
STURGEONS (ACIPENSERIDAE)

Development of Single Nucleotide Polymorphism Markers for the Autotetraploid *Scaphirhynchus* Sturgeons (Acipenseriformes)

T. Zhou^{a, b}, J. Huang^{a, b}, C. B. Dillman^c, Y. He^d, and C. Li^{a, b, *}

^a Shanghai Universities Key Laboratory of Marine Animal Taxonomy
and Evolution, Shanghai Ocean University, Shanghai, China

^b Engineering Research Center of Environmental DNA and Ecological
Water Health Assessment, Shanghai Ocean University, Shanghai, China

^c Cornell University Museum of Vertebrates, Department
of Ecology and Evolutionary Biology, Ithaca, New York, USA

^d Shanghai Synchrotron Radiation Facility, Shanghai Advanced
Research Institute, Chinese Academy of Sciences, Shanghai, China

*e-mail: chli@shou.edu.cn

Received June 6, 2022; revised July 10, 2022; accepted July 14, 2022

Abstract—North American river sturgeons of the genus *Scaphirhynchus* include three species: *S. platyrhynchus*, *S. albus* and *S. suttkusi* that live in the Missouri, Mississippi, and Mobile basin. All species of *Scaphirhynchus* are threatened, endangered or critically endangered due to a combination of factors including of habitat loss and over-harvesting. Genetic tools have been applied for conservation studies in this group, however, the tetraploid nature of the genome of these species have brought a huge challenge to development of nuclear markers for these species and limited knowledge that could be obtained, such as the phylogenetic intrarelationships and population genetics of this genus. Moreover, unintentional hybridization arose from two species of *Scaphirhynchus* (*S. platyrhynchus* and *S. albus*) that share the same spawning space. To address the problem of species identification and provide genetic markers for population genetic studies on *Scaphirhynchus*, we developed a bioinformatics pipeline to find SNP markers, based on comparison between single-copy loci of diploid gar and two released autotetraploid genomes of Acipenseriformes. We found 77 SNPs at single-copy loci and 642 SNPs at double-copy loci after filtering. Both the single-copy and double-copy loci supported the same phylogenetic relationship among the three species, in which *S. albus* and *S. platyrhynchus* were more closely related to each other than either of them to *S. suttkusi*. The principal component analysis using these SNPs also showed that *S. albus* and *S. platyrhynchus* were close to each other. The SNP markers developed in this study should facilitate further researches on population genetics and conservation of the *Scaphirhynchus* sturgeons.

Keywords: autotetraploid genome, Shovelnose sturgeon, Pallid sturgeon, Alabama sturgeon, gene capture, nuclear markers, conservation genetics

DOI: 10.1134/S0032945222060340

INTRODUCTION

Acipenseriformes (sturgeons and paddlefishes) is an ancient lineage of primitive teleosts, dating back to the mid-Jurassic. The Acipenseriformes consists of one extinct “fossil” family (†Peipiaosteidae) and two extant families: Acipenseridae, Polyodontidae. The extant species of Acipenseriformes are distributed in fresh and salt waters of the Palearctic realm (Europe, temperate Asia, North America and North Africa), most of which are endangered worldwide or extinct due to a combination of over-harvesting and habitat loss (Birstein and Bemis, 1997; Haxton et al., 2016), such as the already extinct Chinese paddlefish (*Psephurus gladius*) (Zhang et al., 2020).

The genus *Scaphirhynchus* include three sturgeon species, *S. platyrhynchus* (shovelnose sturgeon), *S. albus*

(pallid sturgeon), and *S. suttkusi* (Alabama sturgeon), all of which are threatened, endangered or critically endangered. The pallid sturgeon has similar characteristics and ecological niches to shovelnose sturgeon (Murphy et al., 2007) and both are found in the Missouri and Mississippi Rivers. Alabama sturgeon is now only believed to exist in the lower Alabama River according to eDNA monitoring (Janosik et al., 2021). The striking morphological similarities between pallid sturgeon and shovelnose sturgeon can make it difficult to differentiate them especially at small size. Moreover, unable to find its own spawning habitat, pallid sturgeon is forced to share spawning ground with the more common shovelnose sturgeon, which can lead to unintentional hybridization (Schrey et al., 2011).

Since the 1990s, a repertoire of molecular markers has been developed for conservation and phylogenetic research of the Acipenseriformes. Molecular markers that have been used on sturgeons include mitochondrial DNA (mtDNA) (Krieger et al., 2000; Birstein et al., 2009), random amplified polymorphic DNA (RAPD) (Comincini et al., 1998), amplified fragment length polymorphic (AFLP) DNA markers (Ludwig, 2008), and microsatellites (McQuown et al., 2000; Ludwig et al., 2001; Schrey and Heist, 2007; Chassaing et al., 2011; Liu et al., 2017). Previous phylogenetic inference and population studies on Acipenseriformes relied heavily on mtDNA markers, using either a single mtDNA gene or whole mitochondrial genome sequences (Shen et al., 2020; Sheraliev and Peng, 2020). Nevertheless, mtDNA has known limitations including maternal inheritance. Sets of microsatellites were developed to distinguish species of sturgeon. Genotyping unhatched sturgeon embryos using microsatellites can identify the parentage of the embryos and may be necessary where multiple acipenseriform species or hybrids occur (Liu et al., 2017).

Single nucleotide polymorphisms (SNPs), one of the latest-generation molecular markers, occur at high frequencies in both animal and plant genomes (Khlestkina and Salina, 2006). The development of SNP markers allows researchers to automatize and also enhance the effectiveness of genotype analysis, but few SNP markers have been developed for sturgeons. Eichelberger et al. (2014) employed three SNP assays to screen sturgeon embryos and larvae collected from the upper Missouri River basin, one mtDNA SNP and two nuclear SNPs. The mtDNA SNP is located in a conserved region of the mtDNA genome and is presumably fixed between paddlefish and sturgeon. Specific multilocus genotypes at two nuclear SNPs occurred in 98.9% of wild adult pallid sturgeon but only in 3% of shovelnose sturgeon sampled in the upper Missouri River potentially providing a diagnostic molecular site to aid in discriminating species.

Polyploidization, a common phenomenon in animals and plants (Van De Peer et al., 2017), makes the use of the nuclear markers in phylogenetic analyses difficult and also increases inefficiency in development of nuclear markers using conventional methods. Searching conserved sequences based on multi-sequence alignment provides a new method to develop large scale nuclear markers (Li et al., 2007, 2017; Rothfels et al., 2013; Zeng et al., 2017). A set of single-copy loci was identified from *Lepisosteus oculatus* by comparing genome sequences of multiple model fish species, comprising 13837 exon-coding loci used in gene enrichment studies of “ancient fishes” (Jiang et al., 2019).

All acipenseriforms are ploidyploid (Ludwig et al., 2001) but lack whole reference genome sequence data. As a result, there has been a strong dependence on haploid mitochondrial data for studying the phyloge-

netic relationships of acipenseriforms, with nuclear data rarely being used (e.g., Shen et al., 2020; Sheraliev and Peng, 2020). As with all lineages, it is difficult to address phylogenetic relationships within the Acipenseriformes with only mitochondrial data (Krieger et al., 2008). Luo et al. (2019) mined transcriptome data comparing a reference genome from spotted gar and identified nuclear protein-coding markers for Acipenseriformes. However, the effect of polyploidy in acipenseriforms on phylogenetic reconstruction and how to differentiate the multi copy loci were not explicitly addressed.

