J. Appl. Ichthyol. 18 (2002), 290–297

ISSN 0175–8659

Received: July 13, 2001
Accepted: May 29, 2002

Evidence of multiple alleles of the nuclear 18S ribosomal RNA gene in sturgeon (Family: Acipenseridae)

By J. Krieger[1] and P. A. Fuerst[1,2]

[1]*Department of Molecular Genetics and* [2]*Department of Evolution, Ecology and Organismal Biology, The Ohio State University, Columbus, OH, USA*

Summary

Nuclear 18S ribosomal RNA (rRNA) gene sequences were determined during a phylogenetic study of the evolutionary relationships of the North American sturgeon and paddlefish. In the course of this study, intra-individual variation of the 18S rRNA gene was discovered in all nine species of sturgeon examined, while no variation was detected for this gene in the North American paddlefish (*Polyodon spathula*) or in other non-acipenseriform fish species included in the study (bichirs, reedfish, gar, bowfin and lungfish). The discovery of multiple sequence variants of the 18S rRNA gene in an individual is unexpected; although most eukaryotes possess many copies of the rDNA unit in their genomes, the copies are normally thought to be homogenized within an individual as well as in a species by concerted evolution. The effect of the variation on the determination of species level relationships within the order and of the relationship of the order to other fish groups is examined. The actual cause(s) for the presence of intra-individual variation of the 18S rRNA gene are as yet unknown, however, it is possible that the origin and maintenance of multiple 18S rDNA sequence variants in sturgeon may be related to the polyploid status of the sturgeon genome.

Introduction

Here we report the first observation of extensive intra-individual variation of the DNA sequences of the 18S rRNA gene in any vertebrate. As an extension of a molecular phylogenetic study of the North American sturgeon and paddlefish (Krieger et al. 2000), sequences of the nuclear 18S rRNA gene were determined. The 18S rRNA gene has been used widely for phylogenetic studies of many organisms (for example, Stock et al. 1991; Stock and Whitt 1992; Wada 1998). Even though the rRNA genes exist in multiple copies within all vertebrate genomes, sequence homogeneity, or near homogeneity, within species has been the rule and a single 18S rRNA gene sequence is usually used to represent a species in molecular phylogenetic studies (Hillis and Davis 1988). The conserved function and structure of the 18S rRNA molecule allow sequences to be aligned, even among divergent species. However, the molecule also possesses phylogenetically informative variable regions that are useful for determining relationships among species (Hillis and Dixon 1991). In eukaryotes, the ribosomal RNA transcription unit usually exists in tandemly repeated arrays, containing three rRNA genes, 18S, 5.8S and 28S RNA genes, with the genes separated by spacer sequences (ITSs) (Fig. 1). Nontranscribed spacers (NTS) separate tandem copies of the entire transcribed unit. The number of copies of the rRNA gene array is species specific; there may be hundreds or thousands of copies of the transcription unit in a particular species (Appels et al. 1980), but there may also be as few as one copy, as in *Tetrahymena* (Yao and Gall 1977). In some species, copies of the rRNA gene array are found on more than one pair of chromosomes.

Because of the large number of copies found in most genomes, various rRNA gene repeats in a chromosome or genome may at first be thought to evolve independently, even though selection acts upon rRNA to retain essential and conserved functions in protein synthesis. Previous studies of rDNA in several organisms have shown, however, that the gene repeats in an array do not evolve independently (Arnheim 1983). The 18S, 28S and 5.8S rRNA genes normally show high, usually complete, homogeneity of sequence among repeats within an organism and even within a species, although differences in gene sequences appear when species are compared. The process of homogenization of repetitive sequences is termed concerted evolution (Arnheim et al. 1980), and may be caused by unequal crossing over, replicative transposition, gene amplification or gene conversion (Dover et al. 1982). The homogenization of rRNA gene repeats within individuals and species by concerted evolution ordinarily allows the species to be represented by a single 18S rRNA gene sequence in molecular phylogenetic studies (Hillis and Davis 1988).

With the assumption that the 18S rRNA genes of acipenseriforms would show no unusual characteristics, we obtained 18S rRNA gene sequences from the North American paddlefish and nine species of North American sturgeon for the purpose of a phylogenetic study. A single homogenous sequence was obtained for the gene in paddlefish, as generally expected, and as seen in non-acipenseriform fish studied in our laboratory and by other researchers. The results, however, were far from expected when sturgeon were studied. Multiple sequence variants were unexpectedly detected for the 18S rRNA genes of all nine species of sturgeon. These findings represent the first discovery of extensive intra-individual variation of the 18S rRNA gene in any vertebrate, and we describe here the relevance of this discovery to molecular phylogenetic analyses and our assumptions about the evolution of repetitive DNA sequences.

Materials and methods

Specimens

Nuclear 18S rRNA gene sequences were examined in one individual for each of 10 acipenseriform taxa: *Polyodon spathula*, *Scaphirhynchus albus*, *S. platorynchus*, *S. suttkusi*,

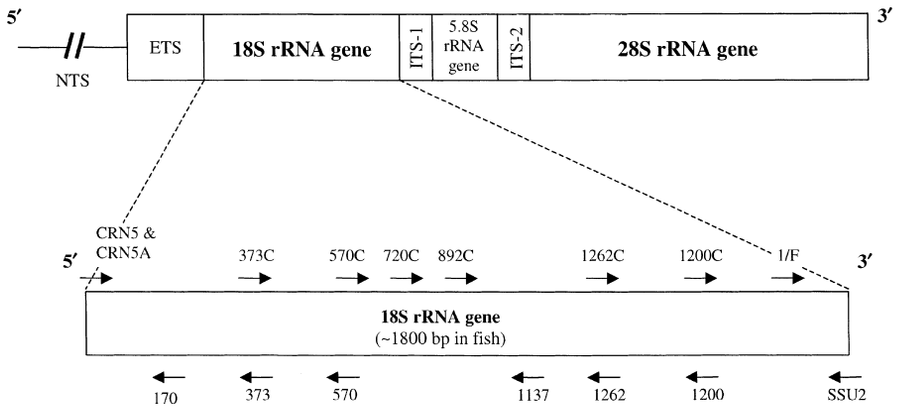

Fig. 1. Diagram of the rDNA array of eukaryotes. The above represents a single transcription unit of rDNA that is found in tandem repeats on various homologous and/or non-homologous chromosomes of an organism. (NTS = non-transcribed spacer; ETS = external transcribed spacer; ITS = internal transcribed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

