Age and biogeography of major clades in sturgeons and paddlefishes (Pisces: Acipenseriformes)

Article *in* Molecular Phylogenetics and Evolution · April 2007

DOI: 10.1016/j.ympev.2006.09.008 · Source: PubMed

6 authors, including:

Zuogang Peng
Southwest University in Chongqing
106 PUBLICATIONS **2,683** CITATIONS
SEE PROFILE

Arne Ludwig
Leibniz Institute for Zoo and Wildlife Research & Humboldt University Berlin
220 PUBLICATIONS **5,561** CITATIONS
SEE PROFILE

Dengqiang Wang
yfi
95 PUBLICATIONS **630** CITATIONS
SEE PROFILE

Rui Diogo
Howard University
351 PUBLICATIONS **4,723** CITATIONS
SEE PROFILE

Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 42 (2007) 854–862
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Short communication

Age and biogeography of major clades in sturgeons and paddlefishes (Pisces: Acipenseriformes)

Zuogang Peng [a,1], Arne Ludwig [b,1], Dengqiang Wang [c], Rui Diogo [d], Qiwei Wei [c], Shunping He [a,*]

[a] *Laboratory of Fish Phylogenetics and Biogeography, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China*
[b] *Department of Evolutionary Genetics, Institute for Zoo and Wildlife Research, Alfred-Kowalke-Str. 17, 10315 Berlin, Germany*
[c] *Key Laboratory of Freshwater Fish Germplasm Resources and Biotechnology, Ministry of Agriculture of China, Yangtze River Fisheries Research Institute, Chinese Academy of Fisheries Science, Jingzhou 434000, China*
[d] *Laboratory of Functional and Evolutionary Morphology, University of Liege, B-4000 Liege, Belgium*

Received 22 January 2006; revised 11 September 2006; accepted 18 September 2006
Available online 27 September 2006

1. Introduction

The order Acipenseriformes includes 25 extant sturgeon species and two extant paddlefish species (e.g., Birstein, 1993; Birstein and Bemis, 1997). Sturgeons and paddlefishes are considered "living fossils"; in fact, although acipenseriform fishes first appeared in the fossil record approximately 200 Mya, they have seemingly not undergone much morphological change since that time (Gardiner, 1984). These "evolutionary relicts" display markedly disjunction distributions with a wide distribution in the northern hemisphere (e.g., Gardiner, 1984; Bemis et al., 1997). Their unique benthic specializations and conserved morphology, their evolutionary age, the variation in their basic diadromous life history, and the large public interest due to their near extinction or critically endangered status, make sturgeons and paddlefishes interesting groups for molecular evolutionary studies (e.g., Birstein, 1993; Choudhury and Dick, 1998).

Moreover, acipenseriform fishes provide an ideal case study for examination of genome duplication events. Multiple levels of ploidy exist among these fishes since they include, for instance, diploidy (species with ~120 chromosomes including all taxa with between 110 and 130 chromosomes, e.g., *Acipenser nudiventris*, *Acipenser oxyrinchus*, *Acipenser ruthenus*, *Acipenser stellatus*, *Acipenser sturio*, *Huso dauricus*, *Huso huso*, *Scaphirhynchus albus*, *Scaphirhynchus platorynchus*, *Scaphirhynchus suttkusi*, *Psephurus gladius*, and *Polyodon spathula*) and tetraploidy (species with ~250 chromosomes including all taxa with between 220 and 276 chromosomes, e.g., *Acipenser baerii*, *Acipenser fulvescens*, *Acipenser gueldenstaedtii*, *Acipenser medirostris*, *Acipenser persicus*, *Acipenser schrenckii*, and *Acipenser transmontanus*) as well as evolutionary octaploidy (species with ~500 chromosomes, e.g., *Acipenser mikadoi*) (reviewed in Birstein et al., 1997). These fishes represent an interesting model for studying lineage specific genome duplication events. Furthermore, polyploidization is believed to be an important evolutionary aspect in animals, as it is in plants (e.g., Soltis and Soltis, 1999; Le Comber and Smith, 2004; Crow et al., 2006). However, the timing of the polyploid events in these acipenseriform fishes remains uncertain.

The phylogenetic relationships among extant acipenseriform groups has been the subject of numerous, detailed works (e.g., Grande and Bemis, 1991; Birstein and DeSalle, 1998; Ludwig et al., 2001; Robles et al., 2004). With a robust hypothesis on the relationships among these groups being available, it is of much interest to date the origin of these groups since this is crucial, in fact, for clarifying acipenseriform evolution and historical biogeography. As stated above, sturgeons and paddlefishes are an interesting group for biogeographical analysis: their wide distribution across North America and Eurasia, for example, raises the question on where the centre of origin is, as well as on whether such wide distribution was mainly due to dispersal, to vicariance, or to a combination of both.

* Corresponding author. Fax: +86 27 68780071.
E-mail addresses: weiqw@yfi.ac.cn (Q. Wei), clad@ihb.ac.cn (S. He).
[1] These authors contributed equally to this work.

1055-7903/$ - see front matter
doi:10.1016/j.ympev.2006.09.008

Therefore, the main aim of the present work is precisely to date the origin of the major groups of the order Acipenseriformes, in order to discuss these and other questions related with the genome duplication events and the biogeographic distribution of these fishes.

2. Materials and methods

2.1. Taxon sampling, DNA extraction, long PCR, and sequencing by primer walking

The Yangtze sturgeon (*Acipenser dabryanus*), Chinese swordfish (*P. gladius*) and Mississippi paddlefish (*P. spathula*) samples were conservation genetic analysis collections deposited in Yangtze River Institute of Fisheries, Chinese Academy of Fisheries Science. The other samples were retrieved from GenBank (Table 1). Total genomic DNA was extracted from the muscle tissue using a QIAamp tissue kit (Qiagen) following the manufacturer's protocol. The mitochondrial genome DNAs of the three acipenseriforms were amplified in its entirety using a long PCR technique (Cheng et al., 1994; Miya and Nishida, 1999). The primers designed to amplify the total mitochondrial genome are given in Appendix Table S1. Long PCR was done in a PTC-100 programmable thermal controller (MJ Research, USA); reactions were carried out in 25 ul reaction volume containing 2.5 ul 10 X LA PCR buffer II (Takara), 0.8 mM dNTPs, 2.5 mM $MgCl_2$, 0.5 um each primer, 0.625 units LA Taq polymerase (Takara) and approximately 20 ng template DNA. The thermal cycle profile was: pre-denaturation at 96 °C for 1 min, and 30 cycles of denaturation at 95 °C for 30 s, annealing at 60 °C for 50 s, extension at 72 °C for 3 min, final extension at 72°C for 15 min, and were electrophoresed on a 0.8% agarose gel (Promega). The long PCR products were diluted in sterilized distilled water for subsequent use.