In this study, we developed a bioinformatics pipeline to extract single-copy and double-copy coding loci from autotetraploid species of *Scaphirhynchus*. We then: (1) compare a published sturgeon and paddlefish genomes to the single-copy loci of the spotted gar (*Lepisosteus oculatus*); (2) compare sequencing reads from our target enrichment of autotetraploid species to the loci identified in the first step; (3) assemble the sorted reads and extracted SNP loci for the targeted species. This approach is used as a test, or proof-of-concept, to develop SNP markers for *Scaphirhynchus*.

MATERIALS AND METHODS

Samples and DNA extraction. Four individuals from the three species of *Scaphirhynchus* were sampled, including one specimen of *S. albus* (coded as *S. albus*), two specimens of *S. platyrhynchus* (coded as *S. platyrhynchus*-1, *S. platyrhynchus*-2) and one specimen of *S. suttkusi* (coded as *S. suttkusi*). All *Scaphirhynchus* samples were preserved in 70% ethanol and archived at -4°C prior to DNA extraction. Genomic DNA extracted from tissue samples was measured using NanoDrop 3300 Fluorospectrometer (Thermo Fisher Scientific, Waltham, MA, USA).

Baits design and synthesis. Markers were identified through comparing eight fish genomes using a bioinformatics tool, EvolMarkers (Li et al., 2012). The sets of single-copy markers were generated and used to design baits to capture species in three different research projects conducted in the authors' laboratory. The baits were designed on sequences from *Lepisosteus oculatus* comprising 13837 loci used in “Basal,” “Acipen” and “Otomor” projects (Jiang et al., 2019). Biotinylated RNA baits of 120 bp were synthesized with $2\times$ tiling for the loci targeted (MYcroarray, Ann Arbor, Michigan).

Library preparation, target enrichment and Illumina sequencing. Library preparation and target enrichment followed Yuan et al. (2016). All final enriched and indexed products were pooled in equimolar concentrations, and the concentration of the pooled sample was measured using a NanoDrop 3300 Fluorospectrometer. The pooled sample was then sequenced on an Illumina HiSeq 2500 sequencer (Illumina, Inc, San Diego, CA, USA).

Identification of target loci. The chromosome-level genome assemblies of *Acipenser ruthenus* (Cheng et al., 2021) (accession number: GCA_010645085.1) and *Polyodon spathula* (Du et al., 2020) (accession number: GCA_017654505.1) were downloaded from the NCBI database. Both *P. spathula* and *A. ruthenus* are autotetraploid species. The single-copy loci of *Lepisosteus oculatus* were aligned to these sturgeon and paddlefish genomes to retrieve conserved single-copy loci and conserved double-copy loci of sturgeon (Fig. 1). To do this, the sequences of target loci of *Lepisosteus oculatus* were first aligned to reference autotetraploid 1 (*P. spathula*) using BLASTN with an e-value cut off of 0.0001. There were three possible outcomes of the alignment: (1) target locus (Figs. 1a, 1a') only aligned to one chromosome in one position; (2) target locus (Fig. 1b) aligned two chromosomes with one position at each corresponding chromosome; (3) target locus (Fig. 1c) aligned to multiple chromosomes (≥ 3) or aligned to multiple positions on one chromosome (not showed in Fig. 1). The loci found from the first outcome were deemed single-copy loci. Loci identified in the second outcome were considered double-copy loci. Other loci from the third outcome were discarded. The single and double-copy loci were aligned to the genome of the second autotetraploid reference genome (*A. ruthenus*) to filter for conserved loci in both autotetraploid species (Fig. 1). Finally, the sequence of conserved single-copy loci and double-copy loci were extracted from the *A. ruthenus* chromosomes (species of *Scaphirhynchus* are closer to *A. ruthenus* than *P. spathula*) for subsequent target gene reads mapping and assembly for the four individuals of *Scaphirhynchus*.

Read assembly. The raw reads of target autotetraploids of the genus *Scaphirhynchus* were parsed to separate files according to the index on the adapter. Adapter sequences and low-quality bases ($Q < 20$) were trimmed using Trim_galore (<https://github.com/FelixKrueger/TrimGalore>). Numbers of clean reads and total bases were calculated using an in-house python script (sum.py). Then, the fastq format of reads was converted to fasta format. The forward strand read (R1) and reverse complement of the reverse strand read (R2) were concatenated with a string of "NNNNN". Then duplicates of the merged reads (R1, R2) were removed using the Seqkit (Shen et al., 2016). Next, the reads were parsed into either single-copy locus (the highest score) or double-copy locus (the top two highest scores) according to the score ($\text{score} = \text{alignment_length} - 2 * \{\text{number of mismatch_bases}\} - 2.5 * \{\text{number of gap_open}\}$) computed based on the results of BLASTN (e-value cut off of 0.001). The parsed reads were assembled into contigs using ABySS (Simpson et al., 2009) with "-k 31". Next, all contigs from each gene of every species were scored by comparing to the reference sequence using the Biopython module written in Python. The high-score contigs were re-aligned to the reference genome

to verify the orthology using BLASTN. Potential paralogs were discarded if the best hit was not aligned to target region in the reference.

Call consensus and SNP calling. The assembled contigs of different samples and species were aligned using MAFFT (Katoh et al., 2002) with default settings (Fig. 2). Consensus was calculated based on the multiple alignment using an in-house python script (consensus.py). Then the corresponding clean reads were separately mapped to the consensus (conserved single-copy loci and conserved double-copy loci) using BWA-MEM (Li, 2013) with default parameter settings. SNP calling was performed using the BCFtools mpileup and the output vcf files were processed with VCFtools (parameter: -max-missing 1 -remove-indels -recode -recode-INFO-all) to remove indels and positions with missing data.

PCA analysis and phylogenetic inference. Filtered vcf files of SNPs were used as input in the Rstudio for PCA analysis with vcfR package (Knaus and Grünwald, 2017). For phylogenetic analysis, the vcf files of SNPs were converted to phylip, fasta and nexus format file using vcf2phylip. The corresponding nucleotide in the genome assembly of *A. ruthenus* at the position of SNPs were concatenated to set as the outgroup during the reconstruction of phylogenetic trees, which were reconstructed using the maximum likelihood method implemented in IQ-TREE (Nguyen et al., 2015) with AUTO model and 1000 bootstrap replicates.

RESULTS AND DISCUSSION

Conserved loci from reference species *Acipenser ruthenus* and *Polyodon spathula*. Using the 60 assembled chromosomes of *Acipenser ruthenus* and *Polyodon spathula*, we identified 1921 conserved single-copy loci and 10914 double-copy loci in *A. ruthenus* (Fig. 3). The top 10 chromosome id with maximum number of loci (total of conserved single-copy loci and conserved double-copy loci) were NC_048323.1, NC_048324.1, NC_048325.1, NC_048326.1, NC_048327.1, NC_048328.1, NC_048329.1, NC_048332.1, NC_048334.1, NC_048335.1.

