In: S. Doroshov, F. Binkowski, T. Thuemler and D. MacKinlay (eds.), *Culture and Management of Sturgeon and Paddlefish, Symposium Proceedings of the International Congress on the Biology of Fishes*, American Fisheries Society (1996)

MOLECULAR PHYLOGENY OF NORTH AMERICAN ACIPENSERIFORMES DERIVED FROM RIBOSOMAL RNA GENE SEQUENCES

Jeannette Krieger
The Ohio State University
Department of Molecular Genetics
484 W. 12th Avenue Columbus, OH 43210
Tel.: (614) 292-4570 Fax: (614) 292-4466 e-mail: krieger.15@osu.edu

Gregory C. Booton, Ted Cavender[1] and Paul A. Fuerst
The Ohio State University
Departments of Molecular Genetics and Zoology[1]

Introduction

The order Acipenseriformes is an ancient group of fish identifiable from the Jurassic (Grande and Bemis, 1991). The group is characterized by a highly cartilaginous skeleton, heterocercal tails and a well developed rostrum with mouth inferior (Sokolov and Berdichevskii, 1989). At present, it consists of approximately twenty-four Recent species of sturgeon and two species of paddlefish. These fish are diadromous or strictly freshwater, with a North American and Eurasian distribution. All species investigated cytologically appear to be polyploid, possessing large numbers of chromosomes. These fish appear to be separated into two groups, one possessing about 120 chromosomes and the other approximately 240. The largest number of chromosomes actually counted is 258 (Birstein and Vasiliev, 1987; Blacklidge and Bidwell, 1993; Dingerkus and Howell, 1976; Fontana, 1994). Sturgeon and paddlefish have long life spans, can grow to become very large in size and, in general, take a long time to reach sexual maturity; females often require ten to twenty years before reproduction in nature (Lauder and Liem, 1983). Exploited for caviar and meat, their popularity as a food source in a number of countries has resulted in frequent over fishing. Recently, the population sizes of many acipenseriform species have decreased to the point where they are threatened or endangered. The decrease can be traced to three major factors: 1) the life histories of these fish; 2) the destruction of spawning habitat by pollution, dams or other human intervention; and 3) over fishing.

Information about the evolutionary relationships of the extant species of sturgeon and paddlefish may be useful in directing conservation efforts towards protecting a maximum diversity of Acipenseriform species. Because these relationships are not well understood, and since the "living fossil" status of the group makes them important for understanding vertebrate evolution in general (Birstein, 1993), a molecular phylogenetic study employing ribosomal RNA (rRNA) genes is being carried out in our laboratory. We have determined sequences from the nuclear 18S rRNA and mitochondrial 12S rRNA genes in order to examine the relationships between nine North American sturgeon and paddlefish taxa: *Polyodon spathula* (North American paddlefish), *Acipenser fulvescens* (lake sturgeon), *Acipenser brevirostrum* (shortnose sturgeon), *Acipenser transmontanus* (white sturgeon), *Acipenser medirostris* (green sturgeon), *Acipenser oxyrhynchus oxyrhynchus* (Atlantic sturgeon), *Acipenser oxyrhynchus desotoi* (Gulf sturgeon), *Scaphirhynchus platorynchus* (shovelnose sturgeon) and *Scaphirhynchus albus* (pallid sturgeon).

Ribosomal DNA (rDNA) has been widely used in phylogenetic studies to resolve the evolutionary relationships between organisms at various levels of taxonomic relationship (Spears et al., 1992; Kranz et al., 1995). The rRNA genes were chosen for the analysis of the Acipenseriformes for several reasons. The 18S rRNA gene produces the cytoplasmic small subunit ribosomal RNA, a molecule with highly conserved function and structure, normally consisting of 1700-1800 nucleotides. This gene contains conserved regions of primary structure that can be used to align sequences from different taxa and which can be exploited to construct "universal" primers for