Long PCR products were purified with a Cycle-Pure Kit (Omega) and subsequently used for direct cycle sequencing with dye-labelled terminators (ABI). Methods of sequencing were the same as above. Initial sequencing primers were the same as those for long PCR, and subsequent sequencing was performed using primers designed with reference to previously-determined sequences (primer-walking). We designed 64 sequencing primers (except for long PCR primers; available from ZP upon request) for determining the complete mtDNA sequences of the three acipenseriforms.

2.2. Sequence analysis and molecular dating

DNA sequences were analyzed using the software Lasergene version 5.0 (DNASTAR). Contig assembly was performed with the program Seqman. Multiple alignments were prepared for two rRNAs and eight protein-coding genes by using CLUSTAL X version 1.83 (Thompson et al., 1997) at default settings, excluding 22 tRNAs, ND3, ND4L, and ATP8, due to their small size (e.g., Peng et al., 2006). The CO2 gene was also excluded in our analyses due to its unavailable in alligator gar (*Lepisosteus spatula*). The nucleotide alignment for ND6 was unreliable due to an unusual substitution pattern (e.g., Yoder and Yang, 2000), and was thus also removed from the analysis. To avoid bias in refining alignments, ambiguous alignment positions (including beginnings and ends of many protein-coding genes and rRNA highly variable regions) were excluded by using the GBLOCKS 0.91b (Castresana, 2000) at default settings.

We have applied two steps for dating the divergence time of sturgeons and paddlefishes. First, we reconstructed

Table 1
Seven complete mitochondrial genome sequences of sturgeons and paddlefishes and related taxa plus one shark as outgroup retrieved from GenBank for the first dating analysis

Family	Species	Common name	GenBank Accession No.	References
Heterodontidae	*H. francisci*	Horn shark	AJ310141	Arnason et al. (2001)
Lepidosirenidae	*Lepidosiren paradoxa*	South American lungfish	AF302934	Brinkmann et al. (2004)
Pelomedusidae	*Pelomedusa subrufa*	African side-necked turtle	AF039066	Zardoya and Meyer (1998)
Bovidae	*Bos taurus*	Cow	V00654	Anderson et al. (1982)
Cercopithecidae	*Macaca mulatta*	Rhesus monkey	AY612638	Gokey et al. (2004)
Hominidae	*Pan troglodytes*	Chimpanzee	D38113	Hixson and Brown (1986)
Hominidae	*Homo sapiens*	Human	V00662	Anderson et al. (1981)
Amiidae	*A. calva*	Bowfin	AB042952	Inoue et al. (2003)
Lepisosteidae	*L. spatula*	Alligator gar	AP004355	Inoue et al. (2003)
Lepisosteidae	*Lepisosteus oculatus*	Spotted gar	AB042861	Inoue et al. (2003)
Acipenseridae	*A. dabryanus*	Yangtze sturgeon	AY510085	This study
Acipenseridae	*Acipenser stellatus*	Stellate sturgeon	AJ585050	Arnason et al. (2004)
Acipenseridae	*Acipenser transmontanus*	White sturgeon	AB042837	Inoue et al. (2003)
Acipenseridae	*H. huso*	Beluga	AY442351	Dunn et al. unpublished
Acipenseridae	*Scaphirhynchus cf. albus*	Pallid sturgeon	AP004354	Inoue et al. (2003)
Polyodontidae	*P. spathula*	Mississippi paddlefish	AY510086	This study
Polyodontidae	*P. gladius*	Chinese swordfish	AY571339	This study
Osteoglossidae	*Osteoglossum bicirrhosum*	Silver arawana	AB043025	Inoue et al. (2001)
Pantodontidae	*Pantodon buchholzi*	Freshwater butterflyfish	AB043068	Inoue et al. (2001)
Catostomidae	*Carpiodes*	River carpsucker	AY366087	Broughton et al. unpublished
Catostomidae	*Myxocyprinus*	Chinese sucker	AY986503	Peng et al. (2006)

the phylogenetic relationships of the order Acipenseriformes with some other basal jawed vertebrates using complete mitochondrial genome data retrieved from GenBank (Table 1). The maximum likelihood (ML) analysis as implemented in PAUP* 4.0b10 (Swofford, 2002) with TBR branch swapping (10 random addition sequences). The best-fitting nucleotide substitution model was selected by using Modeltest 3.7 (Posada and Crandall, 1998; Posada and Buckley, 2004). Bayesian inference of likelihood was implemented by using MrBayes version 3.1 (Ronquist and Huelsenbeck, 2003) with partitioned strategies. Under the partitioning strategy, the data set was divided into 10 partitions: Two rRNAs and eight protein-coding genes. The best-fitting nucleotide substitution models for each of the 10 partitions were selected by using the Akaike Information Criterion (AIC) implemented in MrModeltest version 2.2 (Nylander et al., 2004). Metropolis-coupled Markov chain Monte Carlo (MCMC) analyses (with random starting trees) were run with one cold and three heated chains (temperature set to default 0.2) for one million generations and sampled every 100 generations. To ensure that our analyses were not trapped in local optima, four independent MCMC runs were performed. Topologies and posterior clade probabilities from different runs were compared for congruence.

Then we estimated the origin time of Acipenseriformes and the time for splitting the sturgeons and paddlefishes. The latter combined with fossil records was used as upper time for the next dating analysis. Second, we used Ludwig et al.'s (2001) phylogeny among extant acipenseriform species to estimate the divergence time among these groups using mitochondrial cytochrome *b* gene sequences data retrieved from GenBank (Table 2—including inferred ploidy levels of acipenseriforms). Bayesian approaches (Thorne et al., 1998; Thorne and Kishino, 2002), one of the relaxed molecular clock methods, were used for the two dating analyses, the setting for Bayesian approaches were as follows.