Acipenser brevirostrum, *A. fulvescens*, *A. medirostris*, *A. oxyrinchus oxyrinchus*, *A. o. desotoi,* and *A. transmontanus*. Sequences were also determined for one individual each of *Polypterus senegalus* and *P. delhezi* (bichirs), *Erpetoichthys calabaricus* (reedfish), *Lepisosteus osseus* (longnose gar), *Amia calva* (bowfin) and a member of the lungfish genus *Protopterus*. Tissue samples (fin snips, muscle, blood or barbels) were obtained from researchers and fisheries in the US or collected from fish purchased at a local aquarium store as described in Table 1.

DNA extraction, PCR amplification, cloning and DNA sequencing

Detailed molecular protocols are given in Krieger, 2000. After tissue homogenization and digestion, DNA extraction was carried out following the standard phenol-chloroform method (Maniatis et al., 1982). The Polymerase chain reaction (PCR) amplification (Saiki et al., 1988) of nuclear 18S rRNA genes from total DNA was carried out in 100 μl reactions, using either *Taq* DNA polymerase or Elongase® enzyme mix (Gibco-BRL, Gaithersburg, MD). *Taq* DNA polymerase was used for PCR amplifications from all the non-acipenseriform species and preliminary amplifications from the acipenseriform species. Elongase® Enzyme Mix was used for subsequent amplification of sturgeon and paddlefish genes because the lower error rate of Elongase® reduced the possibility that sequence variation was a result of DNA polymerase errors during amplification. Details of amplification, amplification primers and sequencing primers are provided in Krieger 2000. The total size of the PCR gene products analyzed was approximately 1800 base pairs (bp) in all fish species examined. Cloning of the entire 18S rRNA gene PCR product amplified with the Elongase® enzyme mix was carried out on lake sturgeon (*Acipenser fulvescens*), shortnose sturgeon (*A. brevirostrum*) and North American paddlefish (*Polyodon spathula*) using the TOPO® TA PCR XL cloning kit (Invitrogen, Carlsbad, CA) (Krieger, 2000). Clones were assigned names based on the species plus a number (for example, 18S rDNA clones from lake sturgeon, *A. fulvescens*, were labeled Af-1, Af-2,..., while clones from shortnose sturgeon, *A. brevirostrum*, clones were labeled Ab-1, etc.).

Gene sequences from PCR products and clones were determined (in most cases for both strands) by manual sequencing using the dsCycle Sequencing System Kit (Gibco-BRL, Gaithersburg, MD) with ^{32}P end-labeled sequencing primers. Sequencing products were separated on 6% polyacrylamide gels. Sequences were deposited in GenBank under

Table 1
Species examined in this study

Species name	Tissue sample	Origin of specimen	Collector/Provider of sample
Polyodon spathula	Fin snip	Hatchery fish–Parents originally from Osage River Basin	Blind Pony Lake Conservation Area, Missouri
Scaphirhynchus albus	Fin snip	Hatchery fish–Parents from Mississippi and Missouri rivers [a]	National Biological Services Midwest Science Center, Missouri
Scaphirhynchus platorynchus	Fin snip	Missouri River	National Biological Services Midwest Science Center, Missouri
Scaphirhynchus suttkusi	Muscle	Alabama River (UAIC# 1885.01)	University of Alabama at Tuscaloosa
Acipenser brevirostrum	Barbel	Hatchery fish–Parents from Savannah River	Bears Bluff National Fish Hatchery, South Carolina
Acipenser fulvescens	Fin snip	Wolf River	Wisconsin Department of Natural Resources
Acipenser medirostris	Muscle	Klamath River	Humboldt State University, California
Acipenser oxyrinchus oxyrinchus	Fin snip	Hudson River	Northeast Fishery Center, U.S.F.W.S., Pennsylvania
Acipenser oxyrinchus desotoi	Blood	Hatchery fish–Parents from Suwannee River	University of Florida
Acipenser transmontanus	Fin snip	Snake River	College of Southern Idaho
Lepisosteus osseus	Muscle	Big Darby Creek	The Ohio State University
Amia calva	Fin snip	Pet store	The Ohio State University
Polypterus sengalus	Fin snip	Pet store	The Ohio State University
Polypterus delhezi	Fin snip	Pet store	The Ohio State University
Erpetoichthys calabaricus	Fin snip	Pet store	The Ohio State University
Protopterus sp.	Fin snip	Pet store	The Ohio State University

[a] Male parents were from the Missouri River and female parents were from the Mississippi River.

the following accession numbers: AF188367 through AF188400, AF198114 and AF198115.

Sequence alignment and phylogenetic analyses

DNA sequences were aligned using ESEE (Cabot and Beckenbach 1989), considering information on secondary structure based on currently accepted models (Gutell 1994). One highly variable loop region that could not be reliably aligned (because of differences in length) was omitted from the phylogenetic analyses. The alignment is available from the authors upon request. The 18S rRNA sequence of the tunicate *Styela plicata* (GenBank accession no. M97577) was used as an outgroup for phylogenetic analyses. Also included in the phylogenetic analyses were the non-acipenseriform fish species sequenced for this study, and additional taxa obtained from GenBank: *Petromyzon marinus* (M97575), *Squalus acanthias* (M91179), *Notorhynchus cepedianus* (M91183), *Latimeria chalumnae* (L11288), *Salmo trutta* (X98839), and *Oreochromis esculentus* (AF337051). Neighbor joining analyses were performed in MEGA version 1.01 (Kumar et al., 1993) using the Kimura two-parameter model including both transitions and transversions (Kimura, 1980). Parsimony analyses were performed using DNAPARS in PHYLIP 3.5c (Felsenstein, 1993). A 50% majority-rule consensus of the most parsimonious trees produced by this analysis was then constructed using PAUP 3.0 (Swofford 1990). Maximum parsimony using the branch and bound option of PAUP 3.0 was used to examine the phylogeny of fish sequences including only a single cloned sturgeon sequence.

