# Heteroplasmy in the mtDNA Control Region of Sturgeon (*Acipenser*, *Huso* and *Scaphirhynchus*)

**Arne Ludwig,* Bernie May,† Lutz Debus‡ and Ingo Jenneckens§**

**Institute of Freshwater Ecology and Inland Fisheries, 12561 Berlin, Germany, †Department of Animal Science, University of California, Davis, California 95616, ‡Institute of Zoology, University of Rostock, 18051 Rostock, Germany and §Institute for Animal Breeding and Genetics, University of Göttingen, 37075 Göttingen, Germany*

Manuscript received March 20, 2000
Accepted for publication September 11, 2000

ABSTRACT

Data from 1238 fishes from 19 sturgeon species and 1 paddlefish were used to analyze heteroplasmy in sturgeon. Lengths of central repeat units ranged from 74 to 83 bp among sturgeon species. No repeat sequence was found in the paddlefish, *Polyodon spathula*. A general feature of the repeat units was the presence of termination associated sequence (TAS) motifs. About 50% of 138 interspecific mutations observed among the D-loop sequences are located 10 bp down- and upstream from these TAS motifs. Interestingly, most homoplasmic species showed deletions upstream to the TAS motifs, whereas deletions downstream to the TAS motifs observed in two species do not seem to preclude heteroplasmy. Calculations of secondary structures and thermal stabilities of repeat units showed $\Delta G$ values for all heteroplasmic species to be $<-8$ and for most homoplasmic species $\Delta G$ value to be $>-8$. Most heteroplasmic fishes had two and/or three repeat units. No homoplasmic sturgeon with $>2$ repeat units were observed. Molecular phylogeny based on the entire cytochrome *b* showed that heteroplasmy probably resulted from a single evolutionary event. Our data demonstrate that heteroplasmy is present in most sturgeon species and suggest that the thermal stability of the secondary structure of the repeat unit in combination with mutations downstream of the TAS sequences influences heteroplasmy.

THE family of Acipenseridae (sturgeon) is one of the oldest and most primitive of existing Osteichthyan fishes (Gardiner 1984). However, even such a basic question as the monophyly of the principal genus Acipenser still remains unanswered. Acipenseriformes (including paddlefish) are known to have existed at least since the Lower Jurassic (~200 mybp), and all fossil and recent taxa arose in the Holarctic. Several recent articles have addressed the phylogeny of this group (Bemis *et al.* 1997; Birstein and Bemis 1997; Birstein and DeSalle 1998). The late reproductive maturation, the long life cycles, and the decline of populations from overfishing, river damming, and water pollution during the last century make it difficult to investigate these living fossils.

Similar to several other fish species, sturgeon mitochondrial DNA (mtDNA) has extensive length variation in the D-loop. Individuals may be homo- or heteroplasmic for any number of discrete length variants of mtDNA (Buroker *et al.* 1990; Brown *et al.* 1992, 1996; Miracle and Campton 1995). The precise mechanisms causing mtDNA length variation and heteroplasmy are not known; however, different models have been discussed (Efstratiadis *et al.* 1980; Clark 1988; Rand and Harrison 1989; Buroker *et al.* 1990). In addition, most previous studies addressing heteroplasmy included a limited number of individuals from only a few species (Bentzen *et al.* 1988; Buroker *et al.* 1990; Árnason and Rand 1992; Brown *et al.* 1992, 1996; Nesbø *et al.* 1998). In contrast, we discuss the influence of mutations and thermal stability of predicted secondary structure on heteroplasmy using data from 1238 fishes from 19 species. Moreover, the occurrence of heteroplasmy is discussed in light of the molecular phylogeny of Acipenseriformes.

Earlier work on sturgeon was based on analyses of restriction fragment length polymorphisms (RFLP) fragments between 1.6 and 2.2 kb in length from Southern blots of restricted total mtDNA. Controversy about the presence of heteroplasmy in *Acipenser oxyrinchus* still exists (Brown *et al.* 1992, 1996; Miracle and Campton 1995). Miracle and Campton (1995) found 18.5% heteroplasmic *A. oxyrinchus*. On the other hand, Brown *et al.* (1996) observed no heteroplasmic sturgeon within this species. These differences may be due to the RFLP approach. To circumvent problems associated with RFLP analyses we developed a PCR-based method using primers located upstream in the mitochondrial $tRNA^{Pro}$ gene and downstream from the repeat region in the D-loop (Ludwig and Jenneckens 2000).

## MATERIALS AND METHODS

**Fish samples:** The species of sturgeon, number of specimens, and the geographical origin are given in Table 1. Previous studies showed that *Polyodon spathula* has a basal position

*Corresponding author:* A. Ludwig, Institut für Gewässerökologie und Binnenfischerei, Müggelseedamm 310, 12561 Berlin, Germany. E-mail: al@igb-berlin.de

**TABLE 1**

**Samples of acipenseriform species used in these studies**

| Species | Common name | *n* | Geographic origin of sampled species |
|---|---|---|---|
| Family Polydontidae | | | |
| *P. spathula* | North American paddlefish | 1 | Aquaculture (Mississippi R.) |
| Family Acipenseridae | | | |
| Subfamily Acipenserinae | | | |
| Genus Acipenser | | | |
| *A. baerii* | Siberian sturgeon | 126 | Siberia, Russia |
| *A. brevirostrum* | Shortnose sturgeon | 6 | Aquaculture |
| *A. fulvescens* | Lake sturgeon | 31 | R. Menominee and R. Wolfe, Wisconsin; Lake Winnebago |
| *A. gueldenstaedtii* | Russian sturgeon or osetra[a] | 98 | Caspian Sea; Black Sea; Sea of Azov |
| *A. medirostris* | Green sturgeon | 37 | Columbia R. |
| *A. mikadoi* | Sakhalin sturgeon | 6 | Tumnin R., Russia |
| *A. naccarii* | Adriatic sturgeon | 20 | Buna R., Albania; Po R., Italy |
| *A. nudiventris* | Ship sturgeon | 15 | Caspian Sea; Volga R. |
| *A. oxyrinchus* | Atlantic sturgeon | 90 | Delaware R.; Hudson R.; St. John R. |
| *A. persicus* | Persian sturgeon | 17 | Caspian Sea; Volga R. |
| *A. ruthenus* | Sterlet | 156 | Volga system, Kuban R. and Ob R., Russia; Danube R., Romania |
| *A. stellatus* | Stellate sturgeon or sevruga[a] | 93 | Caspian Sea; Black Sea; Sea of Azov |
| *A. sturio* | European sturgeon | 44 | Gironde R., France; North Sea |
| *A. transmontanus* | White sturgeon | 12 | Kootenai R. |
| Genus Huso | | | |
| *H. huso* | Great sturgeon or beluga[a] | 74 | Caspian Sea; Black Sea |
| Subfamily Scaphirhynchinae | | | |
| *S. albus* | Pallid sturgeon | 6 | Atchafalaya R. |
| *S. platorynchus* | Shovelnose sturgeon | 11 | Mississippi R. and Missouri R. |

[a] Trade name of caviar.

to the genera Acipenser, Huso, and Scaphirhynchus (Bemis *et al.* 1997; Birstein *et al.* 1997). Therefore, a specimen of this species was included.

**DNA extraction and PCR conditions:** Fin or blood samples were taken in the field and stored in 96% ethanol. The DNA was isolated using QIAamp kits (QIAGEN, Hilden, Germany). The repeat region, including partial sequences of the $tRNA^{Pro}$ and the D-loop, was amplified with the following primers: 5′-ACCCTTAACTCCCAAAG-3′ (Hetero I) and 5′-CATTTRATG GTAGATGAAAC-3′ (Hetero II) (Ludwig and Jenneckens 2000) under the following reaction conditions: 20 sec at 94°, 10 sec at 50°, and 1 min at 72° for 30 cycles and a final polymerization for 4 min at 72°. Amplifications were carried out with 1 unit *Taq* DNA polymerase (Oncor-Applegene), 10 pmol of each primer, ~10 ng DNA, 100 μm of each dNTP, 2.5 mm $MgCl_2$ in 2.5 μl incubation buffer and a total volume of 25 μl. No PCR product was obtained using primer Hetero II for *P. spathula*. Therefore, Hetero I and REV-A (5′-ATGTTC GGYTCTTACTACA-3′) were used for amplification of the partial D-loop sequence of this species following the protocol mentioned above.

The entire cytochrome *b* gene (cyt-*b*) was amplified using primers cytb-for1 (5′-CGTTGTHWTTCAACTAYARRAAC-3′) and cytb-rev1 (5′-CTTCGGTTTACAAGACCG-3′) under the following reaction conditions: 60 sec at 94°, 30 sec at 50°, 90 sec at 72° for 30 cycles, and a final polymerization at 72° for 5 min. Amplifications were carried out with 1 unit *Taq* DNA polymerase (Oncor-Applegene), 5 pmol of each primer, 50 ng DNA, 0.10 mm Tris-HCl (pH 8.8 at 25°), 50 mm KCl, 1.5 mm $MgCl_2$, 0.1 μg/μl bovine serum albumin, 0.08% (v/v) Nonidet P40, and 100 μm of each dNTP.

**Screening for intraspecific variability:** Single-strand conformation polymorphism (SSCP) analyses of partial cyt-*b* regions were performed for a rough detection of intraspecific nucleotide variations. Except for a shortened extension time that was reduced to 30 sec in each cycling step, the PCR conditions were the same as for the entire cyt-*b* mentioned above. The following primers were used: cytb-for1, cytb-L113 (5′-GCCTCT GCCTTRTCAC-3′), cytb-L394 (5′-GTACTGCCCTGAGGACA-3′), cytb-L572 (5′-GGAGCTAGCATAATTCACCT-3′), cytb-L823 (5′-CTCTTYGCCTACGCCATYC-3′), cytb-H345 (5′-GAT RTTTCAGGTYTCTTTTTG-3′), cytb-H835 (5′-CGTAGGCRA AGAGRAAG-3′), cytb-H-928 (5′-GTRTTTCCYCGTTGTTTR GAG-3′), and cytb-rev1. Five microliters of each amplification product was mixed with the same amount of a stop-mix solution consisting of 95% (w/v) formamide, 20 mm EDTA, 0.05% (w/v) bromphenol blue, and 0.05% (w/v) xylene cyanol. Furthermore, 1 μl was added to a 10 mm EDTA solution containing 0.1% (w/v) SDS. This mix was heated to 95° for 10 min and cooled on ice for 5 min immediately. Five microliters of each sample was loaded into separate lanes of a 15% (w/v) native polyacrylamide gel prepared according to Orita *et al.* (1989). Electrophoresis was carried out at 120 V for 10 hr. DNA fragments were silver stained using the method described by Bassam *et al.* (1991).