Divergence times were determined using a Bayesian approach that incorporates variation of rates of evolution among genes and among lineages (Thorne et al., 1998; Thorne and Kishino, 2002), which were estimated with MULTIDIVTIME program (available from J. Thorne). This parametric approach relaxes the assumption of a strict

Table 2
Twenty-four mitochondrial cytochrome *b* gene sequences used for dating the divergence time of acipenseriform species retrieved from GenBank for the second dating analysis

Species	Distribution area	GenBank Accession No.	Distribution area code[a]	Ploidy level[b]
Family Polyodontidae				
Genus *Psephurus*				
P. gladius	Yangtze River basin, China	AY571339	D	2n
Genus *Polyodon*				
P. spathula	Mississippi–Missouri basin	AY510086	I	2n
Family Acipenseridae				
Genus *Acipenser*				
A. baerii	Rivers of north coast of Russia	AJ245825	C	4n
A. brevirostrum	North America-Atlantic coast	AJ245828	G	4n or 8n
A. dabryanus	Yangtze River basin, China	AY510085	D	4n
A. fulvescens	North America-Central United States	AJ245829	H	4n
A. gueldenstaedtii	Ponto-Caspian species	AJ245826	A	4n
A. medirostris	North America-Pacific coast	AF184105	F	4n
A. mikadoi	Siberia-Pacific coast	AJ245831	B	8n
A. naccarii	Adriatic Sea	AJ245833	A	4n
A. nudiventris	Ponto-Caspian species	AJ245832	A	2n
A. oxyrinchus	North America-Atlantic coast	AJ245838	GHI	2n
A. persicus	Ponto-Caspian species	AJ245835	A	4n
A. ruthenus	Ponto-Caspian area and some rivers of north coast of Russia	AJ249694	ACE	2n
A. schrenckii	Amur River drainage, Sea of Okhotsk	AJ251451	B	4n
A. sinensis	China, South Japan	AJ252186	D	4n
A. stellatus	Ponto-Caspian species	AJ585050	A	2n
A. sturio	West European Atlantic coast	AJ428497	AE	2n
A. transmontanus	North America-Pacific coast	AB042837	F	4n
Genus *Huso*				
H. huso	Ponto-Caspian species	AY442351	A	2n
H. dauricus	Amur River drainage	AJ252187	B	2n
Genus *Scaphirhynchus*				
S. albus	Mississippi–Missouri basin	AP004354	I	2n
S. platorynchus	Mississippi–Missouri basin	U56986	I	2n
S. suttkusi	Mississippi–Missouri basin	U55994	I	Unknown

[a] Nine distribution areas were coded from A to I according to the analysis of Bemis and Kynard (1997): (A) PC, Ponto-Caspian region; (B) ASJ, Amur River, Sea of Okhotsk and Sea of Japan; (C) SAO, Siberia and Arctic Ocean; (D) CH, China; (E) NEA, Northeastern Atlantic; (F) NEP, North Eastern Pacific; (G) NWA, North Western Atlantic; (H) GL, Great Lakes; (I) MGM, Mississippi River and Gulf of Mexico.

[b] Inferred ploidy levels of acipenseriforms were taken from Zhang et al. (1999), Ludwig et al. (2001), and references therein.

molecular clock with a continuous autocorrelation of substitution rates across the phylogeny, and allows the use of several calibrations/time constraints. This approach involved two steps. First, ESTBRANCHES was run to estimate branch lengths from the data and a fixed tree topology using the F84 + G model of sequence evolution. This allows rates to vary among sites following a discrete gamma distribution with four rate categories (Yang, 1994) along with their variance-covariance matrix. Parameters for the F84 + G model were estimated using the BASEML program in PAML (Yang, 2000).

Next, the species used as outgroups for the first and second dating analysis were horn shark (*Heterodontus francisci*), and bowfin (*Amia calva*), respectively. MULTIDIVTIME was used both for 10 partitions of mitochondrial genome sequences of the first dating analysis with estimating the prior and posterior ages of branching events, their standard deviations, and the 95% credibility intervals via Markov chain Monte Carlo, and for three partitions of cytochrome *b* gene sequences of the second dating analysis. The Markov chain was run for 200,000 generations and sampled every 100 generations after an initial burn-in period of 200,000 cycles for all the analyses. Prior gamma distributions on three parameters of the relaxed clock model were assumed and specified through the mean and SD of the root age, root rate, and rate autocorrelation; 528 Mya (SD = 528 Mya) (first dating analysis), and 200 Mya (SD = 200 Mya) (second dating analysis) for the expected time between tips and root without node time constraints, 0.01 (SD = 0.01) substitutions per site per million years for the rate at the root node, 0.4 (SD = 0.4) for the parameter (v) that controls the degree of rate autocorrelation per Mya along the descending branches of the tree. To check for convergence of the MCMC, analyses were run from at least two different starting points.

The MULTIDIVTIME program allows for both minimum and maximum fossil constraints. Whereas minima are often based on the earliest occurrences in the fossil record, maxima are intrinsically more difficult to estimate. So in the first dating analysis, we used the estimated divergence time between tetrapods and bony fishes versus cartilaginous fishes (528 Mya) for the upper time between tip and root node (C4) (Appendix Table S2, Appendix Fig. S1), and in the second dating analysis, we used the above estimated split time between sturgeons and paddlefishes combined with fossil records (ca. 200 Mya) for the upper time between tip and root node (Appendix Table S3, Fig. 1).

2.3. Biogeography and ancestral areas reconstruction

To assign distributions to the internal nodes in the tree, the dispersal–vicariance analysis (DIVA; Ronquist, 1996; Ronquist, 1997) was used. The program optimizes distributions for each node of the tree by favouring vicariance events and minimizing the number of assumed dispersals and extinctions. Between the nodes of the given tree DIVA assigns a cost to changes in distribution interpreted as extinctions or dispersals but no cost to changes interpreted as vicariance. Optimizations minimizing the cost are favoured. The assigned distributions were then compared with the age estimates and if possible correlated to geological history (see below). Nine main areas of distribution were used and coded from A to I according to the results of Bemis and Kynard (1997): (A) PC, Ponto-Caspian region; (B) ASJ, Amur River, Sea of Okhotsk and Sea of Japan; (C) SAO, Siberia and Arctic Ocean; (D) CH, China; (E) NEA, North Eastern Atlantic; (F) NEP, North Eastern Pacific; (G) NWA, North Western Atlantic; (H) GL, Great Lakes; (I), MGM, Mississippi River and Gulf of Mexico.