Results

18S rRNA gene sequences from PCR products

During direct sequencing of PCR products, no evidence of multiple sequence variants for the 18S rRNA gene was detected in any of the non-acipenseriform species examined. In fact, the gene sequences obtained from the three species of Polypteriformes (*Polypterus senegalus*, *P. delhezi* and *Erpetoichthys calabaricus*) were found to be identical. Subsequently, the sequence from *P. senegalus* was used to represent the order Polypteriformes in phylogenetic analyses, as it was unnecessary to include all three species. There were also no multiple alleles detected in the North American paddlefish. Multiple sequence variants were, however, found in all nine of the sturgeon species examined. The presence of variants in sturgeon but not paddlefish suggests that this characteristic was acquired by the lineage leading to sturgeon after the two families Polyodontidae and Acipenseridae diverged.

Sequence variations included both single nucleotide substitutions and insertion/deletion events (indels). Both types of variation are manifested as multiple bands on sequencing gels. Figures 2 and 3 present examples of the types of sequencing polymorphisms observed in sturgeon. The presence of multiple indels especially made reliable determination of the complete sequence from PCR products impossible without cloning. For this reason, the direct sequences of the sturgeon PCR products are not included here. Nevertheless, regions of the gene that could be read reliably indicate that the 18S rRNA genes from the sturgeon species being studied are in general very similar in sequence. The partial gene sequence from the PCR product of one species (*Scaphirhynchus platorynchus*) with intra-individual variation was used in the phylogenetic analyses because it possesses only two indel sites very near the 3′ end of the gene. Therefore it was possible to determine the sequence of the remainder of the gene by sequencing the PCR product. The regions that could be read for *S. platorynchus* could also be read for *S. albus* and *S. suttkusi*, and from these sequences and the positions of their indels it appears that these three species possess very similar 18S rRNA gene sequence variants in their genomes.

18S rRNA gene sequences from cloned PCR products

In order to confirm the presence of multiple alleles of the 18S rRNA gene in sturgeon, initial attempts were made to clone the 5′ half of the genes from Atlantic sturgeon and white sturgeon. This was followed by other experiments to clone the complete genes from lake sturgeon, shortnose sturgeon and paddlefish. One partial 18S rDNA clone from Atlantic sturgeon (*Acipenser oxyrinchus oxyrinchus*) and six from white sturgeon (*A. transmontanus*) were isolated during the preliminary cloning experiments. The second round of cloning using different methods was more successful and large numbers of clones were obtained. Thirty independent clones from lake sturgeon, eight from shortnose sturgeon and six from paddlefish were selected for sequence analysis.

Screening of 30 lake sturgeon clones by partial sequencing showed that the minimum number of different gene sequences present in the sample of clones was 17, an unexpectedly large figure. The distribution of the numbers of clones potentially sharing the same sequence variants was: one sequence type was represented by six clones; a second sequence type was represented by four clones; two different sequence types were found in three clones each; a fourth was found in two clones, and the remaining 12 sequence types were unique. Nearly complete gene sequences were determined for the 17 clones with unique sequences. Comparison of sequences showed that one of the lake sturgeon clones (Af-7) was almost identical in sequence to the paddlefish 18S rRNA gene sequence, differing by only a single nucleotide substitution. (A more detailed description and analysis of the lake sturgeon 18S rDNA sequence variants and their significance will be presented in a future manuscript.) In comparing the 17 unique cloned sequences, there were 1783 sites which could be aligned, 159 of which are variable. The number of pair-wise differences between various lake sturgeon sequence variants ranged from six (0.003 divergence) to 72 (0.04 divergence). Of the 159 variable sites, 28 involve only indel events. The remaining 135 variable sites primarily involve only substitutions, with some involving a combination of substitution and insertion/deletion.

Eight shortnose sturgeon 18S rRNA gene clones were screened for sequence variation by partial sequencing. Four different sequence variants were detected and subsequently sequenced in their entirety. Shortnose sturgeon variants differed from each other by both indel events and nucleotide substitutions. All four variants characterized from shortnose sturgeon were different from the 17 variants isolated from lake sturgeon, but they did possess polymorphism at many of the same sites as lake sturgeon. This corresponded with observations made when the PCR product sequences of all nine sturgeon species were compared; many of the sites that possessed substitutions and indels in one species also had them in other species.

As a control to ensure that cloning procedures were not responsible for the observed variation in sturgeon 18S rRNA gene sequences, 18S rRNA gene PCR products from the paddlefish were also cloned. Cloning was carried out concurrently with the shortnose sturgeon cloning, but separately from the lake sturgeon cloning. Approximately 300 bp of the 5′ end