amplification by the Polymerase Chain Reaction (PCR) and for DNA sequencing (Hillis and Dixon, 1991). It also contains variable regions that are phylogenetically informative and may be useful for determining relationships between species. The 18S rRNA gene is found universally in eukaryotes along with the 28S (nuclear large subunit) rRNA gene and the 5.8S rRNA gene in a single transcription unit (Figure 1). The ribosomal DNA transcription unit can be found on various chromosomes in different organisms as a tandem repeated sequence. There may be hundreds or even thousands of copies of the rDNA transcription unit in a particular species' individual cells (Appels et al., 1980), but there may also be as few as one copy (Yao and Gall, 1977). The number of copies found in the genomes of sturgeons is at present unknown. Despite the large number of rDNA repeats in an array, the process of concerted evolution is thought to rapidly homogenize all arrays of a particular repetitive sequence within an individual and a species (Arnheim et al., 1983). Such a homogenization process appears to have acted on the rRNA genes, whose tandem repeats form a type of repetitive sequence, in almost all species examined to date. Consequently, it is normally sufficient to sequence a single copy of the 18S rRNA gene in of a species for analysis in molecular phylogenetic studies. We have discovered, however, that this may not be true of the rRNA genes of the sturgeons. Finally, to aid in phylogenetic reconstruction, 18S rRNA gene sequences from several other fish are also available for comparison.

The second gene used in this study, the mitochondrial 12S rRNA gene, also codes for a small subunit ribosomal RNA, one which is restricted to the mitochondrion. The gene is smaller than the nuclear 18S rRNA, being 950 base pairs long. Because it is a mitochondrial gene, it is maternally inherited, and the equivalent of a single 12S rRNA gene sequence should normally be present in an individual, even if the organism is polyploid. In animals, mitochondrial DNA evolves more rapidly than nuclear genes, possibly due to decreased accuracy of the mitochondrial DNA polymerase and less efficient DNA repair mechanisms (Brown et al., 1979; Clayton et al., 1974). PCR primers have been designed which are complementary to transfer RNA gene sequences which flank the genetic region containing the mitochondrial control region and 12S rRNA gene sequences. These primers can be used to amplify the entire 12S rRNA gene from all the North American sturgeon and paddlefish. Although the data base of sequences of the 12S rRNA gene is limited, these other properties make the gene another good candidate to help clarify acipenseriform evolution.

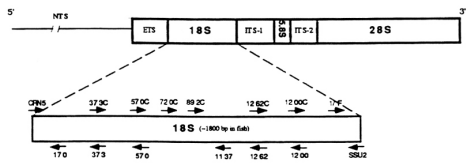

Figure 1. Diagram of the nuclear rDNA array of eukaryotes. The figure represents a single transcription unit of rDNA that is found in tandem repeats on various homologous and/or non-homologous chromosomes of an organism. (NTS = non-transcribed spacer; ETS = external transcribed spacer; ITS = internal transcribed spacer.) The detail of the small subunit (18S) rRNA gene shows the location of PCR primers (and their orientations) used in this study for amplification and sequencing of the gene.

Methods

Tissue samples were obtained from various researchers, fisheries units or from aquarium retailers. Samples included those of the nine sturgeon and paddlefish taxa listed previously, and of several additional species important for determining the position of the Acipenseriformes relative to other fish groups. These other species include *Protopterus annectens* (African lungfish), *Amia calva*

(bowfin), *Lepisosteus osseus* (longnose gar), and three species of polypterids, *Polypterus senegalus*, *Polypterus delhezi*, and *Erpetoichthys calabaricus*. Tissue samples were in the form of muscle, blood, barbels or fin snips, with fin snips being the preferred type because they are the least intrusive to the animals. Genomic DNA was extracted from the tissues using a digestion with proteinase K in ABI lysis buffer followed by phenol-chloroform extraction and ethanol precipitation. DNA was resuspended in TE buffer and quantified using a spectrophotometer. PCR amplification (Saiki et al., 1985) of 18S rRNA or 12S rRNA genes from genomic DNA was carried out with primers designed in our laboratory (Figure 1). Two identical PCR reaction products were pooled for each gene to compensate for possible errors caused by the Taq polymerase that may occur during amplification. DNA sequencing of PCR products was performed either by the direct cycle sequencing (BRL) of the product, or by first cloning the product into a bacterial plasmid using the TA cloning protocol (Invitrogen) followed by cycle sequencing of the clones.