**Sequencing:** PCR products were run on a 1.5% agarose gel at 150 V for 2 hr. PCR products were excised from the gel and extracted using the QIAquick Gel Extraction Kit (QIAGEN). Direct sequencing was performed in both directions using an ABI 310 (Perkin Elmer, Norwalk, CT) following the protocol for cycle sequencing. Sequencing of D-loop fragments was performed under the following conditions: 15–40 ng DNA,

10 pmol of each primer (Hetero I; Hetero II; REV-A), and 2–4 μl BigDye RR terminator cycle sequencing kit (Perkin Elmer) in a Perkin-Elmer 2400 thermocycler programmed for 30 cycles of 10 sec at 94°, 5 sec at 50°, and 4 min at 60°.

In addition to primers cytb-for1 and cytb-rev1 mentioned above, four additional sequencing primers were derived on the basis of partial cyt-*b* sequences of acipenseriform species published recently (Birstein and DeSalle 1998). These were cytb-L113, cytb-L823, cytb-H835, and cytb-H345.

**Data analysis:** Previously published data from 21 *A. fulvescens*, 10 *A. medirostris*, 188 *A. oxyrinchus*, including both subspecies, 174 *A. transmontanus* (Miracle and Campton 1995; Brown *et al.* 1996), and two *Scaphirhynchus suttkusi* (EMBL accession nos. U43679 and U55994) were included in our analyses of heteroplasmy, giving a total of 1238 fishes from 19 species. Secondary structures and thermal stabilities of central repeat units were calculated using an algorithm written for DNA in the program Mfold v. 3.0 (Zuker 1989; Walter *et al.* 1994). Tree calculations in these studies were performed on (1) the entire cyt-*b* gene (1141 bp) and (2) 250 bp of the D-loop including the central repeat unit and an additional 45 bp of the 3′ end of the $tRNA^{Pro}$ gene. Indels showing as gaps in the alignment and only present in D-loop sequences were used as additional characters for tree reconstruction. We used the computer package PAUP* (Swofford 1998) to calculate sequence divergence values among the different haplotypes. Phylogenetic relationships and genetic distances among taxa were estimated using maximum parsimony and neighbor-joining (NJ) algorithms, implemented by PAUP* (Swofford 1998). Bootstrap analyses were based on 1000 replications. The significance of the branch lengths in the NJ tree was also examined by a standard error test using the confidence probability program of MEGA (Kumar *et al.* 1993). The phylogenetic networks were constructed using the median-joining (MJ) algorithm of Bandelt *et al.* (1999). MJ networks include all most-parsimonious trees supported by the data and are particularly appropriate for the low resolution encountered in intra- or closely related interspecific data sets.

## RESULTS

Sequences are archived in the EMBL GenBank (D-loop—AJ275186, AJ275188-9, AJ275194, AJ275205, AJ249660-2, AJ249668, AJ249670-6, AJ275190, AJ275199; cyt-*b*—AJ245825-41 and AJ249692-4).

**Sequence analyses of the $tRNA^{Pro}$ and the D-loop:** In general, organization of the 3′ end of the $tRNA^{Pro}$ gene and the 5′ end of the D-loop from all sturgeon species is similar to that previously reported (Buroker *et al.* 1990; Brown *et al.* 1996). The repeat region, responsible for length variations and heteroplasmy, is located near the 5′ end of the D-loop and separated by only a few nucleotides from the $tRNA^{Pro}$ gene. In contrast to some other fish species (Nesbø *et al.* 1998) these repeat units contain TAS motifs. TAS sequences were also present in *P. spathula*. Overall we observed eight substitutions within the TAS sequences among the sturgeon species and two additional substitutions between all sturgeon species and *P. spathula* (Figure 1). An additional sequence motif (ACATTAAACT) observed in the 3′ end of the $tRNA^{Pro}$ gene in all sturgeon and paddlefish species was highly similar to the TAS motif. In addition to TAS-derived differences among acipenserform species, other interspecific differences were also observed. In detail, a deletion of 36 bp was observed in the 5′ end of the D-loop in the following species: *A. brevirostrum*, *A. persicus*, *A. naccarii*, *A. gueldenstaedtii*, and *A. baerii*. *A. oxyrinchus* and *A. sturio* are characterized by a deletion of 24 bp. Focusing on the molecular phylogeny of the genus Acipenser (Figure 2) on the basis of the entire cyt-*b* data reveals that the deletion of 36 bp was one evolutionary event after the separation of the ancestor of *A. brevirostrum*, *A. persicus*, *A. naccarii*, *A. gueldenstaedtii*, and *A. baerii*. Furthermore, the deletion of 24 bp in *A. oxyrinchus* and *A. sturio* support the monophyly of these two species. In addition, several haplotypes were observed within the flanking D-loop fragment characterized by substitutions and single base insertions or deletions in *A. gueldenstaedtii*, *A. brevirostrum*, *A. transmontanus*, *S. platorynchus*, *A. nudiventris*, *A. persicus*, *A. ruthenus*, and *A. stellatus*.

**Repeat sequences and length variations:** The number of repeat units within sturgeon species ranged between one and seven (Table 2, Figure 3). Length variations of the PCR products resulting from different numbers of repeat units were observed in *S. platorynchus*; however, no heteroplasmy was observed. Individuals with one repeat unit were rare in the genera Acipenser and Huso. In contrast, ~60% of the Scaphirhynchus individuals had one repeat unit. Seven repeat units were observed in one *A. ruthenus* caught in the Ob River, Siberia. Lengths of central repeat units varied as follows: 74 bp in *S. platorynchus* and *S. albus*; 78 bp in *A. mikadoi*; 79 bp in *A. oxyrinchus* and *A. medirostris*; 80 bp in *A. ruthenus* and *A. sturio*; 82 bp in *A. baerii*, *A. fulvescens*, *A. gueldenstaedtii*, *A. naccarii*, *A. nudiventris*, *A. persicus*, *A. stellatus*, *A. transmontanus*, and *Huso huso*; and 83 bp in *A. brevirostrum*. No sequence similar to the repeat sequences found in sturgeon was observed in *P. spathula*. Structures of the central repeat units (Figure 1) showed that most deletions or insertions resulted from single step mutations in the ancestors of the different lineages. For example, all species of the genus Scaphirhynchus showed an identical deletion of 8 bp; *A. sturio* and *A. oxyrinchus* had an identical deletion of 2 bp; and *A. medirostris* and *A. mikadoi* are characterized by a deletion of 3 or 4 bp at the same location, respectively. Overall we observed 138 interspecific repeat sequence substitutions among the sturgeon species. More than 50% of these mutations were ±10 bp of the TAS sequences.

**Heteroplasmy:** The number of different mitochondrial genomes, *i.e.*, mtDNA D-loop types, as distinguished by the length of the PCR products within individual fish, ranged between one (homoplasmic fish) and two to four (heteroplasmic fish; Table 3). No heteroplasmy was observed in *A. nudiventris*, *A. sturio*, *A. oxyrinchus*, *S. albus*, and *S. platorynchus*. In contrast to these homoplasmic species, frequencies of heteroplasmy ranged from 0.064% (*A. fulvescens*) to 0.834% (*A. mikadoi*) in the remaining sturgeon taxa.

```
      1                tRNA^Pro                similar repeat sequences                90
A tra CTAAGATTCTACATTAAACTATTCTCTGACCATGCTATGTTTAATCCACATTAATTTCTAGCCACCATAC-CATAATGCTCACAAGCACA
A med CTAAGATTCTACATTAAACTATTCTCTGACCATACTATGTTTAATCCACATTAATCTCTAGTCACCATAC-CATAATGTTTGTAAATACA
A mik CTAAGATTCTACATTAAACTATTCTCTGACCATACTATGTTTAATCCACATTAATCTCTAGTCACCATAC-CATAATGTTTGTAAATACA
A ful CTAAGATTCTACATTAAACTATTCTCTGAC-ATACTATGTTTAATCCACATTAATTTCTAGTCATCATAC-ATCAATGCTCGTACATACA
S pla CTAAGATTCTACATTAAACTATCCTCTGACCACAT-ATGTTTAATCCACATTAATTTCTAGTCACCATACTCAT-------GCG--TACA
S sut CTAAGATTCTACATTAAACTATCCTCTGACCACAT-ATGTTTAATCCACATTAATTTCTAGTCACCATACTCAT-------GCG--TACA
S alb CTAAGATTCTACATTAAACTATCCTCTGACCACAT-ATGTTTAATCCACATTAATTTCTAGTCACCATACTCAT-------GCG--TACA
A rut CTAAGATTCTACATTAAACTATTCTCTGGTTATACCATGTTTAATCCACATTAACTTCTAGTCACCATA-TCATAATGTTTGCG--TACA
H hus CTAAGATTCTACATTAAACTATTCTCTGACCACACCATGTTTAACCCACACCAATTTCTAGCCACCATACT-ATAATGTTTACATATACA
A ste CTAAGATTCTACATTAAACTATTCTCTGACCATGCTATGTTTAATCCACATTAATTTCTAGCCACCATA-ACATAATGCTTGCACATACA
A nud CTAAGATTCTACATTAAACTATTCTCTGACCA--CCATGTTTAATCCACATTGATTTCTAGTCACCATACATATGATGCTTTCACGTACA
A bre CTAAGATTCTACATTAAACTATTCTCTGACCATACCAC----------GATGCTCACACATACA
A bae CTAAGATTCTACATTAAACTATTCTCTGACCATACCAT----------AATGCTTACATATACA
A per CTAAGATTCTACATTAAACTATTCTCTGACTATACCAT----------AATGTTTGCATGTACA
A nac CTAAGATTCTACATTAAACTATTCTCTGACTATACCAT----------AATGTTTGCATGTACA
A gue CTAAGATTCTACATTAAACTATTCTCTGACTATACCAT----------AATGTTTGCATGTACA
A stu CTAAGATTCTACATTAAACTATTCTCTGACCACATGTCTGA-CCCATAC--CAATGTCTGCAT----------ACA
A oxy CTAAGATTCTACATTAAACTATTCTCTGGCCATATCA-TGCTCGCCTACACCAATGT-T-TAT----------ACA
P spa CTAAGATTCTACATTAAACTATTCTCTGAACAGAC----------ACA

      91                central repeat                180
A tra TTAAATTGTTTAAGTACATAAGACATGCTATGTTTAATCCACATTAATTT-CTAGCCACCA-TAC-CATAATGCTCACAAGCACATTAAA
A med TTAAATT----ACCTATATAGGACATACTATGTTTAATCCACATTAATCT-CTAGTCACCA-TAC-CATAATGTTTGTAAATACATTAAA
A mik TTAGATT----ACCTATATAAGACATACTATGTTTAATCCACATTAATCT-CTAGTCACCA-TAC-CATAATGTTTGTAAATACATTAGA
A ful TTAAATTGTTTAAGTACATAGGACATACTATGTTTAATCCACATTAATTT-CTAGTCATCA-TAC-ATCAATGCTCGTACATACATTAAA
S pla TTAAATTATTGAGGTACATAAGACATGCTATGTTTAATCCACATTAATTT-CTAGTCACCA-TACTCATACGT---------ACATTAAA
S sut TTAAATTATTGAGGTACATAAGAC----------
S alb TTAAATTATTGAGGTACATAAGAC----------
A rut TTAAATTATTTAAGTACATAAGGCATGCTATGTTTAATCCACATTAACTT-CTAGTCACCA-TAT-CATAATGTTTGC--GTACATTAAA
H hus TTAAATTATTCAAGTACATAAGACATGCTATGTATAATCCACATTAGTTT-CTAGCCACCA-TAT-CATAATGTTTACATATACATTAAA
A ste TTAAATTATTTAAGTACATAAGACATGCTATGTTTAATCCACATTAATTT-CTAGCCACCA-TAA-CATAATGCTTGCACATACATTAAA
A nud TTAGATTGTTTAAGTACATAAGACATGCTATGTTTAATCCACATTAACTT-CTAGCCACCA-TAC-TAGAATGTTT-CATCTACATCAAA
A bre TTAAATTGTTTAAGTACATAAGACATACTATGTTTAATCCACATTAATTT-CTAGCCACCA-TAACCATAATGCTTACACATACATTAAA
A bae TTAAATTGTTTAAGTACATAAGACATGCTATGTTTAATCCACATTAATTT-CTAGTCACCA-TAC-CATAATGCTTACATATACATTAAA
A per TTAAATTGTTTAAGTACATAAGGCATACTATGTTTAATCCACATTAATTT-CTAGCCACCA-TAT-CATAATGTTTGCATGTACATTAAA
A nac TTAAATTGTTTAAGTACATAAGGCATACTATGTTTAATCCACATTAATTT-CTAGTCACCA-TAC-CATAATGTTTGCATGTACATTAAA
A gue TTAAATTGTTTAAGTACATAAGGCATACTATGTTTAATCCACATTAATTT-CTAGCCACCA-TAC-CATAATGTTTGCATGTACATTAAA
A stu TTAAATTGTACA--TACATAA-ACATACTATGTTTAATCCCCATTAATTT-CTAGCCACCAATAC-TA-ATGTTTACCTATATATTAAA
A oxy TTAAATTGCTTA--TGCATGG-ACATATTATGTTTAATCCCCATTAATTT-CTAGTCACCA-TAC-CA--ATGTTTATATATACATTAAG
P spa TAAAA-TGTACTA-TGTTTAATCCACATTAATTT-CTAGCCACCATAAATCATTCATCTACTGTATTGAATTTTAAATGCAACAAAAACA
```