It was evident that the number of loci decreased as the order of chromosome id changed from NC_048323.1 to NC_048381.1; concordant with the decrease in chromosome length from longer to shorter. The number of single-copy loci was 17.6% of the number of double-copy loci. The difference in numbers of classified loci (single-copy vs. double-copy) may result from the short history of dipolization of the genome of *A. ruthenus* (Du et al., 2020). Limited by the few available whole genome resources for *Acipenseriformes*, we herein only identified the conserved single-copy loci and conserved double-copy loci using *A. ruthenus* and *P. spathula*.

Target enrichment of *Scaphirhynchus* using conserved loci. The number of clean reads obtained of

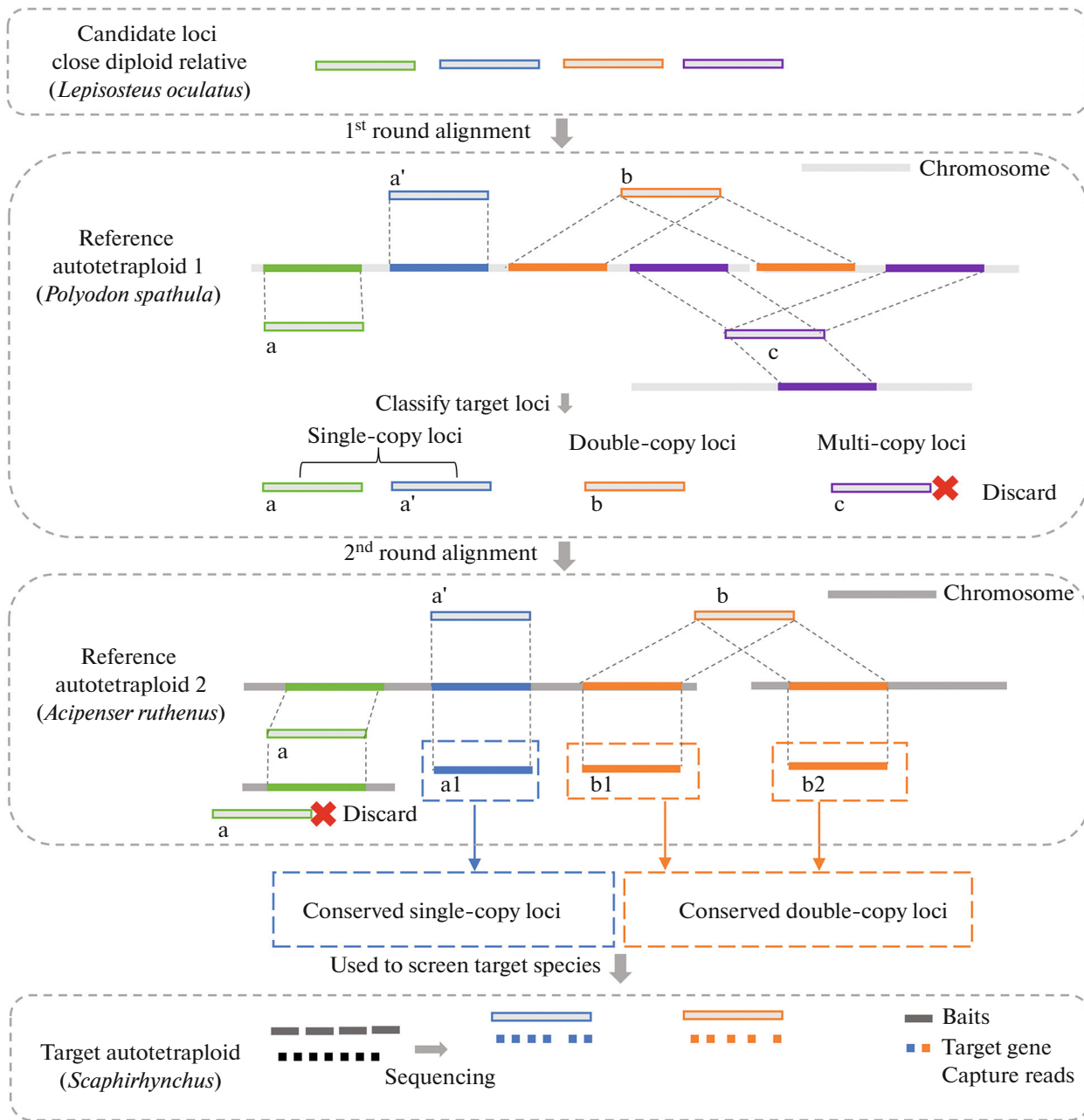


Fig. 1. Scheme of bioinformatics pipeline to identify efficient conserved loci of target autotetraploid *Acipenser ruthenus*.

these four samples ranged from 60792 to 9308420 while the total number of nucleotide bases of clean reads ranged from 9013590 to 499115094 bp after removing adapters and trimming low-quality bases (Table 1). The average number of captured conserved single-copy loci was 779 and conserved double-copy loci was 5241. The sample with the maximum number of captured conserved single-copy loci and conserved double-copy loci was *S. platyrhynchus*-1 while the minimum number of captured single-copy loci and double-

copy loci was *S. suttkusi* (Table 1). Forty-one single-copy loci and 317 double-copy loci were captured in all samples of *Scaphirhynchus*.

The average length of reads in *S. albus*, *S. platyrhynchus*-2 and *S. suttkusi* was about 150 bp while that of *S. platyrhynchus*-1 was approximately 100 bp. There was a positive correlation between the number of captured loci and number of clean reads in these four samples. The small number of captured conserved single-copy loci and conserved double-copy loci resulted