Entire 18S rRNA gene sequences (~1800 base pairs) were determined for *P. delhezi*, *P. senegalus*, *E. calabaricus*, *P. annectens*, *A. calva*, *L. osseus and Polyodon spathula*. Only partial sequences of the gene have been completed for the eight sturgeon species, because of the unexpected discovery of intraindividual variation. The 18S rRNA gene sequences were aligned with other fish sequences which had been completed previously in our laboratory or which were retrieved from the GENBANK database or from other publications. Alignment was carried out with the computer program ESEE (Cabot and Beckenbach, 1989), with the aid of a secondary structure of *P. spathula* small subunit rRNA that was constructed based on accepted secondary structures (Gutell, 1994). Partial 12S rRNA gene sequences were determined for the eight acipenseriform species as well as the three polypterid species, which were used as the outgroup. The partial sequence from the twelve species was aligned using ESEE. Phylogenetic analyses were conducted using MEGA (Kumar et al., 1993) for a neighbor joining analysis and PAUP (Swofford, 1990) for parsimony analyses.

Results

The eight sturgeon taxa were found to possess substantial intraindividual variation in the 18S rRNA gene. This was unexpected, and represents one of the first examples of widespread intraindividual heterogeneity of rRNA gene sequences in any eukaryote. As a result of this finding, it will be necessary to clone variants which are present in each species before sequences can be correctly determined for analysis. It is not clear whether intraindividual heterogeneity will have any effect on estimates of interspecific differentiation. For only one species, the white sturgeon *A. transmontanus*, have we cloned multiple sequences to date, and clones currently involve only the 5' half of the gene (~1200 base pairs). Ten clones have been isolated and sequenced. They represent six different sequences of four distinct types.

Even though all the species of acipenseriforms could not yet be compared using the 18S rRNA gene, the placement of the Acipenseriformes with respect to other orders of fish can be examined. The paddlefish was found to contain only a single 18S rRNA gene sequence. We also used the cloned sequences of the white sturgeon. These cloned sequences were aligned with the 5' segments of the 18S rRNA gene sequences from other fish species and a consensus neighbor-joining tree was constructed from the data using MEGA. This tree is shown in Figure 2. In addition, a consensus parsimony tree was constructed using the same data using the computer program PAUP for comparison (tree not shown). In both cases, *Styela* (a tunicate) was used as the outgroup. All the non-sturgeon species of fish mentioned above could be included in this study. These sequences could be obtained without cloning, because the species possess only one 18S rRNA gene sequence.

Partial mitochondrial 12S rRNA sequences (~250 base pairs) have been obtained and aligned from the acipenseriform species and three polypterid species. As mentioned previously, only one sequence was found to be present for this gene in all the species examined. This allowed a preliminary comparison of the sequences from all the North American sturgeon and paddlefish in an attempt to clarify species relationships within the Acipenseriformes. Shown in Figure 3 is the consensus neighbor-joining tree constructed using MEGA. PAUP was used to construct a consensus parsimony tree based on the same data for comparison. In these analyses, the polypterids were used as the outgroup.