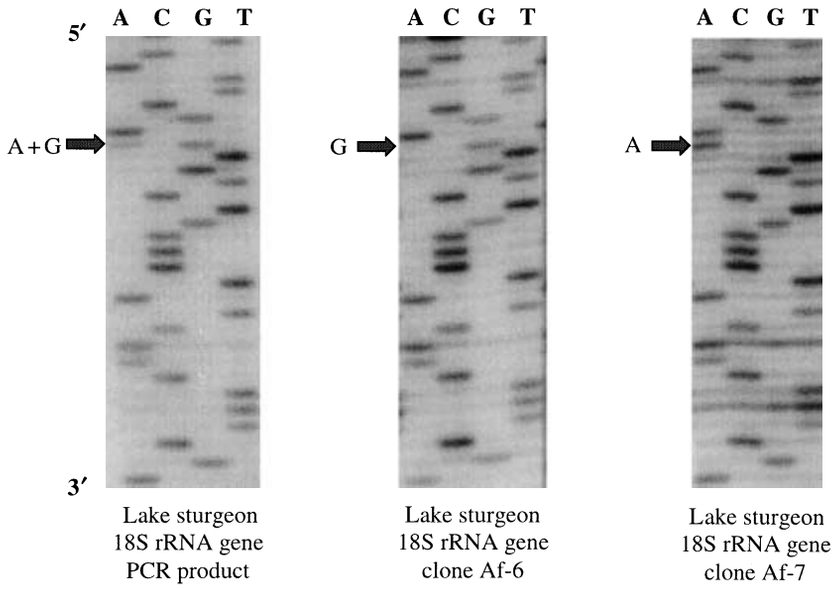

Fig. 2. DNA sequencing gels of lake sturgeon 18S rRNA gene PCR product and two clones, Af-6 and Af-7, for primer 373. Note the two bands found at the site marked with the arrow in the PCR product, indicating a base substitution event. Clone Af-6 has a G at this site while clone Af-7 has an A, illustrating the presence of multiple sequence types within the lake sturgeon genome and their effect on the sequence produced from direct sequencing of PCR products

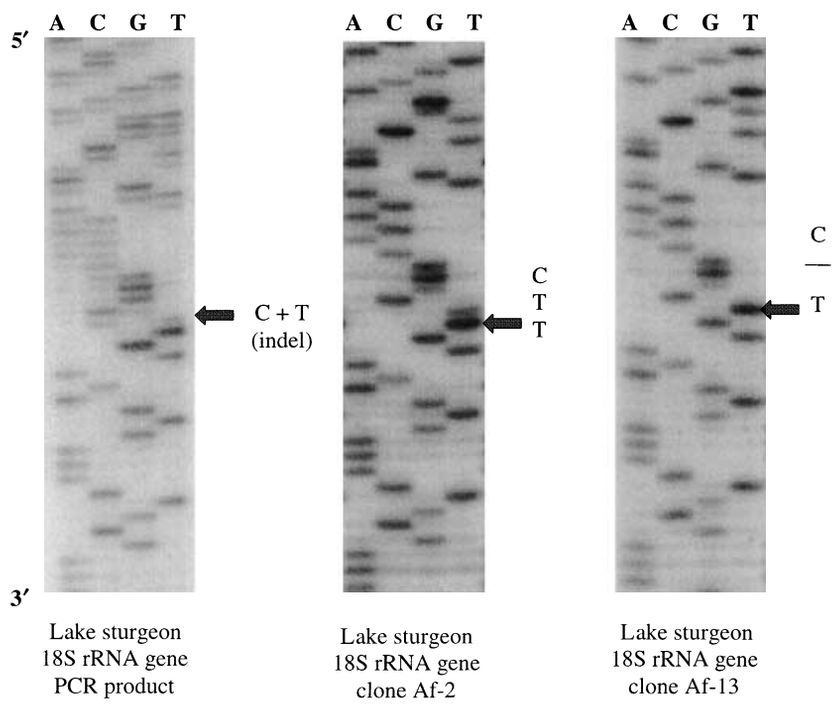

Fig. 3. DNA sequencing gels of lake sturgeon 18S rRNA gene PCR product and two clones, Af-2 and Af-13, for primer 170. Note the two bands found at the site marked with the arrow in the PCR product followed by multiple sequence ladders in the 5′ direction, indicating an indel event. Also note that clone Af-2 has an extra T in this region as compared with clone Af-13, illustrating the presence of multiple sequence types and how they result in the multiple sequence ladders when the PCR product is sequenced directly

of the gene was sequenced for six paddlefish clones. This region was chosen as it contains indels and base substitutions in almost all of the sturgeon species examined, making it a good candidate region for identifying sequence variants that might exist in the paddlefish genome. All six paddlefish clones were identical in sequence and also identical in sequence to the PCR product that was previously directly sequenced from paddlefish, suggesting that the paddlefish genome does not have detectable intra-individual variation for the 18S rRNA gene (at least at the 5′ end).

The lack of variation in paddlefish cloned products also confirms that the sequence variation found in sturgeon is inherent in the genome and not an artifact of the cloning procedures used. Further support for the actual existence of multiple 18S rDNA alleles in the sturgeon genome is provided by a comparison of the sequences of lake sturgeon PCR products that were directly sequenced and clones of the lake sturgeon PCR product (see Figs. 2, 3 for examples). Sites in the PCR product sequence where multiple sequencing ladders appear, and confound sequence determination, directly correspond with sites in the clones where indels occur. Also, for sites in the PCR product sequence where more than one base at a single position is detectable, different clones can be identified that have one or another of the bases found in the PCR product at that site.

Phylogenetic analyses

Extensive investigation of the effect of intra-individual sequence variation on phylogenetic placement of various

sturgeon will be presented elsewhere (Krieger and Fuerst, in preparation). For the purposes of this paper, it is sufficient to note that some of the variants of the 18S rRNA gene which have been isolated from lake sturgeon, short nose sturgeon and white sturgeon are interspersed with each other in a phylogenetic analysis, and do not show monophyletic clusters correlated with species (see Fig. 4). The sequence variants from various sturgeon species, while clustered together with the paddlefish in a distinctive acipenseriform clade, are also mixed in the tree with no clear separation by species. The phylogenetic utility of this gene with respect to its ability to distinguish sturgeon species is affected by the presence of multiple sequence variants in the sturgeon group as a whole. This suggests strongly that many of the gene duplication events and subsequent patterns of divergence predated the divergence of species, and that processes of concerted evolution to bring about sequence homogenization of all 18S rRNA genes has not acted strongly on these sequences since the divergence of the species.

Nevertheless, the sequences retain accurate phylogenetic information with respect to divergence at higher taxonomic levels. We investigated the phylogenetic position of the order Acipenseriformes in relation to other fish groups. Is the position of the order Acipenseriformes stable with respect to other orders if sequence variants representing *Acipenser* changed? Phylogenetic analysis was performed including three representative acipenseriform sequences: paddlefish, shovelnose sturgeon and lake sturgeon clone sequence Af-7, which was the least divergent of the variant *Acipenser* sequences when compared to paddlefish and shovelnose sturgeon. Af-7 differed by only one nucleotide from paddlefish. Figure 5 shows the resulting neighbor-joining tree. The tree produced by maximum parsimony had a similar topology. The topology is in good agreement with relationships based on morphology, with the exception of the placement of the sharks and the relationships among the bowfin, gar and teleosts. In both trees, the three acipenseriform sequences remain clustered and their closest relatives are the Polypteriformes, in agreement with the currently most widely accepted placement (Patterson, 1982; Lauder and Liem, 1983; Bemis et al., 1997).