FIGURE 1.—Alignment of the partial tRNA$^{Pro}$, similar repeat sequences, and the central repeat (A tra, *A. transmontanus*; A med, *A. medirostris*; A mik, *A. mikadoi*; A ful, *A. fulvescens*; S pla, *S. platorynchus*; S sut, *S. suttkusi* (EMBL: U43679); S alb, *S. albus*; A rut, *A. ruthenus*; H hus, *H. huso*; A ste, *A. stellatus*; A nud, *A. nudiventris*; A bre, *A. brevirostrum*; A bae, *A. baerii*; A per, *A. persicus*; A nac, *A. naccarii*; A gue, *A. gueldenstaedtii*; A stu, *A. sturio*; A oxy, *A. oxyrinchus*; P spa, *P. spathula*). TAS motifs are boldface, and repeat units used for calculating secondary structures are italic and underlined. Sequences were aligned with ClustalW (THOMPSON *et al.* 1994).

**Secondary structure:** Previous studies showed that the definition of initial point of repeat units within each species is somewhat arbitrary because the optimal secondary structure may or may not begin at the first nucleotide of a repeat unit (BROWN *et al.* 1996). We found a DNA segment that could form a secondary structure with a terminal hairpin stem in all of the sequences we examined (Figure 4). The stability of the predicted structures varied and the estimated thermal stabilities of the repeat units in homoplasmic species are somewhat lower than in heteroplasmic species. Only *A. nudiventris* showed a thermal stability similar to that of heteroplasmic species. However, $\Delta G$ values $>-10$ were estimated in *S. albus*, *S. platorynchus*, and *S. suttkusi* using the repeat

```
      181                                                                  245
A tra TTGTTTAAGTACATAAGACATGCTATGTTTAATCCACATTAATTTCTAGCCATCATA--TCACAAT
A med TT---GACCTATATAGGACATACTATGTTTAATCCACATTAATCTCTAGTCACCATA--CCATAAT
A mik TT----ACCTATATAAGACATACTATGTTTAATCCACATTAATCTCTAGTCACCATA--CCATAAT
A ful TTGTTTAAGTACATAGGACATATTATGTTTAATCCACATTAATTTCCAGTCACTACA-CACAGAAT
S pla TTATTGAGGTACATAAGACATGCTATGTTTAATCCACATTAATTTCTAGTCAACATA--TCATAAT
S sut ------------------ATGCTATGTTTAATCCACATTAATTTCTAGTCAACATA--TCATAAT
S alb ------------------ATGCTATGTTTAATCCACATTAATTTCTAGTCAACATA--TCATAAT
A rut TTATTTAAGTACATAAGGCATGCTATGTTTAATCCACATTAACTTCTAGTCACCATA--TCATAAT
H hus TTATTCAAGTACATAAGACATGCTATGTATAATCCACATTAGTTTCTAGCCACCATA--CCATAAT
A ste TTATTTAAGTACATAAGACATGCTATGTTTAATCCACATTAATTTCTAGTCAACATA--TCATAAT
A nud TTGTTTAAGTACATAAGGCATACTATGTTTAATCCACATTAATTTCTAGCCACCATA--TCATAAT
A bre TTGTTTAAGTACATAAGACATACTATGTTTAATCCACATTAATTTCTAGCCACCATAACTCACAAT
A bae TTGTTTAAGTACATAAGACATGCTATGTTTAATCCACATTAATTTCTAGTCACCATA--CCATAAT
A per TTGTTTAAGTACATAAGACATACTATGTTTAATCCACATTAATTTCTAGCCACCATA--TCATAAT
A nac TTGTTTAAGTACATAAGGCATACTATGTTTAATCCACATTAATTTCTAGTCACCATA--CCACAAT
A gue TTGTTTAAGTACATAAGGCATACTATGTTTAATCCACATTAATTTCTAGCCACCATA--CCATAAT
A stu TTATCTAAGTACATA-GACATACTATGTTTAATCCCCATTAATTTCTAGTCAACATA--TCAGAAT
A oxy CCATTTAAGTACATG-AACATACTATGTTTTATCCCCATTAACTTCTAGCCACTATA--ACAGAAT
P spa TAGTAT-GTTTAATCC-ACATTAATTTTTATA-CAACATTAAAGA-TAACTTGTATAA-CCAACAA
```

FIGURE 1.—*Continued.*

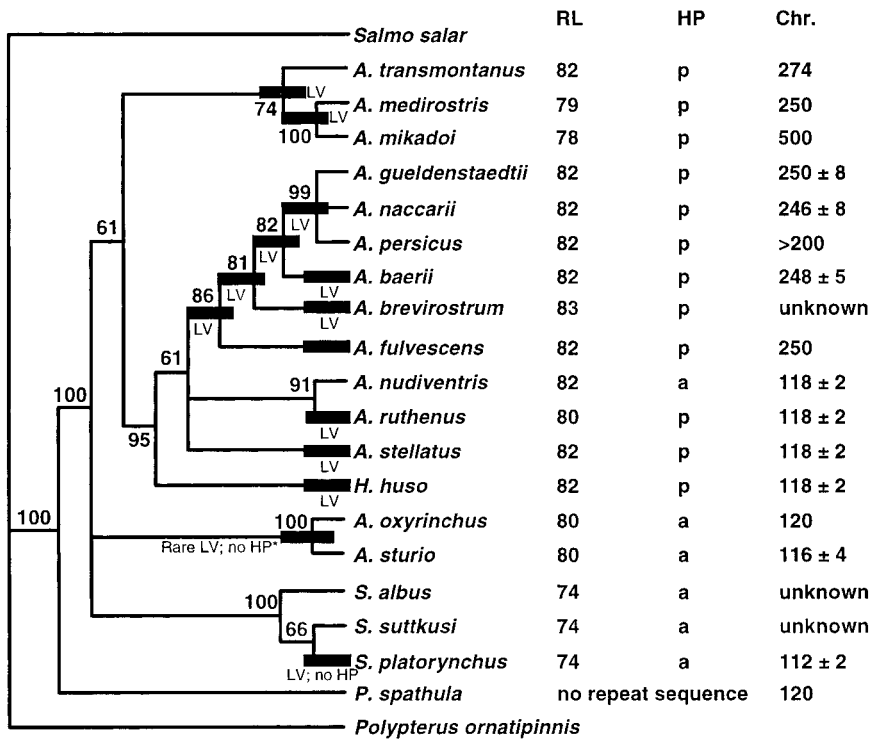



FIGURE 2.—Phylogenetic relationships and the occurrences of length variations (LV), length of central repeat unit in base pairs (RL), heteroplasmy (HP: a, absent; p, present), and chromosome numbers [Chr; reviewed in BIRSTEIN *et al.* (1997), data of *A. transmontanus* from VAN EENNAAM *et al.* (1998)] among sturgeon species. The most parsimonious tree (heuristic search, addition sequence algorithm) was generated using sequences of the entire cytochrome *b* gene (1141 bp, 380 amino acids), calculated in PAUP* (SWOFFORD 1998). Bootstrap values from 1000 replications are shown on the branches. Consensus sequences were used for *A. gueldenstaedtii*, *A. naccarii*, *A. persicus*, *S. albus*, and *S. platorynchus*. The consistency index was 0.639; the retention index was 0.622. One most parsimonious tree was observed with a total length of 1029 steps. The statistical significance of the node joining the species pairs *S. albus*/*S. platorynchus*/*S. suttkusi*, *A. medirostris*/*A. mikadoi*, *A. gueldenstadtii*/*A. naccarii*/*A. persicus*, and *A. oxyrinchus*/*A. sturio* was ≥95% according to bootstrap analyses. Sequences of *S. suttkusi* (EMBL: U55994), *Salmo salar* (SWISSPROT: Q35925), and *Polypterus ornatipinnis* (SWISSPROT: Q95920) were taken from the GenBank. *S. salar* and *P. ornatipinnis* were used as outgroups. BROWN *et al.* (1996) observed a length variation in one *A. oxyrinchus*.