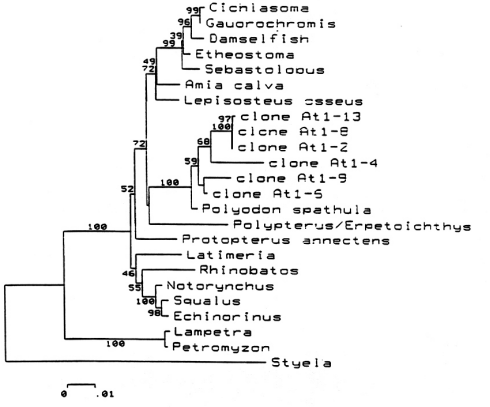

Figure 2. Consensus neighbor-joining tree based on partial (~1200 bp) nuclear 18S rRNA gene sequence including *A. transmontanus* (At1) clones (1000 bootstraps). A parsimony analysis of the same data using PAUP produced a consensus of the four most parsimonious trees which shows the same topology except for the placement of *Latimeria*. Numbers on branches represent bootstrap values.

Discussion

In Figure 2, the consensus neighbor-joining tree demonstrates relationships between the major orders of fish based on ~1200 bases of the 18S rRNA gene sequences. In general, this tree agrees with relationships recently proposed based on morphological data. The lampreys are basal to the entire group, followed by the cartilaginous sharks. The African lungfish is placed outside the Actinopterygii (teleosts, bowfin, gar, Acipenseriforms and polypterids). The gar and bowfin are clustered with the teleosts, with the bowfin slightly closer to the teleosts than the gar. Important for the purposes of this paper, the North American paddlefish clusters with the *A. transmontanus* (At1) clones, as expected since the families Polyodontidae and Acipenseridae are grouped together within the Acipenseriformes.

Variability within *A. transmontanus* clones does not affect their relationship with other species, since the six clones cluster together, separated into four different sequence types. As mentioned, the presence of intraindividual heterogeneity of the 18S rRNA gene in sturgeons is unusual. The sequence of the rDNA region appears to be homogenized to a single sequence type in almost all organisms previously examined, presumably by the process of concerted evolution. One hypothesis to explain the apparent difference between sturgeons and other eukaryotes is that the

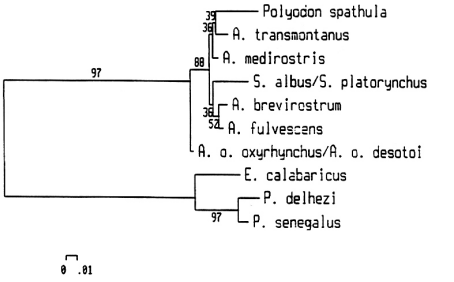

Figure 3. Consensus neighbor-joining tree based on partial (~250 bp) mitochondrial 12S rRNA gene sequence. In a parsimony analysis of the same data, a consensus of the nine most parsimonious trees shows the same basic topology but fails to resolve the branching order for *P. spathula*, *A. transmontanus*, *A. medirostris*, and the *Scaphirhynchus* species.

heterogeneity in these fish results from the high ploidy levels of sturgeon. One point that needs to be stressed here is that although ploidy level may be the cause of the variation, ploidy level does not necessarily always lead to intraindividual heterogeneity. The North American paddlefish is considered to be a tetraploid, according to results of flow cytometry and karyotypic studies (Blacklidge and Bidwell, 1993; Dingerkus and Howell, 1976), but it was not found to possess intraindividual variation of the 18S rRNA gene.

There are two major differences between the 18S rRNA gene tree and some current hypotheses about the evolutionary affinities of the Acipenseriformes. First, the polypterids were found to cluster (although weakly) with the Acipenseriformes, together with which they form a sister group to the clade which contains the teleosts, the bowfin and the gar. This branching pattern occurs in both neighbor-joining and parsimony analyses of the data. Our results support the view that the Polypteridae and the Acipenseriformes should be grouped together within the Chondrostei as part of the Actinopterygii. In other words, the polypterids do not belong in their own subclass. An additional interesting point is that identical 18S rRNA gene sequences were obtained for the three polypterids used in this study, so that the species level relationships could not be resolved with this gene. Second, analysis of the 18S rRNA gene does not result in a clustering of the coelacanth (*Latimeria chalumnae*) with the lungfish. Rather, *Latimeria* clusters with the sharks. However, in the parsimony tree, the coelacanth is shown branching off the line leading to the Actinopterygians before the lungfish, and after the divergence of the Chondrichthyes. This latter tree agrees with most morphological based classifications that place the lungfishes and coelacanth together in the subclass Sarcopterygii. Conflicting results using the two reconstruction methods do not allow us to support or refute the current placement of *Latimeria*.

