findeis (1997) included the genus *Acipenser* in his analysis as a terminal
taxon but noted that it was likely nonmonophyletic as there were no synapomorphies supporting this genus. Subsequent morphological studies have led to various phylogenetic hypotheses. Mayden and Kuhajda (1996) recovered *Acipenser* and *Huso* as sister genera and found *Pseudoscaphirhynchus* and *Scaphirhynchus* to form a clade, though they failed to recover monophyly for *Pseudoscaphirhynchus*. Choudhury and Dick (1998) also recovered Scaphyrhynchini as monophyletic, and found distinct clades within *Acipenser*; *Huso* was not included in their study. Artyukhin (2006) recovered Scaphyrhynchini as sister to the other sturgeons, and recovered distinct subclades within *Acipenser*. *Acipenser oxyrinchus* and *A. sturio* were hypothesized to be sister to the other sturgeon clades, which included *H. huso* and *H. dauricus* as sister taxa, grouped with *A. nudiventris*, *A. schrenckii*, and *A. ruthenus*. Other studies have recovered alternate clades within *Acipenser*. Hilton et al. (2011) also recovered *Scaphyrhinchus* as sister to the other sturgeons, finding (*H. huso, (A. ruthenus, A. baerii)*) to form a monophyletic group with (*A. stellatus, (P. kaufmanni, P. hermanni)*).

Molecular studies of sturgeon phylogenetics have led to additional phylogenetic hypotheses; these studies are based on mitochondrial DNA sequences. In fishes, the mitochondrial genome (also referred to as mitogenome or mitochondrion) is composed of 37 genes that are exclusively maternally inherited as a single linkage-group, and is considered to be a single locus. These genes are distributed amongst three gene families, the tRNAs, rRNAs, and protein coding genes, and evolve at different rates and with different nucleotide substitution patterns. Most animal mitochondrial genomes, those of sturgeons included, contain 22 tRNAs: tRNA-Phe, -Val, -Leu, -Ile, -Gln, -Met, -Trp, -Ala, -Asn, -Cys, -Tyr, -Ser, -Asp, -Lys, -Arg, -His, -Ser (2), -Leu (2), -Glu, -Thr, -Pro. Two genes, 12s and 16s, make up the rRNA (ribosomal ribonucleic acid) gene “family”. The protein coding gene family comprises 13 genes: ATP6 and ATP8, COI, COII, COIII, cyt b, and NADH’s 1-6 and NADH 4L. Mitochondrial DNA, is not protected
from oxidative damage by chromatin, and so is exposed to free radicals and susceptible to injury (Gredilla et al. 2010). Although mechanisms to mitigate this damage have been identified in the mitochondrion, recent studies suggest that these mechanisms cannot repair damage after it is caused (Gredilla et al. 2010). In addition to decreased damage repair as compared to nuclear DNA, proofreading mechanisms during mitochondrial DNA replication are limited. As this portion of the genome does not undergo recombination, mismatches cannot be repaired based on a reference or template sequence. The proofreading mechanisms that are present in mitochondrial DNA have been shown to be inefficient in some mammal cells (Song et al. 2005).

In contrast to this, nuclear DNA is packaged and protected by chromatin and therefore not exposed to the same levels of damage. Recombination allows for corrections to mismatches in nuclear DNA replication, and additional proofreading mechanisms are efficient, leading to high fidelity in nuclear DNA replication (Caravas 2012). This leads to a high rate of substitution and evolution in the mitochondrial genome as compared to the nuclear genome (Avise 1998, Caravas 2012).

Mitochondrial DNA is invaluable to phylogenetic studies, but various concerns have been raised about relying exclusively on mitochondrial genes to elucidate phylogeny (e.g. Caravas 2012, Moore 1995). Pitfalls concerning mitochondrial DNA could lead to the recovery of unresolved or incorrect topologies. Even if a well-supported mitochondrial phylogeny is recovered, in some situations this phylogeny may not represent the species phylogeny (Caravas 2012). The high rate of substitution in mitochondrial DNA could result in saturation or overwriting of phylogenetic signal, weakening the display of this signal (DeFilippis and Moore 2000). Composition bias of adenine and thymine present in mitochondrial DNA may lead to the same issue (DeFilippis and Moore 2000). The reliance on few or short mitochondrial sequences, frequent in past molecular phylogenetic studies of sturgeons, could lead to low resolution in the
resulting phylogeny (Vilstrup et al. 2011). Including multiple genes from the mitogenome, however, only provides a single independent estimate of phylogenetic relationships because these genes are linked and do not evolve strictly independently of one another (Moore 1995).

Additional challenges to the use of mitochondrial DNA in phylogenetic studies include hybridization, or reproduction between different species. The use of mitochondrial DNA in phylogenetic studies of taxa that exhibit hybridization may result in the recovery of a gene tree that is different from the species tree (Baack and Rieseberg 2007, Caravas 2012, Moore 1995). Hybridization can lead to introgression, the incorporation of genetic material from one species into the genome of another (Baack and Rieseberg 2007). Hybridization and introgression are especially problematic when using mitochondrial DNA alone to infer phylogeny because if the mitochondrial DNA of one species is incorporated into a hybrid or an individual of another species, that mitochondrial DNA will not reflect the evolutionary history of the species. If undetected, this will mislead phylogenetic analyses (Moore 1995). If hybridization is detected, it can be accounted for in phylogenetic studies (Moore 1995), but the detection of hybridization and introgression requires comparison between topologies based on unlinked genes (Caravas 2012). Because all mitochondrial genes are linked, hybridization and introgression are undetectable based on mitochondrial sequence data alone. Mitochondrial phylogenies must be compared with nuclear phylogenies to search for congruence. If nuclear and mitochondrial phylogenies present the same topology, introgression is unlikely. If they present conflicting hypotheses, introgression should be explored as one possible cause of incongruence (Duchene 2011, Moore 1995, Caravas 2012, Baack and Rieseberg 2007).

Despite these potential pitfalls, there are several advantages to the use of mitochondrial sequences in phylogenetic analyses. Many copies of the mitochondrion are present in each cell,
making amplification of mitochondrial sequences relatively simpler than amplification of nuclear
genes (Caravas 2012). Mitochondrial gene order is generally conserved in vertebrates, and the
mitochondrion lacks introns and can therefore be sequenced in larger units can than nuclear
genes (Caravas 2012). Recent technological innovations allow for the sequencing of the full
mitogenome at relatively low cost and in relatively short amounts of time (Mardis 2008). Mitochondrial
genes and the mitochondrial genome as a whole have been used to recover the
phylogenies of many taxonomic groups, including Diptera (Caravas 2012), reptiles and
delphinids (Duchene et al. 2011), teleostean and other fishes (Inoue et al. 2003, Ishiguro et al.
2003, Miya et al. 2001, Inoue et al. 2010, Lavoue et al. 2007), and various groups of birds
(DeFilippis and Moore 2000).

In addition to phylogenetic reconstruction, the utility of mitochondrial genes in species
identification, for example, is indisputable (Wolf 1999, Brown 1996, Miya and Nishida 2000). Cytochrome
c oxidase I (COI), for example, is used for DNA barcoding (Ward 2009). Among
fishes, COI has been shown to distinguish among 98% of marine and 93% of freshwater species
(Ward 2009). Additionally, a hypervariable region of rRNA was recently found to successfully
identify, to species, at least 168 taxonomically diverse teleostean fishes, including closely related
congeners (Miya and Nishida 2015). Although individual mitochondrial genes used for species
identification may not carry enough information to resolve phylogenies, they may be useful in
phylogenetic investigations when considered alongside other genes. Genes that differentiate
between species or even subspecies add resolution to “shallow” branches in a phylogeny,
enhancing the phylogeny by providing support for fine-scale relationships (Hajibabaei et al.
2007). Some barcoding genes, COI in particular, have been shown to be extremely useful in
phylogenetic studies focusing on relationships at the species level (e.g. in caddisflies, Hogg et al.

Although mitochondrial genes are linked, they evolve at different rates (Xia 1998, Sloan et al. 2009). Recent studies generally conclude that the common practice of relying on one or only a few mitochondrial genes to inform phylogeny is not ideal, as some genes are misleading and inappropriate for phylogenetic reconstruction because they recover topologies with low resolution (Duchene et al. 2011, Zardoya and Meyer 1996, Miya and Nishida 2000, Rastorguev et al. 2008). Selection of genes to be used in a phylogenetic study is important not only because phylogenetic signal may vary widely from gene to gene, but also because different genes resolve relationships at different evolutionary scales.

Various studies spotlight the different phylogenetic signals carried by different mitochondrial genes by comparing topologies derived using these genes with other topologies recovered by full mitogenomic, morphological, or other data. Zardoya and Meyer (1996) compared tetrapod and mammal topologies recovered by all protein-coding genes combined and individual protein-coding genes. They found that NADH4, NADH5, NADH2, cyt b, and COI recovered topologies consistent with those previously proposed based on morphological and genetic data, whereas topologies based on ATP6, ATP8, NADH3, and NADH4L diverged significantly from previous topologies. Miya and Nishida (2000) evaluated the similarity of single protein-coding gene topologies and a tRNA topology to the full mitogenome topology of eight teleosts and found that only NADH5 and the concatenated tRNA sequences recovered the same topology as the full mitogenome. Rastorguev et al. (2008) similarly compared individual protein coding gene topologies to the full mitogenome topology of 10 sturgeons and found that COI, NADH4, and ATP6 recovered topologies similar to that of the full mitogenome. Analyses of NADH6, in contrast, recovered a very different topology from that of the full mitogenome.
Using the oceanic dolphins (family Delphinidae) and the killer whales (genus Orcinus), Duchene et al. (2011) compared topologies inferred from the full mitogenome to those derived from 12s, 16s, and individual protein coding genes. They found different results for the two different taxonomic groups, with COI, cytb, NADH3 and ATP6 recovering topologies (with varying levels of support) similar to that of the full mitogenome for Orcinus, and NADH1, COI and NADH4 doing so with the delphinids. Clearly, individual mitochondrial genes carry different phylogenetic signal from one another, with some carrying similar signal to that of the full mitogenome.

Efforts to identify factors contributing to the different phylogenetic hypotheses recovered by different genes have found that neither gene length nor evolutionary rate accurately predict how similar a gene’s topology will be to that derived from the full mitogenome (Zardoya and Meyer 1996). The depth of the phylogeny under consideration may play a role, as might the completeness of the taxon sampling (Zardoya and Meyer 1996). Therefore, when selecting genes to be used in a phylogenetic study, the taxonomic group, number of taxa for which data are available, and target resolution of the phylogeny should be considered.

Longer DNA sequences than those that have previously been used may help resolve phylogenetic relationships, particularly higher-level relationships (Miya and Nishida 2015). The use of more data, particularly those from molecular gene regions that evolve at different rates, can provide resolution at multiple phylogenetic levels and reduce biases that may be gene-specific (Hajibabaei et al. 2007). Previous studies were limited because there was no straightforward method to sequence the entire mitogenome until fish-versatile primers were designed to sequence the full mitogenome via a combination of long-and short PCR (Miya and Nishida 1999, 2015). The development of this methodology has allowed exploration of individual mitogenomes, mitochondrial gene arrangement, and reconstruction of phylogeny.
The utility of complete mitogenomic data in phylogenetic analysis of fishes has been demonstrated in studies of both higher-level relationships (e.g. teleostean phylogenetic relationships, Miya and Nishida 2000b) and in population-level relationships (e.g. Jacobsen et al. 2012). More recently, next-generation sequencing (NGS) methods have been applied to full-mitogenome sequencing techniques (Ahn et al. 2009, Groenenberg et al. 2012, Feldemeyer et al. 2010, Jex et al. 2010, Irorizzo et al. 2012, Varshney et al. 2009, Miya and Nishida 2015, Lindqvist et al. 2010, Webb and Rosenthal 2011, Botero-Castro et al. 2013, Shen et al. 2015).

Mitochondrial data have been used frequently in analyses of sturgeon phylogenies. Of ten recent molecular phylogenetic studies of sturgeons (Birstein and DeSalle 1998, Birstein et al. 2002, Dillman et al. 2007, Fontana et al. 2001, Krieger et al. 2000, Krieger et al. 2008, Ludwig et al. 2001, Rastorguev et al. 2008, Simons et al. 2001, Zhang et al. 2000), most have relied on short segments of the mitochondrial genome (<5000 base pairs, or <30% of the mitogenome), with just one (Rastorguev et al. 2008) considering the full mitogenome (but with limited taxon sampling). None of the studies included more than 22 of the 25 extant species, and only three (Birstein et al. 2002, Dillman et al. 2007, Krieger et al. 2008) considered specimens from all four recognized genera. Within these studies, the most commonly used mitochondrial genes were cytB, NADH4, NADH4L, NADH5, the rRNAs, and D-loop. Single-gene trees resulting from analyses of these sequences suggest that each of these genes carries different phylogenetic signals, despite the fact that they all occur in the mitochondrial genome. Further, the strength of phylogenetic signal among these genes varies and when two genes are considered together, the signal carried by one may be stronger than that from another, "swamping" it out. This effect may be reversed or negated when more than two genes are considered together.

Along with recent morphological studies, these molecular analyses have expanded our understanding of sturgeon phylogeny (Hilton et al. 2011). Among the most significant findings of...
these studies include *H. huso* nested within “*Acipenser*” (e.g., Birstein and DeSalle 1998, Birstein et al. 2002, Dillman et al. 2007, Krieger et al. 2008, Hilton et al. 2011), which Vasil'eva et al (2009) formally recognized as *A. huso*, and break up Scaphirhynchini by recovering *A. stellatus* as sister to *Pseudoscaphirhynchus* spp. (e.g., Birstein et al. 2002, Dillman et al. 2007, Krieger et al. 2008, Hilton et al. 2011). However, the hypotheses of relationships among species vary from study to study. For example, various relationships of *Huso* within “*Acipenser*” have been proposed, including the recovery of *H. huso* as sister to *A. ruthenus* (in a redefined Husinae; Hilton et al. 2011) or as sister to other species of “*Acipenser*” (e.g., Fontana et al. 2001). Further, different relationships have been proposed regarding the position of genera formerly in the tribe Scaphirhynchini as well as the monophyly of species groups included in the genus “*Acipenser*.” Molecular-based studies of sturgeon relationships have all used different mitochondrial gene regions or different combinations of gene regions, and therefore each study includes different characters leading to differences in the phylogenies proposed. Nuclear data have not been used in phylogenetic analyses of sturgeons, in part due to difficulties sequencing nuclear genes in these polyploid fishes (Krieger et al. 2006).

The goal of this study was to apply full mitogenome sequencing methods to complement previous studies of sturgeon phylogeny. Combining both long and short PCR methods and NGS methods, I build upon past studies by using more mitochondrial DNA sequence data than has been available previously. This allowed for the comparison of phylogenetic signal carried by individual mitochondrial genes to the signal of the full mitogenome, and led to the resolution of polytomies recovered in some recent sturgeon phylogenetic studies (Birstein et al. 2002, Dillman et al. 2007, Ludwig et al. 2001, and Zhang et al. 2000). I conducted a phylogenetic analysis of Acipenseridae using the full mitogenome of 23 sturgeon species and subspecies representing all four currently recognized sturgeon genera. I
investigated phylogenetic signal present in individual genes, and determined which mitogenome segments exert the strongest signal recovering this phylogenetic hypothesis. To determine the relative phylogenetic signal of these three topologies, data were explored by assessing the topologies based on 1) all protein coding gene regions analyzed together, and sequences of 2) all ribosomal RNA and 3) all transfer RNA sequences. Additionally, I compared my topology to three topologies from a previous study (Birstein et al. 2002).

2. Methods

2.1 Taxon Sampling and Summary

Tissue samples were obtained for vouchered specimens of 22 sturgeon taxa (A. o. desotoi, A. o. oxyrinchus, A. sturio, H. huso, H. dauricus, A. brevirostrum, A. gueldenstaedtii, A. naccarii, A. fulvescens, A. mikadoi, A. medirostris, A. schrenckii, A. transmontanus, A. sinensis, A. baerii, A. rufenus, A. stellatus, S. albus, S. platatorynchus, S. suttkusi, P. Kaufmanni, and P. hermanni) and one outgroup species (P. spathula). Total genomic DNA was extracted from muscle or, in some cases, gill tissue using the Qiagen DNeasy Blood & Tissue Kit following manufacturer’s protocol (Qiagen 2006).

Methods described below were applied in an attempt to obtain full mitogenome sequences for each of these species. Full mitogenome sequencing was successful for all of these taxa except A. baerii, A. rufenus, A. stellatus, and P. hermanni. Full mitogenome sequences for four ingroup species (A. baerii, Chen et al. 2011; A. dabryanus, Peng et al. 2007, A. rufenus, Li et al. 2013; and A. stellatus, Arnason et al. 2007) and four outgroup species including both extant paddlefishes (Polyodon spathula, Inoue et al. 2003 and Psephurus gladius, Peng et al. 2007) and two Polypteriformes (Polypterus delhezi, Chen and Lin 2010 and P. endlicherii, Chen and Lin 2010) were available on GenBank® (Accession numbers are shown in Table 1). To
evaluate the validity of sequences obtained from GenBank®, I sequenced a 16s segment of rRNA from tissue extracted from voucher specimens of *A. baerii* and *A. stellatus*. Sequencing failed for *A. ruthenus* and *P. hermanni*. I aligned the sequences with the full mitogenome sequences on GenBank® and calculated the number of base pair differences. There were no base pair differences between the GenBank® *A. baerii* 16s sequence and the 562 base pair segment I sequenced. The GenBank® 16s sequence from *A. stellatus* was also a 100% match to the 610 base pair sequence I obtained.

### 2.2 DNA Purification and Sequencing of *A. fulvescens* and *H. dauricus*

I attempted to use a long PCR technique (Cheng *et al.* 1994), described below, to amplify the complete mitogenome of the total genomic DNA extractions. This method was successful only for *A. fulvescens* and *H. dauricus*.

Two general fish primers were designed and used to amplify the mitochondrial genome in two long-PCR reactions (Miya and Nishida 1996). Long PCR was done in a BioRad S1000 thermal cycler using a 20-μl reaction (2.4 μl extracted DNA, 2 μl of each primer, 3.2 μl dNTPs, 2 μl MgCl₂, 6.24 μl water, 2 μl TaKaRa LA Taq™ buffer, and 0.16 μl TaKaRa LA Taq Polymerase™) and running 30 cycles of denaturation at 98° C for 10 seconds followed by annealing and extension combined at 60° C for 12 minutes. The two long PCR products from each of these species were mixed together and diluted with TE buffer (1:5) for use as template in subsequent PCR reactions.

Twenty-four sets of general fish primers (Inoue *et al.* 2003) were used for short PCR of long PCR products. PCR was carried out in a BioRad S1000 thermal cycler (Applied Biosystems) using a 10.4-μl reaction (1 μl dilute long PCR product, 1 μl of each primer, 0.8 μl dNTPs, 5.53 μl...
water, 1 μl TaKaRa Z Taq™ buffer, and 0.07 μl TaKaRa Z Taq Polymerase™) and using the following thermal-cycler protocol: initial denaturation at 94°C for 2 minutes, followed by 30 cycles of 98°C for 1 second, 50°C for 5 seconds, and 72°C for 15 seconds. PCR products were purified using a QIAquick PCR Purification Kit and following manufacturers' instructions. Direct-cycle sequencing using dye-labeled terminators (Applied Biosystems Inc.) was conducted using 36 internal sequencing primers. Labeled fragments were analyzed on an Applied Biosystems 3130xl Genetic Analyzer. DNA sequences were edited by eye and contigs were assembled using Sequencher 5.1. Sequences were annotated using mitoannotator in MitoFish (Iwasaki et al. 2013).

2.3 DNA Purification and Sequencing for Other Species

AnNGS method was used to obtain mitogenome sequences from the extracted total genomic DNA of A. o. desotoi, A. o. oxyrinchus, A. sturio, P. kaufmanni, P. hermanni, H. huso, A. brevirostrum, A. gueldenstaedtii, A. naccarii, S. albus, S. platorynchus, S. suttkusi, A. mikadoi, A. medirostris, A. schrenckii, A. transmontanus, and A. sinensis.

Before beginning library preparation for NGS, extracted DNA was quantified using the Qubit 2.0 fluorometer and the Qubit dsDNA BR Assay kit. Samples were prepared for sequencing using the Nextera XT DNA Library Preparation Kit v2 Set A following the manufacturers protocol for tagmentation of input DNA, PCR amplification and cleanup, and library validation, normalization, and pooling (illumine 2012). Tagmentation is an enzymatic reaction in which DNA is simultaneously fragmented and tagged with adapters; the subsequent PCR cycles amplify the tagged segments (illumina 2015). Libraries were sequenced using the Illumina MiSeq System. All MiSeq data were subsequently trimmed and merged in CLC Genomics workbench using the de novo assembly algorithm. This algorithm uses de Bruijn graphs, first identifying contigs (sets of
overlapping sequence segments) in the data read in, creating a sequence of contigs, and mapping the reads through alignment to the contigs (QIAGEN 2016).

2.4 Alignments and Model Tests

MitoAnnotator (Iwasaki et al. 2013) was used to rotate each mitogenome so that all mitogenomes started at tRNA-Phe. Rotated sequences were aligned in MAFFT version 7. The resulting alignment was edited by eye in MacVector6 (Rastogi 1999). The edited alignment was used to generate a consensus sequence using MacVector6. The resulting consensus sequence was then re-annotated in MitoAnnotator to identify individual mitochondrial genes (Iwasaki et al. 2013). Based on this annotation, individual gene alignments were separated.

The program DAMBE was used to test for saturation in the full mitogenome alignment and individual gene alignments (Xia and Lemey 2009). Alignments were imported into DAMBE. First, proportion of invariant sites was estimated for each alignment using the option "Seq.Analysis|Substitution rates over sites|Estimate proportion of invariant sites" using the "new tree" and neighbor joining criteria. Actual tests for saturation were conducted using the option "Seq.Analysis|Measure substitution saturation|Test by Xia et al." and entering the appropriate estimated proportion of invariant sites. Saturation was not detected in the full mitogenome or individual gene alignments.

The best models of nucleotide evolution were determined for each gene using jmodeltest 2.1.3 (Posada 2008) and BIC criteria. The full mitochondrial genome was partitioned by gene, and the best model of evolution for each gene was used in all Bayesian analyses. Models used are listed in Table 2.
2.5 Phylogenetic Analyses

Parsimony analyses were run on the full mitogenome alignment in PAUP* 4.0 (Swofford 2003) with 1000 bootstrap replicates. Gaps were treated as missing data. Stepwise addition was used to obtain starting trees for branch swapping. Only the best trees found were kept. A 50% majority-rule consensus tree was calculated and is shown in the results section.

Maximum likelihood analyses were run on the full mitogenome alignment using the RAxML-HPC2 (Stamatakis 2006) on XSEDE tool on the Cipres Science Gateway (Miller et al. 2011). Gaps were treated as missing data. Nonparametric bootstrapping was enabled and the seed value was set at 12345. A general time-reversible model was identified by jmodeltest 2.1.3 (Posada 2008) as the most appropriate model parameter for the full mitogenome alignment, and was chosen using the setting “GTRCAT” to allow for rate heterogeneity. 1000 bootstrap replicates were run to generate support values. The best-scoring tree was retained, and is presented in the results section.

Bayesian phylogenetic analyses were run for the following four data configurations: 1) the edited alignments of the full mitogenome; and gene family trees made up of edited and concatenated alignments of 2) all protein coding genes; 3) ribosomal RNA sequences, and 4) transfer RNA sequences. Individual gene trees were also constructed using Bayesian methods for each individual mitochondrial gene. Analyses were conducted in Mr.Bayes 3.2.2 (Huelsenbeck and Ronquist 2001). For each analysis, aligned sequences were partitioned by gene, with each gene assigned the best model of evolution as indicated by jmodeltest 2.1.3 (Posada 2008). For each dataset, between 500,000 and 3 million generations were run depending on the necessary number of generations for standard deviation on split frequencies to equal 0.0000 (Huelsenbeck and Ronquist 2001), the full mitogenome analysis required 3 million generations, the protein coding gene analysis was run for 1 million generations, and both
tRNA and rRNA analyses were run for 500,000 generations. All analyses used 4 chains and a burn-in of 30% to allow chain convergence.

In the resulting phylogenetic hypothesis, confidence in each node was expressed in the form of posterior probability values. Distinct taxonomic groupings with relatively high posterior probability support values (>0.90) were recovered by analyses of each dataset. The full mitogenome topology required some re-analysis due to the recovery of several polytomies with low support values within one monophyletic taxonomic grouping (that overall had high support values). Reanalysis conducted for the Bayesian full mitogenome topology involved running analyses for the nine taxa in this group separate from the other sturgeon taxa. Analyses were run for one-million generations, and as with other Bayesian analyses in this study, 4 chains and a burn-in of 30% were used. These analyses recovered a highly supported topology. I constrained the full mitogenome dataset including all taxa to include this topology, re-ran Bayesian analyses using a burn-in of 30% and running for one million generations, and recovered a well-supported, polytomy-free topology for all taxa. Relationships among other taxa did not change as a result of this constraint. The resulting consensus tree is presented as my preferred full mitogenome topology.