When the ancestral distribution was determined by DIVA to be composed of all, or nearly all, possible areas used to define the distribution of acipenseriforms, the "maxareas" option was used to limit the range of ancestral distributions to two areas. This approach determines which is the most likely ancestral distribution if that distribution was restricted to small areas (Ronquist, 1996).

3. Results

3.1. Divergence Time Estimates

The complete mitochondrial L-strand nucleotide sequences of the three acipenseriforms (*A. dabryanus*, *P. gladius*, and *P. spathula*) have been registered in GenBank under the accession numbers AY510085, AY571339, and AY510086, respectively. The organization of the three acipenseriform mitochondrial genome followed that of generalized bony fish and vertebrate mitochondrial genomes. Each of them consists of 13 protein-coding, 22 tRNA, two rRNA genes, and a control region. All genes were of similar length to those in other bony fishes and the gene order was identical to that so far obtained in many other vertebrates.

Maximum likelihood and Bayesian analyses resulted in the same tree topology with slightly supporting value difference based on 11,009 bp mitogenome sequence data set (Appendix Fig. S1). All of the nodes of the Bayesian tree received >96% posterior probabilities.

As explained in the section Materials and methods, divergence times for the nodes in the tree (see Appendix Fig. S1) were estimated using a Bayesian relaxed molecular clock method. The Bayesian approach estimated the origin time for Acipenseriformes at 389.7 Mya with a 95% credibility interval of 361.5–414.2 Mya, for concatenated mitochondrial genome sequence data set. The results were similar to the results of 312.1 Mya for data set #1 and 346.9 Mya for data set #2 of Inoue et al.'s (2005)work. The estimated time for splitting between sturgeons and paddlefishes is at 141.4 Mya with a 95% credibility interval of 132.4–159.5 Mya for the concatenated sequence data set. These results were similar to the results of 114.1 Mya for data set #1 and 145.2 Mya for data set #2 in Inoue et al. (2005).

As also noted in the section Materials and methods, we used the above estimated divergence time (ca. 141 Mya) and the oldest acipenseriform fossil record avaialable

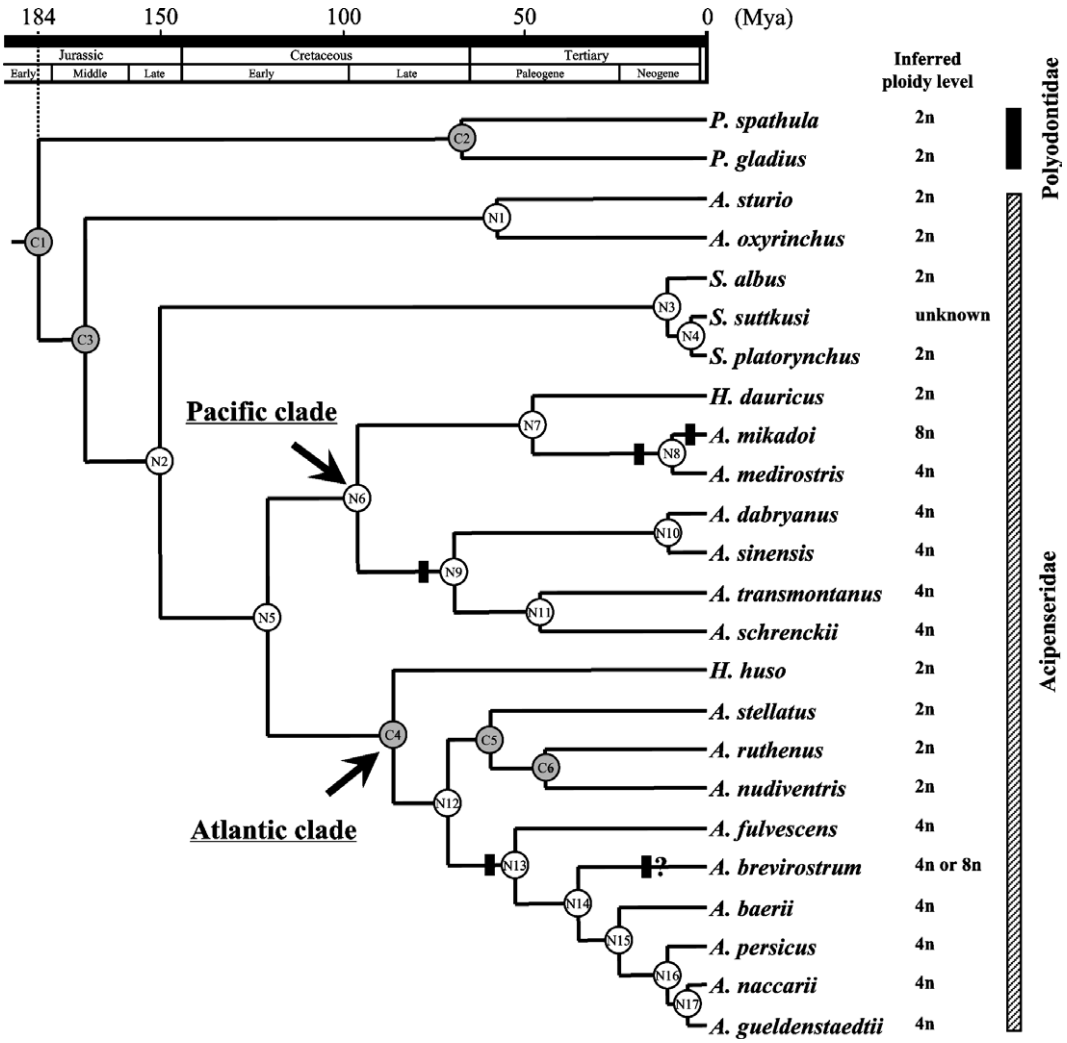

Fig. 1. Rooted ingroup tree assumed in the partitioned Bayesian analysis (Thorne and Kishino, 2002) for cytochrome *b* gene sequences using MULTIDIVTIME. Branches are shown to reflect divergence times estimated. C1–C6 are calibration nodes (see Appendix Table S3 for more detailed information). N1–N17 are nodes of interest for which estimated dates are presented in Table 3. A dash-dotted line indicates the posterior mean for divergence time in millions of years for split the sturgeons and paddlefishes. The cladogram used is the maximum likelihood phylogram of the 24 acipenseriform fishes combined the results from Ludwig et al. (2001) and the present study. Data of inferred ploidy level for corresponding species were given based on the analysis of Zhang et al. (1999), and Ludwig et al. (2001). The chromosome duplication events were mapped on the branches.