Figure 3 presents the consensus neighbor-joining tree based on partial mitochondrial 12S rRNA sequences of North American acipenseriforms and three polypterids. It is evident from this tree

that 12S rRNA gene sequences are sufficiently divergent to distinguish between the polypterid genera and species, which could not be separated using the larger sequences of the 18S rRNA gene. The tree should show even better resolution as additional sequence information from the entire 12S rRNA gene is added. As regards the relationships between sturgeon species illustrated here, the short branch lengths indicate that these species are very closely related. In fact, the two *Scaphirhynchus* species have identical nucleotide sequences for this portion of the gene. This is also true for the Atlantic and Gulf sturgeons (*A. o. oxyrhynchus* and *A. o. desotoi*). The grouping of some *Acipenser* species in this reconstruction appears to be related to their species ranges. The two Eastern American species, *A. fulvescens* and *A. brevirostrum*, cluster and the two Northern Pacific species, *A. transmontanus* and *A. medirostris*, are shown together as well. The Atlantic *A. oxyrhynchus* species is shown basal to the whole group. The relationships between the *Acipenser* species in this tree correlate well with those proposed in a cladogram published in the April 1995 issue of The Sturgeon Quarterly by Evgenii N. Artyukhin in his article "On Biogeography and Relationships Within the Genus *Acipenser*". Artyukhin's cladogram shows the same branching pattern, indicating the same relationships between these six species as shown in our tree. However, the placement of the genera *Scaphirhynchus* and *Polyodon* in the 12S rRNA gene tree is not that predicted by the analysis of morphological data. The *Scaphirhynchus* species are shown within the genus *Acipenser*. These two species are currently classified in two different subfamilies. However, the genus *Scaphirhynchus* is pleisiomorphic, possessing ancestral characters. This would be consistent with the location of these species within the *Acipenser* group as shown in our molecular analysis. *Polyodon*, which is considered a derived form of paddlefish by Grande and Bemis (1991), is shown in the 12S rRNA tree as a derived species within the Acipenseridae. More information is needed to resolve such apparent conflicts between our molecular data and morphological data.

Future work to clarify the relationships between the apparently closely related sturgeon and paddlefish species will concentrate on cloning and sequencing the 18S rRNA genes from the North American sturgeon species with intraindividual variation, completing the 12S rRNA gene sequences for the acipenseriform species as well as the polypterids, and for some other fish, and investigating other potentially more rapidly evolving genes for suitability for use in this phylogenetic study.

Acknowledgments

We would like to thank those who provided the fish tissue samples necessary for this study, including Frank Chapman from the University of Florida, Dave Davies of the Ohio Department of Natural Resources, Jerry Hamilton from the Blind Pony Lake Conservation Area, Ed Little from the National Biological Services Midwest Science Center, Jerre Mohler from the USFWS Northeast Fishery Center, Tim Mulligan from Humboldt State University, Terry Patterson from the College of Southern Idaho, Brady Porter from The Ohio State University and Kent Ware from the Bears Bluff National Fish Hatchery.

This work was partially supported by a grant from the Ohio Sea Grant College Program/NOAA, and by cooperation from the Ohio Department of Natural Resources Division of Wildlife.

References

Appels, R., W.L. Gerlach, E.S. Dennis, H. Swift, and W.J. Peacock. (1980). Molecular and chromosomal organization of DNA sequences coding for the ribosomal RNAs in cereals. Chromosoma, 78: 293-311.

Arnheim, N. (1983). Concerted evolution of multigene families. In: *Evolution of Genes and Proteins*. M. Nei and R.K. Koehn, eds. Sinauer Press, Sunderland, Mass.