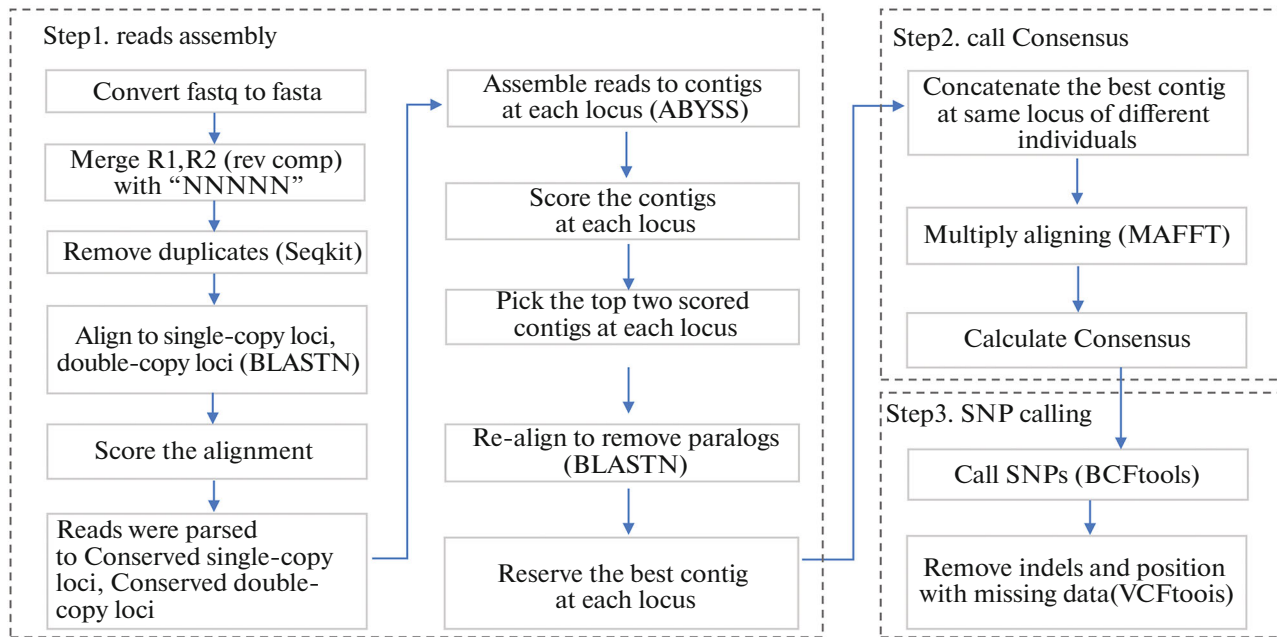


Fig. 2. Scheme of reads assembly, consensus and SNP calling. Program consists of three parts.

from the minimum number of loci captured in *S. suttkusi*. We only identified the target loci in the coding region, though the non-coding flanking regions also have the potential SNPs of utility in conservation genetic and phylogeny reconstruction approaches.

SNP screening for *Scaphirhynchus*. We extracted a total of 7838 SNPs (735 conserved single-copy loci and 7103 conserved double-copy loci) from the bam file mapped from one individual of *S. albus*, two individuals of *S. platyrhynchus* and one individual of *S. suttkusi*. Seventy-seven SNP sites at single-copy loci were chosen by filtering the position with missing data while 642 SNP sites at double-copy loci were kept after filtering (Fig. 4). The top 5 chromosome id regions that contained the maximum number of SNPs (both conserved single-copy loci and conserved double-copy loci) were NC_048324.1, NC_048330.1, NC_048332.1, NC_048335.1, NC_048346.1. The majority of SNPs at conserved single-copy loci were

found in NC_048330.1. Seventy-seven SNPs located in the consensus sequence of conserved single-copy loci correspond to eight different chromosomes in *A. ruthenus* (Fig. 5). The 1 Mb window size screened with the highest density of SNPs were located in the chromosome named NC_048330.1. 642 SNPs from conserved consensus double-copy loci were located on 26 chromosomes in *A. ruthenus* and the screened 1 Mb window size with top two highest density of SNPs were located in the chromosome named NC_048324.1 and NC_048346.1.

Differences in the number of conserved loci changed with the length of chromosomes, number of the selected SNPs at conserved loci was not correlated with the length of corresponding chromosomes. This may result from the missing data in different samples. It was noted that the most of SNPs located at the consensus of conserved single-copy loci corresponding to NC_048330.1 and the screened 1 Mb window size

Table 1. Statistics of clean reads and captured loci

Sample ID	Clean reads number	Total clean reads bases	Captured conserved		Total loci
			single-copy loci	double-copy loci	
<i>S. albus</i>	3 374 408	499 115 094	877	6093	6970
<i>S. platyrhynchus</i> -1	9 308 420	952 215 312	1246	7651	8897
<i>S. platyrhynchus</i> -2	2 314 076	341 121 238	771	5516	6287
<i>S. suttkusi</i>	60 792	9013 590	220	1705	1925

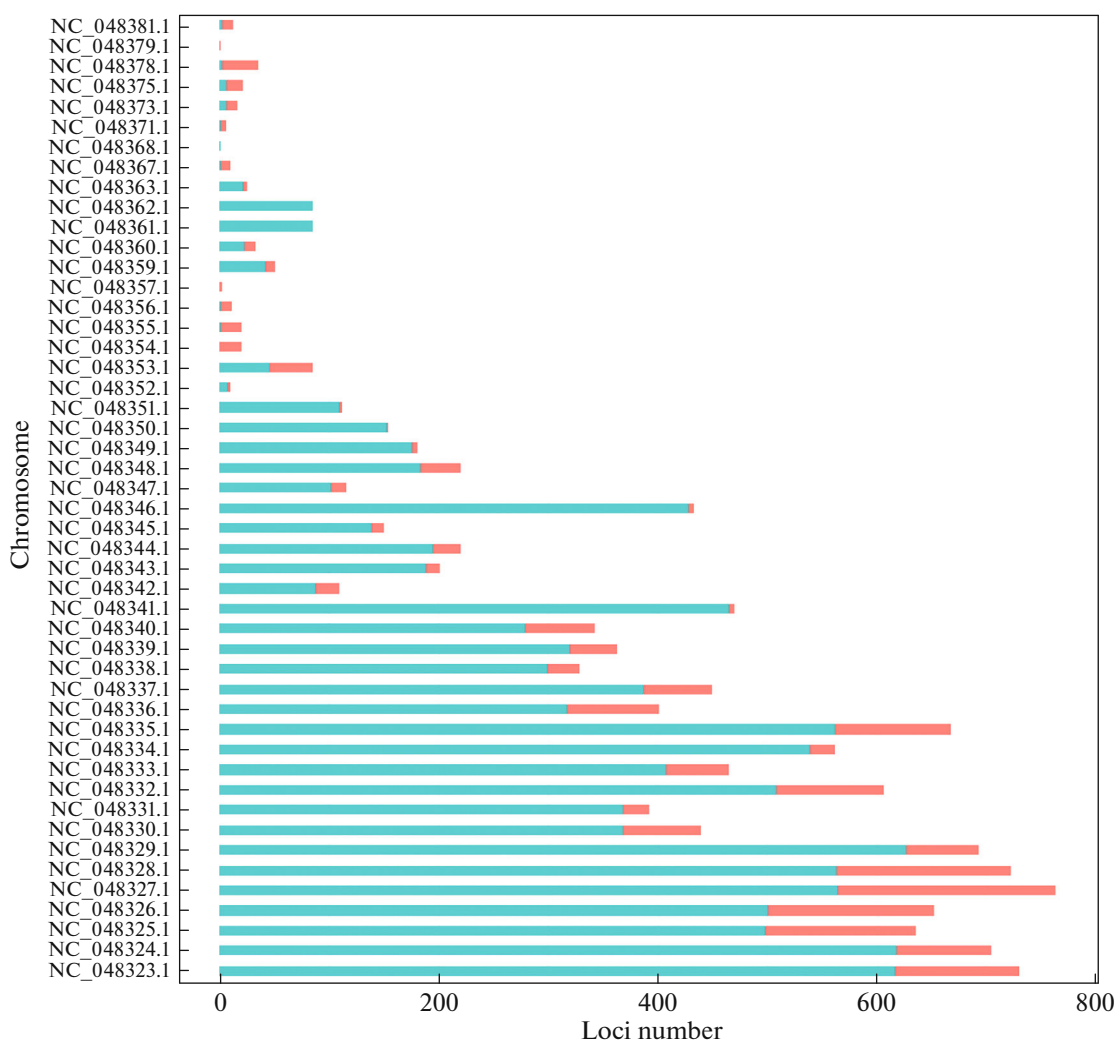


Fig. 3. The identified 1921 conserved single-copy loci (■) and 10914 conserved double-copy loci (■) at different chromosomes of *Acipenser ruthenus*.

with highest density of SNPs was also located at this chromosome. Similarly, for SNPs at conserved double-copy loci, the screened 1Mb window size with top 2 highest density of SNPs were located at two chromosomes with the top 2 highest number of SNPs. Due to the limitation of sample size, the number of SNPs found were also limited and linkage disequilibrium of these SNPs cannot be calculated.