(ca. 200 Mya) (Grande and Bemis, 1991; Jin, 1995) in order to investigate sturgeons and paddlefishes divergence time estimates. The results of Bayesian relaxed molecular clock analysis using mitochondrial cytochrome *b* gene sequence data are given in Table 3 and Fig. 1. The split time between sturgeons and paddlefishes is dated back to Early Jurassic at 184.4 Mya with a 95% credibility interval of 150.0–199.5 Mya, which were slightly older than the result of 141.4 Mya that based on mitogenome sequence data set, but with overlapping credibility interval. Within the Polyodontidae clade the Chinese swordfish (*P. gladius*), with a limited distribution in Yangtze River, splits with the Mississippi paddlefish (*P. spathula*), which also has a limited distribution in the Mississippi–Missouri basin, at about 68 Mya. Within the Acipenseridae clade, the divergence time between the Pacific and the Atlantic clades appears as about 121 Mya (Table 3; Fig. 1).

3.2. Dispersal–vicariance analysis

The historical biogeography of the acipenseriforms was analyzed in terms of the nine main distribution areas referred above (Table 2), with the setting of "maxareas" limiting the number of ancestral areas. A search with DIVA resulted in seven alternative reconstructions of the distribution history, each requiring 15 dispersals between areas. The optimal distributions at each ancestral node are given in Appendix Figure S2. According to the DIVA analysis, the common ancestor of the sturgeons and paddlefishes was widespread on (1) Ponto-Caspian region, and China region, (2) Mississippi river and Gulf of Mexico, (3) Ponto-Caspian region, Mississippi river and Gulf of Mexico, and (4) China region and Mississippi river and Gulf of Mexico. The reconstructions are similar except for some of the basal events. The suggested

Table 3
Posterior means and 95% credibility intervals for divergence times in million year ago (Mya) of nodes in Fig. 1 for the second dating analysis

Node	Cytochrome *b* data (three partitions)
N1	57.9 (23.4, 112.2)
N2	150.8 (104.7, 187.0)
N3	10.7 (1.2, 34.6)
N4	4.1 (0.1, 17.1)
N5	121.3 (76.7, 166.2)
N6	96.2 (55.5, 141.6)
N7	48.0 (17.5, 90.6)
N8	9.6 (0.5, 29.5)
N9	69.7 (34.1, 114.8)
N10	10.5 (0.7, 31.1)
N11	45.8 (18.6, 84.7)
N12	71.4 (37.9, 114.1)
N13	52.7 (24.5, 91.5)
N14	35.4 (13.8, 68.2)
N15	24.0 (7.5, 51.5)
N16	10.8 (1.7, 29.1)
N17	5.3 (0.2, 18.0)
C1	184.4 (150.0, 199.5)
C2	67.7 (50.5, 120.7)
C3	171.6 (132.9, 195.9)
C4	86.4 (48.2, 132.6)
C5	59.7 (29.4, 99.8)
C6	44.5 (18.7, 80.5)

C1–C6 are calibration nodes.

solution implies a centre of origin for Acipenseridae in the Ponto-Caspian region (Fig. 2).

4. Discussion

4.1. Dating the genome duplication events

Since Stephens (1951) and Ohno (1970) introduced their theory of evolution by gene duplication, it is widely accepted that gene duplication has played a crucial role on the origin of metazoans, vertebrates and mammals from unicellular organisms. Nevertheless, the role of ancient and lineage specific duplication events is less investigated until now.

Within all recent vertebrates, sturgeons include some of the species with a larger number of chromosomes. Independently to two ancient genome duplication rounds shared by an early deuterostome ancestor (Meyer and Schartl, 1999) about 590 Mya and a second round up to 150 Myr later (Wang and Gu, 2000), several lineage specific duplication events occurred in sturgeons and paddlefishes (Fig. 1). Therefore, these fishes represent an interesting model for studying these lineage specific genome duplication events. The results of our dating analysis suggest that genome duplication events in the sturgeon Atlantic lineage (53 Mya) and in the sturgeon Pacific lineage (70 Mya) were done at somewhat similar timings (see Fig. 1). Although no clear conclusion can be made about this subject, it is possible to discuss some reasons that might have favoured duplication events at these timings. The possible reason concerns simple random: our general understanding of genome duplication events seems to suggest that these events happen randomly from time to time. The selective advantage resulting from increasing heterozygosity may support a fixation of these new "genotypes". But this process of fixation may be more successfully in times of population expansion (radiation). Based on the outcome of this study (Fig. 1), we conclude that most of the recent sturgeon lineages evolved with the shrank of the Tethys Sea and the creation of new water bodies (for example Caspian, Black, and Aral seas). It is not unlikely that individuals with higher gen(om)e numbers are specially more flexible in their adaptation to new/changing environmental conditions. Independently from this first round of sturgeon lineage specific genome duplications there were a few more recent events, for example in *A. mikadoi* (ca. within 9.6 Myr) as well as in *Acipenser brevirostrum* (ca. within 35.4 Myr) (see Fig. 1). *A. brevirostrum* has an intermediary ploidy level depending on its number of chromosomes ($n = \sim 360$). Kim et al. (2005) classified this species as hexaploid depending on its number of Ag-NORs ($n = 10$) possibly resulting from the functional loss of NORs on more than one pair of chromosomes. Taken together, there is large evidence for an ongoing process of gene silencing in sturgeons but their low evolutionary rate resulted in the "conservation" of lineage specific intermediary character states over a long time.