**TABLE 2**

**Frequency and distribution of repeat units in the sturgeon species analyzed in this study**

| Species | Repeat number | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| *S. albus* | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 |
| *S. platorynchus* | 0.454 | 0.545 | 0 | 0 | 0 | 0 | 0 |
| *A. baerii* | 0 | 0.333 | 0.762 | 0.087 | 0.008 | 0 | 0 |
| *A. brevirostrum* | 0 | 0 | 1.000 | 0.500 | 0.166 | 0 | 0 |
| *A. gueldenstaedtii* | 0.102 | 0.316 | 0.653 | 0.132 | 0.020 | 0.010 | 0 |
| *A. fulvescens* | 0 | 0.581 | 0.387 | 0.129 | 0 | 0 | 0 |
| *A. medirostris* | 0 | 0.811 | 0.676 | 0.405 | 0 | 0 | 0 |
| *A. mikadoi* | 0 | 0.333 | 0.833 | 0.500 | 0.666 | 0 | 0 |
| *A. naccarii* | 0 | 0.150 | 0.350 | 0.900 | 0 | 0 | 0 |
| *A. nudiventris* | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 |
| *A. oxyrinchus* | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 |
| *A. persicus* | 0.059 | 0.353 | 0.294 | 0.294 | 0.176 | 0 | 0 |
| *A. ruthenus* | 0 | 0.339 | 0.666 | 0.416 | 0.122 | 0.013 | 0.006 |
| *A. stellatus* | 0.021 | 0.193 | 0.451 | 0.494 | 0.118 | 0 | 0 |
| *A. sturio* | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 |
| *A. transmontanus* | 0 | 0.166 | 0.187 | 0.583 | 0.333 | 0 | 0 |
| *H. huso* | 0 | 0.040 | 0.932 | 0.067 | 0.027 | 0.013 | 0 |

Frequency values are calculated in relation to number of specimens. Note that heteroplasmic specimens had more than one repeat unit; therefore, all heteroplasmic species had values >1.000.

sequences. The estimated thermal stability of the central repeat unit of *S. platorynchus* ($\Delta G = -7.3$) was similar to the other homoplasmic species. The other homoplasmic species had $\Delta G$ values $> -8.0$ while $\Delta G$ values $< -8.0$ were estimated for all heteroplasmic species. Interestingly, all deletions are located near the TAS sequences. Deletions upstream from the TAS motifs exclude the occurrence of stable secondary structures as observed in homoplasmic species. In contrast, deletions downstream from the TAS motifs seemed to not affect heteroplasmy, as demonstrated by the heteroplasmic *A. medirostris* and *A. mikadoi*.

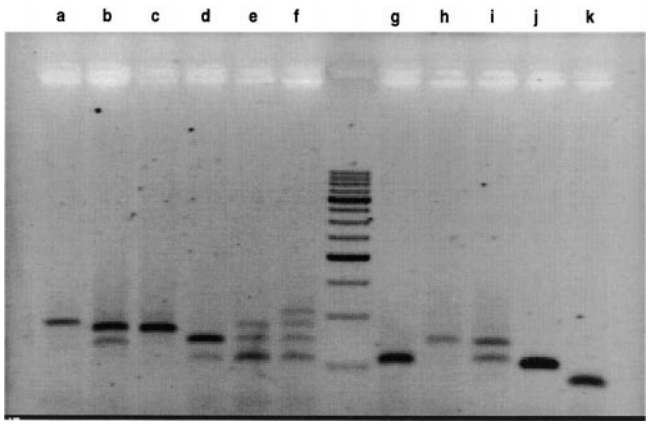



FIGURE 3.—Resolution of PCR products by agarose gel electrophoresis. Homoplasmic band pattern (number of repeats): k (1); g (2); j (2); h (3); a (4); and c (4). Heteroplasmic band pattern (number of repeats): i and d (2, 3); b (3, 4); e (3, 4, 5); and f (2, 3, 4, 5). The marker is a 250-bp ladder (Boehringer, Mannheim, Germany).

**Cyt-*b* data:** The length of the entire cyt-*b* sequence was 1141 bp, encoding 380 amino acids. We found 32 variable sites, most of which are silent mutations. SSCP analyses showed four haplotypes within *A. gueldenstaedtii* (AJ245826–7, AJ249692, and AJ277594). Sequencing results of these haplotypes revealed the following variable nucleotide positions: 195, 318, 342, 393, 429, 483, 498, 705, 840, 891, 893, and 1053 (numbers correspond to the positions in the sequence of the entire cyt-*b*), respectively. Three different haplotypes were observed within *A. persicus* (AJ245835–7), separated by the following variable positions: 654, 831, 1056, and 1058. Both populations of *A. naccarii* from the Po (AJ245834) and the Buna River (AJ245833) are separated by a point mutation (nucleotide position 562). In addition to the stored haplotypes of the genus Scaphirhynchus (GenBank U56983–8, U55994), four haplotypes of *S. platorynchus*, two haplotypes of *S. albus*, and one haplotype of *S. suttkusi* are described. Overall the following nucleotide positions were variable: 87, 111, 136, 303, 597, 891, and 1125. No interspecific substitution was found within the genus Scaphirhynchus or between *A. gueldenstaedtii* and *A. persicus*.

We used a MJ network to explore the phylogenetic relationships between these haplotypes (BANDELT *et al.* 1999). The full MJ networks (Figures 5 and 6) were almost entirely connected by one mutational step in most cases, but there were regions of ambiguity resulting from recurrent mutations at positions 342 and 498 in

**TABLE 3**

**Frequency of homoplasmic and heteroplasmic sturgeons, as well as frequency of copies observed in heteroplasmic species**

| Species | Homoplasmic | Heteroplasmic | | | |
|---|---|---|---|---|---|
| | | Overall | Two copies | Three copies | Four copies |
| *S. albus* | 1.000 | 0 | — | — | — |
| *S. platorynchus* | 1.000 | 0 | — | — | — |
| *S. suttkusi*[a] | 1.000 | 0 | — | — | — |
| *A. baerii* | 0.801 | 0.189 | 0.183 | 0.016 | 0 |
| *A. brevirostrum* | 0.500 | 0.500 | 0.333 | 0.167 | 0 |
| *A. gueldenstaedtii* | 0.857 | 0.143 | 0.113 | 0.020 | 0.010 |
| *A. fulvescens* | 0.935 | 0.064 | 0.032 | 0.032 | 0 |
| *A. fulvescens*[b] | 1.000 | 0 | 0 | 0 | 0 |
| *A. medirostris* | 0.459 | 0.541 | 0.189 | 0.352 | 0 |
| *A. medirostris*[b] | 0.500 | 0.500 | | overall 0.500 | |
| *A. mikadoi* | 0.166 | 0.834 | 0.334 | 0.500 | 0 |
| *A. naccarii* | 0.700 | 0.300 | 0.200 | 0.100 | 0 |
| *A. nudiventris* | 1.000 | 0 | — | — | — |
| *A. oxyrinchus o.* | 1.000 | 0 | — | — | — |
| *A. oxyrinchus o.*[b] | 1.000 | 0 | — | — | — |
| *A. oxyrinchus d.*[c] | 0.815 | 0.185 | 0.072 | 0.048 | 0.065 |
| *A. persicus* | 0.765 | 0.235 | 0.235 | 0 | 0 |
| *A. ruthenus* | 0.532 | 0.468 | 0.372 | 0.090 | 0.006 |
| *A. stellatus* | 0.742 | 0.258 | 0.226 | 0.032 | 0 |
| *A. sturio* | 1.000 | 0 | — | — | — |
| *A. transmontanus* | 0.833 | 0.166 | 0 | 0.166 | 0 |
| *A. transmontanus*[b] | 0.580 | 0.420 | | overall 0.420 | |
| *H. huso* | 0.920 | 0.080 | 0.080 | 0 | 0 |

[a] Data from EMBL (U43679).
[b] Data from Brown *et al.* (1996).
[c] Data from Miracle and Campton (1995).

Figure 5 that need to be identified rather than arbitrarily resolved.

## DISCUSSION

**Addition and deletion of repeat units:** The D-loop is the most variable region of fish mtDNA (Meyer 1993). However, the observed stable secondary structures imply a mechanism that can maintain perfect sequence identity despite the influence of rapid evolutionary change (Buroker *et al.* 1990). Several models have been proposed to be responsible for the occurrence of tandem repeat units in mtDNA: intra- and intermolecular recombination (Rand and Harrison 1989), transposition (Rand and Harrison 1989), slipped mispairing (Efstratiadis *et al.* 1980), and illegitimate elongation (Buroker *et al.* 1990). In contrast to plant mtDNA (Sederoff 1987), limited data for paternal inheritance or recombination of animal mtDNA exist. The low level of paternal mtDNA (0.01–0.1%) in interspecific mitochondrial congenic mice derived from backcrosses between *Mus musculus* and *M. spretus* (Gyllensten *et al.* 1991) and the exclusion of paternal mtDNA from the cytoplasm of intraspecific mice hybrid embryos at the late pronucleus stage (Kaneda *et al.* 1995) do not support an influence of paternal mtDNA on the existence of the heteroplasmy observed in our study. Moreover, paternal mtDNA was not found in any unfertilized eggs (Shitara *et al.* 1998). We suggest that this possibility, which has so far received little attention, should not be dismissed, because a great level of homoplasy and the occurrence of a rare point mutation in individuals of separated human mtDNA lineages could be an indication of recombination in animal mtDNA (Eyre-Walker *et al.* 1999; Hagelberg *et al.* 1999). However, ignoring length variations occurring in the D-loop, intermolecular recombination should yield in molecules of differing sizes as found in plant mtDNA. Intramolecular recombination need not alter molecule size and would thus be rather difficult to detect. It has been shown that DNA could be transferred into mitochondria by a protein (Vestweber and Schatz 1989). Although transposition among mitochondria theoretically could be assumed to account for length variation, this mechanism is not supported by the noninterrupted perfect tandem arrays we observed. Furthermore, transposition should not be responsible for both duplication and deletion, which is what we observed. The slipped mispairing model presented by