As a second test, the ability of a single divergent sequence variant (lake sturgeon clone Af-33) to correctly place the Acipenseriformes in phylogenetic analyses was tested by using it as the sole representative of the order. Clone Af-33 was chosen because it is the most divergent of the lake sturgeon clones from the paddlefish sequence. When compared to other fish taxa, trees (not shown) produced by neighbor-joining and maximum parsimony both place clone Af-33 in a cluster with *Polypterus*, in agreement with the tree in Fig. 5 which includes multiple less divergent sequences from the group. Thus, despite intra-individual variation, acipenseriform 18S rRNA sequences contain information that allows the recovery of the presumptive correct placement of the order.

Discussion

This study is the first to observe extensive intra-individual variation of the 18S rRNA gene in any vertebrate. This was unexpected, since few organisms have been previously found to possess intra-individual variation of small subunit rDNA.

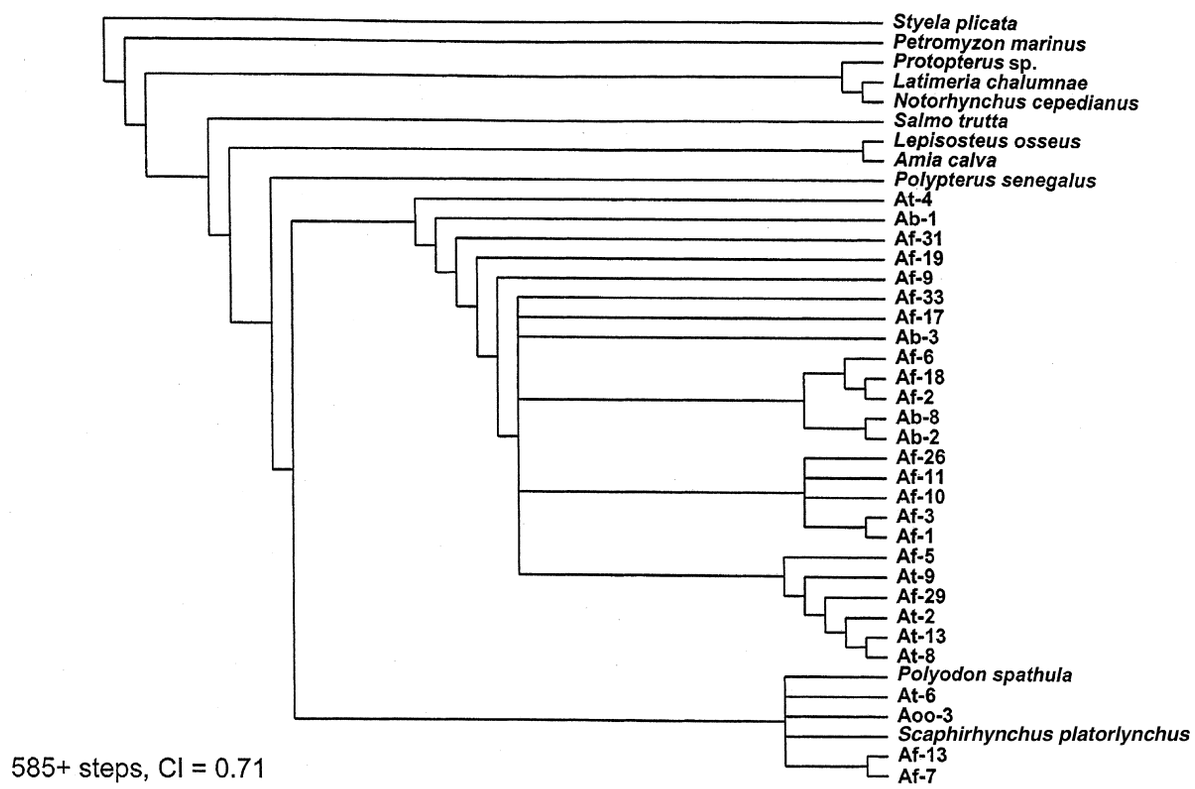

Fig. 4. 50% majority rule consensus of 588 most parsimonious trees constructed using Phylip 3.5c (Felsenstein, 1993) and consensed using Paup 3.0 (Swofford, 1990), based on partial 18S rRNA gene sequences (1212 sites) with *Styela plicata* as the designated outgroup. The sequences labeled Af-# represent the 17 18S rDNA clone sequences isolated from lake sturgeon (*Acipenser fulvescens*). Other labels represent sequences isolated from other sturgeon species: Ab-# = *Acipenser brevirostrum*, At-# = *Acipenser transmontanus*, and Aoo-3 = *Acipenser oxyrinchus oxyrinchus*. The mixing of sequences isolated from different sturgeon species evident in this tree is also observed when phylogenetic reconstruction is carried out using the neighbor-joining method (Saitou and Nei, 1987) (not shown)