4.2. Historical biogeography

Acipenseriformes had existed at least since the Early Jurassic (~200 My BP), and all fossil and recent taxa appeared from the Holartic (Bemis et al., 1997). As above mentioned, there are two extant Acipenseriform families: Polyodontidae, containing two monospecific genera (*Polyodon* and *Psephurus*), and Acipenseridae, containing three genera (*Acipenser*, *Huso*, and *Scaphirhynchus*). The ancient split of these families dated by our analysis of molecular mitochondrial cyt *b* gene data (184 Mya: see Fig. 1) is in good agreement with the paleontological data available. Fossil records of paddlefishes from China date to the Lower Cretaceous (e.g. *Protopsephurus liui*, Grande et al., 2002) and from North America date to the Upper Cretaceous (*Paleopsephurus wilsoni* MacAlpin, 1941). Today only two extant paddlefish species are found: *P. spathula*, inhabiting North American rivers (Mississippi/Missouri basin), and *P. gladius*, inhabiting the Yangtze River in China. Thus, all the fossil and extant paddlefishes described so far are restricted to the northern hemisphere. The Atlantic and Pacific Oceans seemingly began to open during the Jurassic and have continued to do so during the Cretaceous (Laurasia splits in North America and Eurasia) (Smith et al., 1994). About 120 million years ago, the Tethys Sea shrank further, eventually becoming the Black, Caspian and Aral Seas (Smith et al., 1994). These geological events appear to have played an important role in acipenseriform evolution, by: (i) forming the sea sturgeon lineage (~171 Mya) and perhaps by contributing to the subsequent splitting between the North American *A. oxyrinchus* and the European *A. sturio* (see Fig. 1); (ii) contributing to separate

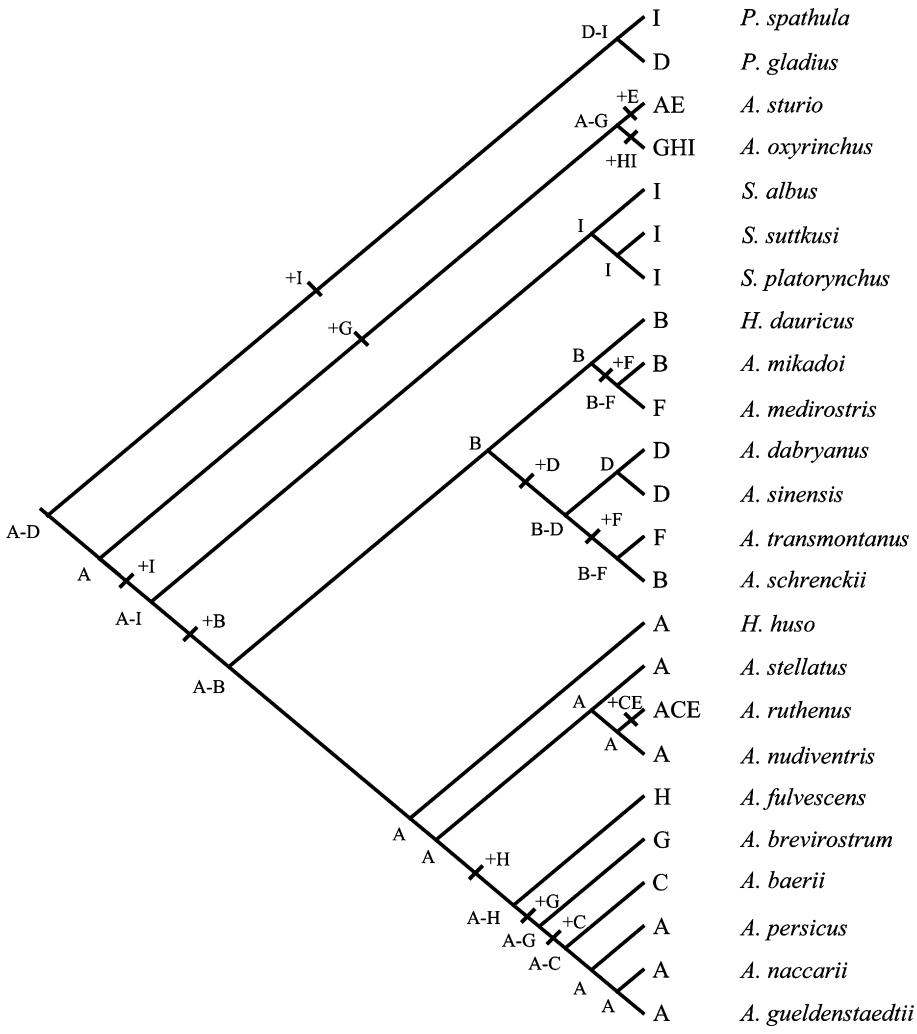

Fig. 2. Suggested biogeographical history of the acipenseriforms (see text for details). The reconstruction is one of the seven alternative, equally parsimonious biogeographical reconstructions (cf. Appendix Fig. S2). Dispersal events are indicated on the branches. Note that the dispersal–vicariance analysis, the cost of dispersal is one per area added to an ancestral distribution. For instance, the cost of assumption that *A. ruthenus* dispersed from Ponto-Caspian region to the other major regions (+CE) is two dispersals. Implied between-area vicariance events are indicated by hyphens in the ancestral distributions. The illustrated reconstruction differs from the other alternatives (cf. Appendix Fig. S2) in postulating that the ancestral species of Acipenseriformes was distributed in the Ponto-Caspian region or China region. In addition, the reconstruction implies a center of origin for the Acipenseridae in the Ponto-Caspian region. Codes A–I are the same as shown in Table 2.

the genus *Scaphirhynchus* from the remaining sturgeons (~151 Mya) (see Fig. 1); (iii) contributing to the Atlantic/Pacific split (~121 Mya) and their further sub-division within both clades (see Fig. 1), and (iv) contributing to the continental subdivision of the two paddlefish species (~68 Mya) (see Fig. 1). Although both these two extant paddlefish species are formed by primary freshwater fishes and all paddlefish fossil records discovered so far are from freshwater deposits, they occasionally stray into costal marine environments (Bemis et al., 1997). For this reason, separation between the two extant paddlefish species (i.e. between Asia and North America) seemingly depended essentially on the increasing geographic distance between both continents during the Late Cretaceous. Nevertheless, one cannot totally exclude other reasons also favoring their splitting, such as, for example the increase of salinity of the waters situated between these continents.

The distribution of the sea sturgeons is a further example seemingly supporting a vicariance scenario based on continental drift. In fact, the opening of the Atlantic Ocean between Europe and North America seemingly began at about 100 Mya (Smith et al., 1994), what thus may have contributed to the divergence between the two major lineages of sea sturgeons (see Fig. 2). Sea sturgeons are anadromous fishes, well adapted to marine conditions and occurring on both shores of the Atlantic Ocean. To provide an exact dating for the final separation of these lineages based on ecological and physiological clues is very difficult due to the eventual tolerance to high values of salinity and the eventually ability for long distance migrations that these fishes might have displayed (Ludwig et al., 2002). But such a separation was probably finished with the closure of the Tethys Sea and the onset of the North Atlantic Gyre (about 15–20 Mya) (Smith et al., 1994). Today, *A. sturio*

inhabits in Europe and *A. oxyrinchus* inhabits in North America. A trans-Atlantic colonization event for *A. oxyrinchus* entering the Baltic Sea was recently proposed (Ludwig et al., 2002). Based on our dating results these species were separated at about 58 Mya - thus somewhat earlier than the final closure of the Tethys Sea (see Fig. 1).