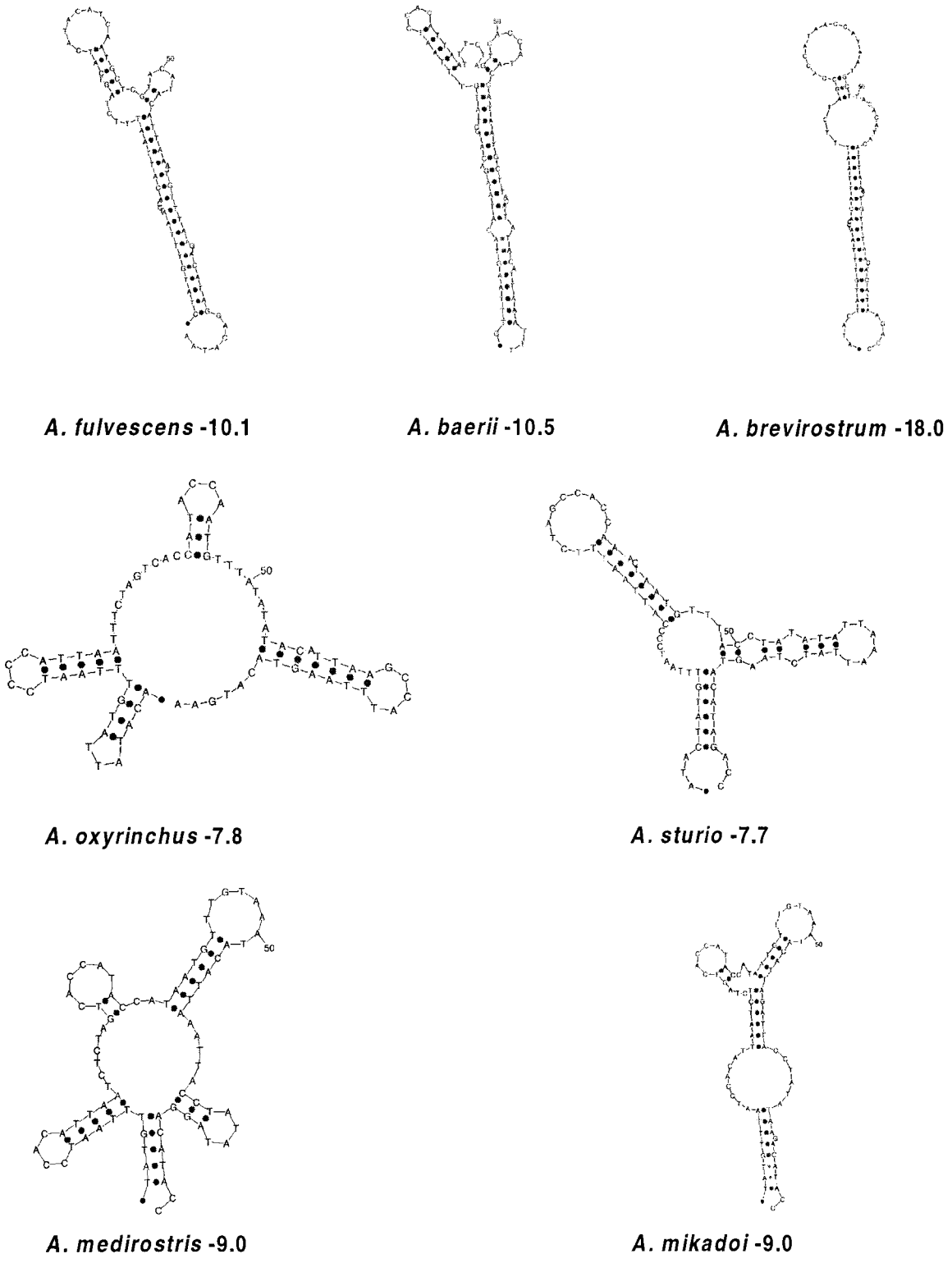
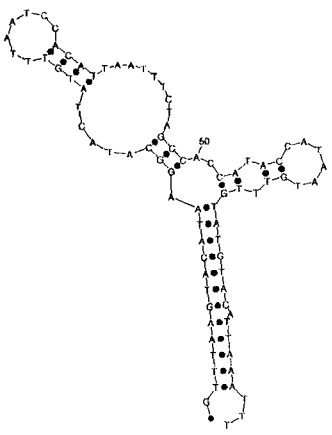
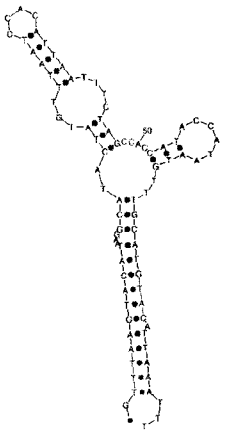



FIGURE 4.—Predicted secondary structures and their $\Delta G$ values (kilocalories per mole). Secondary structures were calculated using the DNA algorithm in the program MFOLD (WALTER *et al.* 1994). Sequences correspond to the L-strand.

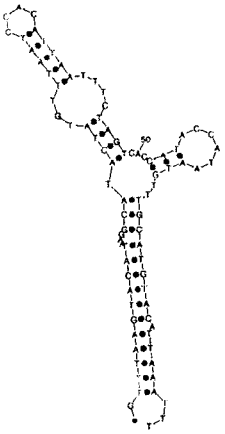
EFSTRATIADIS *et al.* (1980) explained only deletions. This is in contrast to the distribution of repeat units observed in this study (Figure 3). The illegitimate elongation model explains both deletion and duplication events during mtDNA replication (BUROKER *et al.* 1990). mtDNA replication is of triplex nature: the newly synthesized D-loop strand replaces the heavy strand (H-strand) from the light strand (L-strand) (CLAYTON 1982). The L-strand does not begin replicating until H-strand replication reaches the L-strand origin of replication. Thus,



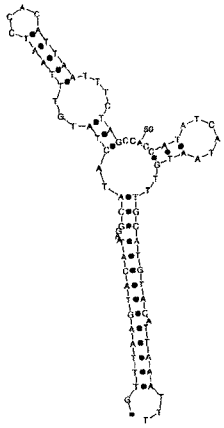
***A. persicus* -14.8**



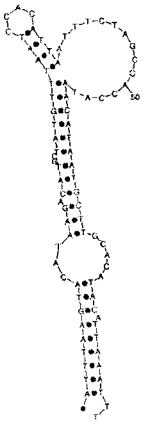
***A. gueldenstaedtii* -14.6**



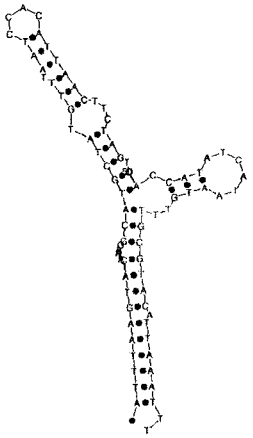
***A. naccarii* -15.8**



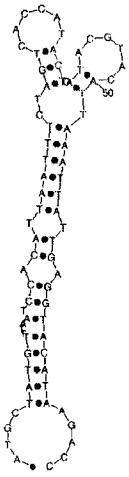
***A. nudiventris* -11.4**



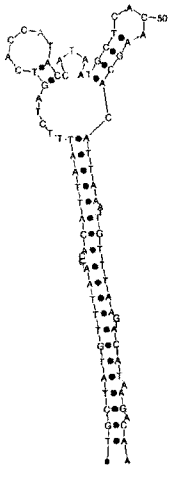
***A. stellatus* -10.9**



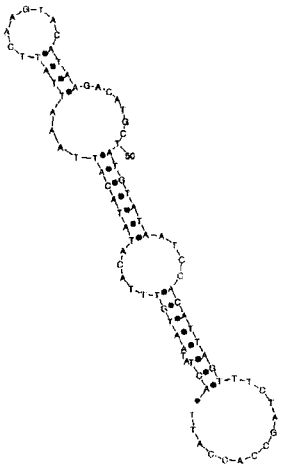
***A. ruthenus* -12.0**



***S. platorynchus* -7.3**



***A. transmontanus* -10.7**



***H. huso* -9.2**

FIGURE 4.—*Continued.*

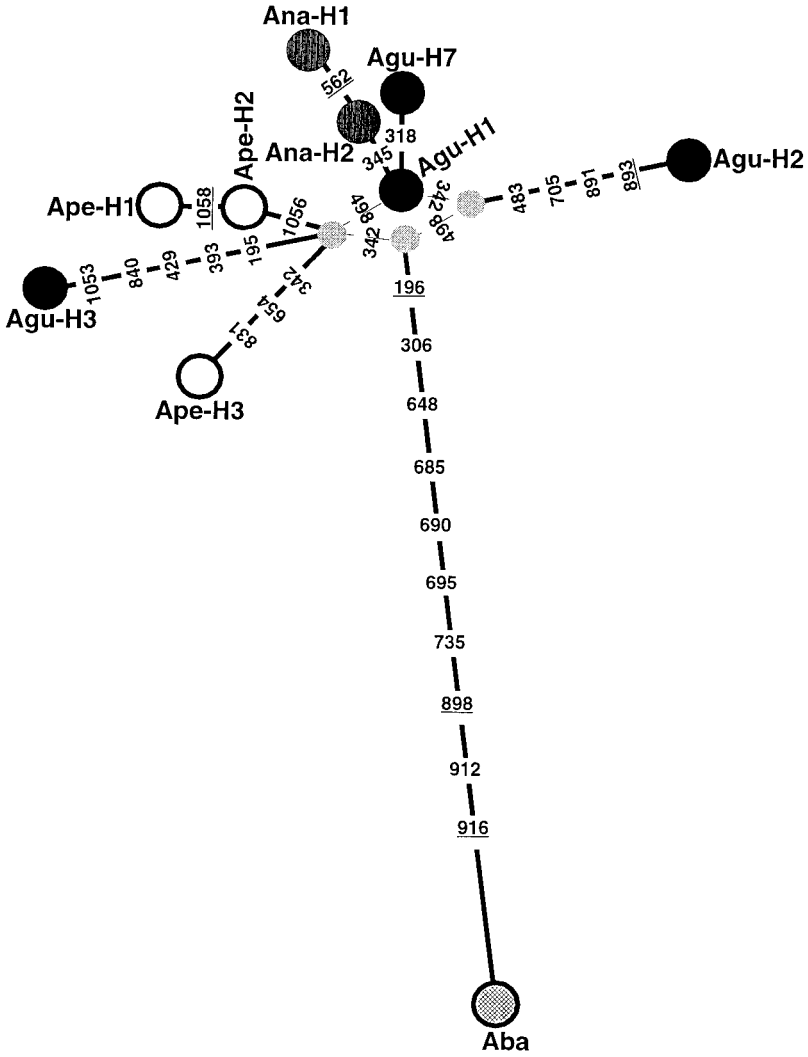



FIGURE 5.—Median-joining network constructed from 10 sturgeon cytochrome *b* sequences (1141 bp) of *A. baerii* (Aba, AJ245825), *A. gueldenstaedtii* (Agu-H7, H2/4, AJ245826/7, AJ249692, and AJ277594), *A. naccarii* (Ana-H1/2, AJ245833–4), and *A. persicus* (Ape-H1/3, AJ245835/7). Branches between haplotypes represent mutations; they are labeled by position of entire cyt-*b* gene. Positions of deduced amino acid exchanges are underlined. Ambiguous links are indicated by dotted lines. The full expanded network ($\varepsilon = 1$) harbors five equally most-parsimonious trees (29 steps of length) that were also observed in the maximum parsimony analysis of the same data set. The small shaded nodes represent possibly extant unsampled sequences or extinct ancestral sequences.