Distinguishing species of *Scaphirhynchus* based on PCA and phylogenetic analysis. From the result of PCA analysis only using the 77 SNPs at single-copy loci (Fig. 5a), we found that two individuals of *S. platyrhynchus* (*S. platyrhynchus*-1, *S. platyrhynchus*-2) clustered closer to one another. One individual of *S. albus* and one individual of *S. suttkusi* were distinct from both *S. platyrhynchus*-1, 2. These three species can be clearly distinguished. When using 642 SNPs at conserved double-copy loci (Fig. 5b), the result was similar.

The maximum likelihood (ML) phylogeny based on the 77 SNPs at conserved single-copy loci (Fig. 6) showed that the two individuals of *S. platyrhynchus* are sister (*S. platyrhynchus*-1, *S. platyrhynchus*-2) to one another and that they are sister to *S. albus*, with *S. suttkusi* sister to both species of *Scaphirhynchus*; though there is very little support at any inferred node. A similar result was found in the tree reconstructed using 642 SNPs from conserved double-copy loci, the exception being the different branch lengths of the species of *S. platyrhynchus* (*S. platyrhynchus*-1, *S. platyrhynchus*-2). The topology of these two ML phylogenetic tree was same. The different branch lengths of two individuals of *S. platyrhynchus* (Fig. 6) may result from the limited SNP numbers at single-copy loci or part of these SNPs were also species-specific SNPs. This result is in agreement with the results of the PCA analysis. The three species of *Scaphirhynchus* can be

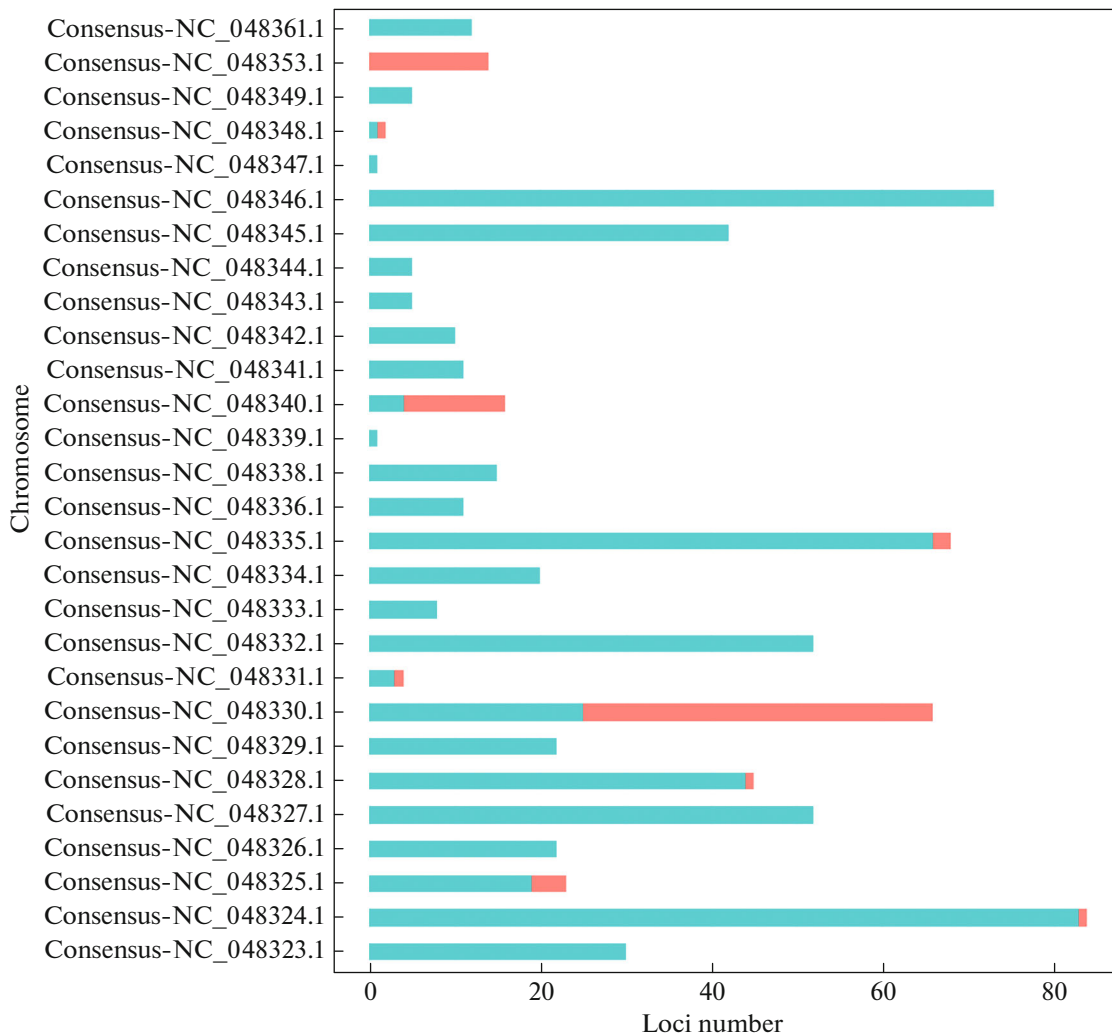


Fig. 4. The identified 7838 SNPs of *Scaphirhynchus* at consensus of conserved single-copy (■) and double-copy loci (■) corresponding to different chromosomes.

distinguished using the target gene capture and bioinformatic analysis pipeline.

Different from the strategy of our method here, Luo et al. (2019) developed an approach to generate new nuclear protein-coding (NPC) markers for phylogenetic analysis using a transcriptome plus reference genome. This strategy of their method may have been the best choice at the time as the chromosome-level genome assemblies of *A. ruthenus* and *P. spathula* hadn't yet been published. They used four new transcriptomes of sturgeon and the *L. oculatus* genome as a database and analyzed this database to screen candidate single-copy genes to develop NPC markers and found the conserved regions of exons based on multi-sequence alignment. The phylogenetic relationships within Order Acipenseriformes were highly resolved according to the novel nuclear markers they found. They made progress in the resolution and reconstruction of phylogenetic relationships of Acipenseriformes, which is very difficult due to the huge chal-

lenge of species with multiple levels of ployploidy in this order. Similarly, we here used chromosome-level genome assembly of reference autotetraploid *A. ruthenus* and *P. spathula* as a database and then searched the single-copy nuclear gene of *L. oculatus* in the database to find conserved single-copy and double-copy loci. This may be better than using transcriptome as expression of some genes in the transcriptome were few and would lessen the total number of single-copy nuclear genes found. Herein we tested this method to find SNP markers using four sampled individuals of *Scaphirhynchus*. Our method successfully distinguished the three species of *Scaphirhynchus* and proposed a resolution for the phylogenetic relationship of the genus *Scaphirhynchus*. The utility of this method is anticipated to facilitate further studies about the population structure, identification of hybrid species, conservation and phylogeny of Acipenseriformes species and other tetraploid fishes.