2.6 Tree Comparisons

To identify which genes most strongly influence the full mitogenome phylogenetic hypothesis, I calculated tree distance metrics between the topology resulting from the full mitogenome Bayesian analysis and those generated through alternate data sets. The Robinson-Foulds metric (Rzhetsky and Nei 1992; similar to the Penny-Hendy distance (Penny and Hendy 1985), which is defined as twice the number of bipartitions that differ between trees, is a tree distance metric used to quantify similarity and differences among trees (Paradis et al. 2004). Duchene et al. (2011) used it to identify which mitochondrial genes carry the same phylogenetic
signal as the full mitogenome in studies of delphinid and killer whale phylogenies (Duchene et al. 2011). Because the Rzhetsky-Nei distance metric is well defined for use only in strictly bifurcating trees and many of my gene trees have polytomies, I employed a version of this metric modified to allow for polytomies, the Rzhetsky-Nei distance (Paradis et al. 2004, Rzhetsky and Nei 1992). Rzhetsky-Nei (R-N) distances were calculated between 1) my Bayesian mitogenome phylogeny and each of the gene-family topologies (protein-coding, tRNA and rRNA topologies), 2) each individual gene tree and the Bayesian full mitogenome tree, and 3) each individual gene tree against the corresponding gene family tree (each individual protein-coding gene tree was run alongside the protein coding family topology). Calculations for all topology comparisons were done using the dist.topo function in the ape package (version 3.4) in R (Paradis et al. 2004, R Core Team 2014). An R-N distance of 0 indicates identical phylogenies. High Rzhetsky-Nei distances indicate divergent trees (Duchene et al. 2011). When conducting multiple tree comparisons, tree pairs with the lowest R-N distances are most similar to one another. Statistical significance tests are not available for R-N distances; this is a method for comparison between topologies only.

2.7 Comparisons to Previous Studies

The phylogenetic analyses of sturgeons conducted by Birstein et al. (2002) included a taxon composition close to that of the present study, and provides three topologies to compare with my results. They conducted parsimony and maximum likelihood analyses of 22 sturgeon species based on NADH5, cytochrome b, 12s, 16s, and D-loop (Figures 1, 2). They also conducted a combined-evidence phylogenetic analysis of 16 sturgeon species, based on the same molecular data and morphological data derived from Mayden and Kuhajda 1996 (Figure 3).
In order to compare Birstein et al.'s (2002) parsimony and maximum likelihood trees to my Bayesian, parsimony, maximum likelihood, 12s, 16s, cyt b, NADH5, and D-loop topologies, I pruned (using the command drop.tip in the ape package, version 3.4, Paradis et al. 2004, for R, R Core Team 2014) outgroups and sturgeon taxa not included in both studies from the trees. I ran Rzhetsky-Nei analyses on the pruned topologies to compare Birstein et al.'s (2002) parsimony and maximum likelihood tree with mine. I performed the same analytical methods using my gene trees and Birstein et al.'s (2002) parsimony and maximum likelihood trees to determine which genes most strongly influence their trees. I also compared my phylogenies and gene trees to Birstein et al.'s (2002) combined-evidence phylogeny. Birstein et al.'s combined evidence phylogeny included fewer taxa than their parsimony and maximum likelihood phylogenies, requiring pruning of additional taxa from my phylogenies and gene trees. I calculated R-N distances to determine the level of similarity between the different studies and to identify the genes most strongly represented in the combined evidence phylogeny of Birstein et al. (2002).

3. Results

3.1 Sequences and Alignment

The final, edited, aligned matrix for the full mitogenome was 14,105 base pairs in length (Appendix 1). It included the control region, or D-loop (1,295 base pairs), 13 protein coding genes (9,614 base pairs total), both rRNAs (2,676 bps long), 22 tRNAs (1,693 bps). Gene order was conserved across sturgeon mitogenomes. Substitution models, gene family, and gene length for each gene are presented in Table 2.
3.2 **Full mitogenome results**

3.2.1 **Phylogenetic Hypothesis Support Values**

The full mitogenome Bayesian phylogenetic hypothesis recovered the sturgeons as a clade with posterior probability (PP) support of 1.0 (Figure 4). Maximum likelihood recovered sturgeons as monophyletic with a bootstrap (BS) value of 100 (Figure 5), while parsimony recovered sturgeons as monophyletic, though with a BS value of only 50 (Figure 6).

3.2.2 **Full Mitogenome Hypotheses**

The Bayesian, maximum likelihood, and parsimony full mitogenome phylogenetic hypotheses had similar topologies, with only a few differences. For ease of discussion, these hypotheses will subsequently be referred to as the Bayesian, maximum likelihood, and parsimony topologies. All three full mitogenome topologies recovered *Scaphirynchus* as a clade in which *S. platatorynchus* and *S. albus* are sister ([*S. suttkusi, (S. platatorynchus, S. albus)*])

(Bayesian PP= (1.0,(.78)), parsimony BS=(100,(88)), maximum likelihood BS=(1.0,(.95))).

*Acipenser sturio* was recovered in a clade with the sister taxa *A. oxyrinchus* and *A. o. desotoi* [[*A. sturio, (A. oxyrinchus, A. o. desotoi)*]] in all three topologies (support values of 100 in all). A clade with two sets of sister species was recovered in all the three analyses: *A. dabryanus* as sister to *A. sinensis* and *A. transmontanus* as sister to *A. schrenckii* [[*A. dabryanus, A. sinensis*, *(A. transmontanus, A. schrenckii)*]] (BS=100, PP=1.0). *Huso dauricus* was consistently recovered in a clade with the sister taxa *A. medirostris* and *A. mikadoi* [[*H. dauricus, (A. medirostris, A. mikadoi)*]] (BS=100, PP=1.0). *Huso huso* was basal in a clade that contained *P. kaufmanni* as sister to *A. stellatus* [[*A. stellatus, P. kaufmanni*]] and also included [[*A. ruthenus, (A. fulvescens, (A. brevirostrum, (A. baerii, (A. naccarii, A. gueldenstaedtii)))*]]. It was in this clade that some differences across topologies were found. Other differences were found in the relationships amongst the clades.
For ease of discussion among topologies, these clades (Table 4) will subsequently be referred to as:

**Clade 1:** *(H. huso, Subclade 1.1, Subclade 1.2)*

- **Subclade 1.1:** *(A. stellatus, P. kaufmanni)*
- **Subclade 1.2:** *(A. ruthenus, (A. fulvescens, (A. brevirostrum, (A. baerii, (A. naccarii, A. gueldenstaedtii))))*

**Clade 2:** *(H. dauricus, (A. medirostris, A. mikadoi))*

**Clade 3:** *((A. dabryanus, A. sinensis), (A. transmontanus, A. schrenckii))*

**Clade 4:** *(A. sturio, (A. oxyrinchus, A. o. desotoi))*

**Clade 5:** *((S. suttkusi, (S. platorynchus, A. albus)))*

Although the three full mitogenome analyses (Bayesian, Parsimony, and Maximum Likelihood, Figures 4-6) all recovered Clade 1, the position of this clade varied among phylogenies. All three topologies recovered Subclade 1.1, *P. kaufmanni* and *A. stellatus* as sister, and Subclade 1.2, *(A. ruthenus, (A. fulvescens, (A. brevirostrum, (A. baerii, (A. naccarii, A. gueldenstaedtii))))*. The Bayesian and maximum likelihood analyses recovered Clade 1 (within which *H. huso* was basal) as sister to Clade 2. The maximum parsimony analysis recovered Subclade 1.1, Subclade 1.2, and *H. huso* to be in a polytomy.

Clades 2 and 3 were consistently recovered as sister to each other in all analyses. The Bayesian analysis recovered Clade 4 as sister to this group and Clade 1 as sister to the Clade 2, 3, and 4 grouping. Parsimony analyses recovered an opposing topology, with Clade 1 as sister to Clades 2 and 3 and Clade 4 as sister to the three clade grouping. Both Bayesian and parsimony phylogenies recovered Clade 5 *((S. suttkusi, (S. platorynchus, S. albus)))* as sister to the other sturgeons. In contrast, the maximum likelihood analysis recovered Clade 5 as sister to Clade 1. Clades 2 and 3 were recovered as sister to Clade 1 + 5 group, with Clade 4 as sister to all other sturgeons.
I favor the topology resulting from Bayesian analyses for comparison to individual gene
trees for several reasons. First, this method allowed me to specify the best model of nucleotide
evolution for the full mitogenome and the individual gene analyses. Second, it showed slightly
more phylogenetic structure than the parsimony phylogeny (which contained a polytomy).
Finally, multiple previous datasets, including molecular, morphological, and combined evidence
data, support a Scaphirhynchus as the sister lineage to all other sturgeons, as recovered in this
topology.

3.3 Gene Tree Analysis and Comparison to the Full Mitogenome

The Rzhetsky-Nei distances between the Bayesian full mitogenome topology and the
Bayesian protein-coding, tRNA, and rRNA topologies were 1, 10, and 12 respectively (Table 3).
This indicates that the protein-coding gene family produced a phylogenetic hypothesis most
similar, but not identical, to that of the full mitogenome. The tRNAs had the second most
similar topology. The rRNA gene family tree was the least similar to the full mitogenome
phylogeny.

3.3.1 tRNA Gene Family Phylogeny

Aligned and edited sturgeon tRNAs ranged from 67 to 165 base pairs in length (Table 2),
and the complete tRNA dataset was 1,693 base pairs long. The Bayesian tRNA gene family
phylogeny (Figure 7) recovered the same composition and structure within Clades 2, 4, and 5
(Table 4) as the full mitogenome Bayesian analysis. Clade 3 was found to be monophyletic, but
the sister grouping of A. transmontanus and A. schrenckii was dissolved, so that these two
species formed a polytomy with (A. dabryanus, A. sinensis). As in the full mitogenome topology,
Clades 2 and 3 were recovered as sister, and Clade 5 (Scaphirhynchus) was sister to all other
extant sturgeon taxon.
In the tRNA phylogeny, Clade 1 was nonmonophyletic. Subclade 1.1 was also nonmonophyletic; *P. kaufmanni* was recovered as sister to *H. huso*. Four species of Subclade 1.2 grouped together, with a structure similar to that recovered in the full mitogenome analysis: (*A. brevirostrum, A. baerii, A. naccarii, A. gueldenstaedtii*)); these four species formed one branch of a five branch polytomy. The other branches are: *A. stellatus; A. fulvescens*, Clades 2 and 3 as sister; and *H. huso* and *P. kaufmanni* as sister. *Acipenser ruthenus* (full mitogenome Subclade 1.2) was recovered as sister to this polytomy. Clade 4 was recovered as sister to all other sturgeons except *Scaphirynchus*.

3.3.2 rRNA Gene Family Phylogeny

The aligned, edited rRNA sequence was 2,676 base pairs long, with a 960 bp 12s and a 1,716 bp 16s. The Bayesian rRNA gene family phylogeny (Figure 8) recovered the same composition and structure within Clades 2-5 (Table 4) as the full mitogenome Bayesian analysis. Clade 1 was monophyletic, but the structure was different from that recovered by the full mitogenome. Clades 2 and 3 were recovered as sister taxa and *Scaphirynchus* (Clade 5) was found to be sister to all other extant sturgeons. Clades 2 and 3 were sister to Clade 1 (these were recovered as sister to Clade 4 in analysis of the full mitogenome).

Within Clade 1, Clades 1.1 and 1.2 were found to be nonmonophyletic in the rRNA phylogeny. *Acipenser brevirostrum* and *P. kaufmanni* were recovered as sister taxa in Clade 1. Rather than being found sister to other Clade 1 taxa, as in the full mitogenome phylogeny, *H. huso* was recovered as sister to *A. ruthenus*, with *A. stellatus* as sister to them. Sister to this grouping was the clade (*A. fulvescens, A. baerii, A. naccarii, A. gueldenstaedtii*))). Although the relationships among *A. baerii, A. naccarii*, and *A. gueldenstaedtii* match those recovered in the full mitogenome analysis, a close relationship of *A. fulvescens* with these taxa does not.
3.3.3 Protein-coding Gene Family Phylogeny

Lengths of the edited, aligned gene sequences varied from 169 to 1422 base pairs, with the protein coding gene family alignment having a total of 9,614 base pairs. The resulting Bayesian protein-coding gene family phylogeny (Figure 9) recovered Clades 1-4 from the full mitogenome topology, with the same within-clade structure (Table 4). A monophyletic Scaphirynchus was recovered, although relationships among the species were unresolved. Clades 2 and 3 were recovered as sister taxa, as in the other topologies. Clade 1, however, was recovered as sister to Clade 5. Further diverging from the full mitogenome topology, analysis of the protein-coding genes recovered Clade 4 as sister to all other extant sturgeons.

3.4 Tree Comparisons: Influence of Individual Gene Regions on Gene Tree topology

Among tRNAs, the lowest R-N distance (indicating the most similar topology) between a tRNA gene tree and the tRNA gene family phylogeny was that of tRNA Asp (R-N=11). The tRNAs Cys, Phe, and Ser (segment 1) all had the next shortest distances (R-N= 12). Of the two rRNAs, 16s produced a tree most similar to that of the rRNA phylogeny, with a R-N distance of 11; that of 12s (R-N= 12) was slightly higher.

Among the protein coding genes, ATP6 had the shortest R-N distance (1) to the full protein coding phylogeny. NADH5 had the second shortest distance of 5. COIII, NADH1, and NADH4L all had distances of 8, and NADH2 had a distance of 9. Both NADH3 and NADH4 had R-N distances of 11 to the protein coding family phylogeny, and ATP8 and COII were each a R-N distance of 12 to the protein-coding phylogeny.

3.5 Gene Influence on Bayesian Full Mitogenome Tree

3.5.1 Gene Families

The R-N distances suggest that the protein coding genes as a group likely influence the topology of the full mitogenome phylogenetic hypothesis most strongly, followed by the tRNAs,
with the rRNA gene family being least influential (Table 3). As a unit, the protein coding genes
recovered, with strong support (posterior probability values >0.98), all five clades that were also
recovered by the full mitogenome. The phylogenetic structure within all but one of these
groups was identical to that recovered by the full mitogenome, although the structure among
the clades differed greatly from that of the full mitogenome. For example, the protein-coding
genes recover *Scaphirhynchus* nested within all other sturgeons, whereas Bayesian analysis of
the full mitogenome found it to be the sister group of all other sturgeons.

In contrast, the tRNA analysis recovered *Scaphirhynchus* as sister to the other sturgeons.
This gene family, however, only recovered four of the five monophyletic groupings (Clades 2-5)
found by the full mitogenome. Structure matching that of the full mitogenome was recovered
with strong support in Clades 2 and 4 (PP >.99) and with weak support in Clade 5
(*Scaphirhynchus*, PP= 0.78). Only partial structure (one of two sister taxa) was recovered in
Clade 3.

The rRNA data recovered all five monophyletic groupings that were recovered by the
Bayesian analysis of the full mitogenome, and identical structure to the full mitogenome in four
of these (PP >0.95) (Clades 2-5). Clade 1 was recovered as monophyletic but the structure
within this grouping does not match that found in analysis of the full mitogenome and support
within Clade 1 was low (average posterior probability <0.80 for nodes within this clade).

3.5.2 Individual Genes

Comparisons between individual gene topologies and the full mitogenome hypothesis
found that the 25% best-scoring genes (i.e., most-similar topologies) were 16s rRNA and eight
protein-coding genes (ATP6, NADH5, CO3, NADH1, NADH4L, NADH2, NADH4, and NADH3)
(Table 5). These genes exhibit strong phylogenetic signal matching that of the full mitogenome.
Protein coding genes made up 13 of the 14 genes recovering topologies most similar to that of the full mitogenome. ATP6 and NADH5 were the two genes with topologies most similar to that of the full mitogenome, having Rzhetsky-Nei scores of 2 and 6, respectively. COIII, NADH1, NADH4L, and 16s, had Rzhetsky-Nei scores of 9. The tRNAs Asp, Cys, Phe, Ser1, and His fell within the 50% best scoring genes, indicating that they had the most similar topologies to the full mitogenome; Asp had the shortest Rzhetsky-Nei distance (15). D-loop had the greatest Rzhetsky-Nei distance (33) from the full mitogenome phylogenetic hypothesis, indicating a lack of similarity between the topologies. This non-coding region recovered only one clade also found by the full mitogenome phylogenetic hypothesis. Additional structure was present in this phylogeny, but this structure is in conflict with that of the full mitogenome and had poor support values (PP <0.90 for more than half the clades).

Evolutionary rate of gene regions does not seem to correlate with the recovery of the same topology as the full mitogenome. Four models of nucleotide rate substitution (Jukes-Cantor, General Time Reversible, Hasegawa-Kishino-Yano, and Kimura 2-parameter) and four models of rate heterogeneity (equal, gamma, some sites invariable, and some sites invariable with the rest drawn from gamma-distributionTable 2) were represented across all genes. Both the most similar (ATP6) and least similar (D-loop) gene trees to the full mitogenome had the same models of nucleotide substitution.

A Pearson's product-moment correlation test between Rzhetsky-Nei distance to the full mitogenome phylogenetic hypothesis and the length (in base pairs) of each gene revealed a significant linear relationship (r=0.59, p<0.05). This test indicates a relationship between the similarity of a single gene tree to the full mitogenome topology and the length of the gene sequence, with longer genes carrying more phylogenetic signal (similar to that of the full
mitogenome) than shorter genes. This contradicts previous studies, where gene length was considered to be a poor estimator of phylogenetic signal (Zardoya and Meyer 1996).

3.6 Comparison to Birstein et al. (2002)

3.6.1 Parsimony

The molecular phylogenetic hypothesis of Birstein et al. (2002) based on a maximum parsimony analysis contains multiple polytomies made up of 9 clades, named here as Polytomy A, B, and C and Clades 1-9 (Figure 1, Table 6). The most derived polytomy, Polytomy A, comprises five clades: A1) (A. brevirostrum, (A. baerii, (A. gueldenstaedtii, (A. persicus, A. naccarii))), A2) (A. stellatus, Pseudoscaphirynchus), A3) (H. dauricus, (A. ruthenus, H. huso), A4) A. fulvescens, and A5) A. nudiventris. Sister to this polytomy is A. mikadoi. This grouping is in another polytomy, Polytomy B, with two additional clades, B1) (A. schrenckii, (A. transmontanus, A. medirostris)) and B2) A. sinensis. A third polytomy, Polytomy C, comprises Polytomy B and two more clades, C1) (A. sturio, A. oxyrinchus) and C2) Scaphirynchus. The clade containing all three species Scaphirynchus species (S. platorynchus, S. albus, and S. suttkusi) recovered these species in a polytomy.

Comparing Birstein et al.’s (2002) parsimony topology with the results of my parsimony analysis reveals broad similarity in the composition of monophyletic groupings (not considering taxa that are missing from either topology). With the exception of H. dauricus, all of the taxa in Birstein et al.’s (2002) Polytomy A were found to form a clade in my topology. All of the taxa included in Polytomy B were also found to be monophyletic in my topology, although H. dauricus was recovered with A. sinensis, A. transmontanus, A. schrenckii, A. medirostris, and A. mikadoi instead of with Polytomy A as in Birstein et al. (2002).
Reviewing the topologies clade-by-clade reveals, on a finer scale, more similarities and several differences. Birstein et al.’s (2002) Clade A1 is nearly identical to Clade 1 in my topology both in composition and structure, except that I did not include A. persicus and A. nudiventris in my study. Birstein et al.’s (2002) Clade A2, the sister-species relationship between A. stellatus and Pseudoscaphirynchus, was also recovered in my topology, although Birstein et al. (2002) included both P. hermanni and P. kaufmanni in their study, and I only included P. kaufmanni. As in Birstein et al.’s study, my results recovered these two clades in a broader monophyletic grouping including H. huso and A. ruthenus. Unlike the topology recovered by Birstein et al. (2002), in my analysis A. fulvescens (Clade A4) was found to be sister to Clade A1, with H. huso sister to that grouping. Huso dauricus was not monophyletic with H. huso and A. ruthenus (Clade A3). Birstein et al.’s (2002) Clade B1 and B2 were not recovered by my analysis. Instead, A. mikadoi (Clade C2), A. medirostris (Clade B2), and H. dauricus (Clade 3) were found to be monophyletic and sister to A. dabryanus (not included by Birstein et al. (2002), A. sinensis (Clade C1), A. transmontanus (Clade B2), and A. schrenckii (Clade B2).

Birstein et al.’s (2002) monophyletic group A. sturio + A. oxyrinchus was reflected in my parsimony topology, which included both subspecies of A. oxyrinchus. Similarly, Birstein et al.’s (2002) monophyletic Scaphirynchus was included in my topology, although they recovered these taxa in a polytomy, whereas my hypothesis found S. platorynchus and S. albus to be sister taxa. Birstein et al. (2002) based their parsimony analysis on five genes: 12s, 16s, D-loop, cyt b, and NADH5. Comparison of the single gene-trees I generated indicated that four of these genes 12s, D-loop, cyt b, and NADH5 also recovered the three Scaphirynchus species in a monophyletic polytomy. A lack of structure among the taxa composing Polytomy A was reflected in the 12s, D-loop, and 16s gene trees. Rzhetsky-Nei distances indicate that of the five individual genes included in Birstein et al.’s analysis, the NADH5 gene tree has the shortest distance from, and is
therefore most similar to, Birstein et al.’s (2002) parsimony topology (R-N = 17), followed by D-loop and 16s (both have R-N = 20). These genes likely influence the topology recovered by Birstein et al. most strongly.

The differences of topology and resolution between my parsimony tree and Birstein et al.’s parsimony Tree may be related to the amount of data used. Of the five genes Birstein et al. (2002) used in their phylogeny, I found two to be inconsistent with the full mitogenomic phylogeny (12s and 16s, discussed above). Rzhetsky-Nei analyses indicate that of the thirteen genes that most strongly influence the results of my parsimony analysis, eleven were found to produce topologies similar to that of the full mitogenome. These results support the finding of previous studies that an increase in data improves resolution in phylogenetic analyses (Hajibabei et al. 2007).

3.6.2 Maximum Likelihood

Birstein et al.’s (2002) maximum likelihood topology recovered several clades referred here as clades ML1-6 (Figure 2, Table 7). Solely Acipenser species made up Clade ML1: (A. fulvescens, (A. brevirostrum, (A. baerii, (A. gueldenstaedtii, (A. naccarii, A. persicus))))). Clade ML1 was recovered as sister to Clade ML2: (A. stellatus, (P. kaufmanni, P. hermanni)), with A. nudiventris as sister to this group. Clade ML3: (H. dauricus, (H. huso, A. ruthenus)) was found to be sister to Clades ML1, ML2, and A. nudiventris; the three species of Scaphirynchus form a polytomy, Clade ML4, which was found to be the sister group of Clades ML1, ML2 and ML3. A. mikadoi is sister group of to Clades ML1-ML4. Four Acipenser species form Clade ML5: (A. mediostris, (A. sinensis, (A. transmontanus, A. schrenckii))). Clade ML6, (A. oxyrinchus, A. sturio), was found to be the sister group of all other sturgeons.
Comparing Birstein et al.'s (2002) maximum likelihood topology to my full mitogenome maximum likelihood topology reveals a few similarities and many differences, with an R-N distance of 27. Clade ML1 is identical to my Subclade 1.2, except that A. persicus was not included in my analyses. Likewise, Clade ML2 is identical to the sister grouping my analyses recover between Pseudoscaphirynchus and A. stellatus, except that my study did not include P. hermanni. The species in Clade ML3 of Birstein et al.'s (2002) phylogeny are not recovered in a clade in my analyses. Instead, in my topology, A. ruthenus was recovered as sister to Clade ML1, and H. huso was sister to Clades ML1 + ML2. Huso dauricus was found to be sister to A. mikadoi and A. medirostris (which was recovered by Birstein et al. (2002) in Clade ML5). Like Birstein et al.'s (2002) maximum likelihood topology, my analyses recovered a monophyletic Scaphirynchus as sister to the taxa in Clades ML1, ML2, and ML3 (except for H. dauricus). Scaphirynchus albus and S. platorynchus were recovered as sister species within this clade in my topology. The taxa Birstein et al. (2002) recovered in Clade 5 were not recovered as monophyletic in my analyses. In my topology, A. medirostris was recovered as sister to A. mikadoi in a clade also containing H. dauricus (Clade ML3). This clade was recovered as sister to a monophyletic group containing three of Birstein et al.'s (2002) Clade ML5 species (A. transmontanus, A. schrenckii, and A. sinensis) along with A. dabryanus, which was not included in Birstein et al.'s (2002) study. As in Birstein et al.'s (2002) topology, though, my phylogenetic hypothesis recovered A. sturio + A. oxyrinchus as the sister group of all other sturgeons.

Of the genes included in Birstein et al.'s (2002) analyses, NADH 5 appears to be the most influential in the maximum likelihood topology (R-N = 17), followed by 16s (R-N= 18), cyt b (R-N= 21), D-loop (R-N= 22), and finally 12s (R-N= 23). D-loop is the only one of these genes that recovered Scaphirynchus within Acipenser. I consider the recovery of Scaphirynchus within Acipenser to be a significant difference between these two topologies and the other gene trees.
3.6.3 Combined Evidence

Birstein et al.'s (2002) combined evidence topology, based on both molecular and morphological data, recovered five distinct clades (Table 8). Clade CE1 included the species (A. baerii, A. brevirostrum) with Clade CE2 (A. stellatus, (P. hermanni, P. kaufmanni) and Clade CE3 (H. dauricus, (A. ruthenus, H. huso)) as sequential sister groups. Acipenser fulvescens was found to be sister to (Clade CE3, (Clade CE2, Clade CE1)). Acipenser sinensis + A. transmontanus (Clade CE4) formed the sister group of the previously described clades. Acipenser oxyrinchus was sister to Clades CE1-4. Scaphirhynchus (Clade CE5), with S. albus and S. platatorynchus as sister taxa, was the sister group of all other sturgeons.