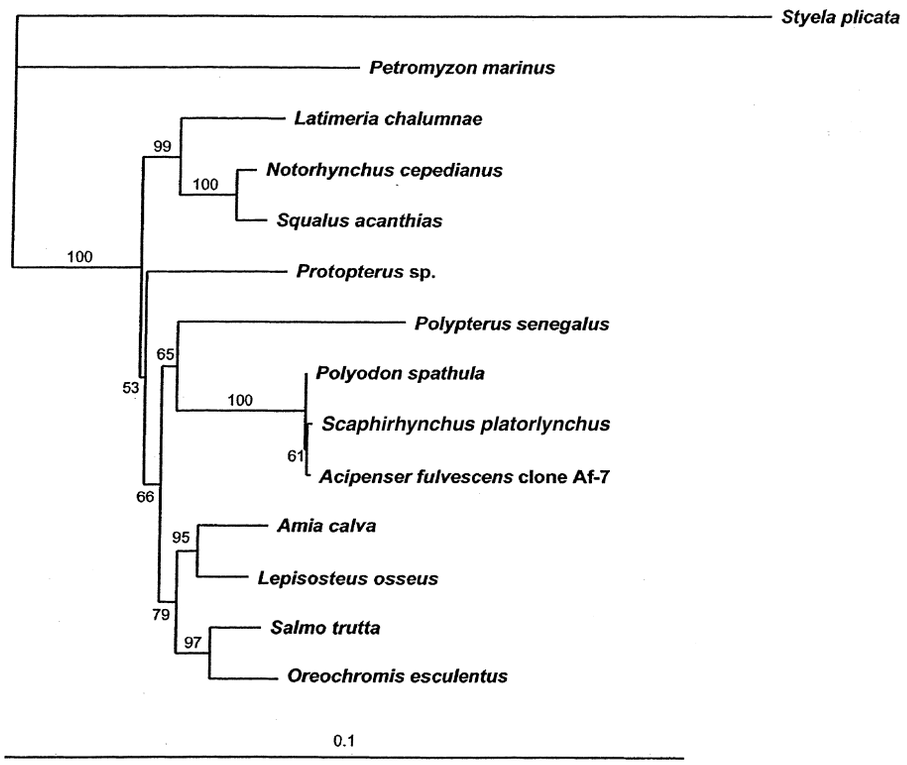

Fig. 5. Phylogenetic tree based on nearly complete 18S rRNA gene sequences (1784 sites), created using the neighbor-joining method (Saitou and Nei, 1987) in MEGA (Kumar et al., 1993). Numbers indicate the bootstrap values for 1000 replications. The topology of this tree is similar to that of the tree produced by the maximum parsimony method (not shown)

Further, when variation exists, usually only two or three sequence variants have been detected (Dame et al. 1984; Carranza et al. 1996, 1999; Stothard et al. 2000). The definite cause(s) for intra-individual variation of 18S rRNA genes in sturgeon (and not in paddlefish) are as yet unknown. Genetic variation within the repeated tandem array of the ribosomal RNA transcription unit has been noted in a number of organisms, although there seldom is variation in the sequences of the 18S rRNA genes. Among the possible causes for variability in the rDNA unit of various organisms, polyploidy has been observed most frequently. For sturgeon also, the polyploid nature of the genome makes it the most likely factor contributing to the origin and maintenance of multiple 18S rRNA gene sequence variants.

Polyploidy is widespread in plants. Campbell et al (1997) noted that in plants most cases of intra-individual ITS variation for the ribosomal RNA repeat unit are associated with polyploidy or multiple NORs, although many polyploid plants do not possess intra-individual ITS sequence variation. Polyploid species in several plant groups seem to have had much greater opportunity than diploid species in the same groups for divergent evolution and variation of rDNA repeat clusters (Suh et al. 1993; Jobst et al. 1998; Gaut et al. 2000; Wendel 2000). In fish, polyploidy has been associated with several groups, including salmonids. In the salmonid *Salvelinus namaycush*, variation in the NTS sequences of the rRNA repeat have been observed, although again without evidence of 18S rRNA sequence variation (Zhuo et al. 1995; Reed and Phillips 2000).

The origin of the multiple sequence variants in sturgeon is likely to have been a result of polyploidization. All species of the order Acipenseriformes that have been cytologically investigated appear to be polyploid, containing large numbers of chromosomes. In sturgeon and paddlefish, both flow cytometry studies to estimate genome size (Blacklidge and Bidwell 1993) and karyotypic studies of chromosomes (Dingerkus and Howell 1976; Birstein and Vasiliev 1987; Birstein et al. 1993; Fontana 1994; Fontana et al. 1998) provide evidence that the group had a polyploid origin. It is still unresolved, however, whether there is a tetraploid–octoploid or a diploid–tetraploid relationship between the two main chromosome groups in sturgeon, 120 and 240 chromosomes (see Lanfredi et al. 2001).

The origin of the multiple sequence variants in sturgeon could arise from polyploidization produced through either of two scenarios. In one case, hybridization could bring together different alleles from the two parental species, thus creating intra-individual variation. Alternatively, either auto- or allopolyploidization could have increased the number of chromosomes, as well as the number of rDNA clusters, compared to the ancestors. Because extra copies of rDNA would then exist in the genome, relaxed selection pressures (as not all copies would be required for proper cellular function) may have allowed mutations to accumulate in some copies, altering them into variant forms. Data exist which suggest that divergent NOR sequences would facilitate divergence of ITS sequences in the rRNA repeat (Arnheim et al. 1982). The diploidization of sequences in the polyploid sturgeon genome would cause differentiation of duplicated homologous chromosomes carrying rRNA arrays, and facilitate sequence drift.

Although it is possible that the large number of sequence variants found in sturgeon are functional and necessary for survival, it is not clear why sturgeon would need different sequence variants of the 18S rRNA gene. If the variants are actually adaptive, we have no evidence of what purpose they might serve. It is more likely that some or many of the sequence variants found in lake sturgeon are pseudogenes. Pseudogenes for small subunit rRNA have been found in diverse organisms such as primates (Brownell et al., 1983), the bacterium *Mycoplasma gallisepticum* (Skamrov et al., 1995) and *Drosophila* (Benevolenskaya et al., 1997). Such pseudogenes are often identified because they are truncated, or they

possess multiple substitutions and indels as compared with their functional counterparts. We observed no truncated versions of the sturgeon 18S rRNA gene. However, our screening procedures were biased to locate near-full length copies of the gene sequence. The results of experiments involving lake sturgeon 18S rDNA which will be presented in a future manuscript do support the idea that at least some of the rRNA variants observed in our study may functionally be pseudogenes.