mtDNA replication is a unidirectional replication process (CLAYTON 1982); so the addition of a repeat is due to the formation of a hairpin structure in the partially displaced and therefore single-stranded, nascent H-strand, followed by re-invasion of this shortened strand and ongoing replication (BUROKER *et al.* 1990). Loss of a repeat results from re-invasion of a partially displaced nascent H-loop strand at a downstream copy of a repeat unit forcing both parental H- and L-strands into hairpin loops. In the illegitimate elongation model both gain and loss of repeat units occur during H-strand synthesis. BUROKER *et al.*'s (1990) model of gain or loss of repeat units proposed a dynamic competitive equilibrium between the D-loop strand and the H-strand during base pairing with the L-strand. This is, however, possible solely if there is at least one array in the D-loop capable of forming a stable secondary structure. That is what we found: all heteroplasmic species showed stable secondary structures of the central repeat units (Figure 7). Summarizing the distribution of repeat units in the 677 sturgeons of the 12 heteroplasmic species observed in this study, 1.9% had one repeat, 30.7% two repeat units, 64.7% three repeat units, 28.8% four repeat units, 6.9% five repeat units, 0.6% six repeat units, and 0.1% seven repeat units (Figure 7). Combining the results of all heteroplasmic sturgeons, we found that >90% had two and/or three repeat units. This indicates a selective mechanism within the distribution of tandem arrays because without such a driving force a more or less even magnitude of distribution of tandem arrays would be expected or theoretically the number of repeat units should be unlimited. Three copies of tandem arrays seemed to be favored in species showing heteroplasmy (Figure 7). Apart from the ratio of mutation rates increasing *vs.* decreasing the copy number, additional forces may act on the distribution of tandem arrays. Directional and purifying selection as proposed by RAND (1993) would lead to a small mtDNA genome size. On the one hand, a so-called race for replication could occur in a heteroplasmic cell. Thus, one mtDNA molecule that is smaller than another could have a substantial temporal advantage during the replication leading to a fixation of smaller mtDNA. Furthermore, a smaller mtDNA molecule is a smaller target for oxidative dam-

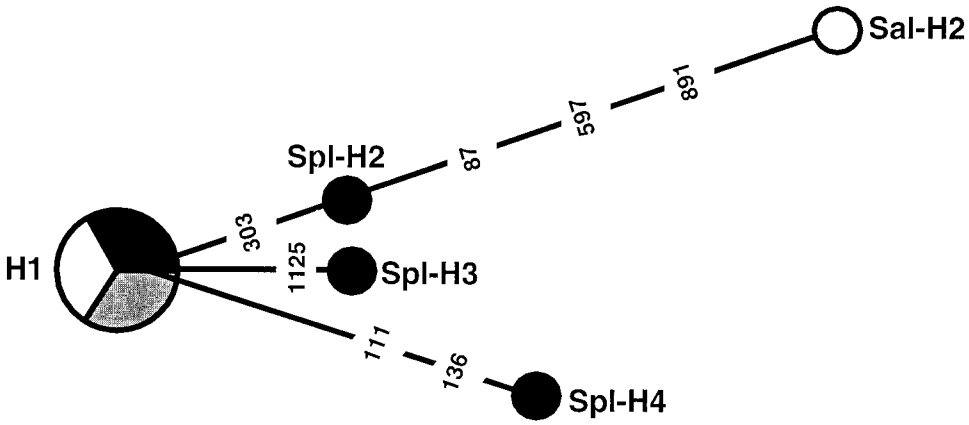



FIGURE 6.—Median-joining network constructed from seven sturgeon cytochrome *b* sequences (1141 bp) of *S. albus* (Sal-H1/2, U56983,7), *S. platorynchus* (Spl-H1/4, U56984/6,8), and *S. suttkusi* (H1, U55994). The haplotype H1 was observed in all three species (U56983–4, U55994). Branches between haplotypes represent mutations; they are labeled by position of entire cyt-*b* gene. No deduced amino acid exchanges and no ambiguous links were observed. The full expanded network ($\varepsilon = 1$) harbors one equally most-parsimonious tree (seven steps of length) that was also observed in the maximum parsimony analysis of the same data set.

age that could stall or block polymerization (RAND 1993). For sturgeons, BROWN *et al.* (1996) suggested another mechanism, the TAS-based replication model. In this model, *trans*-acting proteins are involved in the negative regulation of H-strand synthesis via binding to TAS motifs, which were found in each repeat. Approximately 5% of the initiated D-loop strands will be successfully duplicated during the replication (BOGENHAGEN and CLAYTON 1978). Therefore, multiple copies of TAS-bearing tandem arrays will decrease a successful H-strand replication substantially (BROWN *et al.* 1996) and thus serve as negative selection on the number of tandem arrays. Controversially, RAND (1993) proposed that positive selection pressure toward longer mtDNA molecules may exert stabilized secondary structures that may be favored at the initiation of replication.

In contrast to BUROKER *et al.*'s (1990) prediction that the minimum number of repeat units in a heteroplasmic fish is three, heteroplasmic sturgeons with one and especially with two repeat units were observed at a low level in *A. fulvescens*, *A. gueldenstaedtii*, *A. medirostris*, *A. mikadoi*, *A. naccarii*, *A. persicus*, *A. ruthenus*, *A. stellatus*, *A. transmontanus*, and *H. huso*. A secondary structure including the imperfect unit and one array repeat could have resulted in the deletion of one unit. mtDNA molecules with more than two repeat units were not observed in species showing no heteroplasmy (distribution of repeat units in homoplasmic species was one repeat, 6.6% and two repeat units, 93.4%).

**Evolution of heteroplasmy:** Comparison of the molecular phylogeny of sturgeons based, first, on the entire cyt-*b* gene, second, on the distribution of heteroplasmy, and third, on the length and structure of the repeat units, showed that it is more likely that the existence of heteroplasmy resulted from a single evolutionary event (Figure 2). Contrary explanations like the subsequent loss of heteroplasmy are not well supported by the molecular phylogeny (Figure 2). However, we observed no heteroplasmy within one in-group species: *A. nudiventris*. It is possible that the small sample size and/or the restricted geographic origin are reasons for blurring the existence of heteroplasmy in *A. nudiventris*.

The phylogenetic tree estimated from the cyt-*b* sequences is congruent with the chromosome number, the occurrence of heteroplasmy, and structure of the central repeat units (Figure 2). The assumed monophyly of the clade *A. sturio*/*A. oxyrinchus* (BIRSTEIN *et al.* 1997; CHOUDHURY and DICK 1998) is supported by a bootstrap value of 100%. Both species had a similar repeat length and showed no heteroplasmy in 44 specimens of *A. sturio* and 90 specimens of *A. oxyrinchus* in this study. The position of these species within the molecular phylogeny (Figure 2) refute the hypothesis that *A. oxyrinchus* lost heteroplasmy (BROWN *et al.* 1996). This is in agreement with species of the genus Scaphirhynchus. Both groups (*A. sturio*/*A. oxyrinchus* and *S. albus*/*S. suttkusii*/*S. platorynchus*) had reduced repeat units. Therefore, it is most likely that heteroplasmy evolved in the common ancestor of the remaining sturgeon species. Possibly the insertion of nucleotides upstream of the TAS motifs influence the secondary structure leading to a higher thermal stability and at least to the occurrence of heteroplasmy in most in-group species (*H. huso*, *A. stellatus*, *A. fulvescens*, *A. brevirostrum*, *A. baerii*, *A. persicus*, *A. naccarii*, *A. gueldenstaedtii*, and *A. transmontanus*). The absence of 8 bp in the repeat region of the genus Scaphirhynchus and of 2 bp in *A. sturio* and *A. oxyrinchus* (both closely related to the TAS motifs) influence the thermal stability and may thus prevent the presence of heteroplasmy in these species. However, a deletion of 2 bp upstream of the TAS motifs was also observed in one heteroplasmic species: *A. ruthenus*. *A. medirostris* (the length of the central repeat is 79 bp) and *A. mikadoi* (the length of the central repeat is 78 bp) showed the same insertion of 8 bp upstream of the TAS motifs and a deletion of 3 or 4 bp downstream from the TAS motifs. Our data showed

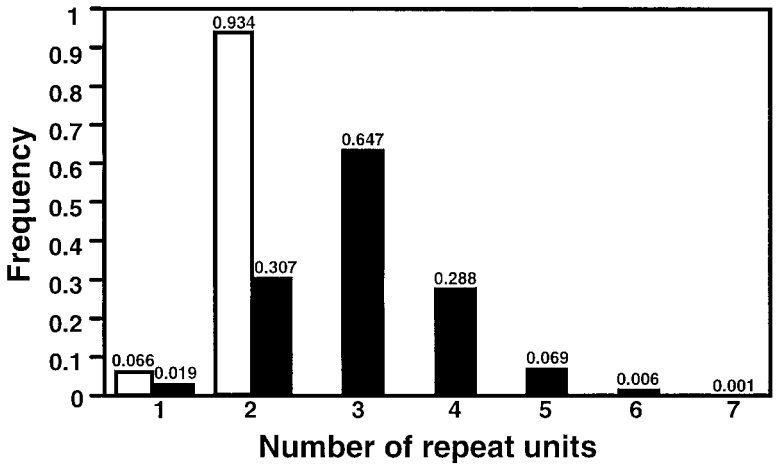



FIGURE 7.—Frequency of repeat units in homoplasmic species (open bar) and heteroplasmic species (solid bar) summarizing data from Table 2.

that these deletions do not influence the existence of heteroplasmy within these species. Deletions of 36 bp observed in the 5′ end of the D-loop in *A. brevirostrum*, *A. baerii*, *A. persicus*, *A. naccarii*, and *A. gueldenstaedtii* do not preclude the existence of heteroplasmy. Therefore, it is likely that the deletion of 24 bp seen in the 5′ end of the D-loop in *A. sturio* and *A. oxyrinchus* is also of no importance to the absence of heteroplasmy. The absence of heteroplasmy in these species as well as in the genus Scaphirhynchus may be supported by the $\Delta G$ value $>-8.0$. To the contrary, all heteroplasmic species had $\Delta G$ value $<-8.0$. Our conclusion is that a critical $\Delta G$value $<-8.0$ is a necessary prerequisite for the occurrence of heteroplasmy in sturgeons. (See Figure 4.)