Based on Rzhetsky-Nei analyses, the Birstein et al. (2002) combined evidence phylogeny (molecular and morphological data), is most similar to the topology based on my parsimony analysis (R-N= 9), than to the results of my Bayesian analysis (R-N= 12), and least similar to my maximum likelihood results. The results of my parsimony analysis suggests that A. oxyrinchus and A. sturio form the sister group of all sturgeons other than Scaphirhynchus, as does the Birstein et al. combined evidence phylogeny. This similarity is absent from my Bayesian results. This illustrates that the algorithm selected for phylogenetic analysis dictates a role in the relationships recovered.

Rzhetsky-Nei analyses suggest that, of the genes considered, Birstein et al.'s (2002) combined evidence phylogeny is most strongly influenced by NADHS (R-N=10), then by 16s (R-N=11), followed by 12s and cyt b (R-N=16), with D-loop being the least influential (R-N=28).

48
4. Discussion

4.1 Phylogenetic structure within Acipenseridae

Several studies have proposed varying relationships regarding \textit{Huso huso}. Based on skeletal characters, Findeis (1997) recognized a subfamily Husinae, containing the two species in the genus \textit{Huso} (\textit{H. huso} and \textit{H. dauricus}), with this subfamily as sister to all other Acipenseridae. \textit{Huso} since has been found to be nonmonophyletic (Krieger et al. 2008), with \textit{H. huso} included in \textit{Acipenser} (Vasil’eva et al. 2009). Relationships between \textit{H. huso} and various species or species groups have been recovered in past analyses, including those similar to the monophyletic grouping I recovered here and discussed above. For example, based on morphological data, Hilton et al. (2011) defined the subfamily Husinae as including only \textit{H. huso} and \textit{A. ruthenus}, with \textit{A. baerii} as its sister-group. Although I recover \textit{H. huso} in a clade with these species, these data and analyses do not support Husinae \textit{sensu} Hilton et al. (2011).

My recovery of \( (H. huso, ((Acipenser stellatus, Pseudoscaphirynchus kaufmanni), (A. ruthenus, (A. fulvescens, (A. brevirostrum, (A. baerii, (A. naccarii, A. gueldenstaedtii))))) \)
illustrates a potential new sturgeon clade containing significant phylogenetic structure. This result supports the previous recovery of \textit{Pseudoscaphirynchus} and \textit{A. stellatus} as a clade in both morphological (Hilton et al. 2011) and molecular studies (Birstein et al. 2002, Dillman et al. 2007, Krieger et al. 2008). My results support the recognition of Pseudoscaphirhynchinae, proposed by Hilton et al. 2011, that includes \textit{Pseudoscaphirynchus} and \textit{A. stellatus}.

Recovery of \textit{H. huso} within a group of \textit{Acipenser} species is not surprising, as \textit{H. huso} has been frequently recovered within \textit{Acipenser} (e.g. Ludwig et al. 2001, Zhang et al. 2000, Fontana et al. 2001, Hilton et al. 2011, Dillman et al. 2007, Krieger et al. 2008). Krieger et al. noted that two positions of \textit{Huso} within \textit{Acipenser} have been proposed: 1) \textit{Huso} clustered with \textit{A. ruthenus}; and
2) *Huso* basal to other *Acipenser* species. Krieger et al. (2008) first recovered *H. huso* clustered with *A. rathenus*. After examination of sequences used in analyses, Krieger et al. (2008) suggested that this position could be a result of past introgression. When Krieger et al. (2008) excluded *A. rathenus* from analyses, *H. huso* was found as sister to other *Acipenser* species. Krieger et al. (2008) analyzed sequences of specific gene segments used in their and Birstein et al.'s. (2002) analyses, and found that the 12s sequences of the two taxa had only three differences. In my analyses *A. rathenus* was recovered in the same clade as *H. huso* by full mitogenome and protein-coding data, but *H. huso* was found to be basal to the other taxa within this clade, with *A. rathenus* more deeply nested in this clade. My rRNA topology, however, recovered *H. huso* and *A. rathenus* as sister taxa within a broader *Acipenser* clade. My 12s topology recovered *H. huso* clustered with *A. rathenus* and *A. stellatus*, whereas 16s recovered *H. huso* as sister to an *Acipenser* species clade. None of the other gene trees recovered a close relationship between *H. huso* and *A. rathenus*. Although further investigation might reveal more regarding the position of *Huso*, my study suggests that a close or sister relationship with *A. rathenus* is likely influenced by 12s. My analyses ultimately support an *Acipenser* species clade in which *H. huso* occupies a sister position.

My full mitogenome and protein coding data configurations recovered a core group of taxa clustered together: *Acipenser gueldenstaedtii, A. naccarii, A. baerii, A. brevirostrum,* and *A. fulvescens*. These species have been recovered together by several previous studies. Ludwig et al. (2001) and Fontanta et al. (2001) used molecular analyses and recovered these taxa (along with *A. persicus*) with the same structure as my full mitogenome and protein coding phylogenies. This clade is not novel, but my data support previous phylogenetic hypotheses including this configuration of taxa.
The grouping of \textit{A. medirostris}, \textit{A. mikadoi}, and \textit{H. dauricus} was recovered here and previously in analyses by Krieger \textit{et al.} (2008) and Ludwig \textit{et al.} (2001). The separation between \textit{H. dauricus} and \textit{H. huso} provides further support for the dissolution of a monophyletic \textit{Huso}.


All four of the main data configurations recovered \textit{Scaphirhynchus} as a monophyletic genus including \textit{S. albus}, \textit{S. platorhynchus}, and \textit{S. sukkusi}. My recovery of \textit{Scaphirhynchus} as the sister group of all other sturgeons (by all data configurations except for the protein-coding sequences) supports the dissolution of the tribe Scaphirhynchini (sensu Findeis, 1997), as previously proposed (Birstein \textit{et al.} 2002). \textit{Scaphirhynchus} has been recovered as basal to other sturgeons by molecular analyses, morphological studies, and studies combining both types of data. This topology was recovered in a strictly molecular analysis based on \textit{cytb}, 12s, and 16s, and in a combined-data study using these genes, NADH5, D-loop, and morphological data. A
morphological study including extant and extinct sturgeons recovered two equally most parsimonious trees, one with each topology (Hilton et al. 2011). The authors favored the topology recovering *Scaphirhynchus* as sister to all other extant sturgeons. Some previous studies have instead considered *Scaphirhynchus* to be within *Acipenser*; these include studies based on cyt* and studies using cytochrome *b* with cytochrome oxidase 2, NADH5, D-loop, both rRNAs, and the tRNAs Phe and Asp (Fontana et al. 2001, Krieger et al. 2008).

Most of my analyses recover *S. albus* and *S. platorhynchus* as sister taxa within the monophyletic genus *Scaphirhynchus*. Support for this relationship is especially strong (e.g., PP=1) in the full mitogenome and rRNA analyses. This is consistent with previous studies combining morphological and molecular data (Birstein et al. 2002), but conflicts with studies relying exclusively on protein-coding genes or morphological data. Based on morphological characters, Mayden and Kuhajda (1996) recovered *S. suttkusi* and *S. platorhynchus* as sister taxa relative to *S. albus*. Ludwig et al. (2001) recovered that same sister relationship using exclusively cyt* but Fontana et al. (2001), who also used cyt* instead found *S. suttkusi* to be sister to *S. albus* relative to *S. platorhynchus*. My data support a monophyletic *Scaphirhynchus* sister to other extant sturgeons, with a sister relationship between *S. platorhynchus* and *S. albus*, as previously recovered by Birstein et al. (2002).

My full mitogenome analysis recovered a large clade containing *A. oxyrinchus, A. sturio, A. schrenckii, A. transmontanus, A. sinensis, A. dabryanus, A. mikadoi, A. medirostris*, and *H. dauricus*. Within this group are three well-supported, previously defined clades. Two of these clades, (*H. dauricus, (A. medirostris, A. mikadoi) and (A. dabryanus, A. sinensis), (A. transmontanus, A. schrechnii)*), were recovered as sister by all of my major data configurations. This sister relationship has been previously recovered by Birstein and DeSalle (1998) based on
protein coding and rRNA sequences. A sister relationship between these grouping (excluding *A. dabryanus*) was found by Ludwig *et al.* (2001).

### 4.2 Influence of Individual Genes on Phylogeny

All of the clades recovered by the full mitogenome phylogenetic hypothesis were also recovered by ATP6, and all but one had the same structure exhibited in the full mitogenome phylogenetic hypothesis. It is unsurprising that NADH5 also yielded one of the most similar gene trees to the full mitogenome phylogenetic hypothesis; previous studies have found this gene to recover mammal, dolphin, orca, teleostean, and sturgeon phylogenies similar to that recovered by the full mitogenome (Zardoya and Meyer 1996, Miya and Nishida 2000, Rastorguev *et al.* 2008, Duchene *et al.* 2011). COIII, NADH1, and NADH4L recovered topologies similar to that found by the full mitogenome phylogenetic hypothesis. CO3 was previously found to recover teleostean phylogenies similar to that of the full mitogenome (Miya and Nishida 2000) and NADH1 was useful in testing the phylogeny of Delphinae (Duchene *et al.* 2011). Previously, Rastorguev *et al.* (2008) found that NADH4L did not recover a phylogeny very similar to that of the full mitogenome topology of Russian sturgeons. The relatively high (compared to other protein trees) tree distance results from this study for ATP8, COII, NADH2, NADH3, and NADH6 (R-N = 10 to 14) correspond with conclusions of similar studies that although these genes carry similar signal to that of the full mitogenome in various taxa, the signal carried by other protein-coding genes is more similar (Zardoya and Meyer 1996, Miya and Nishida 2000, Rastorguev *et al.* 2008, Duchene *et al.* 2011). Surprisingly, I obtained relatively high Rzhetsky-Nei distances between the full mitogenome and both cyt b and COI. These genes have been consistently found by previous studies to recover topologies similar to that of the full mitogenome (in studies of sturgeons, teleosts, dolphins, and terrestrial mammals) (Zardoya and Meyer 1996, Miya and Nishida 2000, Rastorguev *et al.* 2008, Duchene *et al.* 2011). Although cyt b and COI did recover
similar topologies to the full mitogenome in this study, the topologies recovered by 16s and the other protein coding genes were more similar to that of the mitogenome.

Results of my tree comparison indicate that individual mitogenome segments with varying functions strongly influence the topology of the full mitogenome phylogenetic hypothesis. Further, segments in the same gene family (rRNA, tRNA, protein-coding) do not necessarily carry the same phylogenetic signal. Therefore, I cannot recommend restricting mitochondrial phylogenetic studies by gene family. Instead, I underscore conclusions of previous studies indicating that as a group, the protein coding genes carry similar signal to the full mitogenome, but this signal is not identical to that of the full mitogenome (when all genes are considered). In addition, 16s rRNA recovered a topology similar to that of the full mitogenome. Further studies, however, are necessary to determine the relative influence of signal carried by 16s rRNA, as it relate to the full mitogenomic signal, beyond this family.

4.3 Taxon coverage

This study represents almost complete taxon coverage for extant sturgeons. The three extant (P. fedtschenkoi is presumed extinct) sturgeons missing from this analysis are P. hermanni, A. persicus, and A. nudiventris (IUCN 2014). While the absence of species from a topology represents missing phylogenetic signal that, if present, might change the topology, the omission of P. hermanni and A. persicus from my analyses might be of little consequence. *Pseudoscaphirhynchus hermanni* has been recovered, by both morphological (Hilton *et al.* 2011) and molecular datasets (e.g. Dillman *et al.* 2007, Birstein *et al.* 2002), as sister to *P. kaufmanni*, which was included in my study. Molecular characters recovering these taxa as sister taxa include three of the genes I found to carry similar phylogenetic signal to the full mitogenome phylogenetic hypothesis (NADH5, 16s, and cytb; Birstein *et al.* 2002). Because previous studies based on protein coding genes have recovered *P. hermanni* and *P. kaufmanni* as sister, I predict
that if *P. hermanni* had been included in my study the same sister relationship would have been recovered. This would likely not change the rest of the topology.

Also missing from my analyses is *A. persicus*, which has been regarded as a subspecies of *A. gueldenstaedtii* rather than a valid species (Birstein *et al.* 2005). A recent study focused on the taxonomic status of *A. persicus* employed morphological data and two mitochondrial genes, *cytb* and D-loop, and concluded that *A. persicus* is not a valid species (Ruban *et al.* 2008). The authors of the study (Ruban *et al.* 2011) argued, based on both morphological and molecular data, that these are conspecific. Molecular studies based on *cytb*, 12s, 16s, NADH5, and D-loop failed to recover *A. persicus* as a monophyletic species separate from *A. gueldenstaedtii* (Birstein and DeSalle 1998; Birstein *et al.* 2000; Birstein and Doukakis 2001, and Birstein *et al.* 2005). My gene tree comparison analyses found that three of the genes supporting this conclusion (NADH5, 16s, and *cytb*) reflect the full mitogenome phylogenetic hypothesis. I hypothesize that if included in my study, *A. persicus* would fall as a sister species to *A. gueldenstaedtii*, and would not significantly alter the overall topology.

*Acipenser nudiventris* has been recovered as sister to various taxa across phylogenetic studies. Ludwig *et al.* (2001) and Fontana *et al.* (2001) both recovered it as sister to *A. ruthenus* based on *cytb* data. Using 12s and 16s sequences, Birstein and DeSalle (1998) found *A. nudiventris* to be sister to *A. dabryanus*. Employing more sequence data, however, both Krieger *et al.* (2008; using *cytb*, 12s, cytochrome c, the tRNAs Phe and Asp, NADH5, and D-loop) and Birstein *et al.* (2002; in their combined data set) found *A. nudiventris* to be sister to (*P. kaufmanni, A. stellatus*). Both studies used *cytb*, which I found to carry a signal similar to that of the full mitogenome. Birstein *et al.* (2002) also employed NADH5, the second best scoring gene in my study for recovering full-mitogenome topology, and 16s, which also recovered a topology similar to that of the full mitogenome analysis. I consider it likely that had I been able to include
A. nudiventris in my analyses, it would have been recovered as closely related to both Pseudoscaphirhynchus and A. stellatus. I recommend that future studies include A. nudiventris, ideally with full mitogenomic data or, at least, with the full protein-coding family and 16s data, to resolve its relationship.

4.4 Mitogenome Coverage

Although two other studies included more sturgeon species than mine (Birstein et al. 2002, Krieger et al. 2008), I included the greatest amount of molecular data applied to sturgeon phylogeny to date, over 14,000 base pairs. More data may not always be better, but various studies have shown that longer mitochondrial sequences produce higher resolution phylogenies (Vilstrup et al. 2011).

4.5 Use of Mitochondrial Sequences: Benefits and Drawbacks

Most concerns regarding the use of mitochondrial DNA sequences in phylogenetic reconstruction have been addressed in this study. Although the high substitution rate is of concern due to potential saturation and subsequent loss of lineage information, genes with high substitution rates are particularly useful in determining relationships between genetically similar species (Moore 1995). Nuclear genes are also at risk for loss of evolutionary information, though; recombination can lead to loss of the phylogenetic signal reflecting true evolutionary history (Rautenberg et al. 2008). The issue of recovering poorly resolved phylogenies due to use of short sequence segments is addressed easily. Addition of more sequences, for example the use here of the full mitogenome, leads to higher resolution in the phylogenies produced. This has been illustrated in studies across taxa, and is illustrated here by the number of polytomies present in single gene trees compared to the lack of polytomies in the phylogenetic hypothesis presented here (Vilstrup et al. 2011). The concern that mitochondrial DNA is linked and therefore only provides a single estimate of phylogeny, making introgression and hybridization
undetectable, is valid. However, mitochondrial gene regions evolve at different rates and carry different phylogenetic signals, illustrated by the difference in topology among gene trees. Therefore, including the full mitogenome in this study ensures that various signals are considered in the analysis, even if they are from linked genes. Additionally, various studies have shown that the full mitogenome phylogenetic hypothesis is more likely to reflect the actual presumed species phylogeny than one nuclear gene alone (Moore 1995).

I am unable to account for potential hybridization and introgression in sturgeons, because nuclear sequences are not available for phylogenetic analysis and comparison to my phylogeny. Sturgeon are polyploid, making efforts to sequence nuclear DNA for phylogenetic studies extremely difficult (Krieger et al. 2006). This is a particular concern because both historic and recent hybridization is known in sturgeons (Peng 2007). To address this, future work should focus on sequencing nuclear DNA and producing sturgeon phylogenies for comparison to my mitochondrial phylogeny.

5. Concluding Remarks

Based on more data than has been previously included in any phylogenetic analysis of Acipenseridae, the full mitogenome of 22 sturgeon species (plus one subspecies), I propose a new phylogeny for the Acipenseridae. This topology, like many recent studies, recovers Scaphirhynchus as sister to all other sturgeons. These data support three previously described clades: (A. sturio, {A. oxyrinchus, A. o. desotoi}); ((A. dabryanus, A. sinensis), (A. transmontanus, A. schrenckii)); and (H. dauricus, {A. medirostris, A. mikadoi}). Pseudoscaphirhynchinae sensu Hilton et al. (2011) is supported, but falls within a broader clade containing additional Acipenser species and H. huso.
The monophyly of Husinae sensu Hilton et al. (2011) is not supported. Husinae, if restricted to only *H. huso*, is included in a clade with Pseudoscaphirhynchinae and a clade comprising *A. gueldenstaedtii, A. naccarii, A. ruthenus, A. baerii, A. fulvescens*, and *A. brevirostrum*; additional data are required to determine whether *A. persicus* and *A. nudiventris* (not included in this study) belong within this group.

Past molecular studies of sturgeon phylogeny have frequently employed a single or few mitochondrial genes. There is strong support for the ability of protein coding genes, individually and as a unit, to recover the same monophyletic groupings, and structure within these groupings, as the full mitogenome. My results suggest that the protein coding gene family along with the rRNA 16s may reflect a topology more similar to that of the full mitogenome than any other mitochondrial gene grouping. My results reinforce the conclusion drawn by previous authors that evolutionary rate does not account for gene performance in topology recovery, but contradict the assertion that gene length cannot be used to infer utility of a gene to recover the full mitogenome phylogenetic hypothesis. Although past studies have not found a trend between gene length and reliability in recovering phylogeny, this study did recover such a trend. I suggest that future studies consider addressing this further. The use of the full mitogenome along with other characters (nuclear genes, morphological data) may recover differing or better supported topologies.
Acknowledgements

I would like to acknowledge Dr. Masaki Miya for his tremendous mentorship, support, and input during all aspects of this research. Also, many thank to Mr. Tetsuya Sado for his assistance and kindness, and to everyone at the Natural History Museum and Institute in Chiba, Japan for welcoming me during my time there. Dr. Jan McDowell provided much appreciated assistance and advice, and allowed me space in her and Dr. John Graves’ lab. I’d like to also thank Dr. John Graves for his mentorship and support, and comments on early drafts of this chapter. Thanks also to Casey Dillman for help in development of the chapter and in the lab, and for comments on early drafts. I acknowledge all of my committee members for support and input during development of this Chapter. Also, thanks to all of my Molecular Labmates for general support and inspiration- Nadya Mamoozedeh, Brendan Turley, Hamish Small, and Heidi Brightman.

Funding sources included National Science Foundation (DEB 0841691), the NSF GK-12 Program, the NSF East Asia and Pacific Summer Institute Program, the Japan Society for Promotion of Science Summer Program, and the Virginia Institute of Marine Science Dean and Director’s Office and Fisheries Department.
References

Korean genome sequence and analysis: Full genome sequencing for a socio-ethnic

gnathostome divergences: a fish is a fish. Gene, 333: 61-70.

Artyukhin E. N. 1995. On biogeography and relationships within the genus Acipenser. Sturgeon


Avise J.C. 2010. Perspective: conservation genetics enters the genomics era. Conservation

Baack E.J. and Rieseberg L.H. 2007. A genomic view of introgression and hybrid


Caravas J. 2012. Phylogenetic Utility of Mitochondrial and Nuclear Genes: A Case Study in the Diptera (True Flies). Wayne State University Dissertations. Available at: DigitalCommons@WayneState.


62


illumin. 2012. Nextera™ XT DNA Sample Preparation Guide. ILLUMINA PROPRIETARY Part #15031942 Rev. C. Available at:


Lavoué S., Miya M., Saitoh K., Ishiguro N.B., and Nishida M. 2007. Phylogenetic relationships among anchovies, sardines, herrings and their relatives (Clupeiformes), inferred from


Qiagen. 2006. DNeasy® Blood & Tissue Handbook. In DNeasy Blood & Tissue Kit. Available at:


mitochondrial substitution rate variation in the angiosperm tribe *Sileneae*. BMC Evolutionary Biology, 9: 260. Available at:

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2777880/>


<http://www.biomedcentral.com/1471-2148/11/65>


**Tables**

**Table 1. Tissue samples used in this study.** GenBank accession numbers of the tissue specimens sequenced are listed (or are pending); collection numbers for voucher specimens are provided (in some cases these are pending cataloging in a home institution).

<table>
<thead>
<tr>
<th>Species</th>
<th>Source</th>
<th>Voucher</th>
<th>Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acipenser baerii (full mitogenome)</td>
<td>GenBank *</td>
<td>n/a</td>
<td>JQ045341.1</td>
</tr>
<tr>
<td>Acipenser baerii (16s rRNA, spot checking)</td>
<td>novel sequence</td>
<td>FMNH 117783</td>
<td>pending</td>
</tr>
<tr>
<td>Acipenser dabryanus</td>
<td>GenBank *</td>
<td>n/a</td>
<td>AY510085.1</td>
</tr>
<tr>
<td>Acipenser ruthenus</td>
<td>GenBank *</td>
<td>n/a</td>
<td>NC_022453.1</td>
</tr>
<tr>
<td>Acipenser stellatus (full mitogenome)</td>
<td>GenBank *</td>
<td>VIMS 13552</td>
<td>NC_005795.1</td>
</tr>
<tr>
<td>Acipenser stellatus (16s rRNA, spot checking)</td>
<td>novel sequence</td>
<td>pending</td>
<td>pending</td>
</tr>
<tr>
<td>Acipenser oxyrinchus desotai</td>
<td>novel sequence</td>
<td>VIMS uncataloged</td>
<td>pending</td>
</tr>
<tr>
<td>Acipenser brevirostrum</td>
<td>novel sequence</td>
<td>VIMS 12073</td>
<td>pending</td>
</tr>
<tr>
<td>Acipenser fulvescens</td>
<td>novel sequence</td>
<td>VIMS uncataloged</td>
<td>pending</td>
</tr>
<tr>
<td>Acipenser gueldenstaedtii</td>
<td>novel sequence</td>
<td>VIMS uncataloged</td>
<td>pending</td>
</tr>
<tr>
<td>Acipenser medirostris</td>
<td>novel sequence</td>
<td>VIMS 17715</td>
<td>pending</td>
</tr>
<tr>
<td>Acipenser mikadoi</td>
<td>novel sequence</td>
<td>n/a*</td>
<td>pending</td>
</tr>
<tr>
<td>Acipenser naccarii</td>
<td>novel sequence</td>
<td>n/a**</td>
<td>pending</td>
</tr>
<tr>
<td>Acipenser oxyrinchus oxyrinchus</td>
<td>novel sequence</td>
<td>VIMS 12207</td>
<td>pending</td>
</tr>
<tr>
<td>Acipenser ruthenus</td>
<td>novel sequence</td>
<td>n/a***</td>
<td>pending</td>
</tr>
<tr>
<td>Acipenser schrenckii</td>
<td>novel sequence</td>
<td>VIMS uncataloged</td>
<td>pending</td>
</tr>
<tr>
<td>Acipenser sinensis</td>
<td>novel sequence</td>
<td>VIMS 13834</td>
<td>pending</td>
</tr>
<tr>
<td>Acipenser sturio</td>
<td>novel sequence</td>
<td>VIMS 13756</td>
<td>pending</td>
</tr>
<tr>
<td>Acipenser transmontanus</td>
<td>novel sequence</td>
<td>VIMS uncataloged</td>
<td>pending</td>
</tr>
<tr>
<td>Huso dauricus</td>
<td>novel sequence</td>
<td>NSMT P 91146</td>
<td>pending</td>
</tr>
<tr>
<td>Huso huso</td>
<td>novel sequence</td>
<td>n/a***</td>
<td>pending</td>
</tr>
<tr>
<td>Pseudoscaphirhynchus kaufmanni</td>
<td>novel sequence</td>
<td>PKN-9</td>
<td>pending</td>
</tr>
<tr>
<td>Scaphirhynchus albus</td>
<td>novel sequence</td>
<td>VIMS 12193</td>
<td>pending</td>
</tr>
<tr>
<td>Scaphirhynchus platatorynus</td>
<td>novel sequence</td>
<td>VIMS 13515</td>
<td>pending</td>
</tr>
<tr>
<td>Scaphirhynchus suttkusi</td>
<td>novel sequence</td>
<td>SUT-3</td>
<td>pending</td>
</tr>
<tr>
<td>Polyodon spathula</td>
<td>GenBank *</td>
<td>n/a</td>
<td>NC_004419.1</td>
</tr>
<tr>
<td>Polypterus delhezi</td>
<td>GenBank *</td>
<td>NTNU LSP136</td>
<td>NC_020652</td>
</tr>
<tr>
<td>Polypterus endlicherii</td>
<td>GenBank *</td>
<td>NTNU LSP117</td>
<td>NC_020791</td>
</tr>
<tr>
<td>Psephurus gladius</td>
<td>GenBank *</td>
<td>n/a</td>
<td>AY571339.1</td>
</tr>
<tr>
<td>Polyodon spathula</td>
<td>novel sequence</td>
<td>n/a</td>
<td>pending</td>
</tr>
</tbody>
</table>