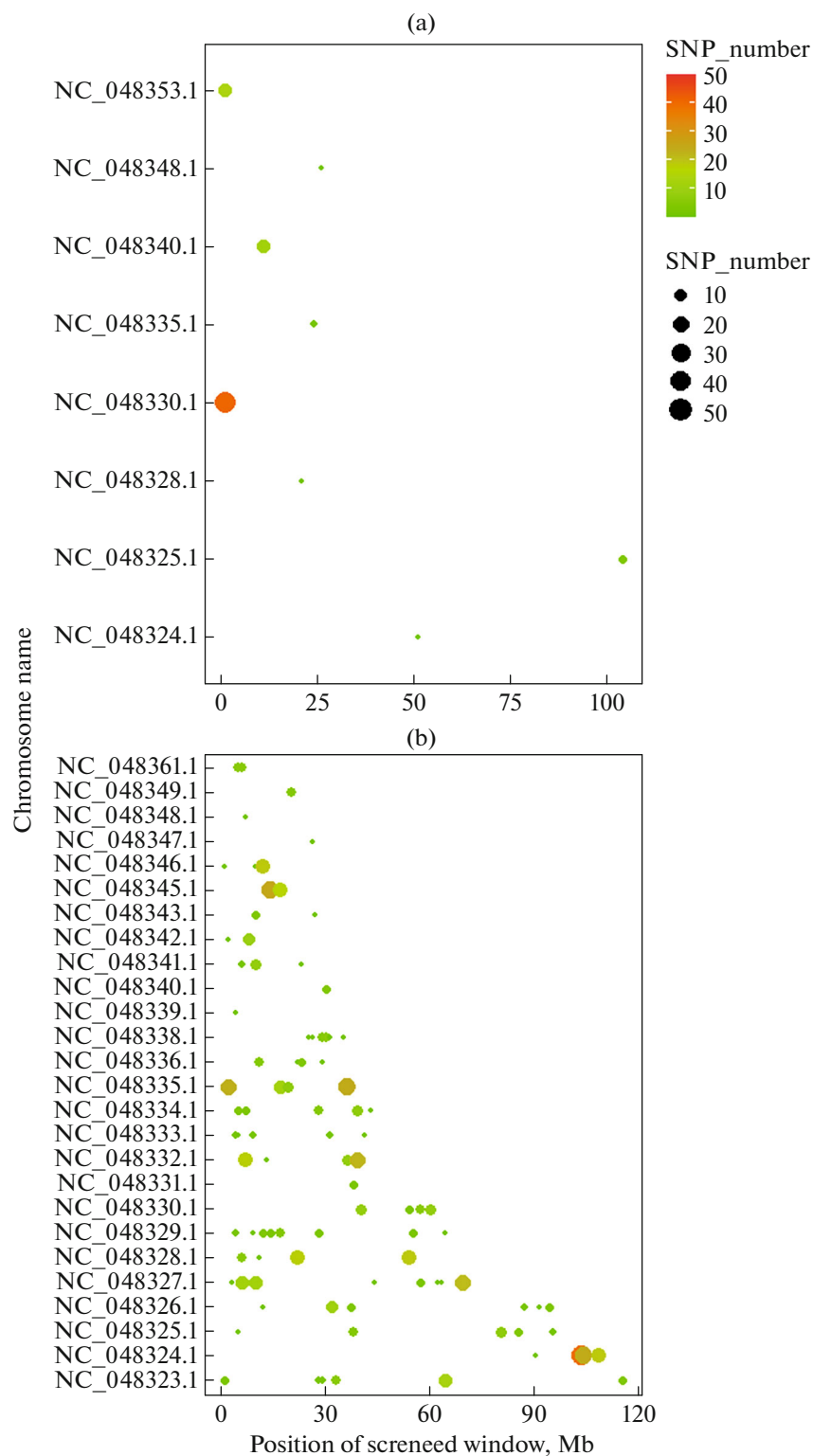


Fig. 5. Density of 77 SNPs of *Scaphirhynchus* at consensus of conserved single-copy loci and conserved double-copy loci corresponding to different chromosomes in *A. ruthenus*. The color bar represents the number of SNPs within 1Mb window size; (a) SNP density of 77 SNPs located at conserved single-copy loci; (b) SNP density of 642 SNPs located at conserved double-copy loci.