**Molecular phylogeny:** Contrary to the well-known phylogenetic position of Acipenseriformes, phylogenetic relationships of species within this order are still controversial (MAYDEN and KUHAJDA 1996; BIRSTEIN and BEMIS 1997; FINDEIS 1997; BIRSTEIN and DESALLE 1998; CHOUDHURY and DICK 1998). Cladistic analyses of FINDEIS (1997) indicated that Acipenser is the sister taxon to the monophyletic lineage of Scaphirhynchus. Including the genus Pseudoscaphirhynchus, these three genera represent the subfamily Acipenerinae with Huso as its sister taxon (CHOUDHURY and DICK 1998). In contrast, on the basis of morphological and meristic characters, MAYDEN and KUHAJDA (1996) found an autapomorphy for Acipenser and postulated that Acipenser and Huso form a monophyletic clade that is separated from Scaphirhynchus and Pseudoscaphyrinchus. In a sequence data set combining different partial mitochondrial genes BIRSTEIN and DESALLE (1998) assigned *S. albus* a sister-specific status to all species of Acipenser and Huso. This is in agreement with our results based on sequence data of the entire cyt-*b* gene. As shown in Figure 2, the three species of Scaphirhynchus occurred as a sister clade to all other sturgeons.

Concerning the phylogenetic position of the genus Huso, the representative *H. huso* implemented in our study is clustered among species of genus Acipenser. This is in agreement with Linneaus (1758). BIRSTEIN and DESALLE (1998) also called the validity of Huso as a separated taxonomic unit into question. Therefore, the taxonomic status of Huso should be critically reevaluated in further comprehensive investigations.

*A. sturio* and *A. oxyrinchus* form a separate clade within the genus Acipenser and appear to be a sister clade with respect to all other sturgeon species. In the chronology of acipenseriform evolution, *A. sturio* and *A. oxyrinchus* form the oldest clade within the genus Acipenser. Using paleontological and molecular data the origin of the sturio-lineage was dated back 95 mya (BIRSTEIN and DESALLE 1998). Cyt-*b* data as well as the tree based on D-loop sequences provide evidence for a more distant relationship to the other species in the genus Acipenser (Figures 2 and 7).

Focusing on the cyt-*b* tree, a separation into two in-groups emerged. The first in-group contained the following species inhabiting the Atlantic region: *A. baerii*, *A. brevirostrum*, *A. fulvescens*, *A. gueldenstaedtii*, *A. naccarii*, *A. nudiventris*, *A. persicus*, *A. ruthenus*, *A. stellatus*, and *H. huso*, respectively. Several exchanges of the freshwater fish fauna (families Polyodontidae, Hiodontidae, and Catostomidae) took place between Siberia and North America (BANARESCU 1992). Some of these exchanges have an old Cretaous or Palaeogene origin. This could explain the inclusion of the North American species *A. brevirostrum* and *A. fulvescens* among the remaining species mentioned above occurring in Eurasia. Interestingly, *A. baerii* is restricted solely to Siberian rivers and Lake Baikal whereas the other species are found mainly in the Ponto-Caspian region. It might be assumed that the ancestral *A. baerii* could have migrated from the Ponto-Caspian area to Siberia or may have survived in the Beringian refugium. This refugium remained unglaciated during the last ice ages as a restricted area to Alaska and easternmost Siberia (BANARESCU 1992). Lakes connecting Siberian rivers emerged ~10,000 to 15,000 years ago (ARKHIPOV 1998). This riverine network may be responsible for the spread of *A. baerii* into

Siberia. The sympatric occurrence of *A. ruthenus*, a Ponto-Caspian species, in the Ob River supports this theory of a post-glacial newcomer.

The second in-group comprised the three species inhabiting the Pacific region: *A. transmontanus* (North America), *A. medirostris* (North America), and *A. mikadoi* (northern East Asia). *A. medirostris* and *A. mikadoi* are almost indistinguishable morphologically and were considered as members of one species (Artyukhin and Andronov 1990). A taxonomic separation of the North American conspecifics from the ones occurring at the northern East Asian Pacific coast resulted from chromosomal studies: *A. mikadoi* exhibited twice as many chromosomes as *A. medirostris* (Birstein *et al.* 1993; Birstein and Bemis 1997). Apart from two silent nucleotide exchanges observed within entire cyt-*b* sequences, both species differed by only one nucleotide insertion/deletion in the central repeat units. In contrast to the molecular phylogeny presented by Birstein and DeSalle (1998), in our study both species were closely related. The inclusion of *A. schrencki*, *A. sinensis*, and *H. dauricus* inhabiting the Pacific area may provide more detailed information concerning the phylogenetic relationships among sturgeons in future studies.

A comparison of the mtDNA tree of Birstein and DeSalle (1998) to the cyt-*b* tree obtained in our study revealed some differences. Russo *et al.* (1996) showed that phylogenetic trees reconstructed on the basis of the entire cyt-*b* gene were close to the true tree. However, in the tree presented by Birstein and DeSalle (1998) short partial sequences of different mitochondrial genes were used for reconstruction of phylogenetic relationships. Due to the fact that these genes evolve at different rates, such a tree may not reflect the most likely phylogeny. Additionally, our tree is well supported by both number of chromosomes and repeat lengths. The previous phylogeny of the Acipenseridae included five changes of the ploidy levels (120 → 250 → 500 → 250 → 120 → 250) (Birstein and DeSalle 1998). In contrast, our phylogeny was based on three changes only (120 → 250—Atlantic clade and 120 → 250 → 500—Pacific clade).

Phylogenetic analyses using D-loop sequences of nonconserved regions have to be handled with care (Figure 8). The high level of intraspecific variation of the flanking region of the repeat units and of interspecific variation of the repeat units itself implies that this part of the D-loop contains a large amount of homoplasmic characters and therefore should not be used in phylogenetic reconstructions of interspecific relationships. On the basis of the sequence data obtained in this study, the taxonomic status of the three species composed of *A. gueldenstaedtii*, *A. naccarii*, and *A. persicus* (Figure 5) and within the genus Scaphirhynchus (Figure 6) may be questionable. The high genetic similarity of cyt-*b* as well as D-loop sequences between *A. gueldenstaedtii*, *A. naccarii*, and *A. persicus* indicated a very recent separation of this species complex. The interruption of the connection between Black Sea and Adriatic Sea resulted in an isolation of the ancestral *A. naccarii* after the last glaciation ~8000 years ago (Zonneveld 1996). The classification of *A. gueldenstaedtii* and *A. persicus* as separate species is still under discussion (Birstein and Bemis 1997). No interspecific substitutions were found in the entire cyt-*b* or partial D-loop sequences. On the one hand, diagnostic morphological characters are rare and both species inhabit the same distribution areas. On the other hand, a more detailed view on the cyt-*b* sequences within this species complex showed a great intraspecific variability (Figure 5). Therefore, a critical examination of the *A. gueldenstadtii* complex seems to be necessary. While different intraspecific haplotypes were observed for *S. albus* and *S. platorynchus*, individuals of *S. suttkusi* showed no variability. There were no species-specific sequence differences between the three species within cyt-*b* sequences.

**Concluding remarks:** Repeat units were observed in all sturgeon species examined. Length of the repeat units ranged from 74 to 83 bp. The array and the associated length variations are general features of sturgeon mtDNA. In contrast, no similar repeat sequence was observed in *P. spathula*. One hundred and thirty-eight

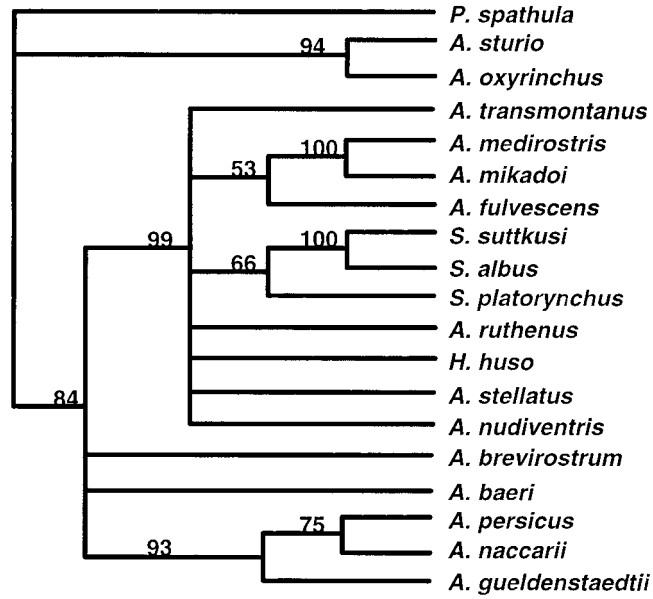



Figure 8.—Maximum parsimony tree (heuristic search, addition sequence algorithm) based on partial D-loop and $tRNA^{Pro}$ sequences calculated in PAUP* (Swofford 1998); bootstrap values of 1000 replications are shown on the branches. The consistency index was 0.707; the retention index was 0.687. Eight most-parsimonious trees were observed with a total length of 493 steps. In agreement with Figure 2 the statistical significance of the node separating the species pairs *S. albus*/*S. suttkusi*, *A. medirostris*/*A. mikadoi*, and *A. oxyrinchus*/*A. sturio* was ≥95% according to bootstrap analyses. Sequence of *S. suttkusi* was taken from the GenBank (U43679). The sequence of *P. spathula* was used as an outgroup.

interspecific mutations were observed among the repeat units. About 50% of these mutations are located ±10 bp down- and upstream from TAS sequences. Deletions were observed downstream from TAS sequences in most homoplasmic species (*A. sturio*, *A. oxyrinchus*, *S. albus*, *S. platorynchus*, and *S. suttkusi*). Deletions upstream from TAS sequences were found in *A. medirostris*, *A. mikadoi*, and *A. ruthenus* and do not seem to have precluded heteroplasmy. Calculations of secondary structures and thermal stabilities of repeat units lead to the conclusion that a $\Delta G$ value $<-8.0$ is necessary for heteroplasmy in sturgeons. The distribution of repeat units showed that three repeat units is the most common number in heteroplasmic fish, and no homoplasmic sturgeon with >2 repeat units was observed. Combining our results, the large data set shows that heteroplasmy correlates with a combination of several factors: $\Delta G$ values $<-8.0$ and a set of mutations especially near the TAS sequences. The existence of heteroplasmy in the ancestor of Acipenserformes is unknown. Further research activities should be addressed to analysis of heteroplasmy and D-loop structure in the family Polyodontidae, especially *Psephurus gladius*.