With respect to our results concerning the divergence of the Atlantic/Pacific sturgeon split (121 Mya), this is in good agreement with the shrank of the Tethys Sea (120 Mya) separating these two oceans. Probably, the Aral, Black and Caspian seas evolved from this Tethys Sea (Smith et al., 1994). This might thus have contributed to the emergence of several lineages of the Atlantic group. This Atlantic group contains Eurasian (mainly Ponto-Caspian species) and two North American species (*A. brevirostrum* and *A. fulvescens*). The separation between the Eurasian and the North American species was very likely related to the increasing geographic distance between these landmasses and the closure of the Tethys Sea, and this seemingly occurred after the genome duplication (53 Mya) shown in Fig. 1. Nevertheless, the limited adaptation of some species to seawater conditions probably also played an important role in this separation. For example, *A. fulvescens*, *A. baerii*, and *A. naccarii* are predominantly freshwater species.

Based on the outcome of our study, we can thus say that sturgeon adaptation on seawater was very likely a primitive feature for this group of fishes. Later, some lineages lost/reduced their ability to enter seawater, depending on the realization of ecological niches situated in freshwater reservoirs. Due to different geological events the distribution of some lineages was limited to freshwater lakes, which resulted in a reduced level of seawater adaptation. The genome duplication event dated about 53 Mya (see Fig. 1) can probably explain the founding of new lineages of Atlantic sturgeons. Such lineages include the most diverse sturgeon lineage ("*gueldenstaedtii*-complex": see Birstein et al., 2005). Several rounds of radiation and refugial differentiation were seemingly caused by glaciations resulting in the complicated genotype patterns observed for this species complex (Birstein et al., 2005). Moreover, introgression events caused by river captures or release programs blurred the original subdivision of lineages/species. For these reasons, molecular dating within this complex should be discussed with extreme caution. Furthermore, one should note that in the present study we used only "reference" genotypes for each species analyzed. Nevertheless, the present study supports a rather ancient origin of the different lineages/species.

Patterns similar to those described above for the Atlantic group seemingly also apply to the Pacific sturgeon species. In fact, the founding of many of these species is very likely related to genome duplications events associated with continental drift (see Fig. 1). The following species pairs have a trans-Pacific (American/Asian) distribution: *A. transmontanus* versus *A. schrenckii* and *A. medirostris* versus *A. mikadoi*. Speciation of *A. mikadoi* was very likely related to a second genome duplication event within the Pacific group (see Fig. 1). Taken together, largest species diversity of Acipenseridae clade was observed in Ponto-Caspian region, especially in the region occupied by the former Tethys Sea, which implies a possible origin of most of these lineages in this region (Fig. 2), and which scenario had previously been suggested by Birstein and DeSalle (1998).

Acknowledgments

We thank S.Y.W. Ho, University of Oxford, for his critical reading of this manuscript and for his helpful comments that greatly improved the manuscript. We also thank Associate Editor G. Bernardi, and the anonymous referee for insightful comments on earlier versions of this manuscript. The research was funded and supported by grants from National Science Fund for Distinguished Young Scholars (30225008) and the Knowledge Innovation Project of the Chinese Academy of Sciences (KSCX2-SW-101B) to S.H.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2006.09.008.

References

Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J., Staden, R., Young, I.G., 1981. Sequence and organization of the human mitochondrial genome. Nature 290, 457–465.

Anderson, S., de Bruijn, M.H., Coulson, A.R., Eperon, I.C., Sanger, F., Young, I.G., 1982. Complete sequence of bovine mitochondrial DNA: conserved features of the mammalian mitochondrial genome. J. Mol. Biol. 156, 683–717.

Arnason, U., Gullberg, A., Janke, A., 2001. Molecular phylogenetics of gnathostomous (jawed) fishes: old bones, new cartilage. Zool. Scr. 30, 249–255.

Arnason, U., Gullberg, A., Janke, A., Joss, J., Elmerot, C., 2004. Mitogenomic analyses of deep gnathostome divergences: a fish is a fish. Gene 333, 61–70.

Bemis, W.E., Kynard, B., 1997. Sturgeon rivers: an introduction to acipenseriform biogeography and life history. Environ. Biol. Fishes 48, 167–183.

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

Birstein, V.J., 1993. Sturgeons and paddlefishes: threatened fishes in need of conservation. Cons. Biol. 7, 773–787.

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

Birstein, V.J., DeSalle, R., 1998. Molecular phylogeny of Acipenserinae. Mol. Phylogenet. Evol. 9, 141–155.

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

Birstein, V.J., Ruban, G., Ludwig, A., Doukakis, P., DeSalle, R., 2005. The enigmatic Caspian Sea Russian sturgeon: how many cryptic forms does it contain? Syst. Biodiv. 3, 203–218.

Brinkmann, H., Denk, A., Zitzler, J., Joss, J.J., Meyer, A., 2004. Complete mitochondrial genome sequences of the South American and the Australian lungfish: testing of the phylogenetic performance of mitochondrial data sets for phylogenetic problems in tetrapod relationships. J. Mol. Evol. 59, 834–848.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540–552.

Cheng, S., Fockler, C., Barnes, W., Higuchi, R., 1994. Effective amplification of long targets from cloned inserts and human genomic DNA. Proc. Natl. Acad. Sci. USA 91, 5695–5699.

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

Crow, K.D., Stadler, P.F., Lynch, V.J., Amemiya, C., Wagner, G.P., 2006. The "Fish-Specific" hox cluster duplication is coincident with the origin of teleosts. Mol. Biol. Evol. 23, 121–136.

Gardiner, B.G., 1984. Sturgeons as Living Fossils. In: Eldredge, N., Stanley, S.M. (Eds.), Living Fossils. Springer-Verlag, New York, pp. 148–152.

Gokey, N.G., Cao, Z., Pak, J.W., Lee, D., McKiernan, S.H., McKenzie, D., Weindruch, R., Aiken, J.M., 2004. Molecular analyses of mtDNA deletion mutations in microdissected skeletal muscle fibers from aged Rhesus monkeys. Aging Cell 3, 319–326.