* * tissue sent courtesy of Shinji Adachi at Hokkaido University.
** tissue sent courtesy of Dr. Leonardo Congiu, University of Padova.
*** tissue sent courtesy of Dr. Radu Suciu, Danube Delta National Institute, Tulcea, Romania; photo available upon request.
Table 2. Gene Families, Lengths, and Best Models. Mitochondrial genes are listed, with the gene family to which each belongs and the length of each in base pairs. The best evolutionary model for each gene as determined by jModelTest is listed with the abbreviation for that model and the model name spelled out. These evolutionary models were used in all Bayesian analyses.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Gene Family</th>
<th>Length (in base pairs)</th>
<th>Best Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-loop</td>
<td>control region</td>
<td>348</td>
<td>GTR+G General time reversible</td>
</tr>
<tr>
<td>ATP6</td>
<td>protein coding</td>
<td>674</td>
<td>GTR+H+G General time reversible</td>
</tr>
<tr>
<td>ATP8</td>
<td>protein coding</td>
<td>169</td>
<td>HKY+H+G General time reversible</td>
</tr>
<tr>
<td>COI</td>
<td>protein coding</td>
<td>1422</td>
<td>GTR+H+G General time reversible</td>
</tr>
<tr>
<td>COII</td>
<td>protein coding</td>
<td>691</td>
<td>HKY+H+G General time reversible</td>
</tr>
<tr>
<td>COIII</td>
<td>protein coding</td>
<td>785</td>
<td>GTR+H+G General time reversible</td>
</tr>
<tr>
<td>Cytb</td>
<td>protein coding</td>
<td>562</td>
<td>GTR+H+G General time reversible</td>
</tr>
<tr>
<td>ND1</td>
<td>protein coding</td>
<td>750</td>
<td>GTR+H+G General time reversible</td>
</tr>
<tr>
<td>ND2</td>
<td>protein coding</td>
<td>756</td>
<td>GTR+H+G General time reversible</td>
</tr>
<tr>
<td>ND3</td>
<td>protein coding</td>
<td>348</td>
<td>GTR+H+G General time reversible</td>
</tr>
<tr>
<td>ND4</td>
<td>protein coding</td>
<td>1374</td>
<td>GTR+H+G General time reversible</td>
</tr>
<tr>
<td>ND4L</td>
<td>protein coding</td>
<td>297</td>
<td>GTR+H+G General time reversible</td>
</tr>
<tr>
<td>ND5</td>
<td>protein coding</td>
<td>1383</td>
<td>GTR+H+G General time reversible</td>
</tr>
<tr>
<td>ND6</td>
<td>protein coding</td>
<td>403</td>
<td>HKY+G Hasegawa-Kishono-Yano</td>
</tr>
<tr>
<td>12s</td>
<td>rRNA</td>
<td>960</td>
<td>GTR+H+G Hasegawa-Kishono-Yano</td>
</tr>
<tr>
<td>16s</td>
<td>rRNA</td>
<td>1716</td>
<td>GTR+H+G Hasegawa-Kishono-Yano</td>
</tr>
<tr>
<td>Ala</td>
<td>tRNA</td>
<td>74</td>
<td>HKY+I Hasegawa-Kishono-Yano</td>
</tr>
<tr>
<td>Arg</td>
<td>tRNA</td>
<td>70</td>
<td>K80 Hasegawa-Kishono-Yano</td>
</tr>
<tr>
<td>Asn</td>
<td>tRNA</td>
<td>75</td>
<td>HKY+G Hasegawa-Kishono-Yano</td>
</tr>
<tr>
<td>Asp</td>
<td>tRNA</td>
<td>87</td>
<td>HKY+G Hasegawa-Kishono-Yano</td>
</tr>
<tr>
<td>Cys</td>
<td>tRNA</td>
<td>67</td>
<td>K80+G Hasegawa-Kishono-Yano</td>
</tr>
<tr>
<td>Gln</td>
<td>tRNA</td>
<td>69</td>
<td>HKY+G Hasegawa-Kishono-Yano</td>
</tr>
<tr>
<td>Glu</td>
<td>tRNA</td>
<td>165</td>
<td>HKY+I Hasegawa-Kishono-Yano</td>
</tr>
<tr>
<td>Gly</td>
<td>tRNA</td>
<td>73</td>
<td>HKY Hasegawa-Kishono-Yano</td>
</tr>
<tr>
<td>His</td>
<td>tRNA</td>
<td>69</td>
<td>HKY Hasegawa-Kishono-Yano</td>
</tr>
<tr>
<td>Ile</td>
<td>tRNA</td>
<td>72</td>
<td>K80+G Jukes-Cantor</td>
</tr>
<tr>
<td>Leu</td>
<td>tRNA</td>
<td>75</td>
<td>K80+G Kimura</td>
</tr>
<tr>
<td>Leu2</td>
<td>tRNA</td>
<td>75</td>
<td>HKY+I Kimura</td>
</tr>
<tr>
<td>Lys</td>
<td>tRNA</td>
<td>75</td>
<td>K80+I Kimura</td>
</tr>
<tr>
<td>Met</td>
<td>tRNA</td>
<td>70</td>
<td>JC Kimura</td>
</tr>
<tr>
<td>Phe</td>
<td>tRNA</td>
<td>71</td>
<td>K80 Kimura</td>
</tr>
<tr>
<td>Pro</td>
<td>tRNA</td>
<td>71</td>
<td>HKY+G Kimura</td>
</tr>
<tr>
<td>Ser1</td>
<td>tRNA</td>
<td>78</td>
<td>K80+G Kimura</td>
</tr>
<tr>
<td>Ser2</td>
<td>tRNA</td>
<td>70</td>
<td>K80+G Kimura</td>
</tr>
<tr>
<td>Thr</td>
<td>tRNA</td>
<td>72</td>
<td>K80 Kimura</td>
</tr>
<tr>
<td>Trp</td>
<td>tRNA</td>
<td>73</td>
<td>K80 Kimura</td>
</tr>
<tr>
<td>Tyr</td>
<td>tRNA</td>
<td>71</td>
<td>K80+G Kimura</td>
</tr>
<tr>
<td>Val</td>
<td>tRNA</td>
<td>71</td>
<td>K80 Kimura</td>
</tr>
</tbody>
</table>
Table 3. Rzhetsky-Nei Distances: Bayesian Mitogenome Topology and Gene-Family Topologies. Rzhetsky-Nei Distances, signifying the level of similarity to that of the Bayesian full mitogenome topology, are listed for each gene-family tree. A distance of 0 signals identical topologies.

<table>
<thead>
<tr>
<th>Gene Family</th>
<th>Rzhetsky-Nei distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein-coding</td>
<td>1</td>
</tr>
<tr>
<td>tRNA</td>
<td>10</td>
</tr>
<tr>
<td>rRNA</td>
<td>12</td>
</tr>
</tbody>
</table>

Table 4. Full Mitogenome Hypothesis Clade Constituents, and Presence in Gene-Family Trees. Names assigned to clades and subclades recovered in the full mitogenome Bayesian phylogeny are listed, with the constituents of each clade and their presence (indicated by a check mark) or absence in the three gene-family topologies.

<table>
<thead>
<tr>
<th>Full Mitogenome Clade numbers</th>
<th>Clade Constituents</th>
<th>Present in Gene-Family Topologies?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>tRNA</td>
<td>rRNA</td>
</tr>
<tr>
<td>Clade 1 (H. huso(Subclade 1.1, Subclade 1.2))</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Subclade 1.1 (A. stellatus, P. kaufmanni)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subclade 1.2 (A. ruthenus, (A. fulvescens, (A. brevirostrum, (A. baerii, (A. naccarii, A. gueldenstaedtii))))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clade 2 (H. dauricus, (A. medirostris, A. mikadoi))</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Clade 3 ((A. dabryanus, A. sinensis), (A. transmontanus, A. schrenckii))</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Clade 4 (A. sturio, (A. oxyrinchus, A. o. desotoi))</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Clade 5 (S. suttkusi, (S. platorarynchus, S. albus))</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Table 5. Rzhetsky-Nei Distances between Bayesian Mitogenome Topology and Gene Trees.

Rzhetsky-Nei Distances, signifying the level of similarity to that of the Bayesian full mitogenome topology, are listed for each individual gene phylogeny. A distance of 0 signals identical topologies. The most similar 25% and 50% of genes to the full mitogenome are indicated.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Rzhetsky-Nei distance</th>
<th>Gene length</th>
<th>Top-scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP6</td>
<td>2</td>
<td>674</td>
<td></td>
</tr>
<tr>
<td>ND5</td>
<td>6</td>
<td>1383</td>
<td></td>
</tr>
<tr>
<td>16s</td>
<td>9</td>
<td>1716</td>
<td></td>
</tr>
<tr>
<td>CO3</td>
<td>9</td>
<td>785</td>
<td></td>
</tr>
<tr>
<td>ND1</td>
<td>9</td>
<td>750</td>
<td></td>
</tr>
<tr>
<td>ND4L</td>
<td>9</td>
<td>297</td>
<td></td>
</tr>
<tr>
<td>ND2</td>
<td>10</td>
<td>756</td>
<td></td>
</tr>
<tr>
<td>ND4</td>
<td>10</td>
<td>1374</td>
<td></td>
</tr>
<tr>
<td>ND3</td>
<td>12</td>
<td>348</td>
<td></td>
</tr>
<tr>
<td>ATP8</td>
<td>13</td>
<td>169</td>
<td></td>
</tr>
<tr>
<td>CO2</td>
<td>13</td>
<td>691</td>
<td></td>
</tr>
<tr>
<td>CO1</td>
<td>14</td>
<td>1422</td>
<td></td>
</tr>
<tr>
<td>cytb</td>
<td>14</td>
<td>562</td>
<td></td>
</tr>
<tr>
<td>ND6</td>
<td>14</td>
<td>403</td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>15</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td>Cys</td>
<td>16</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>Phe</td>
<td>16</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>Ser1</td>
<td>16</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>His</td>
<td>17</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>12s</td>
<td>18</td>
<td>960</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>18</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>18</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>18</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>18</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>18</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>18</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>19</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>19</td>
<td>165</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>19</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>Leu2</td>
<td>19</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>19</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>Ser2</td>
<td>19</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>Trp</td>
<td>19</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>20</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>20</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>Leu1</td>
<td>20</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>21</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>D-loop</td>
<td>28</td>
<td>348</td>
<td></td>
</tr>
</tbody>
</table>
Table 6. Birstein et al. 2002 Parsimony Clade Constituents. Names assigned to polytomies and clades recovered in Birstein et al.'s (2002) parsimony phylogeny are listed, along with the constituents of each polytomy or clade.

<table>
<thead>
<tr>
<th>Polytomy Name (or sister species designation)</th>
<th>Clade Name (or sister species position)</th>
<th>Polytomy/Clade Constituents</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>(A. brevirostrum, (A. baerii, (A. gueldenstaedtii, (A. persicus, A. naccarii))))</td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>(A. stellatus, (P. kaufmanni, P. hermannii))</td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>(H. dauricus, (A. ruthenus, H. huso))</td>
<td></td>
</tr>
<tr>
<td>A4</td>
<td>A. fulvescens</td>
<td></td>
</tr>
<tr>
<td>A5</td>
<td>A. nudiventris</td>
<td></td>
</tr>
<tr>
<td>sister taxon Polytomy A</td>
<td>A. mikadoi</td>
<td></td>
</tr>
<tr>
<td>B1</td>
<td>(A. schrenckii, (A. transmontanus, A. mediayris))</td>
<td></td>
</tr>
<tr>
<td>B2</td>
<td>A. sinensis</td>
<td></td>
</tr>
<tr>
<td>Polytomy B</td>
<td>see above</td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>(A. sturio, A. oxyrinchus)</td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>(S. platerynchus, S. albus, and S. suttkusi)</td>
<td></td>
</tr>
</tbody>
</table>
Table 7. Birstein et al. 2002 Maximum Likelihood Clade Constituents. Names assigned to clades recovered in Birstein et al.’s (2002) maximum likelihood phylogeny are listed, along with the constituents of each polytomy or clade. Species not belonging to any of the named clades (ML1-6Z) are listed according to their relationship with named clades.

<table>
<thead>
<tr>
<th>Clade Name (or sister species position)</th>
<th>Clade Constituents</th>
</tr>
</thead>
<tbody>
<tr>
<td>ML1</td>
<td>(A. fulvescens, (A. brevirostrum, (A. baeri, (A. gueldenstaedtii, (A. naccarii, A. persicus))))</td>
</tr>
<tr>
<td>ML2</td>
<td>(A. stellatus, (P. kaufmanni, P. hermanni))</td>
</tr>
<tr>
<td>sister to Clades ML1 and 2</td>
<td>A. nudiventris</td>
</tr>
<tr>
<td>ML3</td>
<td>(H. dauricus, (H. huso, A. ruthenus))</td>
</tr>
<tr>
<td>ML4</td>
<td>(S. platorynchus, S. albus, and S. suttkusi)</td>
</tr>
<tr>
<td>ML5</td>
<td>(A. medirostris, (A. sinensis, (A. transmontanus, A. schrenckii)))</td>
</tr>
<tr>
<td>ML6</td>
<td>(A. oxyrinchus, A. sturio)</td>
</tr>
</tbody>
</table>

Table 8. Birstein et al. 2002 Combined Evidence Clade Constituents. Names assigned to clades recovered in Birstein et al.’s (2002) combined evidence (molecular and morphological) phylogeny are listed, along with the constituents of each polytomy or clade. Species not belonging to any of the named clades (CE1-5) are listed according to their relationship with named clades.

<table>
<thead>
<tr>
<th>Clade Name (or sister species position)</th>
<th>Clade Constituents</th>
</tr>
</thead>
<tbody>
<tr>
<td>CE1</td>
<td>(A. baeri, A. brevirostrum)</td>
</tr>
<tr>
<td>CE2</td>
<td>(A. stellatus, (P. hermanni, P. kaufmanni))</td>
</tr>
<tr>
<td>CE3</td>
<td>(H. dauricus, (A. ruthenus, H. huso))</td>
</tr>
<tr>
<td>sister to Clades CE1, 2, and 3</td>
<td>Acipenser fulvescens</td>
</tr>
<tr>
<td>CE4</td>
<td>(A. sinensis, A. transmontanus)</td>
</tr>
<tr>
<td>sister to Clades CE1-CE4</td>
<td>Acipenser oxyrinchus</td>
</tr>
<tr>
<td>CE5</td>
<td>(S. suttkusi, (S. albus, S. platorynchus))</td>
</tr>
</tbody>
</table>
Figures

Figure 1: Parsimony tree based on 12s, 16s, cytb, D-loop, and NADH 5; modified from Birstein et al. (2002). Clades and clade numbers are in boxes with dashed lines; polytomy constituents and polytomy names are to the right of the phylogeny. Clades and polytomies are listed in Table 6.
Figure 2: Maximum Likelihood tree based on 12s, 16s, cytb, D-loop, and NADH 5; modified from Birstein *et al.* 2002. Clades and clade numbers are in boxes with dashed lines; clades are listed in Table 7.
Figure 3: Combined Evidence tree based on 12s, 16s, cytb, D-loop, NADH 5, and morphological data; modified from Birstein et al. (2002). Clades and clade numbers are in boxes with dashed lines; clades are listed in Table 8.
Figure 4: Bayesian full mitogenome phylogeny. Posterior probability values are to the left of and below each node. Branch lengths indicate estimated amount of evolutionary (or genetic) change. Long branches (pertaining to outgroups) have been broken to allow for better visualization of ingroup relationships. Clades and clade numbers are in boxes with dashed lines; Clade 1, which is composed of units 1.1, 1.2, and H. huso, is labeled to the right of the phylogeny. Clade constituents are listed in Table 4.
**Figure 5: Maximum Likelihood Full mitogenome phylogeny.** Bootstrap values are below and to the left of each node. Branch lengths indicate estimated amount of evolutionary (or genetic) change. Long branches (pertaining to outgroups) have been broken to allow for better visualization of ingroup relationships. Clades and clade numbers are in boxes with dashed lines; Clade 1, which is composed of units 1.1, 1.2, and *H. huso*, is labeled to the right of the phylogeny. Clade constituents are listed in Table 4.
Figure 6: Parsimony full mitogenome phylogeny. Bootstrap values are below and to the left of each node. Branch lengths indicate estimated amount of evolutionary (or genetic) change. Clades and clade numbers are in boxes with dashed lines; clade constituents are listed in Table 4.
Figure 7: Bayesian phylogenetic hypothesis based on concatenated tRNA sequences. Posterior probability values are to the left of and below each node. Branch lengths indicate estimated amount of evolutionary (or genetic) change. Long branches (pertaining to outgroups) have been broken to allow for better visualization of ingroup relationships. Clades that match those recovered by full mitogenome topologies are shown in boxes with dashed lines and labeled with clade numbers listed in Table 4.
Figure 8: Bayesian phylogenetic hypothesis based on concatenated rRNA sequences. Posterior probability values are to the left of and below each node. Branch lengths indicate estimated amount of evolutionary (or genetic) change. Long branches (pertaining to outgroups) have been broken to allow for better visualization of ingroup relationships. Clades that match those recovered by full mitogenome topologies are shown in boxes with dashed lines and labeled with clade numbers listed in Table 4.
Figure 9: Bayesian phylogenetic hypothesis based on concatenated protein-coding sequences.

Posterior probability values are to the left of and below each node. Branch lengths indicate estimated amount of evolutionary (or genetic) change. Long branches (pertaining to outgroups) have been broken to allow for better visualization of ingroup relationships. Clades (except for Clade 1) that match those recovered by full mitogenome topologies are shown in boxes with dashed lines and labeled with clade numbers listed in Table 4. Clade 1 is made up of groups 1.1, 1.2, and *H. huso*, and is identified to the right of the phylogeny.
CHAPTER 2.

The Biogeography of Extant Sturgeons
Abstract

Previous studies have investigated the biogeography of sturgeons (Acipenseridae), but the presentation of a new, full mitogenome phylogeny for the family presents the opportunity for a revised study. Here, I consider sturgeon phylogeny and their current distributions, along with phylogenetic and distribution information from lampreys and pinnipeds, to explore relationships among geographic areas inhabited by sturgeons. Three-item analysis, a comparative biogeography method, is employed. Eleven areas of endemism are identified, and relationships amongst these areas are proposed based on phylogenetic and species range data. These areas of endemism are recovered in three broad clades illustrating area relationship hypotheses: a Pacific clade, and Atlantic clade, and the Bering Sea. The Atlantic and Pacific clades are recovered as sister, reflecting previous biogeographic hypotheses. Relationships within each clade also follow previously described hypotheses and are explored in the context of geologic history. The Bering Sea area falls outside of this Atlantic-Pacific relationship.
1.0 Introduction

1.1 Sturgeons and Biogeography

Sturgeons (Acipenseridae) are a relatively old (at least 85 million years) group of ray-finned fishes, with two known fossil genera: *†Protoscaphirhynchus* and *†Priscosturion* (Grande and Hilton 2006). The 25 broadly recognized extant sturgeon species exhibit a Holarctic distribution, inhabiting wide and disjunct ranges exclusively in the temperate zones of Europe, Asia, and North America (Bemis and Kynard 1997, Choudhury and Dick 1998, Hilton and Grande 2006). Most extant sturgeons are anadromous, moving between fresh and saltwater to feed (primarily downstream movement) and to reproduce (upstream movement) (Bemis and Kynard 1997, Peng 2007). Sturgeon taxa are distributed among four commonly recognized genera (*Acipenser*, *Huso*, *Scaphirhynchus*, and *Pseudoscaphirhynchus*) in the monophyletic family Acipenseridae (Findeis 1997, Bemis et al., 1997). Sturgeons, paddlefishes (Polyodontidae), and their fossil relatives form the order Acipenseriformes, which occupies a basal position among the other ray-finned fishes (Actinopterygii), the clade containing over half of all vertebrate diversity (Hilton et al. 2011). Across their ranges, sturgeons face threats such as habitat loss, overfishing, and other fishery interactions (Holzkamm and Waisberg 2004). These threats have led to the decline of sturgeon populations worldwide and the implementation of conservation measures that afford them protection (Holzkamm and Waisberg 2004). The disjunct distributions of sturgeon, their long evolutionary history, and their anadromous behavior have made this group of particular interest for biogeographic studies (Bemis and Kynard 1997).

Biogeography is the study of the relationship between taxa that inhabit the Earth and Earth's geography, with the goal of understanding the shared relationships between taxa and their distributions (Nelson and Platnick 1981, Posadas et al. 2006, Parenti and Ebach 2009). Several methods exist for biogeographic studies, and many of them consider phylogeny and use...
cladistic principles (Parenti and Ebach 2009). Comparative biogeography, a field that combines biogeography with systematics (Nelson and Platnick 1981), is defined by Parenti and Ebach (2009) as the use of hierarchical phylogenetic relationships of taxa to identify area relationships among biogeographic regions. Area relationships may be considered alongside geologic history and compared to historical biogeographic patterns (sets of area relationships shared by more than one taxon) previously revealed. Only after patterns of area relationships are discovered, should the process by which these relationships evolved, be interpreted (Parenti and Ebach 2009).

The first explicitly cladistic study of relationships within Acipenseridae (Findeis 1997) recovered *Huso* as the sister taxon of a clade comprising the other three genera, and recovered *Scaphirhynchus* and *Pseudoscaphirhynchus* as sister taxa forming Scaphirhynchini, which is recognized as Scaphirhynchinae by some authors (Myden and Kuhajda 1996). Many studies since Findeis (1997) have suggested broad changes to this topology. Widely accepted changes indicate that the tribe “Scaphirhynchini” (*sensu* Findeis) is paraphyletic, and suggest nonmonophyly of the genus *Huso* (Dillman *et al.* 2007, Krieger *et al.* 2008, Hilton *et al.* 2011). A phylogenetic analysis of the full mitogenome (see Chapter 1) supports these changes and provides structure within the nonmonophyletic “Acipenser.”

Many biogeographic studies have examined geographical distributions of sturgeons using the phylogenetic context available at the time. The many recent hypotheses of sturgeon phylogeny encourage a reconsideration of the historical biogeography of sturgeons. For example, Bemis and Kynard (1997) considered Scaphirhynchini to be monophyletic. They also considered *Huso* to be monophyletic and basal to other sturgeons, and noted that were *Huso* nested within “Acipenser,” biogeographical hypotheses would likewise change. Choudhury and Dick (1998) also considered Scaphirhynchini to be monophyletic, and based their biogeographic
analyses on a modified version of a morphological phylogeny proposed by Artyhukin (1995).
Peng et al. (2007) considered *Huso* to be nonmonophyletic and nested within *Acipenser*. Many studies (e.g. Vasileva et al. 2009) have confirmed the paraphyly of *Huso*, and, subsequently, its classification within *Acipenser*. Following the suggestion of Bemis and Kynard (1997), we propose a new analysis of sturgeon biogeography, based on recent, strongly supported changes to phylogeny.

Traditional biogeographic studies of sturgeons have focused on identifying a center of origin for sturgeons from which dispersal routes are inferred. The Ponto-Caspian Region, which includes the Mediterranean, Black, Caspian, and Aral seas and their tributaries, has the highest diversity of extant sturgeons (Bemis and Kynard 1997, Peng et al. 2007). This region has frequently been inferred to be the sturgeon center of origin (Bemis and Kynard 1997, Choudhury and Dick 1998, Peng et al. 2007), although some studies have ruled it out (Choudhury and Dick 1998). Conflicting hypotheses regarding the timing of divergences among sturgeon taxa have further led to the suggestion of varying and conflicting mechanisms and paths of dispersal from the Ponto-Caspian (Choudhury and Dick 1998, Peng et al. 2007, Bemis and Kynard 1997). Nevertheless, most studies agree that geologic events of the evolution of the Tethys Sea are critical to sturgeon biogeography (Bemis and Kynard 1997, Choudhury and Dick 1998, Peng et al. 2007).

The Tethys Sea was one of two large water bodies during the Triassic, and was situated in between Gondwana and Laurasia (Zonenshain and Pichon 1986); the Panthalassic Ocean covered the rest of the non-terrestrial world (Algeo et al. 2011). As geologic events shaped the earth, the Tethys Sea began to narrow during the Late Triassic, and eventually became separated from the Paratethys Sea (Bartock 1993, Rogl 1999). The Mediterranean Sea is the last remnant of the Tethys, and the other Ponto-Caspian Seas- the Black, Caspian, and Aral seas-
formed from the Paratethys Sea (Rogl 1999), giving rise to the frequently presumed sturgeon center of origin. The proposal of a center of origin of a taxon has been criticized for many reasons, chief among them that the hypotheses are untestable and applied frequently to only one taxon at a time, and hence are not comparative (Croizat et al. 1974).

Past biogeographic studies of sturgeons have also focused on the disjunct distributions of closely related species. For example *A. oxyrinchus*, which is sympatric with *A. brevirostrum* inhabits the Atlantic coast and rivers of North America from Canada to Florida, while its sister species *A. sturio*, is distributed in the North Sea, Mediterranean, and Black Sea (Ludwig et al. 2002). Although *A. oxyrinchus* and *A. brevirostrum* are sympatric, phylogenetic studies have not found them to be closely related to one another (e.g. Birstein and DeSalle 1998; Laumann et al., unpublished, Chapter 1). Another pair of sympatric sturgeon species, *A. transmontanus* and *A. medirostris*, are distributed along the North American Pacific Coast. Some recent studies have recovered a sister relationship between these two species and found them to be closely related to *A. schrenckii* (e.g., Birstein et al. 2002), whereas others have recovered them within a monophyletic group that also includes *A. dabryanus* (e.g., Dillman et al. 2007). Still others have not found them to be closely related (e.g., Krieger et al. 2008). In Chapter 1, I recovered these two species to be within a broad clade also including *A. dabryanus*, *A. sinensis*, *A. schrenckii*, *A. mikadoi*, and *H. dauricus*. Within this clade, *A. medirostris* is closely related to *A. mikadoi* and *H. dauricus*, whereas *A. transmontanus* is sister to *A. schrenckii*, a clade which is sister to *A. sinensis* and *A. dabryanus*. Therefore, based on this and other topologies, there are trans-Pacific relationships reflected within Acipenseridae.