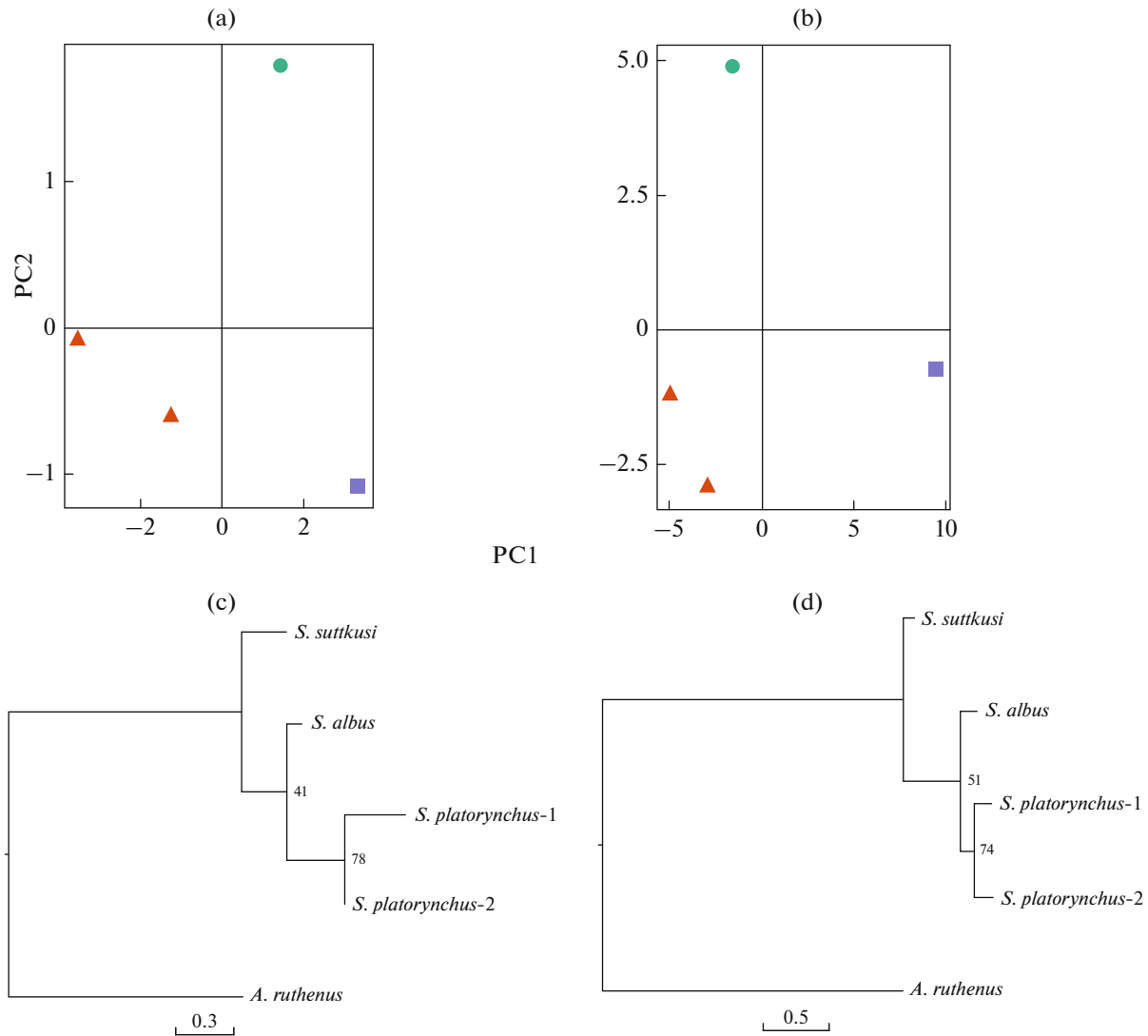


Fig. 6. Clustering of *Scaphirhynchus* species based on PCA analysis of SNPs and reconstructed ML phylogenetic trees based on SNPs of *Scaphirhynchus*: (●) *S. albus*, (▲) *S. platyrhynchus*, (■) *S. suttkusi*. (a) PCA result based on 77 SNPs at conserved single-copy loci, (b) PCA result based on 642 SNPs at conserved double-copy loci, (c) ML trees using 77 SNPs at conserved single-copy loci, (d) ML trees using 642 SNPs at conserved double-copy loci.

CONCLUSIONS

We developed a bioinformatics pipeline for finding SNP markers from polyploid *Scaphirhynchus* using target enrichment data. SNP markers found in this pipeline can be applied to a range of species identification and phylogenetic questions, and for conservation genetics of endangered sturgeons and paddlefishes.

ACKNOWLEDGMENTS

We are grateful to E.J. Hilton, Virginia Institute of Marine Science, Gloucester Point, VA, USA, R.M. Wood, Saint Louis University, St. Louis, MO, USA, B.R. Kuha-jda, University of Alabama, Tuscaloosa, AL, USA and R.L. Mayden, Saint Louis University, St. Louis, MO, USA

for tissues of *Scaphirhynchus*. We thank the anonymous reviewers for their constructive suggestions.

FUNDING

This work was supported by the Science and Technology Commission of Shanghai Municipality (19050501900). C.B. Dillman was supported by U.S. National Science Foundation (DEB-0841691).

AUTHOR CONTRIBUTIONS

C. Li and T. Zhou conceived the research. T. Zhou and J. Huang performed experiment and carried the analysis. T. Zhou, J. Huang, C.B. Dillman, Y. He, and C. Li drafted

the manuscript. All authors have edited and approved the final version of the manuscript.

COMPLIANCE WITH ETHICAL STANDARDS

Conflict of interests. The authors declare that they have no conflicts of interest.

Statement on the welfare of animals. The animal study was reviewed and approved by the Animal Ethics Committee of Shanghai Ocean University, China (no. 22, Aug. 2014).

SUPPLEMENTARY INFORMATION

Custom scripts used for generating the results are available at GitHub (https://github.com/TaoZhou2021/SNP_marker). Supplementary file of sequences of conserved loci (fasta) and filtered SNPs (vcf) were available by searching the <https://data.mendeley.com/datasets/77z94-jv27v/2>.

REFERENCES

- Birstein, V.J. and Bemis, W.E., How many species are there within the genus *Acipenser*?, *Environ. Biol. Fishes*, 1997, vol. 48, no. 1–4, pp. 157–163.
<https://doi.org/10.1023/A:1007354827093>
- Birstein, V.J., Desalle, R., Doukakis, P., et al., Testing taxonomic boundaries and the limit of DNA barcoding in the Siberian sturgeon, *Acipenser baerii*, *Mitochondrial DNA*, 2009, vol. 20, no. 5–6, pp. 110–118.
<https://doi.org/10.3109/19401730903168182>
- Chassaing O., Hänni C., and Berrebi P., Distinguishing species of European sturgeons *Acipenser* spp. using microsatellite allele sequences, *J. Fish Biol.*, 2011, vol. 78, no. 1, pp. 208–226.
<https://doi.org/10.1111/j.1095-8649.2010.02852.x>
- Cheng, P.L., Huang, Y., Lv, Y.Y., et al., The American Pad-dlefish Genome Provides Novel Insights into Chromosomal Evolution and Bone Mineralization in Early Vertebrates, *Mol. Biol. Evol.*, 2021, vol. 38, no. 4, pp. 1595–1607.
<https://doi.org/10.1093/molbev/msaa326>
- Comincini, S., Lanfredi, M., Rossi, R., and Fontana, F., Use of RAPD markers to determine the genetic relationships among sturgeons (Acipenseridae, Pisces), *Fish. Sci.*, 1998, vol. 64, no. 1, pp. 35–38.
<https://doi.org/10.2331/fishsci.64.35>
- Du, K., Stock, M., Kneitz, S., et al., The sterlet sturgeon genome sequence and the mechanisms of segmental rediploidization, *Nat. Ecol. Evol.*, 2020, vol. 4, no. 6, pp. 841–852.
<https://doi.org/10.1038/s41559-020-1166-x>
- Eichelberger, J.S., Braaten, P.J., Fuller, D.B., et al., Novel Single-Nucleotide Polymorphism Markers Confirm Successful Spawning of Endangered Pallid Sturgeon in the Upper Missouri River Basin, *Trans. Am. Fish. Soc.*, 2014, vol. 143, no. 6, pp. 1373–1385.
<https://doi.org/10.1080/00028487.2014.935479>
- Haxton, T.J., Sulak, K., and Hildebrand, L., Status of scientific knowledge of North American sturgeon, *J. Appl. Ichthyol.*, 2016, vol. 32, pp. 5–10.
<https://doi.org/10.1111/jai.13235>
- Janosik, A.M., Whitaker, J.M., VanTassel, N.M., and Rider, S.J., Improved environmental DNA sampling scheme for Alabama sturgeon provides new insight into a species once presumed extinct, *J. Appl. Ichthyol.*, 2021, vol. 37, no. 2, pp. 178–185.
<https://doi.org/10.1111/jai.14183>
- Jiang, J., Yuan, H., Zheng, X., et al., Gene markers for exon capture and phylogenomics in ray-finned fishes, *Ecol. Evol.*, 2019, vol. 9, no. 7, pp. 3973–3983.
<https://doi.org/10.1002/ece3.5026>
- Katoh, K., Misawa, K., Kuma, K., and Miyata, T., MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform, *Nucleic Acids Res.*, 2002, vol. 30, no. 14, pp. 3059–3066.
<https://doi.org/10.1093/nar/gkf436>
- Khlestkina, E. K., and Salina, E. A., SNP markers: Methods of analysis, ways of development, and comparison on an example of common wheat, *Russ. J. Genet.*, 2006, vol. 42, no. 6, pp. 585–594.
<https://doi.org/10.1134/S1022795406060019>
- Knaus, B. J., and Grünwald, N. J., VCFR: a package to manipulate and visualize variant call format data in R, *Mol. Ecol. Res.*, 2017, vol. 17, no. 1, pp. 44–53.
<https://doi.org/10.1111/1755-0998.12549>
- Krieger, J., Fuerst, P. A. and Cavender, T. M., Phylogenetic relationships of the North American sturgeons (order Acipenseriformes) based on mitochondrial DNA sequences, *Mol. Phylogenet. Evol.*, 2000, vol. 16, no. 1, pp. 64–72.
<https://doi.org/10.1006/mpev.1999.0743>
- Krieger, J., Hett, A.K., Fuerst, P.A., et al., The molecular phylogeny of the order Acipenseriformes revisited, *J. Appl. Ichthyol.*, 2008, vol. 24, no. s1, pp. 36–45.
<https://doi.org/10.1111/j.1439-0426.2008.01088.x>
- Li, H., Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM, *arXiv pre-print server*, 2013.
<https://doi.org/10.48550/arXiv.1303.3997>
- Li, C. H., Orti, G., Zhang, G., and Lu, G.Q., A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study, *BMC Evol. Biol.*, 2007, vol. 7, Article 44.
<https://doi.org/10.1186/1471-2148-7-44>
- Li, C.H., Riethoven, J.J.M., and Naylor, G.J.P., Evol-Markers: a database for mining exon and intron markers for evolution, ecology and conservation studies, *Mol. Ecol. Res.*, 2012, vol. 12, no. 5, pp. 967–971.
<https://doi.org/10.1111/j.1755-0998.2012.03167.x>
- Li, Z., De La Torre, A. R., Sterck, L., et al., Single-Copy Genes as Molecular Markers for Phylogenomic Studies in Seed Plants, *Genome Biol. Evol.*, 2017, vol. 9, no. 5, pp. 1130–1147.
<https://doi.org/10.1093/gbe/evx070>
- Liu, Y., Chen, Y.Y., Gong, Q., et al., Paternity assignment in the polyploid *Acipenser dabryanus* based on a novel microsatellite marker system, *PLoS One*, 2017, vol. 12, no. 9, Article e0185280.
<https://doi.org/10.1371/journal.pone.0185280>
- Ludwig, A., Identification of Acipenseriformes species in trade, *J. Appl. Ichthyol.*, 2008, vol. 24, pp. 2–19.
<https://doi.org/10.1111/j.1439-0426.2008.01085.x>