Grande, L., Bemis, W.E., 1991. Osteology and phylogenetic relationships of fossil and recent paddlefishes (Polyodontidae) with comments on the interrelationships of Acipenseriformes. J. Vertebr. Paleontol. 11 (Suppl. 1), 1–121.

Grande, L., Jin, F., Yabumoto, Y., Bemis, W.E., 2002. *Protopsephurus liui*, a well-preserved primitive paddlefish (Acipenseriformes: Polyodontidae) from the lower Cretaceous of China. J. Vertebr. Paleontol. 22, 209–237.

Hixson, J.E., Brown, W.M., 1986. A comparison of the small ribosomal RNA genes from the mitochondrial DNA of the great apes and humans: sequence, structure, evolution, and phylogenetic implications. Mol. Biol. Evol. 3, 1–18.

Inoue, J.G., Miya, M., Tsukamoto, K., Nishida, M., 2001. A mitogenomic perspective on the basal teleostean phylogeny: resolving higher-level relationships with longer DNA sequences. Mol. Phylogenet. Evol. 20, 275–285.

Inoue, J.G., Miya, M., Tsukamoto, K., Nishida, M., 2003. Basal actinopterygian relationships: a mitogenomic perspective on the phylogeny of the 'ancient fish'. Mol. Phylogenet. Evol. 26, 110–120.

Inoue, J.G., Miya, M., Venkatesh, B., Nishida, M., 2005. The mitochondrial genome of Indonesian coelacanth *Latimeria menadoensis* (Sarcopterygii: Coelacanthiformes) and divergence time estimation between the two coelacanths. Gene 349, 227–235.

Jin, F., 1995. Late Mesozoic acipenseriforms (Osteichthyes: Actinopterygii) in Central Asia and their biogeographical implications. In: Sun, A., Wang, Y. (Eds.), Sixth symposium Mesozoic terrestrial ecosystems and biota, short papers. China Ocean Press, Beijing, pp. 15–22.

Kim, D.S., Nam, Y.K., Noh, J.K., Park, C.H., Chapman, F.A., 2005. Karyotype of North American shortnose sturgeon *Acipenser brevirostrum* with the highest chromosome number in the Acipenseriformes. Ichthyol. Res. 52, 94–97.

Le Comber, S., Smith, C., 2004. Polyploidy in fishes. Biol. J. Linnean Soc. 82, 432–442.

Ludwig, A., Belfiore, N.M., Pitra, C., Svirsky, V., Jenneckens, I., 2001. Genome duplication events and functional reduction of ploidy levels in sturgeon (*Acipenser*, *Huso* and *Scaphirhynchus*). Genetics 158, 1203–1215.

Ludwig, A., Debus, L., Lieckfeldt, D., Wirgin, I., Benecke, N., Jenneckens, I., Williot, P., Waldman, J.R., Pitra, C., 2002. When the American sea sturgeon swam east. Nature 419, 447–448.

MacAlpin, A., 1941. *Paleopsephurus wilsoni*, a new polyodontid fish from the Upper Cretaceous of Montana, with a discussion of allied fishes, living and fossil. Geol. Soc. Am. Bull. 52, 1989.

Meyer, A., Schartl, M., 1999. Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. Curr. Opin. Cell Biol. 11, 699–704.

Miya, M., Nishida, M., 1999. Organization of the mitochondrial genome of a deep sea fish, *Gonostoma gracile* (Teleostei: Stomiiformes): first example of transfer RNA gene rearrangements in bony fishes. Mar. Biotechnol. 1, 416–462.

Nylander, J.A.A., Ronquist, F., Huelsenbeck, J.P., Nieves-Aldrey, J.L., 2004. Bayesian phylogenetic analysis of combined data. Syst. Biol. 53, 47–67.

Ohno, S., 1970. Evolution by gene duplication. Springer, Verlag, New York, NY.

Peng, Z., He, S., Wang, J., Wang, W., Diogo, R., 2006. Mitochondrial molecular clocks and the origin of the major Otocephalan clades (Pisces: Teleostei): a new insight. Gene 370, 113–124.

Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: advantages of the AIC and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53, 793–808.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Robles, F., De la Herran, R., Ludwig, A., Ruiz Rejon, C., Ruiz Rejon, M., Garrido-Ramos, M.A., 2004. Evolution of ancient satellite DNAs in sturgeon genomes. Gene 338, 133–142.

Ronquist, F., 1996. DIVA, version 1.1. Computer program and manual available by anonymous ftp from Uppsala University.

Ronquist, F., 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. Syst. Biol. 46, 193–201.

Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Smith, A.G., Smith, D.G., Funnell, B.M., 1994. Atlas of Mesozoic and Cenozoic Coastlines. Cambridge University Press, Cambridge.

Soltis, D.E., Soltis, P.S., 1999. Polyploidy: recurrent formation and genome evolution. Trends Ecol. Evol. 14, 348–352.

Stephens, S.G., 1951. Possible significance of duplication in evolution. Adv. Genet. 4, 247–265.

Swofford, D.L., 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer Associates, Sunderland, MA.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. Clustal x windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Thorne, J.L., Kishino, H., 2002. Divergence time estimation and rate evolution with multilocus data sets. Syst. Biol. 51, 689–702.

Thorne, J.L., Kishino, H., Painter, I.S., 1998. Estimating the rate of evolution of the rate of molecular evolution. Mol. Biol. Evol. 15, 1647–1657.

Wang, Y., Gu, X., 2000. Evolutionary patterns of gene families generated in the early stage of vertebrates. J. Mol. Evol. 51, 88–96.

Yang, Z., 1994. Maximum likelihood phylogenetic estimations from DNA sequences with variable rates over sites: approximate method. J. Mol. Evol. 39, 306–314.

Yang, Z., 2000. Phylogenetic analysis by maximum likelihood (PAML), 3rd ed. University College London, London.

Yoder, A.D., Yang, Z., 2000. Estimation of primate speciation dates using local molecular clocks. Mol. Biol. Evol. 17, 1081–1090.

Zardoya, R., Meyer, A., 1998. Cloning and characterization of a microsatellite in the mitochondrial control region of the African side-necked turtle, *Pelomedusa subrufa*. Gene 216, 149–153.

Zhang, S.M., Yan, Y., Deng, H., Wang, D.Q., Wei, Q.W., Wu, Q.J., 1999. Genome size, ploidy characters of several species of sturgeons and paddlefishes with comment on cellular evolution of Acipenseriformes. Acta Zool. Sin. 45, 200–206.