Sturgeons have also been of particular interest for biogeographic analysis due to their anadromous life history (Bemis and Kynard 1997). Diadromous taxa have frequently been the subject of biogeographic studies seeking to identify explanations for their distributions, to
identify whether diadromy is a derived or ancient trait, and to identify the ancestral habitat of the taxa involved as freshwater or marine (Parenti 2008). Anadromy in sturgeons has been considered a derived trait with the assumption that the ancestral habitat was freshwater (Sulak and Randall 2002). As stated by Parenti (2008), identifying the ancestral habitat as either freshwater or marine relies on assuming that the habitat has (or has not) changed; this assumption is not necessarily supported (Parenti and Ebach 2009). Any number of events, from stranding (becoming stuck in a newly isolated habitat) to invasion, may be used to explain distributions of anadromous taxa, as they could be used to explain biogeographic distributions of any taxonomic group (Parenti 2008).

In this chapter, I employ comparative biogeographic methods to identify patterns of relationships among regions inhabited by sturgeons. I use a new phylogenetic hypothesis based on full mitogenomic data (Figure 1; Chapter 1) and current sturgeon distributions. Mechanisms leading to current distributions of sturgeons are not discussed. The aim of this study was to identify area relationships (relationships between geographic regions) using comparative biogeographic methods, consider them alongside geologic history, and discuss them in relation to area relationships suggested by other taxa distributed throughout the northern hemisphere. Broad biogeographic patterns that are considered alongside the area relationships hypothesized in this study include those involving the five areas of global distribution identified by Croizat (1958): Atlantic, Indian, Pacific, boreal, and austral (Figure 2). Historically the boreal and austral regions have been associated with the distributions of holarctic taxa like sturgeons (Nelson and Ladiges 2001, Parenti and Ebach 2009). These regions have been repeatedly linked with the Pacific by biogeographic studies, while the Atlantic has been linked with the Indian regions. Nelson and Ladiges (2001) distilled these frequently proposed area relationships into a “cladistic summary of global patterns” (Figure 3).
1.2 History of Areas Inhabited by Sturgeons

One important aspect of comparative biogeographic analysis is the consideration of the geologic history of areas under consideration, in this case the areas inhabited by extant sturgeon taxa. The geologic history of these areas is discussed below.

1.2.1 Geologic History: Formation of the Atlantic and Pacific Oceans and Ponto-Caspian Region

During the Triassic, Pangea enclosed the Tethys Sea on three sides (Zonenshain and Pichon 1986) and an archipelgo comprising portions of China defined the eastern border of the Tethys (Yang et al. 1982). The Panthalassic Ocean covered the rest of the Earth (Algeo et al. 2011). The oldest of the world’s oceans, the Pacific arose from the Panthalassic Ocean at least 167 million years ago (Neall and Trewick 2008). In the Late Triassic and Early Jurassic, Pangea began breaking into two super-continents, Laurasia and Gondwana (Golonka 2007). Laurasia, to the north, comprised present day North America and Eurasia (excluding the Indian subcontinent). Gondwana, the southern supercontinent, was composed of South America, Africa, Australia, India, and the Arabian Peninsula. The shift of Laurasia away from Gondwana opened the North Atlantic, then an arm of the Tethys Sea (McHone 2000, Thiede 1979). The rifting of Pangea also resulted in the formation of the Gulf of Mexico and the movement of Eurasia southwards, narrowing the Tethys Sea (Bartock 1993). Uplift of Great Britain and Ireland began, as did a cycle of marine incursions, flooding what is now the North Sea (Ziegler 1975, Torsvik et al. 2002).

Rising sea levels inundated much of Laurasia so that large portions of North America and Europe were submerged by 100 million years ago (Cretaceous) (Schlee 1999). During the Cretaceous, the South Atlantic opened and the Atlantic Ocean as a whole continued to widen (Thiede 1979, Fairhead 1988). Connections between the narrowing Tethys and the Atlantic closed, with the Tethys separating Laurasia from Gondwana (Rogl 1999). Modern day
continents that formed Laurasia began to separate, as did Gondwana (Thiede 1979). During this period, much of North America and Eurasia were submerged, with Eurasia separated from Gondwana by the Tethys Sea (Hay et al. 1993). North America was inundated by the Western Interior Seaway (Hay et al. 1993).

In the Paleocene, the continents composing Laurasia continued to move away from one another, but North America was still connected to Asia by the Bering Land Bridge and Europe by Greenland; this separated the future Pacific from the Bering Sea (Hopkins 1973). Likewise, the continents that composed Gondwana were shifting, with India moving toward Asia (Tapponier et al. 1986). Africa shifted toward Europe, continuing the narrowing of the Tethys. In the Eocene, North America became more separated from Eurasia as it broke apart from Greenland and Europe; this movement opened the North Sea, allowing it to deepen (O’Leary et al. 2004; Kender et al. 2012). Also during the Eocene, India collided with Asia forming the Himalayas (Tapponier et al. 1986).

During the Oligocene, the Atlantic continued to widen (Fairhead 1988). The continued growth of the Himalayas separated a section of the Tethys Sea, forming the Paratethys Sea (Rogl et al. 1999). The Paratethys Sea was fed with water from the North Sea and was connected to the Mediterranean (Rogl 1999). At this time, the African and Eurasian plates converged, isolating what remained of the Tethys Sea from other water bodies except for the Indian Ocean. The isolation of the Tethys from other water bodies was completed in the Miocene, sealing it into today’s Mediterranean Basin (Rogl et al. 1999). Today’s Mediterranean Sea is the last remnant of the Tethys Sea (Rogl et al. 1999), and the Paratethys Sea formed the Black, Caspian, and Aral seas; these together with the Mediterranean Sea form the Ponto-Caspian Region (Hakanson et al. 2002). The isolation of the North Sea (from the Tethys) was followed by
Miocene shallowing of the North Sea and the formation of deltas in the area (Torsvik et al. 2002).

1.2.2 Geologic History: The Arctic Ocean and Bering Sea

The geologic history of the Arctic is poorly known and contentious (Clark 1974, 1975, Shepard et al. 2013, Moran et al. 2006). In the early Miocene, around 18 million years ago, a connection between the Arctic Ocean and the Atlantic was established in the form of the Fram Strait (Jacobson et al. 2007). Initially shallow, the strait allowed exchange of water between the two water bodies, transitioning the Arctic through an estuarine phase into the ocean of today (Thompson et al. 2010; Engen et al. 2008). Fluctuations in sea level led to the opening and closing of the Bering Sea, alternately linking and isolating the Arctic from the Pacific (Elias et al. 1996). About 11,000 years ago, sea level rise caused the final inundation of the Bering land bridge, connecting the Bering Sea to the Pacific via the Bering Strait (Elias et al. 1996). Today, the Arctic and Atlantic oceans are linked by the Barents, Greenland, and Norwegian seas and the Denmark Strait. The Arctic Ocean remains connected to the Pacific via the Bering Sea.

1.2.3 Recent Geologic History: Baltic and North Seas

Pleistocene glaciations and interglacial periods caused several periods of drying and subsequent filling of the North Sea. During one of the interglacial periods, the Eemian (around 130,000 and 115,000 years ago), the North Sea expanded (Torvick et al. 2002). During this time, the first body of water known from the Baltic Basin, the Eemian Sea, formed (Torvick et al. 2002, Miettinen 2003). The Eemian Sea was larger than the Baltic Sea is today and had connections to the Atlantic Ocean and the North Sea (Miettinen 2003). The transformation from the Eemian Sea to today's Baltic Sea occurred through several stages, involving freezing, melting, different salinity regimes, and varying connections to the Atlantic Ocean and North Sea. The Weischlian Glaciation, the last glaciation of Earth, was characterized by the coverage of much of Eurasia by
ice sheets (Siegert et al. 2001). This resulted in the drying of large portions of the North Sea to form plains (Torsvik et al. 2002, Konradi et al. 2005).

About 13,000-14,000 years ago, receding glaciers isolated sections of the Eemian Sea, which merged to form the Baltic Ice Lake (Morner 1995, Vassiljev and Saarse 2013). During the time of the Baltic Ice Lake, cooling and warming periods resulted in expansion and subsequent regression of ice sheets as well as uplift of land, changing the shape of the Baltic Ice Lake (Morner 1995, Vassiljev and Saarse 2013). The exact timing of the formation of the Baltic Ice Lake is controversial, but it is hypothesized to have been in existence between 13,800 and 10,300 years ago (Morner 1995, Vassiljev and Saarse 2013).

Toward the end of the Weisclian Glaciation, warming temperatures caused the final retreat of ice sheets and increased sea level. Water again began filling the North Sea, and once again a connection between the North Sea and the Baltic Ice Lake formed (Lepparanta and Myreberg 2009). North Sea waters inundated the lake, increasing salinity and forming the Yoldia Sea, which persisted until about 9,500 years ago (Andren et al. 2002, Lepparanta and Myreberg 2009).

Continuing uplift of land separated the Yoldia and North seas, resulting in the formation of the freshwater Ancylus Lake from the Yoldia Sea (Morner 1995). About 8,150 years ago, a Tsunami hit the North Sea, inundating the remaining dry areas (Weninger et al. 2008). Portions of Ancylus Lake were still ice-covered and uplift continued, gradually forming a new river (Dana River) between Ancylus Lake and the North Sea (Lepparanta and Myreberg 2009). The river widened into a strait, introducing saltwater into Ancylus Lake once again. This led to the formation of the Littorina Sea about 7,500 years ago (Sandgren et al. 2004). The exchange of water between the North and Littorina seas and sea level rise led to higher levels of salinity than present in the Baltic Sea (Sandgren et al. 2004). Influx of seawater into the Littorina Sea
continued until about 4,000 years ago, when sea level rise ended and salinity began decreasing in what is now known as the Baltic Sea (Sandgren et al. 2004, Berglund et al. 2005). The Baltic Sea is brackish and continues to shrink due to uplift (Walday and Krogland 2011). It is connected to and exchanges water with the North Sea (Walday and Krogland 2011).

1.2.4 Recent Geologic History: East China, Yellow, and Japan Seas

The Sea of Japan first opened 32 million years ago, a process that continued until about 15 million years ago at which point it was an isolated body of water (Tada et al. 2013). About 2.6 million years ago, the oldest portion of the East China Sea began opening (Gallagher et al. 2015). Around this time, the Sea of Japan became connected to the Pacific Ocean by the opening of the Tsugaru and Tsushima straits (Uda 2016, Gallagher et al. 2015). Linkages between the Sea of Japan and the Pacific continued intermittently with sea level rise and fall (Liu and Milliman 2004). Today, the Sea of Japan maintains a connection to the Pacific Ocean via the Tsugaru Strait.

The East China Sea and Yellow Sea were still primarily terrestrial or subtidal areas 20,000 years ago (Lee 2012). Multiple phases of sea level rise inundated these areas and led to both the East China and Yellow Seas as we know them today. By 15,000 years ago, rising sea levels filled the East China Sea and reached the southern portion of the Yellow Sea (Liu and Milliman 2004, Yi et al. 2003). Subsequent periods of rapid sea level rise led to the gradual filling of the Yellow Sea, from the south to the north, over the next 5,000 years (Liu and Milliman 2004, Yi et al. 2003). The Tsushima Strait connects the East China Sea to the Sea of Japan today.

1.2.5 Recent Geologic History: Interior North America, The Mississippi and Missouri Rivers and their tributaries

The Mississippi River and its tributaries existed before the Pleistocene, but their courses were so different from those seen today that the history of these rivers is commonly considered
to begin with the end of the last glacial stage (Leverett 1921). The Missouri River also formed before the last ice age, but its course was significantly diverted by glaciation (Trimble 1980). Much of North America was covered with ice sheets and glaciers during the Pleistocene (Leverett 1921, Trimble 1980). 13,000 years ago, glacial melt flooded the area now occupied by Lake Champlain and the Gulf of Saint Lawrence, forming the Champlain Sea (Russell and Cummings 2009, Mather 1917). As glacial retreat continued in the Holocene, continental rebound caused this arm of the Atlantic Ocean to shrink (Mather 1917). Further melting formed the Great Lakes Erie and Michigan, which drained into the Mississippi, about 10,000 years ago (Larson and Schaetzl 2001, UWSG 2013). By one thousand years later, Lake Superior, which also drained into the Mississippi, had formed. Sea level rise about 7,500 years ago flooded the Great Lakes, connecting them to form one large lake, Lake Nipissing, which drained into the Mississippi and the Atlantic (Larson and Schaetzl 2001, UWSG 2013). By 3,000 years ago, the level of the lake had fallen, resulting in the current configuration of the Great Lakes (Farrand 1988, UWSG 2013). The final retreat of glaciation also established the Mississippi and Missouri Rivers in the (approximate) courses they occupy today, draining into the Gulf of Mexico (Wickert et al. 2013, Trimble 1980, Leverett 1921).

2.0 Methods

In keeping with the cladistic nature of my dissertation and following the detailed example of Hoagstrom et al. (2014), I selected a comparative biogeographical method of analyses based on hierarchical relationships (Parenti and Ebach 2009). This method is rooted in the use of three-item analysis to infer relationships among geographic areas (Parenti and Ebach 2009). In three-item analysis, phylogenies of taxa that have overlapping distributions are converted to areagrams. An areagram is a cladogram on which species names have been
replaced by the areas the taxon inhabits. It is the overlap of multiple species (with hypothesized evolutionary history) that provides the shared history informing relationships of the area-termed areas of endemism. Analyses are then conducted to extract three-area relationships (i.e., hierarchical relationships reflecting the relationships among sets of three areas). Three-area relationships are combined to form optimal area cladograms, which are further broken down into an intersection tree, which displays the three area relationships recovered by all optimal area cladograms. These relationships may be explored by examining the geologic history of areas of endemism; geologic patterns considered alongside biological ones allow for reciprocal illumination—each method sheds light on the other. (Parenti and Ebach 2009).

2.1 Phylogenies for Comparison

The full mitogenome sturgeon (Figure 4) phylogeny was used for the exploration of area relationships based on sturgeon biogeography, as were phylogenies for lampreys (Figure 5; Gill et al. 2013), and pinnipeds (Figure 6; Arnason et al. 2006). These taxa have overlapping ranges with sturgeons, and phylogenetic hypotheses were available for them.

The lamprey phylogeny (Gill et al. 2003) was based on 32 morphological and karyological characters. The original phylogeny included 18 lamprey species. Three of these species (Mordacia mordax, M. lapicida, and Geotria australis) do not have ranges that overlap with those of sturgeons or within my proposed areas of endemism and were pruned from the phylogeny, resulting in a 15 terminal taxa topology (Figure 5).

The pinniped phylogeny (Arnason et al. 2006) was based on protein-coding sequences of the mitogenome, and included 28 species. Sixteen of these species overlapped with sturgeons in the proposed areas of endemism, and the remaining species (Otaria byronia, Arctocephalus pusillus, Phocarctos hookeri, Neophoca cinerea, A. forsteri, A. australis, Monachus schauinslandi,
Ommatophoca rossi, Lobodon carcinophagus, Leptonychotes weddelli, Hydrurga leptonyx, and Pusa sibirica) were pruned from the topology (Figure 6).

2.2 Species Ranges and Phylogenies

I conducted literature and fishery landing searches to map the ranges of extant sturgeon species in my full mitogenome Bayesian phylogeny of extant sturgeons (Figure 7). This sturgeon phylogeny (Figure 1; Chapter 1) was used in analyses described below. Maps were superimposed upon one another and were used, along with Bemis and Kynard’s (1997) sturgeon biogeography study, to identify potential areas of endemism (Figure 4). Once these areas were identified, the ranges of lamprey and pinniped species were identified from the literature (Gill et al. 2003, Arnason et al. 2006), IUCN RedList species range maps and descriptions (IUCN 2014) and the global fish species database FishBase (Froese and Pauly 2016). Lamprey and pinniped taxa that do not inhabit the areas of endemism identified for sturgeons were pruned (using the command drop.tip in the ape package, version 3.4, Paradis et al. 2004, for R, R Core Team 2014) from the available phylogenies.

2.3 Areas of Endemism and Three-Item Analysis

Areas of endemism and pruned phylogenies were converted into areagrams and coded for LISBETH version 1.3 (Bagils et al. 2012). LISBETH conducts three-item analysis using fractional weighting and paralogy free sub-tree analysis (Nelson and Ladiges 1996). Three-area statements were extracted from LISBETH and imported into PAUP*4b10 (Swofford 2003). PAUP was used to identify optimal taxon-area cladograms (by combining TACs). Optimal TACs were then imported into LISBETH, where an intersection tree including all three-area relationships was constructed. Results are reported with a “completeness index”, which is the percentage of three-area statements in both the intersection trees and the paralogy free sub-trees. A high completeness index is indicative of agreement on area relationships among all taxa considered.
Results are discussed in comparison to geologic history of the areas of endemism considered.

3.0 Results

3.1 Areas of Endemism

Proposed areas of endemism followed “provinces” used by Bemis and Kynard (1997) to explore sturgeon biogeography. Modifications were made to Bemis and Kynard’s (1997) provinces, and two additional areas of endemism were proposed. Proposed areas of endemism are given different names here to reflect that the regions have been modified from those of Bemis and Kynard (1997). Eleven putative areas of endemism were proposed (Figure 4) and are defined as follows.

1) Japan (JPN) Region
   a. Includes the Sea of Japan, Amur Basin, Sea of Okhotsk, and Tatar Strait,
   b. Corresponds to Bemis and Kynard’s (1997) ASJ province, but is expanded to include the entire Sea of Okhotsk,
   c. Inhabited by A. mikadoi, A. schrenckii, and H. dauricus.

2) China (CHN) Region
   a. Includes the Yangtze Basin and the southern portion of the East China Sea,
   b. Corresponds to Bemis and Kynard’s (1997) CH province, but does not include the part of the coast along the Yellow Sea,
   c. Inhabited by A. dabryanus and A. sinensis.

3) Yellow Sea (YLW) Region
   a. Includes the Yellow Sea itself, and the coast of China along the Yellow Sea,
b. The Yellow Sea was not included in Bemis and Kynard's (1997) study, but part of the coast of China along the Yellow Sea was included in their CH province,
c. Inhabited by *A. sinensis*.

4) Pacific Coast of North America (PNA) Region
   a. Includes the Pacific Coast and rivers draining into the Pacific,
   b. Corresponds to Bemis and Kynard's (1997) NEA province, but is slightly expanded (here) to the south,
   c. Inhabited by *A. mediorestris* and *A. transmontanus*.

5) Bering Sea (BRG) Region
   a. Includes the Bering Strait,
   b. Was not included in Bemis and Kynard's (1997) study,
   c. Inhabited by *A. mikadoi*.

6) Baltic and North Seas (BNS) Region
   a. Includes the Gulf of Bothnia,
   b. Corresponds to Bemis and Kynard's (1997) NEA province, but is expanded to include the Gulf of Bothnia,
   c. Inhabited by *A. sturio*.

7) Ponto-Caspian (PTC) Region
   a. Includes the Black, Azov, Caspian, Mediterranean and Adriatic seas,
   b. Corresponds to Bemis and Kynard's (1997) PT province,
   c. Inhabited by *A. gueldenstaedtii*, *A. naccarii*, *A. ruthenus*, *A. stellatus*, *A. sturio*, *H. huso*, and *P. kaufmanni*.

8) Arctic (AES) Region
a. Includes the Arctic Ocean and East Siberian Sea,
b. Corresponds to Bemis and Kynard’s (1997) SAO province, expanded to include Arctic North America,
c. Inhabited by *A. baerii*, *A. mikadoi*, *A. ruthenus*, and *A. stellatus*.

9) Atlantic Coast of North America (ANA) Region
a. Includes the Atlantic and rivers draining into it,
b. Corresponds to Bemis and Kynard’s (1997) NWA province, expanded slightly northward
c. Inhabited by *A. brevisotrum* and *A. o. oxyrinchus*.

10) Gulf of Mexico Region (MGM)
a. Includes the Mississippi and Mobile basins,
b. Corresponds to Bemis and Kynard’s (1997) MGM province,
c. Inhabited by *A. fulvescens*, *A. o. desotoi*, and *Scaphirhynchus*.

11) North American Lakes (NAL)
a. Includes the Great Lakes, Hudson Bay, and Lake Champlain,
b. Corresponds to Bemis and Kynard’s (1997) GL province,
c. Inhabited by *A. fulvescens* and *A. oxyrhynchus*.

Each sturgeon (Table 1), lamprey (Table 2), and pinniped (Table 3) species was assigned to all areas of endemism that they inhabit (see above). Areagrams were constructed using the phylogenies of each taxonomic group and the areas of endemism they inhabit; thees areagram is shown in Figure 8, the lamprey areagram in Figure 9, and the Pinniped areagram in Figure 10.
3.2 Intersection Tree

The Sturgeon-Lamprey-Pinniped intersection tree had a completeness index of 92.7%. This intersection tree (Figure 11) suggested two clades of related areas, a Pacific clade and a primarily Atlantic clade. The Pacific clade consisted of China and the Yellow Sea areas of endemism grouping as sister areas, with Pacific North America outside of them. Outside of this relationship was the Japan area of endemism, with the Arctic area just outside of that grouping of four areas. The Atlantic clade recovered the Ponto-Caspian and Baltic/North Sea areas of endemism as sister areas. The North American Lakes area of endemism was sister to the Gulf of Mexico area, with the Atlantic coast of North America area falling just outside. The Atlantic and Pacific clades grouped together, with the Bering Sea area of endemism outside of that sister-area relationship.

4.0 Discussion and Conclusions

4.1 Anadromy

Although I do not attempt to identify the ancestral habitat of sturgeons as marine or freshwater, or to identify the processes by which sturgeons came to inhabit their current ranges, it is interesting to consider the areas of endemism occupied by the strictly freshwater taxa (Figure 1). Exclusively freshwater sturgeons are present in the Japan, China, Ponto-Caspian, Arctic, Gulf of Mexico, and North American Lakes Regions. These regions are also occupied by anadromous species. Except for Scaphirinchus, none of the sturgeon clades in this analysis is exclusively freshwater. This may be due to taxon sampling. For example, only one species of Pseudoscaphirhynchus was included in this study. Pseudoscaphirhynchus is generally considered to be a clade also containing P. kaufmanni, P. hermanni (Birstein et al. 2002, Dillman et al. 2007, Hilton et al. 2011) and the presumed recently extinct P. fedtschenkoi (IUCN 2014). The three Pseudoscaphirhynchus taxa are freshwater species, and as such represent an entire freshwater
clade (IUCN 2014). Parenti (2008) noted that, without making assumptions that the habitat of a taxon has changed from either marine to freshwater or vice versa, the ancestral habitat can be reconstructed as one that includes both fresh and saltwater areas. This analysis does not contradict this statement, and, in fact, supports it.

4.2 Area Relationships

Many of the area relationships hypothesized via three-item-analysis mirror those suggested by geologic history and by past biogeographic studies. Area relationships are divided into two primary area clades: 1) the Atlantic Clade containing three North American Regions (Lakes, Atlantic Coast, and Gulf of Mexico), the Baltic/North Sea Region, and the Ponto-Caspian Region; and 2) the Pacific Clade, composed of Pacific North America, China, Yellow Sea, Japan, and Arctic Regions. The Bering Sea Region was recovered separate from these two clades.

4.2.1 Atlantic Clade

4.2.1a Atlantic Clade: North American and Baltic/North Sea/Ponto-Caspian Regions

Connections between the two Atlantic Region clades are exemplified in the relationships between two clades of sturgeons. In the first clade, *Acipenser sturio*, which ranges throughout the Ponto-Caspian and Baltic/North Sea regions, is sister to the two subspecies of *A. oxyrinchus*, which inhabit the North American Atlantic Regions (including Gulf of Mexico). The second clade contains six Ponto-Caspian species (*H. huso, A. ruthenus, A. naccarii, A. gueldenstaedtii, A. stellatus, and P. kaufmanni*), *A. brevirostrum* from the Atlantic Coast of North America, and *A. fulvescens*, which ranges in the Gulf of Mexico Region and North American Lakes Regions. Of interest, two of the Ponto-Caspian species, *A. stellatus* and *A. ruthenus*, are also found in the Arctic Region, as is *A. baerii* which is also part of this sturgeon clade.

Current geography also supports a relationship between the two Atlantic Region Area Clades. The North Sea connects the Atlantic Ocean to the Baltic Sea today; these waters have
been intermittently connected since at least 115,000 years ago. The Atlantic is also currently connected to the Mediterranean (included in the Ponto-Caspian Region) via the Strait of Gibraltar. Although the other seas in the Ponto-Caspian region do not currently connect to the Atlantic Ocean, the origin of the Atlantic as part of the Tethys Sea in the Jurassic, and continued connection to the Tethys until the Cretaceous, supports these area relationships.