The presence of multiple sequence variants in a sturgeon individual indicates that the mechanisms of concerted evolution have been unsuccessful at homogenizing the sequence variants found within the sturgeon genome. Our results suggest that care should be taken when using repeated nuclear genes for phylogenetic studies within this group (and, by extension, other polyploid groups). The presence of multiple sequence variants for repetitive genes within individuals may obscure the actual relationships among species (Buckler et al. 1997).

For the Acipenseriformes, any of the lake sturgeon variant 18S rDNA sequences described here, even when used alone, appear to be able to accurately represent the order in phylogenetic analyses compared to more divergent taxa. The level of variation among the lake sturgeon sequence variants ranged from 0.3 to 4%. Although the upper limit represents considerable divergence for the normally conserved 18S rRNA gene, it is still smaller than the difference (approx. 5%) found between the Acipenseriformes (represented by *Polyodon*) and their nearest relatives (the Polypteriformes). This implies that the presence of multiple variants in this family may not confound the elucidation of relationships among major fish groups using the 18S rRNA gene. However, intra-individual variation confuses any attempts to determine species relationships within the order. Variant sequences from different sturgeon species do not form monophyletic groups associated with a particular species.

The Acipenseriformes cluster with the Polypteriformes in all trees, although the support for this grouping is not very strong. This agrees with the hypotheses based on morphological data that place the order Polypteriformes as part of the subclass Actinopterygii, and that it also is the closest extent sister group of all the other Actinopterygii, within which the Acipenseriformes are the next basal group (Patterson 1982; Lauder and Liem 1983; Bemis et al. 1997). That the Polypteriformes are the closest extent relatives of the Acipenseriformes has also been supported by phylogenetic analysis of partial 28S rRNA gene sequences (Lê et al. 1993).

In conclusion, we report the first extensive intra-individual variation of the 18S rRNA gene in any organism. These findings indicate that the assumption may no longer be valid that patterns of gene sequence divergence reflect patterns of species divergence. The definite cause for intra-individual variation of 18S rRNA genes in sturgeon and not in paddlefish is as yet unknown. It is possible that the origin and maintenance of multiple 18S rRNA gene sequence variants in sturgeon is most related to the polyploid nature of the sturgeon genome. The discovery of this unusual phenomenon in the Acipenseridae is just one example of how much is yet to be discovered about this endangered group.

Acknowledgements

The authors would like to thank those whose efforts helped locate fish tissue samples as well as those who generously provided the samples necessary for this study, including Greg Booton, James Brown, Ron Bruch, Frank Chapman, David Davies, Carlos Echevarria, Kim Graham, Jerry Hamilton, Leroy Heman, Bernie Kuhadja, Ron Larson, Edward Little, Richard Mayden, Alan Miller, Jerre Mohler, Tim Mulligan, Terry Patterson, Brady Porter, Andy Vick and Kent Ware. This work was supported by a Theodore Roosevelt Memorial Fund Grant from the American Museum of Natural History and a Dissertation Improvement Grant from the National Science Foundation.

References

Appels, R.; Gerlach, W. L.; Dennis, E. S.; Swift, H.; Peacock, W. J., 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, Sunderland, Mass. pp. 38–61.

Arnheim, N.; Krystal, M.; Schmickel, R.; Wilson, G.; Ryder, O.; Zimmer, E., 1980: Molecular evidence for genetic exchanges among ribosomal genes on nonhomologous chromosomes in man and apes. Proc. Natl. Acad. Sci. USA **77,** 7323–7327.

Arnheim, N.; Treco, D.; Taylor, B.; Eicher, E. M., 1982: Distribution of ribosomal gene length variants among mouse chromosomes. Proc. Natl. Acad. Sci. USA **79**(15), 4677–4680.

Bemis, W. E.; Findeis, E. K.; Grande, L., 1997: An overview of Acipenseriformes. Environ. Biol. Fish. **48,** 25–71.

Benevolenskaya, E. V.; Kogan, G. L.; Tulin, A. V.; Philipp, D.; Gvozdev, V. A., 1997: Segmented gene conversion as a mechanism of correction of 18S rRNA pseudogene located outside of rDNA cluster in *D. melanogaster*. J. Mol. Evol. **44,** 646–651.

Birstein, V. J.; Vasiliev, V. P., 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.; Poletaev, A. I.; Goncharov, B. F., 1993: DNA content in Eurasian sturgeon species determined by flow cytometry. Cytometry **14,** 377-383.

Blacklidge, K. H.; Bidwell, C. A., 1993: Three ploidy levels indicated by genome quantification in Acipenseriformes of North America. J. Hered. **84,** 427–430.

Brownell, E.; Krystal, M.; Arnheim, N., 1983: Structure and evolution of human and African ape rDNA pseudogenes. Mol. Biol. Evol. **1**(1), 29–37.

Buckler, E. S., IV; Ippolito, A.; Holtsford, T. P., 1997: The evolution of ribosomal DNA: Divergent paralogues and phylogenetic implications. Genetics **145,** 821–832.

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

Campbell, C. S.; Wojciechowski, M. F.; Baldwin, B. G.; Alice, L. A.; Donoghue, M. J., 1997: Persistent nuclear ribosomal DNA sequence polymorphism in the *Amelanchier* agamic complex (Rosaceae). Mol. Biol. Evol. **14,** 81–90.

Carranza, S.; Giribet, G.; Ribera, C.; Baguñà, J.; Riutort, M., 1996: Evidence that two types of 18S rDNA coexist in the genome of *Dugesia (Schmidtea) mediterranea* (Platyhelminthes, Turbellaria, Tricladida). Mol. Biol. Evol. **13**(6), 824–832.

Carranza, S.; Baguñà, J.; Riutort, M., 1999: Origin and evolution of paralogous rRNA gene clusters within the flatworm family Dugesiidae (Platyhelminthes, Tricladida). J. Mol. Evol. **49**(2), 250–259.

Dame, J. B.; Sullivan, M.; McCutchan, T. F., 1984: Two major sequence classes of ribosomal RNA genes in *Plasmodium berghei*. Nuc. Acid. Res. **12**(14), 5943–5952.