Birstein, V.J. and V.P. Vasiliev. (1987). Tetraploid-octoploid relationships and karyological evolution in the order Acipenseriformes (Pisces): Karyotypes, nucleoli, and nucleolus-organizer regions in four acipenserid species. Genetica, 72: 3-12.

Birstein, V.J. (1993). Sturgeons and paddlefishes: Threatened fishes in need of conservation. Conservation Biology, 7: 773-787.

Blacklidge, K.H. and C.A. Bidwell. (1993). Three ploidy levels indicated by genome quantification in Acipenseriformes of North America. Journal of Heredity, 84: 427-430.

Brown, W.M., M.J. George and A.C. Wilson. (1979). Rapid evolution of animal mitochondrial DNA. Proc. Natl. Acad. Sci. USA, 76: 1967-1971.

Cabot, E.L. and A.T. Beckenbach. (1989). Simultaneous editing of multiple nucleic acid and protein sequences with ESEE. Comput. Appl. Biosci., 5: 233-234.

Clayton, D.A., J.N. Doda and E.C. Friedberg. (1974). The absence of a pyrimidine dimer repair mechanism in mammalian mitochondria. Proc. Natl. Acad. Sci. USA, 71: 2777-2781.

Dingerkus, G. and W.M. Howell. (1976). Karyotypic analysis and evidence of tetraploidy in the North American paddlefish, *Polyodon spathula*. Science, 194: 842-843.

Fontana, F. (1994). Chromosomal nucleolar organizer regions in four sturgeon species as markers of karyotype evolution in Acipenseriformes (Pisces). Genome, 37: 888-892.

Grande, L. and W.E. Bemis. (1991). Osteology and phylogenetic relationships of fossil and Recent paddlefishes (Polyodontidae) with comments on the interrelationships of Acipenseriformes. Journal of Vertebrate Paleontology, 11(suppl. 1): 1-121.

Gutell, R.R. (1994). Collection of small subunit (16S- and 16S-like) ribosomal RNA structures: 1994. Nucleic Acids Research, 22(17): 3502-3507.

Hillis, D.M. and M.T. Dixon. (1991). Ribosomal DNA: Molecular evolution and phylogenetic inference. The Quarterly Review of Biology, 66(4): 411-453.

Kranz, H.D., D. Miks, M.-L. Siegler, I. Capesius, C.W. Sensen, and V.A.R. Huss. (1995). The origin of land plants: Phylogenetic relationships among Charophytes, Bryophytes, and vascular plants inferred from complete small-subunit ribosomal RNA gene sequences. Journal of Molecular Evolution, 41: 74-84.

Kumar, S., K. Tamura and M. Nei. (1993). MEGA: (Molecular Evolutionary Genetics Analysis), version 1.01. University of Molecular Evolutionary Genetics, The Pennsylvania State University, University Park, PA 16802.

Lauder, G.V., and K.F. Liem. (1983). The evolution and interrelationships of the Actinopterygian fishes. *Bull. Mus. Comp. Zool.*, 150(3): 95-197.

Saiki, R.K., S. Scharf, F. Faloona, K. Mullis, G. Horn, H.A. Erlich and N. Arnheim. (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science, 230: 1350-1354.

Sokolov, L.I. and L.S. Berdichevskii. (1989). B. Systematic Part: Acipenseriformes Berg, 1940. In: *The Freshwater FIshes of Europe: Volume I, Part II*. J. Holcik, editor. AULA-Verlag GmbH, Wiesbaden.

Spears, T., L.G. Abele, and W. Kim. (1992). The monophyly of brachyuran crabs: A phylogenetic study based on 18S rRNA. Systematic Biology, 41(4): 446-461.

Swofford, D.L. (1990) PAUP- Phylogenetic Analysis Using Parsimony (Version 3.0) Computer program distributed by the Illinois Natural History Survey, Champaign, Illinois.

Yao, M.-C., and J.G. Gall. (1977). A single integrated gene for ribosomal RNA in a eucaryote, *Tetrahymena pyriformis*. Cell, 12: 121-132.