4.2.1b Atlantic Clade: North American Regions

The recovery of three area relationships among the Atlantic Coast of North America, the North American Lakes, and the Gulf of Mexico Regions, in which the latter two areas are most closely related to one another, is supported by three sturgeon taxa and geologic history. *Acipenser fulvescens* is found in both the Gulf of Mexico and North American Lakes Regions. Linkages between these two regions have occurred intermittently, with the Great Lakes draining into the Mississippi at some points over the last 10,000 years. None of the sturgeon taxa considered in this analysis inhabits all three North American regions, but *A. o. oxyrinchus* lives in the North American Lakes region and along the Atlantic Coast. Its subspecies *A. o. desotoi* is endemic to the Gulf of Mexico. These evolutionary relationships and the area relationships they suggest are strengthened by the geologic connections between the Atlantic Ocean and the other two North American Regions. The North American Lakes Region was historically connected to the Atlantic Ocean via the Champlain Sea, and has a current connection to the Atlantic through the Gulf of Saint Lawrence. The Gulf of Mexico is connected to the Atlantic Ocean by the Florida Straits, and the Mississippi River had historic connections to the Atlantic.

4.2.1c Atlantic Clade: Baltic/North Sea and Ponto-Caspian Regions

The Baltic/North Sea and Ponto-Caspian regions are hypothesized to have a sister-area relationship to one another. A single sturgeon species in this analyses, *Acipenser sturio*, inhabits
the Baltic/North Sea Region; it is also found in the Ponto-Caspian Region. The Oligocene connection between the North Sea and Paratethys Sea supports this area relationship.

4.3 Pacific Clade

4.3.1 Pacific Clade: Asia and Pacific North America

A seven-species clade supports the area relationships among several Pacific regions: (((China Region + Yellow Sea Region) Pacific North America Region) Japan Region). The taxa reflecting this clade include two sister species, *A. dabryanus* and *A. sinensis*, that inhabit the East China Sea. *Acipenser sinensis* is also found in the Yellow Sea. Sister to *A. dabryanus* and *A. sinensis* is another pair of sister-species: *A. transmontanus*, which inhabits the Pacific North America region and *A. schrenckii*, which is from the Japan Region. Another pair of Pacific North America-Japan Region taxa, *Acipenser medirostris* and *A. mikadoi*, respectively, form a sister species relationship. Together with the Japan Region species *H. dauricus*, these taxa form a sister clade to the four taxa listed above.

The shared geologic history between the East China and Yellow Seas, which were filled sequentially by an increase in sea level, further supports the hypothesized biogeographic relationship between these two areas. Although it might seem that these areas should be most closely related to the Sea of Japan, as they are all marginal seas of the western Pacific, this analysis and geologic history do not support such a relationship.

In this analysis, the East China and Yellow seas are more closely related to Pacific North America than they are to the Sea of Japan. Although the Sea of Japan was connected to the Pacific long before the East China Sea was flooded, about 2.6 million years ago, these connections have been ephemeral. The Sea of Japan has also maintained intermittent connections with the East China Sea. This long history has provided for periods of exchange of water and taxa between the Sea of Japan and the Pacific, illustrated by the sister species
relationships between *A. transmontanus* and *A. schrenckii*, and *A. medirostris* and *A. mikadoi*. The East China Sea and Yellow Sea, however, have had a constant connection to the Pacific Ocean, and therefore provided more consistent opportunity for taxonomic connections since they formed.

Current and past connections between the East China Sea and the Yellow Sea, along with the fact that the only sturgeon taxon that inhabits the Yellow Sea (*A. sinensis*) also inhabits the China Region, suggest that subsequent analyses should consolidate these into a single region or area of endemism.

### 4.3.2 Pacific Clade: Pacific and Arctic

Four turgeon species, *A. baerii*, *A. mikadoi*, *A. ruthenus*, and *A. stellatus* occur in the Arctic region. The relationships hypothesized in this analysis among the Arctic and other Pacific Clade Regions are not strongly supported by the phylogeny of sturgeons alone. A single taxon occurring in the Arctic Region (*A. mikadoi*) also occurs in other Pacific Clade Regions. This species is sister to a Pacific North American species, and part of a clade with other exclusively Pacific Region taxa. Two other taxa from in the Arctic (*A. ruthenus* and *A. stellatus*) occur in the Ponto-Caspian Region of the Atlantic Area clade. Another species (*A. baerii*) is sister to two Ponto-Caspian Region species (*A. naccarii* and *A. gueldenstaedtii*).

The examination of areagrams based on phylogenetic relationships of other taxa included in three-item analysis help clarify the history of the Arctic Region. Two lamprey taxa inhabit the Arctic region- *Lethenteron camtschaticum* and *Petromyzon marinus* (IUCN 2014). The first of these also inhabits the Pacific region, and is phylogenetically most closely related to two pairs of sister species; one species in each pair inhabits the Atlantic region, and one in each pair the Pacific (IUCN 2014, Gill *et al.* 2003). Along with the two species pairs, *L. camtschaticum*
is part of a larger clade composed of five other Pacific and one Atlantic species (IUCN 2014, Gill et al. 2003). *Petromyzon marinus*, on the other hand, inhabits the Atlantic as well as the Arctic (IUCN 2014). This taxon is in a clade with three Atlantic species (Gill et al. 2003). Of the pinnipeds included in three-item analysis, *Odobenus rosmarus, Histriophoca fasciata, Pagophilus groenlandicus, Pusa hispida* and *Phoca vitulina* range in the Arctic (Arnason et al. 2006, IUCN 2014). *Pusa hispida* and *P. vitulina* are also be found in regions in the Atlantic and Pacific Regions of this analysis (Arnason et al. 2006, IUCN 2014). *Histriophoca fasciata* is found in the Pacific as well as the Arctic, while its sister species *P. groenlandicus* is found in the Atlantic (Arnason et al. 2006, IUCN 2014). *Phoca vitulina* is sister to a Pacific Clade species, and these sister species form a polytomy with *H. grypus*, an Atlantic species, and *P. hispida* (Arnason et al. 2006, IUCN 2014). These phylogenetic relationships suggest area relationships among the Arctic, Pacific, and Atlantic, but the evolutionary history of *O. rosmarus* supports inclusion in the Pacific Clade (IUCN 2014). *Odobenus rosmarus* can be found in the Bering Sea Region as well as the Arctic (IUCN 2014). Although it is not found in the Atlantic or Pacific, it is part of a clade including four other pinnipeds that inhabit only Pacific Regions in our analysis (Arnason et al. 2006, IUCN 2014).

Although the Arctic became geographically connected to the Atlantic in the Cenozoic, long before its connection to the Pacific, other studies show a similarly close area relationship between the Arctic and Pacific as opposed to an Arctic-Atlantic area relationship. Early Cenozoic taxa in the Arctic and East Siberian Seas (my Arctic Region) also inhabited the Atlantic (Marincovich and Gladenkov 2001). During the mid-to-late Cenozoic, with the opening of the Bering Strait (Brigham-Grette 2001, Marincovich and Gladenkov 2001), the fauna of the Arctic became dominated by taxa common in the Pacific, indicating migration from the Pacific to the Arctic (Marincovich and Gladenkov 2001). This relationship also mirrors past studies that
recover Croizat's "boreal" region, which encompasses my Arctic Region, with the Pacific Ocean rather than the Atlantic.

4.4 Bering Sea

Three-item analysis recovered the Atlantic and Pacific Area clades as "sister", with the Bering Sea Region outside of these clades, although this is not strongly supported by geologic history. Geographically, the Bering Sea provides a connection between the Pacific and Arctic oceans. The Bering and Arctic seas are also an aquatic "bridge" between the Pacific and Atlantic. Therefore, one might predict that taxa in the Bering Sea Region would be more closely related to the taxa in the Pacific Area Clade than to those of the Atlantic Clade. Past studies also support inclusion of the Bering Sea with the Pacific; Croizat's "boreal" region includes my Bering Sea Region, and has been repeatedly associated with the Pacific.

Among sturgeons, only A. mikadoi occurs in the Bering Sea; it also occurs in the Pacific and is closely related to other Pacific taxa (IUCN 2014). The two lamprey species occurring in the Bering Sea Region, *Lethenteron camtschaticum* and *Entosphenus tridentatus*, also occur in the Pacific Region (IUCN 2014). *Entosphenus tridentatus* groups in a clade with other Pacific species, while *L. camtschaticum* is in a clade with two Pacific and two Atlantic taxa (IUCN 2014, Gill *et al.* 2003). Of the pinniped taxa that occur in the Bering Sea, two occur in both the Atlantic and Pacific regions and four occur in Pacific Regions (including the Arctic) (IUCN 2014). The area relationships of the Bering Sea may be strongly influenced by widespread and highly migratory pinniped taxa that inhabit this region. These factors suggest re-examination of whether the Bering Sea Region should be a distinct area for future analyses. Further analyses including more taxa may help clarify the issue.
4.5 Conclusions

With the exception of the Bering Sea, the area relationships proposed here are supported by geologic history. Excluding the Bering Sea Region, broad area relationships, namely the inclusion of the Arctic Region within the Pacific Area Clade, reflect patterns that have been proposed in past studies (Marincovich and Gladenkov 2001). Future studies should consider different or additional taxa to clarify the issues with the hypothesized Bering Sea Region area relationships. They should also consider consolidation of the Yellow Sea Region into the China Region.

Mechanisms leading to current distributions of sturgeons were not discussed here for multiple reasons. First, the geologic history of some of the regions considered is contentious (e.g. the Arctic, see Clark 1974). Second, even with reliable dates for geologic events, a well-supported, time-calibrated sturgeon phylogeny would be needed to hypothesize potential specific mechanisms such as invasions or strandings. Finally, the methods of biogeography that I endorse (comparative biogeography) do not include selection of ancestral habitats as either strictly marine or freshwater. Because distributional mechanisms are not discussed, the evolution of migratory habits and how they relate to geologic history (previously discussed by Bemis and Kynard 1997) are not discussed here. Despite the lack of speculation on the ancestral habitat and on potential mechanisms for the evolution of anadromy, the occurrence of both anadromous and strictly freshwater taxa in nearly every region, as well as the occurrence of anadromy and freshwater habits in taxa at varying levels of the phylogeny, supports previous hypotheses that the ancestral habitat likely included both freshwater and marine environments.
Acknowledgements

Many thanks to my committee member Lynne Parenti, whose advice, support, and help with analyses has been invaluable from the development of this chapter onward. Thanks also to my advisor Eric Hilton for helpful comments and advice on the chapter, and to the rest of my committee, John Graves, Casey Dillman, and Emmett Duffy. Funding sources included National Science Foundation (DEB 0841691) and the GK-12 Program.
References


(Osteichthyes: Acipenseridae): A Synthesis of Phylogenetics, Palaeontology and

Clark D. L. 1974. Late Mesozoic and early Cenozoic sediment cores from the Arctic Ocean.
Geology, 2: 41-44.

Croizat L. 1958. Panbiogeography or an introductory synthesis of zoogeography,
phytogeography, and geology; with notes on evolution, systematics, ecology,
anthropology, etc. Vol. 1—The New World, Vol. 2a- The Old World, Vol. 2b- The Old
World. L Croizat: Caracas.


systematics of Scaphirhynchinae: an assessment of North American and Central Asian

Lampetra inferred from mitochondrial cytochrome b and ND3 gene sequences.
Canadian Journal of Fisheries and Aquatic Sciences, 56: 2340-2349.


Torsvik T.H., Carlos D., Mosar J., Cocks L.R.M. and Malme T., 2002. Global reconstructions and
North Atlantic palaeogeography 400 Ma to Recent. In: Eide E.A., BATLAS – Mid Norway
plate reconstructions atlas with global and Atlantic perspectives. Geological Survey of


University of Wisconsin Sea Grant Institute (UWSG) 2013. Great Lakes: The Formation of the
Great Lakes: How They Were Made. Available at:

Vasil’eva E.D., Vasil’ev V.P., Shedko S.V., Novomodny G.V. 2009. The Revision of the Validity of
Genus Huso (Acipenseridae) Based on Recent Morphological and Genetic Data with

Vassiljev J. and Saarse L/ 2013. Timing of the Baltic Ice Lake in the eastern Baltic. Bulletin of the

Agency Publication. Available at:
around-europe/page141.html>


### Tables

**Table 1: Sturgeon species ranges.** Ranges of sturgeon species in areas of endemism. JPN=Japan Region, CHN=China Region, YLW=Yellow Sea Region, PNA=Pacific Coast of North America Region, AES=Arctic Region, BNS=Baltic and North Sea Region, PTC=Ponto-Caspian Region, ANA=Atlantic Coast of North America Region, MGM=Gulf of Mexico Region, NAL=North American Lakes Region, BRG=Bering Sea Region

| Species                        | JPN | CHN | YLW | PNA | AES | BNS | PTC | ANA | MGM | NAL | BRG |
|--------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Acipenser fulvescens           |     |     |     |     |     |     |     | ✓   |     |     |     |
| Acipenser baerii               |     |     |     |     |     |     | ✓   |     |     |     |     |
| Acipenser brevirostrum         |     |     |     |     |     | ✓   |     |     |     |     |     |
| Acipenser dabryanus            |     |     |     |     |     | ✓   |     |     |     |     |     |
| Acipenser gueldenstaedtii      |     |     |     |     |     |     | ✓   |     |     |     |     |
| Acipenser mediostris           |     |     |     |     |     |     |     | ✓   |     |     |     |
| Acipenser mikadoi              | ✓   |     |     |     |     |     | ✓   | ✓   |     |     |     |
| Acipenser naccarii             |     |     |     |     |     |     |     |     |     | ✓   |     |
| Acipenser oxyrinchus desotoi   |     |     |     |     |     |     |     |     |     |     | ✓   |
| Acipenser oxyrinchus oxyrinchus|     |     |     |     |     |     |     |     | ✓   | ✓   |     |
| Acipenser ruthenus             | ✓   |     |     |     |     |     |     | ✓   |     |     |     |
| Acipenser schrenckii           |     |     |     |     |     |     |     |     |     |     | ✓   |
| Acipenser sinensis             |     | ✓   | ✓   |     |     |     |     |     |     |     |     |
| Acipenser stellatus            |     |     |     |     |     |     | ✓   | ✓   |     |     |     |
| Acipenser sturio               |     |     |     |     |     |     |     | ✓   |     |     | ✓   |
| Acipenser transmontanus        |     |     |     |     |     |     |     |     |     |     | ✓   |
| Huso dauricus                  |     |     |     |     |     |     |     |     |     |     | ✓   |
| Huso huso                      |     |     |     |     |     |     |     |     |     |     |     |
| Pseudoscaphirhynchus kaufmanni |     |     |     |     |     |     |     |     |     |     | ✓   |
| Scaphirhynchus albus           |     |     |     |     |     |     |     |     |     |     |     |
| Scaphirhynchus platorynchus    |     |     |     |     |     |     |     |     |     |     |     |
| Scaphirhynchus suttkusi        |     |     |     |     |     |     |     |     |     |     |     |
Table 2: Lamprey species ranges. Ranges of lamprey species in areas of endemism. JPN=Japan Region, CHN=China Region, YLW=Yellow Sea Region, PNA=Pacific Coast of North America Region, AES=Arctic Region, BNS=Baltic and North Sea Region, PTC=Ponto-Caspian Region, ANA=Atlantic Coast of North America Region, MGM=Gulf of Mexico Region, NAL=North American Lakes Region, BRG=Bering Sea Region

<table>
<thead>
<tr>
<th>Species</th>
<th>JPN</th>
<th>CHN</th>
<th>YLW</th>
<th>PNA</th>
<th>AES</th>
<th>BNS</th>
<th>PTC</th>
<th>ANA</th>
<th>MGM</th>
<th>NAL</th>
<th>BRG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caspiomyzon wagneri</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entosphenus macrostomus</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entosphenus minimus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entosphenus similis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entosphenus tridentatus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Eudontomyzon danfordi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eudontomyzon morii</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ichthyomyzon unicuspis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Ichthyomyzon bdellium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Ichthyomyzon casteneus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Lampetra ayresii</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Lampetra fluviatilis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Lethenteron camtschaticum</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Petromyzon marinus</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Tetrapeurodon spadiceus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

131
Table 3: Pinniped species ranges. Ranges of pinniped species in areas of endemism. JPN=Japan Region, CHN=China Region, YLW=Yellow Sea Region, PNA=Pacific Coast of North America Region, AES=Arctic Region, BNS=Baltic and North Sea Region, PTC=Ponto-Caspian Region, ANA=Atlantic Coast of North America Region, MGM=Gulf of Mexico Region, NAL=North American Lakes Region, BRG=Bering Sea Region

| Species                          | JPN | CHN | YLW | PNA | BNS | PTC | ANA | MGM | NAL | AUS | BRG |
|----------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| *Monachus monachus*               |     |     |     | X   |     |     |     |     |     |     |     |     |
| *Arctocephalus townsendi*         |     |     |     |     | X   |     |     |     |     |     |     |     |
| *Callorhinus ursinus*             | X   |     |     |     |     |     |     |     | X   |     |     |     |
| *Cystophora cristata*             |     |     | X   |     |     | X   | X   |     |     |     |     |     |
| *Erignathus barbatus*             | X   |     |     |     |     | X   |     |     |     |     |     |     |
| *Eumetopias jubatus*              | X   |     |     |     | X   |     |     |     |     |     |     |     |
| *Halichoerus grypus*              |     |     |     |     |     |     |     | X   |     |     |     |     |
| *Histriophoca fasciata*           | X   |     |     |     |     |     |     |     |     |     | X   | X   |
| *Mirounga angustirostris*          |     |     |     |     |     |     |     |     |     | X   |     |     |
| *Odobenus rosmarus*               |     |     |     |     |     |     |     |     |     |     | X   | X   |
| *Pagophilus groenlandicus*        |     |     |     |     |     |     |     |     |     |     |     | X   |
| *Phoca largha*                    | X   | X   |     |     |     |     |     |     |     |     |     |     |
| *Phoca vitulina*                  | X   |     |     |     |     |     | X   | X   |     |     |     |     |
| *Pusa caspica*                    |     |     |     |     |     |     |     |     |     | X   |     |     |
| *Pusa hispida*                    | X   | X   |     |     |     |     | X   | X   | X   |     | X   |     |
| *Zalophus californianus*           |     |     |     |     |     |     |     |     |     |     | X   |     |
Figure 1: Bayesian full mitogenome phylogeny. Posterior probability values are to the left of and below each node. Branch lengths indicate estimated amount of evolutionary (or genetic) change. Long branches (pertaining to outgroups) have been broken to allow for better visualization of ingroup relationships. Exclusively freshwater taxa are indicated by an asterisk (*)
Figure 2: Croizat's global biogeography map. Austral and Boreal regions are depicted with hatched lines.
Figure 3: Areagram of Croizat’s (1958) biogeographic regions following Nelson and Ladiges (2001) and Parenti (2008). Asterisks indicate regions inhabited by sturgeons.
Figure 4: Map of sturgeon Areas of Endemism; Regions are abbreviated as follows BNS: Baltic and North Sea, PTC: Ponto-Caspian, JPN: Japan, YLW: Yellow Sea, CHN: China, AES: Arctic, BRG: Bering Sea, PNA: Pacific North America, MGM: Gulf of Mexico, NAL: North American Lakes, and ANA: Atlantic North America.
Figure 5: Morphological Lamprey Phylogeny modified from Gill et al. 2003; species that do not share area of endemism with sturgeons were pruned from the original topology.
Figure 6: Molecular Pinniped Phylogeny modified from Arnason et al. 2006; species that do not share area of endemism with sturgeons were pruned from the original topology.
7s. *P. kaufmanni*

71. *S. albus*
Figure 7: Range maps of sturgeon species included in biogeographic analyses. Red indicates areas inhabited by each species.
Figure 8: Sturgeon Areagram; species names are replaced on the full mitogenome phylogeny with areas of endemism.
Figure 9: Lamprey Areagram. Species names are replaced on the pruned Gill et al. phylogeny with areas of endemism.
Figure 10: Pinniped Areagram; species names are replaced on the pruned Arnason et al. 2006 phylogeny with areas of endemism
Figure 11: Sturgeon-Lamprey-Pinniped Intersection Tree showing area relationships for sturgeon, lamprey, and pinniped areas of endemism. Box A indicates the primarily Pacific clade, while box B indicates the Atlantic-Arctic clade. The completeness index 92.7%.
CHAPTER 3.

Utility of Ontogenetic Characters in Phylogenetic Studies
Abstract

Form, which is represented by morphology, and function, which is exemplified through behavior, are said to be inextricably linked. Sturgeons (Acipenseridae) exhibit dramatic morphological and behavioral shifts during ontogeny: developing then resorbing teeth and beginning their exogenous feeding behavior by swimming up into the water column before eventually becoming benthic feeders. Researchers often use behavioral and ontogenetic characters in the study of phylogeny, but the timing of morphological and behavioral shifts in sturgeons has not been investigated for phylogenetic signal. The primary goal of this chapter is to identify links between morphological and behavioral ontogeny, and to determine whether the timing of ontogenetic changes in their morphology and behavior carries phylogenetic signal. Correlation is found between one set of behavioral-morphological characters, the age at which migration of early life stage sturgeons is completed and the length at which tooth development is complete. One ontogenetic character, the age at which teeth are completely resorbed, shows phylogenetic signal.
1. Introduction

It is widely understood that form (morphology) and function (exemplified in behavior) are often inextricably linked (e.g. Russell 1982, Arnold 1983, Zelditch et al. 2004, Bertossa 2011). Comparative analyses among species, which have traditionally been used to test for correlations between morphological traits (Garland et al. 1992), may also be used to test for links between form (morphological traits) and function (behavioral traits) (Jordana et al. 1999). Traditional comparative methods, such as Pearson product moment correlations, assume that trait data under consideration are not statistically linked and that all taxa included in analysis are equally related to one another (Felsenstein 1985, Garland et al. 2005). However, the hierarchical nature of species relationships violates this assumption; traits are linked through the shared evolutionary history of their taxa (Garland et al. 2005). To remove the confounding factor of hierarchical relationships from comparative analyses and avoid making an assumption that is sure to be violated, phylogenetic relationships must be taken into account in comparative analyses. Phylogenetic comparative methods such as phylogenetically independent contrasts (PIC) use phylogenies to control for the confounding phylogenetic relationships. This allows for the straightforward testing for correlations among species traits. The 25 species of sturgeon, distributed amongst the genera *Acipenser*, *Scaphirhynchus*, *Pseudoscaphirhynchus*, and *Huso*, exhibit behavioral and morphological changes during development that may be linked, and may exhibit phylogenetic signal.

Frequently described as “ancient” fishes (Inoue et al. 2003, Birstein 1993, Ludwig et al. 2007, Katsu et al. 2008), sturgeons exhibit dramatic morphological and behavioral shifts between hatching and maturity (Kirschbaum and Williot 2011, Nelson et al. 2013, Dang and
Many sturgeon species are anadromous, moving between fresh and saltwater during their lives (Bemis and Kynard 1997, Peng 2007). They spawn and hatch as free embryos in riverine systems, where they develop into larvae and begin foraging (Kynard and Parker 2005, Kynard 1997, Kynard and Horgan 2002, Kynard et al. 2005, Parker 2007, Kynard and Parker 2010). The anadromous species move into the ocean by adulthood, but return to rivers to spawn (Nelson et al. 2013). Sturgeon larval development includes major changes to jaw morphology, the appearance and proliferation of electroreceptors on the rostrum, and the development of bony scutes in rows along the sides of the body (although scute development continues after the larval stage) (Parker 2007). Sturgeons hatch with forward facing, subterminal jaws, which shift to a ventral orientation during early stages of development (i.e., before the juvenile stage). Oral jaws (subsequently referred to as jaws) in juvenile and adult sturgeons are highly protrusible, enabling suction feeding. The jaws of sturgeons are toothless at hatching, but teeth develop on the dentary, dermopalatines, and palatoptyygoid during the larval stage. The teeth are subsequently resorbed and the jaws of juvenile and adult sturgeons are edentulous. It is likely that teeth assist in prey capture while the jaws are still terminally positioned (B. Kuhajda, personal communication, 2011).

Sturgeons also display shifts in behavior during larval and early juvenile life stages. The links between behavioral and morphological development in sturgeons have been the focus of a variety of studies, and morphological and behavioral changes during early life history stages of sturgeons are known to occur concurrently (Dang and Zhang 2014, Parker 2007, Gisbert and Williot 1999). Dang and Zhang (2014) found that behavioral ontogeny, in general, correlated with morphological development in *A. fulvescens*. Parker (2007) showed that the peak of
migration in A. brevirostrum is linked with the completion of development of the pectoral and pelvic fins and the presence of a pre-anal fin fold. Gisbert and Williot (1999) showed that the beginning of free feeding in A. baerii is accompanied by changes in digestive enzymes present in the gut. The ability to perform behaviors such as swimming downstream requires that certain morphological structures be developed (e.g., fins; Parker 2007). It follows, then, that the timing of tooth and oral jaw development should correlate with changes in feeding behavior.

Several studies have documented early life history changes in sturgeons (e.g. Buckley and Kynard 1981, Kynard 1997, Kynard and Horgan 2002, Kynard et al. 2005, Parker 2007, Richmond and Kynard 1995). The goals of these studies included informing conservation and restoration efforts and describing various aspects of behavior (feeding behavior, habitat use, upstream and downstream movement, activity patterns at different time of day and different light levels night, and position in the water column). They have also shown that different sturgeon species exhibit different behaviors during the larval stage. For example, the larvae of different species lie on the substrate, seek cover in rocky habitat, or swim up above the bottom during foraging migrations (Parker 2007, Gisbert and Williot 1997). Behaviors that are shared among taxa in the larval stage may diverge as the species develop; most species first feed on plankton and eventually progress to a diet of benthic organisms, but some species (e.g., S. albus) may become piscivorous as adults (Grohs et al. 2009, Parker 2007).