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

Dover, G.; Brown, S.; Coen, E.; Dallas, J.; Strachan, T.; Trick, M., 1982: The dynamics of genome evolution and species differentiation. In: Genome evolution. G. Dover and R. B. Flavell (Eds.). Academic Press, London. pp. 343–372.

Felsenstein, J., 1993: PHYLIP (Phylogeny Inference Package), version 3.5c. Department of Genetics, University of Washington, Seattle, USA.

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

Fontana, F.; Tagliavini, J.; Congiu, L.; Lanfredi, M.; Chicca, M.; Laurente, C.; Rossi, R., 1998: Karyotype characterization of the great sturgeon, *Huso huso*, by multiple staining techniques and fluorescent in situ hybridization. Mar. Biol. **132,** 495–501.

Gaut, B. S.; Le Thierry d'Ennequin, M.; Peek, A. S.; Sawkins, M. C., 2000: Maize as a model for the evolution of plant nuclear genomes. Proc. Natl. Acad. Sci. USA **97**(13), 7008–7015.

Gutell, R. R., 1994: Collection of small subunit (16S- and 16S-like) ribosomal RNA structures: 1994. Nuc. Acid. Res. **22**(17), 3502–3507.

Hillis, D. M.; Davis, S. K, 1988: Ribosomal DNA: Intraspecific polymorphism, concerted evolution, and phylogeny reconstruction. Syst. Zool. **37**(1), 63–66.

Hillis, D. M.; Dixon, M. T., 1991: Ribosomal DNA: Molecular evolution and phylogenetic inference. Quart. Rev. Biol. **66**(4), 411–453.

Jobst, J.; King, K.; Hemleben, V., 1998: Molecular evolution of the internal transcribed spacers (ITS1 and ITS2) and phylogenetic relationships among species of the family Cucurbitaceae. Mol. Phylogenet. Evol. **9**(2), 204–219.

Kimura, M., 1980: A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. **16,** 111–120.

Krieger, J., 2000: Molecular phylogenetics and evolution of the North American sturgeon and paddlefish (Order Acipenseriformes). Doc. diss., Ohio State Univ., Columbus, Ohio, USA.

Krieger, J., Fuerst, P. A.; Cavender, T. M., 2000: Phylogenetic relationships of the North American sturgeons (order Acipenseriformes) based on mitochondrial DNA sequences. Mol. Phylog. Evol. **16**(1), 64–72.

Kumar, S.; Tamura, K.; Nei, M., 1993: MEGA: (Molecular Evolutionary Genetics Analysis), version 1.01. University of Molecular Evolutionary Genetics, Penn. State Univ., University Park, PA, USA.

Lanfredi, M.; L. Congiu, L.; Garrido-Ramos, M. A.; De la Herrán, R.; Leis, M.; Chicca, M.; Rossi, R.; Tagliavini, J.; Ruiz Rejón, C.; Fontana, F., 2001: Chromosomal location and evolution of a satellite DNA family in seven sturgeon species. Chromosome Res. **9,** 4–52.

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

Lê, H. L. V.; Lecointre, G.; Perasso, R., 1993: A 28S rRNA based phylogeny of the gnathostomes: First steps in the analysis of conflict and congruence with morphologically based cladograms. Mol. Phylog. Evol. **2**(1), 31–51.

Maniatis, T.; Fritsch, E. F.; Sambrook, J., 1982: Molecular Cloning: A laboratory manual. Cold Spring Harbor Lab., Cold Spring Harbor, NY. 545 pp.

Patterson, C., 1982: Morphology and interrelationships of primitive Actinopterygian fishes. Am. Zool. **22,** 241–259.

Reed, K. M.; Phillips, R. B., 2000: Structure and organization of the rDNA intergenic spacer in lake trout (*Salvelinus namaycush*). Chromosome Res. **8**(1), 5–16.

Saiki, R. K.; Gelfand, D. H.; Stoffel, S.; Scharf, S. J.; Higuchi, R.; Horn, G. T.; Mullis, K.B.; Erlich, E.H., 1988: Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239,** 487–491.

Saitou, N.; Nei, M., 1987: The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**(4), 406–425.

Skamrov, A.; Goldman, M.; Klasova, J.; Beabealashvilli, R., 1995: *Mycoplasma gallisepticum* 16S ribosomal RNA genes. FEMS Microbiol. Letters **128,** 321–325.

Stock, D. W.; Moberg, K. D.; Maxon, L. R.; Whitt, G. S., 1991: A phylogenetic analysis of the 18S ribosomal RNA sequence of thecoelacanth *Latimeria chalumnae*. Environ. Biol. Fish. **32,** 99–117.

Stock, D. W.; Whitt, G. S., 1992: Evidence from 18S ribosomal RNA sequences that lampreys and hagfishes form a natural group. Science **257**(5071), 787–789.

Stothard, J. R.; Frame, I. A.; Carrasco, H. J.; Miles, M. A., 2000: Analysis of genetic diversity of *Trypanosoma cruzi*: an application of riboprinting and gradient gel electrophoresis methods. Mem. Inst. Oswaldo Cruz **95**(4), 545–551.

Suh, Y.; Thien, L. B.; Reeve, H. E.; Zimmer, E. A., 1993: Molecular evolution and phylogenetic implications of internal transcribed spacer sequences of ribosomal DNA in Winteraceae. Am. J. Bot. **80,** 1042–1055.

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

Wada, H., 1998: Evolutionary history of free-swimming and sessile lifestyles in urochordates as deduced from 18S rDNA molecular phylogeny. Mol. Biol. Evol. **15**(9), 1189–1194.

Wendel, J. F., 2000: Genome evolution in polyploids. Plant Mol. Biol. **42**(1), 225–249.

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

Zhuo, L.; Reed, K. M.; Phillips, R. B., 1995: Hypervariability of ribosomal DNA at multiple chromosomal sites in lake trout (*Salvelinus namaycush*). Genome **38,** 487–496.

Author's address: Paul A. Fuerst, The Ohio State University, Department of Molecular Genetics and Department of Evolution, Ecology and Organismal Biology, 484 West Twelfth Avenue, Columbus, OH 43210, USA. E-mail: fuerst.1@osu.edu