The authors recognize numerous staff members and professional fishermen from Albania, Canada, France, Georgia, Germany, Iran, The Netherlands, Romania, Russia, and the United States game fish agencies for logistic support, obtaining specimens, and permitting catches during protected spawning periods (P. Anders, E. N. Artyukhin, V. Birstein, H. Bolig, P. Bronzi, M. Chebanov, B. Elvira, J. Fletcher, J. Gessner, O. Goemann, R. Gross, T. Gulyas, H. Kincaid, F. Kirschbaum, S. LaPan, M. Lepage, N. Patriche, M. Pourkazemi, E. Rochard, C. Tracy, and P. Williot). Special thanks go to H.-J. Bandelt, C. Pitra and J. Felsenstein for the helpful comments to our statistics and phylogenetic reconstruction. We thank K. Annet, N. Belfiore, D. J. Grunwald, and two anonymous referees for comments on earlier drafts of this manuscript and J. Becker and C. Kaltwasser for technical assistance. The studies were possible only with grants from the DFG (KI 189/11-1 and ME 712/12) and the Federal Agency for Nature Conservation and the Ministry of Environment, Nature Conservation and Nuclear Safety, Germany (FKZ-No. 808 05 078).

## LITERATURE CITED

Arkhipov, S. A., 1998 Stratigraphy and palaegeography of the Sartan glaciation in west Siberia. Quart. Int. **45/46:** 29–42.

Àrnason, E., and D. M. Rand 1992 Heteroplasmy of short tandem repeats in mitochondrial DNA of Atlantic cod, *Gadus morhua*. Genetics **132:** 211–220.

Artyukhin, E. N., and A. E. Andronov, 1990 A morphological study of the green sturgeon *Acipenser medirostris* (Chondrostei, Acipenseridae), from the Tumnin (Datta) River and some aspects of the ecology and zoogeography of Acipenseridae. J. Ichthyol. **30:** 11–21.

Banarescu, P., 1992 Distribution and dispersal of freshwater animals in North-America and Eurasia, pp. 519–1091 in *Zoogeography of Fresh Waters*, edited by P. Banarescu. Aula-Verlag, Wiesbaden, Germany.

Bandelt, H.-J., P. Forster and A. Röhl, 1999 Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. **16:** 37–48.

Bassam, B. J., G. Caetano-Annolles and P. M. Gresshoff, 1991 Fast and sensitive silverstaining of DNA in polyacrylamide gels. Anal. Biochem. **196:** 80–83.

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

Bentzen, P., W. C. Leggett and G. G. Brown, 1988 Length and restriction site heteroplasmy in the mitochondrial DNA of American Shad (*Alosa sapidissima*). Genetics **118:** 509–518.

Birstein, V. J., and W. E. Bemis, 1997 How many species are there within the genus *Acipenser*? Environ. Biol. Fish **48:** 157–163.

Birstein, V. J., and R. DeSalle, 1998 Molecular phylogeny of acipenserinae. Mol. Phylogenet. Evol. **9:** 141–155.

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

Birstein, V. J., R. Hanner and R. DeSalle, 1997 Phylogeny of the acipenseriformes: cytogenetic and molecular approaches. Environ. Biol. Fish **48:** 127–155.

Bogenhagen, D., and D. A. Clayton 1978 Mechanism of mitochondrial DNA replication in mouse L-cells: introduction of superhelical turns into newly replicated molecules. J. Mol. Biol. **119:** 69–81.

Brown, J. R., A. T. Beckenbach and M. J. Smith, 1992 Mitochondrial DNA length variation and heteroplasmy in populations of white sturgeons (*Acipenser transmontanus*). Genetics **132:** 221–228.

Brown, J. R., K. Beckenbach, A. T. Beckenbach and M. J. Smith, 1996 Length variation, heteroplasmy and sequence divergence in the mitochondrial DNA of four species of sturgeon (*Acipenser*). Genetics **142:** 525–535.

Buroker, N. E., J. R. Brown, T. A. Gilbert, P. J. O'Hara, A. T. Beckenbach *et al.*, 1990 Length heteroplasmy of sturgeon mitochondrial DNA: an illegitimate elongation model. Genetics **124:** 157–163.

Choudhury, A., and T. A. Dick, 1998 The historical biogeography of sturgeons (Osteichthyes: Acipenseridae): a synthesis of phylogenetics, palaeontology and palaeogeography. J. Biogeogr. **25:** 623–640.

Clark, A. G., 1988 Deterministic theory of heteroplasmy. Evolution **42:** 621–626.

Clayton, D. A., 1982 Replication of animal mitochondrial DNA. Cell **28:** 693–705.

Efstratiadis, A., J. W. Posakony, T. Maniatis, R. W. Lawn, C. O'Connell *et al.*, 1980 The structure and evolution of the human β-globin gene family. Cell **21:** 653–668.

Eyre-Walker, A., N. H. Smith and J. Maynard Smith, 1999 How clonal are human mitochondria? Proc. R. Soc. Lond. Ser. B **266:** 477–483.

Findeis, E. K., 1997 Comperative osteology and relationships of recent sturgeons. Environ. Biol. Fish **48:** 73–126.

Gardiner, B. G., 1984 Sturgeons as living fossils, pp. 148–152 in *Living Fossils*, edited by N. Eldredge and S. M. Stanley. Springer Verlag, New York.

Gyllensten, U., D. Wharton, A. Josefsson and A. C. Wilson, 1991 Paternal inheritance of mitochondrial DNA in mice. Nature **352:** 255–257.

Hagelberg, E., N. Goldman, P. Lio, S. Whelan, W. Schiefenhövel *et al.*, 1999 Evidence for mitochondrial DNA recombination in a human population of island Melanesia. Proc. R. Soc. Lond. Ser. B **266:** 485–492.

Kaneda, H., J.-I. Hayashi, S. Takahama, C. Taya, K. Fischer Lindahl *et al.*, 1995 Elimination of paternal mitochondrial DNA in intraspecific crosses during early mouse embryogenesis. Proc. Natl. Acad. Sci. USA **92:** 4542–4546.

Kumar, S., K. Tamura and M. Nei, 1993 MEGA: Molecular Evolutionary Genetics Analysis, Version 1.01. Pennsylvania State University, University Park.

Ludwig, A., and I. Jenneckens, 2000 A PCR test for heteroplasmy in sturgeon. Anim. Genet. **31:** 153–154.

Mayden, R. L., and B. R. Kuhajda, 1996 Systematics, taxonomy, and conservation status of the endangered Alabama sturgeon, *Scaphirhynchus suttkusi* Williams and Clemmer (Actinopterygii, Acipenseridae). Copeia **1996:** 241–273.

Meyer, A., 1993 Evolution of mitochondrial DNA of fishes, pp. 1–38 in *The Biochemistry and Molecular Biology of Fishes*, edited by P. W. Hochachka and P. Mommsen. Elsevier Press, Amsterdam.

Miracle, A. L., and D. E. Campton, 1995 Tandem repeat sequence variation and length heteroplasmy in the mitochondrial DNA D-Loop of the threatened Gulf of Mexico sturgeon, *Acipenser oxyrhynchus desotoi*. J. Hered. **86:** 22–27.

Nesbø, C. L., M. O. Arab and K. S. Jakobsen, 1998 Heteroplasmy,

length and sequence variation in the mtDNA control regions of three percid fish species (*Perca fluviatilis*, *Acerina cernua*, *Stizostedion lucioperca*). Genetics **148:** 1907–1919.

Orita, M., T. Suzuki, T. Sekiya and K. Hayashi, 1989 Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. Genomics **5:** 874–879.

Rand, D. M., 1993 Endotherms, ectotherms, and mitochondrial genome-size variation. J. Mol. Evol. **37:** 281–295.

Rand, D. M., and R. G. Harrison, 1989 Molecular population genetics of mtDNA size variation in crickets. Genetics **121:** 551–569.

Russo C. A., N. Takezaki and M. Nei, 1996 Efficiencies of different genes and different tree-building methods in recovering a known vertebrate phylogeny. Mol. Biol. Evol. **13:** 525–536.

Sederoff, R. R., 1987 Molecular mechanism of mitochondrial genome evolution in higher plants. Am. Nat. **130:** S30–S45.

Shitara, H., J.-I. Hayashi, S. Takahama, H. Kaneda and H. Yonekawa, 1998 Maternal inheritance of mouse mtDNA in interspecific hybrids: segregation of the leaked paternal mtDNA followed by the prevention of subsequent paternal leakage. Genetics **148:** 851–857.

Swofford, D. L., 1998 *PAUP* (Phylogenetic Analyses Using Parsimony): Preliminary Test Version.* Sinauer, Sunderland, MA.

Thompson, J. D., D. G. Higgins and T. J. Gibson, 1994 CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:** 4673–4680.

Van Eenennaam, A. L., J. D. Murray and J. F. Medrano, 1998 Mitotic analysis of the North American white sturgeon, *Acipenser transmontanus* Richardson (Pisces, Acipenseridae), a fish with a very high chromosome number. Genome **41:** 266–271.

Vestweber, D., and G. Schatz, 1989 DNA-protein conjugates can enter mitochondria via the protein import pathway. Nature **338:** 170–172.

Walter, A. E., D. H. Turner, J. Kim, M. H. Lyttle, P. Muller *et al.*, 1994 Coaxial stacking of helixes enhances binding of oligoribonucleotides and improves predictions of RNA folding. Proc. Natl. Acad. Sci. USA **91:** 9218–9222.

Zonneveld, K. A. F., 1996 Paleoclimatic reconstruction of the last deglaciation (18-8 ka B.P.) in the Adriatic Sea region; a land-sea correlation based on palynological evidence. Palaeogr. Palaeoclimatol. Palaeoecol. **122:** 89–106.

Zuker, M., 1989 On finding all suboptimal foldings of an RNA molecule. Science **244:** 48–52.

Communicating editor: D. J. Grunwald