Developmental studies have noted a general trend for post-hatch embryonic (=free embryo) A. oxyrinchus and A. brevirostrum to seek and maintain cover until the larval stage is reached, upon which the fish emerge to migrate and forage (Buckley and Kynard 1981, Kynard 1997, Kynard and Horgan 2002, Richmond and Kynard 1995). Acipenser transmontanus,
however, begins a short migration immediately upon hatching (Kynard and Parker 2005 and 2010). This species migrates during the day and night, but exhibits nocturnal foraging habits (Kynard and Parker 2005, 2010). *Acipenser medirostris* cannot swim well upon hatching, but can move short distances to reach cover (Kynard *et al.* 2005). Like *A. oxyrinchus* and *A. brevirostrum*, *A. medirostris* emerges from cover-providing habitat at the larval stage and begins migrating downstream, with brief foraging periods during migration (Kynard *et al.* 2005). Conversely, free embryos and larvae of *A. medirostris* exhibit primarily nocturnal migration and foraging habits (Kynard *et al.* 2005). *Acipenser oxyrinchus* is nocturnal during the early stages of its 12-day migration period, and decidedly diurnal during the later migration period (Kynard and Horgan 2002). *Acipenser brevirostrum* migrates diurnally (Kynard and Horgan 2002).

Patterns of development of sturgeons, including behavioral and morphological ontogeny, may reflect phylogenetic signal. For example, migration and habitat preference have been used as characters in phylogenetic studies of sturgeons (Mayden and Kuhajda 1996). Some researchers have suggested that similarities in behavior in sturgeon development is the result of convergent evolution rather than shared evolutionary history (Kynard and Horgan 2002). Morphological characters have been used in phylogenetic studies of sturgeons (e.g. Findeis 1997; Artyukhin 2006, Hilton *et al.* 2011), but the timing of development of these characters has not been explored in the context of the phylogenetic history of sturgeons. The goals of this study are twofold. First, I test whether the timing of specific behavioral and morphological shifts carries phylogenetic signal. Results are discussed in terms of which characters carry phylogenetic signal matching that of a full mitogenome Bayesian phylogeny (Figure 1). Second, I test for correlation between the timing of morphological and behavioral
ontogenetic shifts related to feeding structures (teeth and oral jaw elements, specifically Meckel's cartilage and cartilage of the palatoquadrate) and feeding, respectively.

2. Methods

2.1 Taxon Sampling and Data Collection

2.1.1 Morphological Characters

Developmental series of known-age specimens (from one to 65 days post-hatch) of *Acipenser transmontanus* (VIMS 17716), *A. oxyrinchus* (VIMS uncataloged), *A. brevirostrum* (VIMS 12091; also VIMS 12076-12082, VIMS 12086, VIMS12088, VIMS 12089, VIMS, 12101), *A. medirostris* (VIMS 17715), and *A. fulvescens* (VIMS 13577), were obtained from hatcheries and collections; all specimens were deposited in the VIMS Ichthyology Collection. Specimens were preserved in 4% paraformaldehyde and transferred to 70% ethanol. Specimens were cleared and stained for bone (stained red) and cartilage (stained blue) based on a protocol modified from Dingerkus and Uhler (1977) (examples of cleared and stained specimens are shown in Figure 2). All specimens were photographed and total length (TL) measured using a Zeiss Discovery V20 microscope. Specimens were examined to identify the earliest appearance feeding-related structures (teeth and oral jaw elements) in each species. Additionally, the development of teeth was tracked, as was the resorption of the yolk sac. The earliest age (days posthatch) and shortest length (total length) at which elements were first observed was recorded for each species. Age at occurrence for some elements is unknown. This resulted in 10 continuous morphological ontogenetic characters:

1. Age at which first appearance of mandibular arch (jaw) elements (specifically Meckel’s cartilage and cartilage of the palatoquadrate) occurs;
2. Length at which first appearance of mandibular arch elements occurs;
3. Age at which teeth first appear;
4. Length at which teeth first appear;
5. Age at which all teeth are completely developed, and after which teeth begin to resorb, becoming thinner, shorter, and fewer (this milestone is identified by the maximum counts and lengths of teeth);
6. Length at which all teeth are completely developed;
7. Age when teeth are completely resorbed, defined as the presence of tooth sockets but absence of mineralized tissue protruding from sockets;
8. Length at which teeth are completely resorbed;
9. Age at which the yolk sac is completely digested; and
10. Length at which the yolk sac is completely digested.

2.1.2 Behavioral Characters
Extensive ontogenetic behavioral data for five species were available from the literature: *A. brevirostrum* (Richmond and Kynard 1995, Kynard 1997, Kynard and Horgan 2002), *A. medirostris* (Kynard et al. 2005), *A. oxyrinchus* (Kynard and Horgan 2002), *A. fulvescens* (Smith and King 2005) and *A. transmontanus* (Kynard and Parker 2005, 2010). These data include information about habitat use, preference for illumination, foraging behavior, migratory behavior, preference for dark or light substrate, swimming height at different life stages, nocturnal and diurnal behavior, and timing of the onset of arval migration. Data for some of these characters were not available for *A. fulvescens* (see Smith and King 2005). Habitat use was characterized by the percent of fish using rocky cover, the open bottom, the water column, and the water surface on each day for *A. transmontanus*, *A. brevirostrum*, and *A. medirostris*; this
has not been investigated for *A. oxyrinchus* (see Kynard and Horgan 2002, Kynard *et al.* 2005, Kynard and Parker 2005). Preference for light or dark habitats was measured by the mean percent of fish in illuminated sections of a tank as opposed to covered, dark sections each day for *A. transmontanus*, *A. brevirostrum*, *A. medirostris*, and *A. oxyrinchus* (see Kynard and Horgan 2002, Kynard *et al.* 2005, Kynard and Parker 2005). Preference for different color substrates was determined by the mean percent of fish on white, rather than black, substrate in *A. transmontanus*, *A. oxyrinchus*, and *A. medirostris* (see Kynard and Horgan 2002, Kynard *et al.* 2005, Kynard and Parker 2005). It was measured as the percent of individuals on black, gray, and white substrates for *A. brevirostrum* (Richmond and Kynard 1995). Swimming height was calculated as the mean swimming height of fish, in centimeters, each day for *A. transmontanus* and *A. medirostris* (Kynard *et al.* 2005, Kynard and Parker 2005). For *A. brevirostrum*, swimming height was measured as the mean swimming height of fish introduced at the water surface and after swimming to the bottom (Kynard and Horgan 2002). Information regarding *A. oxyrinchus* was not available for this trait. Migratory behavior was determined for *A. transmontanus*, *A. brevirostrum*, *A. medirostris*, and *A. oxyrinchus* by counting the number of times fish passed a set point moving upstream during daylight hours (Kynard and Horgan 2002, Kynard *et al.* 2005, Kynard and Parker 2005). Nocturnal migratory behavior was measured for all species in the same way as migratory behavior, but at night (Kynard and Horgan 2002, Kynard *et al.* 2005, Kynard and Parker 2005).

From these studies, I identified and analyzed five continuous behavioral ontogenetic characters that were measured in a standard format across studies, and comparable across four
of the included species- *A. transmontanus*, *A. oxyrhynchus*, *A. brevirostrum*, and *A. mediostris*.

These characters are:

1. The earliest age (in degree-days posthatch) at which fish begin to swim above the substrate;
2. The age at which downstream migration begins, regardless of whether migration is during the day or night (foraging begins at the same time);
3. The age at which migration ends;
4. The age at which fish first use open habitat, including the open bottom, the water column, and the water surface regardless of whether they previously hid in rocky substrate; and
5. The earliest age at which fish use open habitat after a period of time hiding in rocky substrate.


1. Nocturnality: categorized as primarily nocturnal or diurnal during the larval stage;
2. Migration: timing of migration was determined to be either early (as free embryos) in development or late (as larvae) for all five species; and
3. Larval behavior related to phototaxis: photopositive or photonegative.
2.2 Metrics for Phylogenetic Signal in Individual Traits

To test for phylogenetic signal in individual ontogenetic characters, I calculated two autocorrelation indices (Moran’s I and Abouheif’s C-mean) following Munkemuller et al. (2012). A full mitogenome Bayesian phylogeny, originally including 22 sturgeon species (Figure 1), was used as the reference phylogeny for these tests. All species except *A. transmontanus*, *A. oxyrinchus*, *A. fulvescens*, *A. brevirostrum*, and *A. transmontanus* were pruned (using the command drop.tip in the ape package, version 3.4, Paradis et al. 2004, for R, R Core Team 2014) from this topology prior to comparison (Figure 3). For tests of traits for which data were missing for *A. fulvescens*, this species was also pruned from the phylogeny (Figure 4).

Moran’s I was originally designed to measure spatial autocorrelation (Moran 1950), but was modified for use with phylogenies and trait data by Gittleman and Kot (1990). It measures phylogenetic signal based on the correlation between variation of the trait in question and the given phylogeny. To calculate Moran’s I, I ran 999 observations under the Moran.I function in the ape package (version 3.4, Paradis et al. 2004) for R (R Core Team 2014). This package calculates the observed value of Moran’s I and the value that would be expected (in the case of no correlation between the trait and the phylogeny in question). If the observed value is greater than the expected value, the trait in question positively correlated with the phylogeny, indicating phylogenetic signal. I used a 95% confidence interval to determine significance in the difference between the observed and expected values.

Abouheif’s C-mean tests for independence between traits and a given phylogeny by squaring the differences between the trait values of sister species as defined by the phylogeny (Jombart and Dray 2008). I calculated this parameter with the abouheif.moran function in the
adephylo package for R, using the method “oriAbouheif” and running 999 iterations (Jombart and Dray 2008). As with Moran’s I, observed and expected (if no phylogenetic signal is present) values of Abouheif’s $C_{mean}$ are calculated, and an observed value greater than the expected value indicates phylogenetic signal. Again, I used a 95% confidence interval to determine significance.

2.3 Tree Comparison Metrics

A character matrix including *A. transmontanus*, *A. fulvescens*, *A. medirostris*, *A. oxyrinchus*, and *A. brevirostrum* was constructed. Characters included in the matrix were coded as categorical (nocturnality, migration, and phototaxic habit; Table 1) or continuous (ages at which teeth first appear, mandibular arch elements appear, teeth are fully developed, teeth are completely resorbed, and the yolk sac is completely absorbed; Table 2). Single-character trees were constructed using the “Heuristic Add and Rearrange Tree Search” function in Mesquite v. 3.04 (Maddison and Maddison 2015). Criteria used in tree searches were Parsimony Character Steps, using the SPR Tree Rearranger and storing 100 maximum trees. The trees were unrooted.

Rzhetsky-Nei (R-N) distances are a metric used to calculate the difference between two topologies using the number of bipartitions that differ between trees (Rzhetsky and Nei 1992). High R-N distances indicate divergent trees, low distances indicate similarity between trees, and an R-N distance of 0 indicates identical topologies (Duchene et al. 2011). Statistical significance tests are not available for R-N distances; this is a method for comparison between topologies only. Rzhetsky-Nei distances were calculated to examine similarity in phylogenetic signal between single-character trees and a reference phylogeny (a full mitogenome Bayesian sturgeon phylogeny pruned (using the command drop.tip in the ape package, version 3.4,
Paradis et al. 2004, for R, R Core Team 2014) to include only the species under consideration) using the dist.topo function in the ape package in R (Paradis et al. 2004, R Core Team 2014).

2.4 Phylogenetic Comparative Tests

To test whether morphological characters correlate with behavioral characters (i.e., are linked), I conducted Phylogenetically Independent Contrast (PIC) and Pearson’s Product Moment of Correlation analyses. PIC allows for tests of correlation with the confounding factor of shared evolutionary history removed, but when few taxa are under consideration, it is recommended that correlation tests in which no branch transformations occur (e.g., Pearson’s Product Moment of Correlations) are run alongside PIC (Swenson 2014, Paradis 2006). Therefore, Pearson’s Correlations were also calculated, and correlations were considered significant if they were found to be so by both PIC and Pearson’s Correlations.

Phylogenetically Independent Contrasts were run for each continuous ontogenetic character (10 morphological and five behavioral characters) using the pic function in the ape package (version 3.4, Paradis et al. 2004) for R (R Core Team 2014) and the pruned full mitogenome phylogeny. PICs were plotted between every possible morphological-behavioral character pair. Because plots revealed that observed correlations were, for the most part, linear, a simple linear regression with the regression line forced through the origin was performed for each pair (Swenson 2014, Paradis 2006). Correlation coefficients were calculated on the PIC-transformed data for morphological and behavioral character pairs using the cor.test function in R, and a 95% confidence interval was used to determine significance. Pearson’s product moment correlation coefficients were calculated for all character pairs using the cor.test function in R and a 95% confidence interval to determine significance.
3. Results

3.1 Traits

3.1.1 Morphological Ontogenetic Traits: Age

Morphological milestones (Table 2) in *A. oxyrinchus* and *A. transmontanus* occurred in the same order, as did traits in *A. medirostris* and *A. fulvescens*. *Acipenser brevirostrum* reached the measured milestones in a unique order.

In *A. oxyrinchus* and *A. transmontanus*, the first appearance of mandibular arch elements occurs at two and three days posthatch, respectively, followed by the appearance of teeth at four days posthatch in both species. The yolk sac is completely resorbed at six days posthatch in *A. oxyrinchus* and 9 days posthatch in *A. transmontanus*. Teeth are completely developed at 19 and 20 days posthatch in *A. oxyrinchus* and *A. transmontanus*, respectively. After the development of teeth is complete, they are lost via resorption; teeth gradually become thinner and shorter until the only evidence they ever existed is the presence of shallow tooth sockets in the underlying bones. This occurs in *A. oxyrinchus* at 23 days posthatch, and slightly earlier, at 21 days posthatch, in *A. transmontanus*. Eventually the tooth sockets become smoothed out, leaving no evidence of teeth; the available developmental series ended prior to some species reaching this stage, so it was not included in the analyses.

In *A. medirostris* and *A. fulvescens*, teeth are completely developed before the yolk sac is completely resorbed. In *A. fulvescens*, first appearance of mandibular arch elements occurs at day seven and teeth first appear at day eight, whereas both of these are reached at seven days posthatch in *A. medirostris*. Completion of tooth development (13 days) and yolk sac resorption (15 days) occur earlier in *A. fulvescens* than in *A. medirostris* (14 and 18 days posthatch,
respectively). Although tooth resorption begins later in *A. medirostris* than in *A. fulvescens*, teeth are completely resorbed in *A. medirostris* earlier than in *A. fulvescens* (21 and 44 days post hatching, respectively).

Mandibular arch elements first appear in *A. brevirostrum* at six days posthatch. Eight days after hatching, the teeth first appear. The complete development of teeth and completion of yolk sac resorption occur at 13 days posthatch. Teeth are completely resorbed, with tooth sockets present, 52 days after hatching.

3.1.2 Morphological Ontogenetic Traits: Length

Lengths at which milestones are reached are provided in Table 2. Commonalities and differences in the order in which these milestones are reached are discussed here.

In *A. oxyrinchus*, *A. transmontanus*, *A. brevirostrum*, and *A. fulvescens*, the mandibular arch elements appear first, followed by first appearance of teeth. In *A. medirostris*, the first appearance of teeth and the oral jaws occur at the same size and mark the earliest milestone reached. *Acipenser oxyrinchus* and *A. transmontanus* both have a completely resorbed yolk sac prior to complete tooth development. In contrast, the completion of tooth development occurs prior to complete resorption of the yolk sac in *A. medirostris*, *A. fulvescens*, and *A. brevirostrum*. The complete resorption of the teeth is the final milestone reached in all five taxa.

3.1.3 Behavioral Ontogenetic Traits: Age

Four of the five sturgeon species under consideration, *A. oxyrinchus*, *A. medirostris*, *A. fulvescens*, and *A. brevirostrum*, migrate as larvae, whereas *A. transmontanus* undergoes a short migration during the free embryo phase (Table 2). *Acipenser transmontanus*, *A. medirostris*, and
A. fulvescens are primarily nocturnal during their larval migration phases (Smith and King 2005, Kynard et al. 2005, Kynard and Parker 1995 and 2010). Acipenser oxyrinchus and A. brevirostrum are diurnal during migratory life stages (Kynard and Horgan 2002).

Phototaxic behavior was characterized for free embryos and larval A. oxyrinchus, A. brevirostrum, A. transmontanus, and A. medirostris based on tests for preference for illuminated or dark habitat (Kynard 1997, Kynard et al. 2005, Kynard and Horgan 2002, Kynard and Parker 2005 and 2010); phototaxis data were not available for A. fulvescens (Table 1). All four species are photonegative as free embryos. Both A. oxyrinchus and A. brevirostrum become photopositive when they reach the larval stage. Acipenser transmontanus and A. medirostris continue photonegative behavior into the larval stage of life.

3.2 Pruned (reference) topologies

When pruned (using the command drop.tip in the ape package, version 3.4, Paradis et al. 2004, for R, R Core Team 2014) to include only the five taxa in these analyses, the full mitogenome Bayesian topology included two monophyletic groups: one in which A. oxyrinchus is sister to (A. transmontanus, A. medirostris), and one in which A. fulvescens and A. brevirostrum are sister species (Figure 3). When A. fulvescens was pruned out of the topology for analyses including only four taxa, A. brevirostrum fell outside of the A. oxyrinchus, (A. transmontanus, A. medirostris) clade (Figure 4). These topologies were used as the reference phylogenies in tree comparison and PIC.

3.3 Phylogenetic Signal in Individual Traits

Moran’s I and Abouheif’s Cmean both returned statistically significant results for one morphological character (Table 3). The age at which teeth are completely resorbed but tooth
sockets are still present carries phylogenetic signal matching that of the reference phylogeny. Neither length (Table 4) nor behavioral characters exhibited phylogenetic signal.

3.4 Tree Comparisons

3.4.1 Morphological Character Tree Comparisons

The topology recovered by the tree based on age at which teeth are completely resorbed but tooth sockets are present exactly matched the reference phylogeny (Rzhetsky-Nei distance = 0; Figure 3).

The trees based on the number of days at which teeth first appear, at which teeth are completely developed, and at which the yolk sac is completely resorbed also recovered A. brevirostrum and A. fulvescens as sister taxa, although they recovered A. medirostris as sister to these species, and (A. transmontanus, A. oxyrinchus) as sister to this clade (Rzhetsky-Nei distance = 2; Figure 5). The topology based on first appearance of mandibular arch elements also recovered A. transmontanus and A. oxyrinchus as sister taxa, but with A. medirostris sister to A. fulvescens, and A. brevirostrum outside of that group (Rzhetsky-Nei distance = 2; Figure 6).

3.4.2 Behavioral Character Tree Comparisons

The tree based on the timing of initial migration in sturgeon larvae recovered a polytomy consisting of A. brevirostrum, A. fulvescens, A. medirostris, and A. oxyrinchus, to which A. transmontanus is sister (Rzhetsky-Nei distance = 2; Figure 7). Two different topologies, both with a Rzhetsky-Nei distance = 3, were recovered by nocturnal behavior (Figure 8) and phototaxis (Figure 9).
Nocturnal behavior recovered a polytomy composed of *A. fulvescens*, *A. medirostris*, and *A. transmontanus*, with *A. brevirostrum* and *A. oxyrinchus* as sister taxa (Figure 8). Data on the phototaxic habits of *A. fulvescens* were not available. This trait recovers a topology with two sister groups: *A. medirostris* and *A. transmontanus*; and *A. brevirostrum* and *A. oxyrinchus* (Figure 9). No behavioral traits recovered the same topology as the full mitogenome.

3.5 Correlations between Characters

Correlation was shown in one pair of behavioral and morphological characters, the age at which migration ends and the length at which teeth are completely developed. Both PIC and Pearson's product moment correlation tests identified a negative correlation between these characters (Pearson: $c=-0.97$, $p=0.03$; PIC: adjusted $R^2=0.92$, $p=0.03$; Figures 7, 5, and 10), showing that correlation is present independent of phylogeny. This indicates that species whose teeth are fully developed at smaller sizes end their larval migration at older ages than those whose teeth finish developing later. No other morphological-behavioral trait pair showed evidence of correlation in both transformed (PIC) and non-transformed (Pearson) analyses.

4. Discussion and Conclusion

4.1 Phylogenetic Signal

The phylogenetic signal shown, by autocorrelation indices, in the age at which teeth are completely resorbed but tooth sockets are still present, is reflected by the fact that this trait recovered the same topology as the full mitogenome. The failure to detect phylogenetic signal in other traits does not necessarily mean that phylogenetic signal is not present. Autocorrelation tests do not pick up on all aspects of phylogenetic signal, so signal could be present but undetected by the methods used. Additionally, few species were included in these
analyses, and tests for phylogenetic signal perform better when more species are considered (Blomberg et al. 2003, Munkemuller et al. 2012). Previous evaluations of the effect of sample size (number of nodes in a phylogeny) found that autocorrelation tests perform better at larger sample sizes (Munkemuller et al. 2012); these tests have not been evaluated for use with fewer than eight species (Pavoine et al. 2008). Still, my results suggest further consideration of the patterns detected and factors other than phylogenetic signal that may drive these patterns.

Considering behavioral characters in the context of the full mitogenome phylogeny revealed that both the full mitogenome and phototaxic behavior link *Acipenser medirostris* and *A. transmontanus* as sister species. This might suggest some phylogenetic signal in phototaxis even though signal was not detected by autocorrelation metrics. However, factors other than common ancestry may contribute to this similarity between these species. Although these species are sympatric, the specimens examined in behavioral analyses were not from the same river of origin so I cannot conclude that habitat-related pressures drive the similarity.

Convergent evolution, as suggested by previous studies (Kynard and Horgan 2002), may play a role in shared behavioral traits. Further analyses with data on more species could help shed light on this similarity.

Behavioral trait similarities between other species are influenced by non-phylogenetic factors; *Acipenser oxyrinchus* and *A. brevirostrum* are not closely related (based on the full mitogenome phylogeny), but they share several larval behavioral traits. Similarity in their diurnal/nocturnality behavior and phototactic habits, coupled with the fact that these characters did not appear to exhibit phylogenetic signal, suggests that factors other than phylogeny influence these behaviors. Kynard and Horgan (2002) concluded that similarities in the larval
behavior of these species was likely due to common adaptations. Kynard et al. (2005) described predation, foraging, competition, and visual acuity as factors that contribute to nocturnal behavior. While these species are sympatric, individuals used in behavioral studies did not come from the same river of origin. Pressures influencing these behaviors cannot, therefore, be directly linked to shared habitat but may reflect similarities between habitats.

4.2 Correlations between Characters

Because teeth may play a role in the foraging that occurs during migration, the negative correlation between the age at which migration of early life stage sturgeons is completed and the length at which tooth development is complete is somewhat counter-intuitive (Kynard 1997, Kynrd and Horgan 2002, Kynard and Park 2005 and 2010, Kynard et al. 2005). Species such as A. oxyrinchus, in which tooth development is complete at relatively small sizes, also complete their migration later than those (e.g. A. mediostris) that develop teeth at larger sizes. Although the end of migration does not correlate with the age at which tooth development is complete, and length data are unavailable for the individuals used in behavioral studies, teeth are present when foraging begins (Richmond and Kynard 1995). The developmental stage of the teeth during this migratory and foraging behavior is unknown.

Although the correlation between the age at which migration ends and completion of tooth development supports my hypothesis that the feeding-related morphological characters I measured would be correlated to the behavioral traits examined, the lack of evidence for correlation between most morphological and behavioral characters must be considered.
Factors other than an actual lack of correlation might have influenced my analyses. Behavioral characters obtained from the literature (Kynard et al. 1997, 2002, 2005, 2010) were measured as age in degree days at which milestones occurred. Age in degree-days was calculated by Kynard et al. (1997, 2002, 2005, 2010) as the sum of 0.5 x temperature °C for all days post-hatch. Because data on the temperature at which all specimens in my developmental series were not available, morphological traits were measured in days post-hatch rather than degree-days post-hatch. This difference in age measurement may have contributed to the lack of correlation between age-based morphological and behavioral characters.

Similar to autocorrelation metrics, PICs are sensitive to sample size. While Ackerly (2000) found that PIC is relatively reliable for the exploration of links between continuous traits even without complete taxon sampling, the use of non-phylogenetically-transformed data necessarily accompanied PIC due to the small sample size included in this study. The inclusion of additional taxa in future analyses would make for more reliable analyses.

The timing of behavioral and morphological shifts may be influenced by different factors; predation risk may influence behavior while temperature influences morphological development. Some of the behavioral characters I evaluated, such as use of open habitat, are likely linked to predation risk. For example, the manner in which larval sturgeons migrate has been hypothesized to be a habit used to find cover and avoid predation (Richmond and Kynard 1995). The use of open habitat is linked to predation and food availability—sturgeons begin using open habitat and foraging when predation risk is not prohibitively high (Parker 2007).

In addition to helping shed light on the lack of correlation between the timing of morphological and behavioral shifts, the influence of different factors on growth may explain
discrepancies in the age and length at which morphological milestones are reached. While food availability affects both behavior and growth (Gisbert and Williot 1997), the exact nutritional components of the food available can strongly impact growth and skeletal development of larval fishes (Gawlicka et al. 2002, Cahu et al. 2003). Temperature is widely known to impact growth of both fresh and saltwater fishes, and colder temperatures have been linked to slow growth in *A. brevirostrum* and *A. oxyrinchus* (see Houde 1989, Claramunt and Wahl 2000, Hardy and Litvak 2004). Light and day length have also been shown to impact larval fish growth, with longer days increasing growth (Boeuf and Bail 1999).

### 4.3 Concluding Remarks

Overall, I did not find the timing of the behavioral ontogenetic milestones considered here to carry phylogenetic signal. The timing of one morphological ontogenetic milestone did carry phylogenetic signal; the use of this type of data needs to be considered carefully in phylogenetic analyses. This study included few behavioral and morphological milestones in the context of a small (five taxon) phylogeny. Other studies have found these types of characters useful in the recovery (e.g. Tucker et al. 1993), resolution (e.g. Scholtz et al. 2009), and polarization (Bryant 2001) of phylogenies. The timing of the aspects of morphological and behavioral development in larval sturgeons that were considered here are likely impacted by factors other than common ancestry, such that any phylogenetic signal is swamped by selection pressures. Food availability and quality, temperature, and other habitat factors are known to influence early life history (Houde 1989, Claramunt and Wahl 2000, Gawlicka et al. 2002, Cahu et al. 2003, Hardy and Litvak 2004). Therefore, signal may be present in these character, but that signal may be swamped out by other factors that play a role in the timing of development.
The correlation between one morphological and one behavioral trait support the fact that form and function are inextricably linked, but lack of correlation among other traits suggests that behavior and morphological development are impacted by differing factors. While form and function are definitely linked (larval sturgeons feed in the water column until their oral jaws are ventrally oriented to allow for benthic feeding), the characters considered here appear to be influenced more strongly by outside factors than by one another.

Finally, my findings may be biased by the small number of taxa considered in analysis. The best metrics available for the number of taxa included were selected for phylogenetic signal testing, and the restricted taxon sampling was accounted for in correlation tests by including non-transformed data along with PIC. However, the inclusion of more taxa and consideration of more traits is recommended in future studies of this kind.

Acknowledgements

Thanks to the Hilton lab, specifically Todd Clardy and Alison Deary for help with brainstorming and implementing methods. Thanks to Eric Hilton for support throughout this chapter, and to Bernie Kuhajda for insightful correspondence. Peter Struffenegger and Sterling Caviar and Dave Sien and Manning/Cedarville Hatcheries, Maryland Department of Natural Resources provided specimens for study. Funding sources included National Science Foundation (DEB 0841691), the GK-12 Program, Virginia Institute of Marine Science Dean and Director’s Office and Fisheries Department, and John E. Olney Sr. Ichthyology Award.
References


179


Parker E.L. 2007. Ontogeny and Life History of Shortnose Sturgeon (Acipenser brevirostrum Leseur 1818): Effects of Latitudinal Variation and Water Temperature. Dissertation; Graduate School of the University of Massachusetts Amherst, Amherst, Massachusetts.

clades in sturgeons and paddlefishes (Pisces: Acipenseriformes). Molecular Phylogenetics
and Evolution, 42: 854-862.


Chicago: University of Chicago Press.

evolution" In: Decapod Crustacean Phylogenetics (eds. Martin, J.W., Crandall, K.A. and
Felder D.L.) Taylor & Francis, CRC Press, Boca Raton pp. 31-46


Tables

Table 1: Character matrix of three behavioral traits. For migration, character state 1 indicates that first migration occurs during the free embryo stage, 2 indicates first migration at larval stage. For nocturnality, 1 indicates nocturnal habits during the larval stage, 2 indicates a diurnal habit. For Phototaxis, 0 indicates no data, 1 indicates photonegative behavior, and 2 photopositive behavior.

<table>
<thead>
<tr>
<th>TRAITS</th>
<th>CHARACTER STATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Migration</td>
<td>1</td>
</tr>
<tr>
<td>Nocturnality</td>
<td>1</td>
</tr>
<tr>
<td>Phototaxis</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 2: Earliest occurrences of ontogenetic changes in four species. KEY: A.t.=A. transmontanus; A.o.=A. oxyrinchus; A.b.=A. brevirostrum; A.m.=A. medirostris; A.f.=A. fulvescens

<table>
<thead>
<tr>
<th>TRAITS</th>
<th>Days posthatch</th>
<th>Size (mm TL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st appearance of mandibular arch elements</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>1st appearance of teeth</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Yolk sac completely resorbed</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>Teeth completely developed</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>Teeth completely resorbed</td>
<td>21</td>
<td>23</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Fishes swim above benthos</td>
<td>1</td>
<td>8</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>Downstream migration begins</td>
<td>0</td>
<td>8</td>
<td>18</td>
<td>10</td>
</tr>
<tr>
<td>Downstream migration ends</td>
<td>7</td>
<td>27</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>Fishes us open habitat</td>
<td>0</td>
<td>8</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Downstream migration ends</td>
<td>7</td>
<td>27</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>Fishes us open habitat</td>
<td>0</td>
<td>8</td>
<td>9</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 3: Calculations of phylogenetic signal in morphological characters based on age at which milestones first occur. *s indicate significance.

<table>
<thead>
<tr>
<th>TRAITS</th>
<th>Moran's I</th>
<th></th>
<th></th>
<th>Abouheif's Cmean</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I observed</td>
<td>I expected</td>
<td>p-value</td>
<td>observed</td>
<td>p-value</td>
<td>observed</td>
</tr>
<tr>
<td>1st appearance of mandibular arch elements</td>
<td>0.02</td>
<td>-0.22</td>
<td>0.31</td>
<td>-0.01</td>
<td>-0.01</td>
<td>0.51</td>
</tr>
<tr>
<td>1st appearance of teeth</td>
<td>0.73</td>
<td>0.05</td>
<td>0.30</td>
<td>1.38</td>
<td>0.23</td>
<td>0.21</td>
</tr>
<tr>
<td>Yolk sac completely digested</td>
<td>-0.43</td>
<td>-0.38</td>
<td>0.31</td>
<td>-0.67</td>
<td>-0.13</td>
<td>0.80</td>
</tr>
<tr>
<td>Teeth completely developed</td>
<td>0.52</td>
<td>-0.03</td>
<td>0.27</td>
<td>1.05</td>
<td>0.15</td>
<td>0.21</td>
</tr>
<tr>
<td>Teeth completely resorbed</td>
<td>2.43</td>
<td>0.43</td>
<td>.00*</td>
<td>2.81</td>
<td>0.56</td>
<td>0.04*</td>
</tr>
</tbody>
</table>
Table 4. Calculations of phylogenetic signal in morphological characters based on total length (mm) at which milestones first occur.

<table>
<thead>
<tr>
<th>TRAITS</th>
<th>Moran's I</th>
<th></th>
<th>Abouheif's Cmean</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I observed</td>
<td>I expected</td>
<td>p value</td>
<td>Cmean observed</td>
</tr>
<tr>
<td>1st appearance of mandibular arch elements</td>
<td>-0.42</td>
<td>-0.48</td>
<td>0.53</td>
<td>-0.17</td>
</tr>
<tr>
<td>1st appearance of teeth</td>
<td>-0.29</td>
<td>-0.13</td>
<td>0.35</td>
<td>-0.04</td>
</tr>
<tr>
<td>Yolk sac completely resorbed</td>
<td>-0.33</td>
<td>-0.27</td>
<td>0.47</td>
<td>-0.04</td>
</tr>
<tr>
<td>Teeth completely developed</td>
<td>-0.17</td>
<td>0.36</td>
<td>0.33</td>
<td>0.04</td>
</tr>
<tr>
<td>Teeth completely resorbed</td>
<td>-0.34</td>
<td>-0.53</td>
<td>0.77</td>
<td>-0.01</td>
</tr>
</tbody>
</table>
Figure 1: Bayesian full mitogenome phylogeny. Posterior probability values are to the left of and below each node. Branch lengths indicate estimated amount of evolutionary (or genetic) change. Long branches (pertaining to outgroups) have been broken to allow for better visualization of ingroup relationships.
<table>
<thead>
<tr>
<th>Lateral View of Head</th>
<th>Upper Jaw</th>
<th>Lower Jaw</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a.</td>
<td>1b.</td>
<td>1c.</td>
</tr>
<tr>
<td>13.56 mm TL</td>
<td>10 mm</td>
<td>10 mm</td>
</tr>
<tr>
<td>4 days</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.6 mm</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1a.</td>
<td>2b.</td>
<td>2c.</td>
</tr>
<tr>
<td>21.39 mm TL</td>
<td>10 mm</td>
<td>10 mm</td>
</tr>
<tr>
<td>10 days</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5 mm</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1a.</td>
<td>3b.</td>
<td>3c.</td>
</tr>
<tr>
<td>55 days</td>
<td>0.2 mm</td>
<td>2.0 mm</td>
</tr>
<tr>
<td>122.40 mm TL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.0 mm</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2: Example of cleared and stained specimens (*A. transmontanus*) illustrating various stages of tooth and mandibular arch elements. Boxes 1a, 1b, and 1c illustrate appearance of teeth and mandibular arch elements. Boxes 2a, b, and c show developing teeth. Boxes 3a, b, and c show a specimen with teeth completely resorbed.
Figure 3: Full mitogenome Bayesian phylogeny, pruned so that only Acipenser oxyrinchus, A. brevirostrum, A. transmontanus, A. medirostris, and A. fulvescens appear. This topology is also shown by single-character tree based on the age at which complete resorption of teeth except for sockets occur.
Figure 4: Full mitogenome Bayesian phylogeny, pruned so that only Acipenser oxyrinchus, A. brevirostrum, A. transmontanus, and A. medirostris appear.
Figure 5: Topology shown by single-character (yolk sac and teeth) trees based on age at which first appearance of teeth, complete development of teeth, and complete resorption of yolk sac occur.
Figure 6: Topology shown by single-character (jaws elements) tree based on age at which first appearance of mandibular arch elements occurs.
Figure 7: Topology shown by single-character (migration) tree based on age at which migration begins.
Figure 8: Topology shown by single-character (nocturnality) trees based on larval nocturnal behavior.
Figure 9: Topology shown by single-character (phototaxis) trees based on larval phototaxic behavior.
Full Mitogenome Five Sturgeon Species Phylogeny

-4.28
-14.58
-50.57

A. oxyrinchus

48.61
68.05

A. transmontanus

29.16
16.61

A. medirostris

0.94
20.33

A. fulvescens

-20.33
53.1
16.61

A. brevirostrum

Key:
Dashed lines indicate species included only in strictly morphological analyses.
Posterior probability values: italics to the left of each node
PIC values for traits:
Black- length at which teeth are completely developed
Red- age at which migration begins
Green- age at which migration ends

Figure 10: Phylogenetic Independent Contrast values used to calculate correlations between traits.
CONCLUSIONS
Evolutionary relationships of sturgeons have been contentious, as phylogenetic studies based on different datasets have resulted in varied hypotheses of sturgeon evolution. Although general trends hold from study to study, specific relationships, namely the position of the genus *Scaphirhynchus* and the composition of subfamilies, have remained questionable. I used the most comprehensive dataset available thus far, over 14,000 base pairs of the sturgeon mitogenome for 22 species and subspecies, to produce a new sturgeon phylogenetic hypothesis. This hypothesis indicates, with strong support, a basal position for *Scaphirhynchus* and proposes the subfamily Husinae to be monotypic, including just *H. huso*.

My new phylogenetic hypothesis provided a framework for the exploration of phylogenetic signal among individual mitochondrial genes and gene families. Recent studies of the evolutionary relationships of various taxa (sturgeons, fungi, dolphins) have investigated signal carried by genes and gene families in an attempt to identify which genes are most useful to phylogenetic analyses. I sought to determine which genes most strongly influence phylogenetic signal in the full mitochondrial genome of sturgeons, and how this has impacted previous phylogenetic studies. I found that the protein coding genes and the rRNA 16s have signal similar to that of the full mitogenome, while D-loop, 12s rRNA, and the tRNAs have signal less similar to that of the full mitogenome. Previous phylogenetic analyses of sturgeons have commonly relied heavily upon the protein coding genes, the rRNAs, and D-loop. The fact that some of these regions (D-loop and 12s) have topologies that conflict with that of the full mitogenome while others have signal matching it 1) illustrates that different signal is carried by different genes and 2) indicates why studies using different genes have proposed many different phylogenetic hypotheses for sturgeons.
Sturgeon evolutionary relationships proposed in the full mitogenome phylogeny provided insight into the relationships among regions inhabited by these fishes. A clear division between Pacific and Atlantic regions was supported, and patterns of relationships within each of those groupings were further supported by the exploration of geologic history. These area relationships were further supported by those proposed by past biogeographic studies of other taxa. These results may be added to by future analyses incorporating additional sturgeon species and other taxa that inhabit similar ranges.

Finally, morphological and behavioral ontogenetic characters present in sturgeons were explored. A link between the age at which migration ends and the length at which teeth are completely developed supported the often-made statement that “form and function are inextricably linked”. Although links between other morphological-behavioral characters were not found, they may be present but undetectable. Additionally, ontogenetic characters were evaluated for phylogenetic signal, with signal matching the full mitogenome identified in one character - the age at which the resorption of teeth is complete. Many factors contribute to the timing of development in fishes - from temperature and food availability to predation risk. Ontogenetic characters for which signal was not found may be more influenced by these or other environmental factors, they may be present but not detected, or signal differing from that of the full mitogenome might be present.

This new, well-supported sturgeon phylogenetic hypothesis provided a framework for many analyses, from exploration of phylogenetic signal to the investigation of relationships among areas sturgeons inhabit. These analyses provide new information not only about the interrelationships of sturgeons, but about the type of data that may be useful in phylogenetic
analyses. My findings may be useful in future studies of sturgeons, including those on biogeography and early life history, and the data used may be incorporated into future phylogenetic analyses. The utility of this study does not end here. As imperiled, ancient fishes, sturgeons are frequently the subject of conservation efforts. The information revealed by my research may be useful to these efforts. For example, sturgeons are more and more frequently raised in aquaculture facilities, and hybridization between captive-raised escaped fishes has been identified as a threat to wild populations (Ludwig et al. 2009). Understanding species relationships and habitat needs of sturgeon species may help to identify where hybridization between species is likely and to prevent or mitigate this threat (e.g. identify if escaped fish would be able to 1: survive in habitat available and 2: hybridize with wild sturgeon(s) present). This study may further aid in conservation efforts by providing information that will help prioritize species or geographic areas for conservation (e.g. 1: the evolutionarily “distinctiveness” of species and 2: regions with the highest overall diversity of sturgeons). Ultimately, the data, analyses, and interpretations presented here have added to our knowledge of sturgeons, and it is my hope that these data and my findings will be useful in future research and in the conservation of these remarkable fishes.

References

Appendix 1. Alligned data matrix used in phylogenetic analyses presented in Chapter 1
<table>
<thead>
<tr>
<th>String 1</th>
<th>String 2</th>
<th>String 3</th>
<th>String 4</th>
<th>String 5</th>
<th>String 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCAATACCTCA ATTATCTTCG</td>
<td>CACCCGTAGGC</td>
<td>CCGCTACTCTA ATTATCCTTG</td>
<td>TATCATCCTG</td>
<td>ATTCCTGTTCT</td>
<td>ATTCCTGTTCT</td>
</tr>
<tr>
<td>CCAATACCTCA ATTATCTTCG</td>
<td>CACCCGTAGGC</td>
<td>CCGCTACTCTA ATTATCCTTG</td>
<td>TATCATCCTG</td>
<td>ATTCCTGTTCT</td>
<td>ATTCCTGTTCT</td>
</tr>
<tr>
<td>CCAATACCTCA ATTATCTTCG</td>
<td>CACCCGTAGGC</td>
<td>CCGCTACTCTA ATTATCCTTG</td>
<td>TATCATCCTG</td>
<td>ATTCCTGTTCT</td>
<td>ATTCCTGTTCT</td>
</tr>
<tr>
<td>CCAATACCTCA ATTATCTTCG</td>
<td>CACCCGTAGGC</td>
<td>CCGCTACTCTA ATTATCCTTG</td>
<td>TATCATCCTG</td>
<td>ATTCCTGTTCT</td>
<td>ATTCCTGTTCT</td>
</tr>
<tr>
<td>CCAATACCTCA ATTATCTTCG</td>
<td>CACCCGTAGGC</td>
<td>CCGCTACTCTA ATTATCCTTG</td>
<td>TATCATCCTG</td>
<td>ATTCCTGTTCT</td>
<td>ATTCCTGTTCT</td>
</tr>
<tr>
<td>CCAATACCTCA ATTATCTTCG</td>
<td>CACCCGTAGGC</td>
<td>CCGCTACTCTA ATTATCCTTG</td>
<td>TATCATCCTG</td>
<td>ATTCCTGTTCT</td>
<td>ATTCCTGTTCT</td>
</tr>
<tr>
<td>CCAATACCTCA ATTATCTTCG</td>
<td>CACCCGTAGGC</td>
<td>CCGCTACTCTA ATTATCCTTG</td>
<td>TATCATCCTG</td>
<td>ATTCCTGTTCT</td>
<td>ATTCCTGTTCT</td>
</tr>
<tr>
<td>CCAATACCTCA ATTATCTTCG</td>
<td>CACCCGTAGGC</td>
<td>CCGCTACTCTA ATTATCCTTG</td>
<td>TATCATCCTG</td>
<td>ATTCCTGTTCT</td>
<td>ATTCCTGTTCT</td>
</tr>
<tr>
<td>CCAATACCTCA ATTATCTTCG</td>
<td>CACCCGTAGGC</td>
<td>CCGCTACTCTA ATTATCCTTG</td>
<td>TATCATCCTG</td>
<td>ATTCCTGTTCT</td>
<td>ATTCCTGTTCT</td>
</tr>
<tr>
<td>CCAATACCTCA ATTATCTTCG</td>
<td>CACCCGTAGGC</td>
<td>CCGCTACTCTA ATTATCCTTG</td>
<td>TATCATCCTG</td>
<td>ATTCCTGTTCT</td>
<td>ATTCCTGTTCT</td>
</tr>
<tr>
<td>CCAATACCTCA ATTATCTTCG</td>
<td>CACCCGTAGGC</td>
<td>CCGCTACTCTA ATTATCCTTG</td>
<td>TATCATCCTG</td>
<td>ATTCCTGTTCT</td>
<td>ATTCCTGTTCT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TGCCTCATTCT</th>
<th>TACATACCTA</th>
<th>CGGACACAT</th>
<th>TTATTAACCA</th>
<th>TTTTCTAAAAT</th>
<th>ATCTCCTAAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGCCTCATTCT</td>
<td>TACATACCTA</td>
<td>CGGACACAT</td>
<td>TTATTAACCA</td>
<td>TTTTCTAAAAT</td>
<td>ATCTCCTAAT</td>
</tr>
<tr>
<td>TGCCTCATTCT</td>
<td>TACATACCTA</td>
<td>CGGACACAT</td>
<td>TTATTAACCA</td>
<td>TTTTCTAAAAT</td>
<td>ATCTCCTAAT</td>
</tr>
<tr>
<td>TGCCTCATTCT</td>
<td>TACATACCTA</td>
<td>CGGACACAT</td>
<td>TTATTAACCA</td>
<td>TTTTCTAAAAT</td>
<td>ATCTCCTAAT</td>
</tr>
<tr>
<td>TGCCTCATTCT</td>
<td>TACATACCTA</td>
<td>CGGACACAT</td>
<td>TTATTAACCA</td>
<td>TTTTCTAAAAT</td>
<td>ATCTCCTAAT</td>
</tr>
<tr>
<td>TGCCTCATTCT</td>
<td>TACATACCTA</td>
<td>CGGACACAT</td>
<td>TTATTAACCA</td>
<td>TTTTCTAAAAT</td>
<td>ATCTCCTAAT</td>
</tr>
<tr>
<td>TGCCTCATTCT</td>
<td>TACATACCTA</td>
<td>CGGACACAT</td>
<td>TTATTAACCA</td>
<td>TTTTCTAAAAT</td>
<td>ATCTCCTAAT</td>
</tr>
<tr>
<td>TGCCTCATTCT</td>
<td>TACATACCTA</td>
<td>CGGACACAT</td>
<td>TTATTAACCA</td>
<td>TTTTCTAAAAT</td>
<td>ATCTCCTAAT</td>
</tr>
<tr>
<td>TGCCTCATTCT</td>
<td>TACATACCTA</td>
<td>CGGACACAT</td>
<td>TTATTAACCA</td>
<td>TTTTCTAAAAT</td>
<td>ATCTCCTAAT</td>
</tr>
<tr>
<td>TGCCTCATTCT</td>
<td>TACATACCTA</td>
<td>CGGACACAT</td>
<td>TTATTAACCA</td>
<td>TTTTCTAAAAT</td>
<td>ATCTCCTAAT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ATTCCTGATTCT</th>
<th>GCGGACATTTA</th>
<th>TCCCTAGTACG</th>
<th>TGCCAAGAT</th>
<th>ATATTTTCAAC</th>
<th>TCTTATTGAG</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATTCCTGATTCT</td>
<td>GCGGACATTTA</td>
<td>TCCCTAGTACG</td>
<td>TGCCAAGAT</td>
<td>ATATTTTCAAC</td>
<td>TCTTATTGAG</td>
</tr>
<tr>
<td>ATTCCTGATTCT</td>
<td>GCGGACATTTA</td>
<td>TCCCTAGTACG</td>
<td>TGCCAAGAT</td>
<td>ATATTTTCAAC</td>
<td>TCTTATTGAG</td>
</tr>
<tr>
<td>ATTCCTGATTCT</td>
<td>GCGGACATTTA</td>
<td>TCCCTAGTACG</td>
<td>TGCCAAGAT</td>
<td>ATATTTTCAAC</td>
<td>TCTTATTGAG</td>
</tr>
<tr>
<td>ATTCCTGATTCT</td>
<td>GCGGACATTTA</td>
<td>TCCCTAGTACG</td>
<td>TGCCAAGAT</td>
<td>ATATTTTCAAC</td>
<td>TCTTATTGAG</td>
</tr>
<tr>
<td>ATTCCTGATTCT</td>
<td>GCGGACATTTA</td>
<td>TCCCTAGTACG</td>
<td>TGCCAAGAT</td>
<td>ATATTTTCAAC</td>
<td>TCTTATTGAG</td>
</tr>
<tr>
<td>ATTCCTGATTCT</td>
<td>GCGGACATTTA</td>
<td>TCCCTAGTACG</td>
<td>TGCCAAGAT</td>
<td>ATATTTTCAAC</td>
<td>TCTTATTGAG</td>
</tr>
<tr>
<td>ATTCCTGATTCT</td>
<td>GCGGACATTTA</td>
<td>TCCCTAGTACG</td>
<td>TGCCAAGAT</td>
<td>ATATTTTCAAC</td>
<td>TCTTATTGAG</td>
</tr>
<tr>
<td>ATTCCTGATTCT</td>
<td>GCGGACATTTA</td>
<td>TCCCTAGTACG</td>
<td>TGCCAAGAT</td>
<td>ATATTTTCAAC</td>
<td>TCTTATTGAG</td>
</tr>
<tr>
<td>ATTCCTGATTCT</td>
<td>GCGGACATTTA</td>
<td>TCCCTAGTACG</td>
<td>TGCCAAGAT</td>
<td>ATATTTTCAAC</td>
<td>TCTTATTGAG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GGCTCAGTCTA</th>
<th>TACATACCTA</th>
<th>CCGACACAT</th>
<th>TTATTAACCA</th>
<th>TTTTCTAAAAT</th>
<th>ATCTCCTAAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>GGCTCAGTCTA</td>
<td>TACATACCTA</td>
<td>CCGACACAT</td>
<td>TTATTAACCA</td>
<td>TTTTCTAAAAT</td>
<td>ATCTCCTAAT</td>
</tr>
<tr>
<td>GGCTCAGTCTA</td>
<td>TACATACCTA</td>
<td>CCGACACAT</td>
<td>TTATTAACCA</td>
<td>TTTTCTAAAAT</td>
<td>ATCTCCTAAT</td>
</tr>
<tr>
<td>GGCTCAGTCTA</td>
<td>TACATACCTA</td>
<td>CCGACACAT</td>
<td>TTATTAACCA</td>
<td>TTTTCTAAAAT</td>
<td>ATCTCCTAAT</td>
</tr>
<tr>
<td>GGCTCAGTCTA</td>
<td>TACATACCTA</td>
<td>CCGACACAT</td>
<td>TTATTAACCA</td>
<td>TTTTCTAAAAT</td>
<td>ATCTCCTAAT</td>
</tr>
<tr>
<td>GGCTCAGTCTA</td>
<td>TACATACCTA</td>
<td>CCGACACAT</td>
<td>TTATTAACCA</td>
<td>TTTTCTAAAAT</td>
<td>ATCTCCTAAT</td>
</tr>
<tr>
<td>GGCTCAGTCTA</td>
<td>TACATACCTA</td>
<td>CCGACACAT</td>
<td>TTATTAACCA</td>
<td>TTTTCTAAAAT</td>
<td>ATCTCCTAAT</td>
</tr>
<tr>
<td>GGCTCAGTCTA</td>
<td>TACATACCTA</td>
<td>CCGACACAT</td>
<td>TTATTAACCA</td>
<td>TTTTCTAAAAT</td>
<td>ATCTCCTAAT</td>
</tr>
<tr>
<td>GGCTCAGTCTA</td>
<td>TACATACCTA</td>
<td>CCGACACAT</td>
<td>TTATTAACCA</td>
<td>TTTTCTAAAAT</td>
<td>ATCTCCTAAT</td>
</tr>
<tr>
<td>GGCTCAGTCTA</td>
<td>TACATACCTA</td>
<td>CCGACACAT</td>
<td>TTATTAACCA</td>
<td>TTTTCTAAAAT</td>
<td>ATCTCCTAAT</td>
</tr>
<tr>
<td>GGCTCAGTCTA</td>
<td>TACATACCTA</td>
<td>CCGACACAT</td>
<td>TTATTAACCA</td>
<td>TTTTCTAAAAT</td>
<td>ATCTCCTAAT</td>
</tr>
</tbody>
</table>