- Ludwig, A., Belfiore, N.M., Pitra, C., et al., Genome duplication events and functional reduction of ploidy levels in sturgeon (*Acipenser*, *Huso* and *Scaphirhynchus*), *Genetics*, 2001, vol. 158, no. 3, pp. 1203–1215.
<https://doi.org/10.1093/genetics/158.3.1203>
- Luo, D.H., Li, Y.P., Zhao, Q.Y., et al., Highly resolved phylogenetic relationships within order acipenseriformes according to novel nuclear markers, *Genes*, 2019, vol. 10, no. 1, Article 38.
<https://doi.org/10.3390/genes10010038>
- McQuown, E.C., Sloss, B.L., Sheehan, R.J., et al., Microsatellite analysis of genetic variation in sturgeon: New primer sequences for *Scaphirhynchus* and *Acipenser*, *Trans. Am. Fish. Soc.*, 2000, vol. 129, pp. 1380–1388.
[https://doi.org/10.1577/1548-8659\(2000\)129<1380:MAO-GVI>2.0.CO;2](https://doi.org/10.1577/1548-8659(2000)129<1380:MAO-GVI>2.0.CO;2)
- Murphy, C.E., Hoover, J.J., George, S.G., and Killgore, K.J., Morphometric variation among river sturgeons (*Scaphirhynchus* spp.) of the Middle and Lower Mississippi River, *J. Appl. Ichthyol.*, 2007, vol. 23, no. 4, pp. 313–323.
<https://doi.org/10.1111/j.1439-0426.2007.00883.x>
- Nguyen, L.T., Schmidt, H.A., von Haeseler, A., and Minh, B., Q. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies, *Mol. Biol. Evol.*, 2015, vol. 32, no. 1, pp. 268–274.
<https://doi.org/10.1093/molbev/msu300>
- Rothfels, C.J., Larsson, A., Li, F.W., et al., Transcriptome-mining for single-copy nuclear markers in ferns, *PLoS One*, 2013, vol. 8, no. 10, Article e76957.
<https://doi.org/10.1371/journal.pone.0076957>
- Schrey, A.W. and Heist, E.J. Stock structure of pallid sturgeon analyzed with microsatellite loci, *J. Appl. Ichthyol.*, 2007, vol. 23, no. 4, pp. 297–303.
<https://doi.org/10.1111/j.1439-0426.2007.00881.x>
- Schrey, A.W., Boley, R., and Heist, E.J., Hybridization between pallid sturgeon *Scaphirhynchus albus* and shovelnose sturgeon *Scaphirhynchus platyrhynchus*, *J. Fish Biol.*, 2011, vol. 79, no. 7, pp. 1828–1850.
<https://doi.org/10.1111/j.1095-8649.2011.03123.x>
- Shen, W., Le, S., Li, Y., and Hu, F.Q., SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation, *PLoS One*, 2016, vol. 11, no. 10, Article e0163962.
<https://doi.org/10.1371/journal.pone.0163962>
- Shen, Y., Yang, N., Liu, Z. et al., Phylogenetic perspective on the relationships and evolutionary history of the Acipenseriformes, *Genomics*, 2020, vol. 112, no. 5, pp. 3511–3517.
<https://doi.org/10.1016/j.ygeno.2020.02.017>
- Sheraliev, B. and Peng, Z.G., Complete mitochondrial genome sequence and phylogenetic position of the Amu Darya sturgeon, *Pseudoscaphirhynchus kaufmanni* (Acipenseriformes: Acipenseridae), *J. Appl. Ichthyol.*, 2020, vol. 36, no. 4, pp. 389–392.
<https://doi.org/10.1111/jai.14043>
- Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J. M., and Birol, I. ABySS: A parallel assembler for short read sequence data, *Genome Res.*, 2009, vol. 19, no. 6, pp. 1117–1123.
<https://doi.org/10.1101/gr.089532.108>
- Van De Peer, Y., Mizrachi, E., and Marchal, K., The evolutionary significance of polyploidy, *Nat. Rev. Genet.*, 2017, vol. 18, no. 7, pp. 411–424.
<https://doi.org/10.1038/nrg.2017.26>
- Yuan, H., Jiang, J.M., Jimenez, F.A., et al., Target gene enrichment in the cyclophyllidean cestodes, the most diverse group of tapeworms, *Mol. Ecol. Resour.*, 2016, vol. 16, no. 5, pp. 1095–1106.
<https://doi.org/10.1111/1755-0998.12532>
- Zeng, L.P., Zhang, N., Zhang, Q.A., et al., Resolution of deep eudicot phylogeny and their temporal diversification using nuclear genes from transcriptomic and genomic datasets, *New Phytol.*, 2017, vol. 214, no. 3, pp. 1338–1354.
<https://doi.org/10.1111/nph.14503>
- Zhang, H., Jarić, I., Roberts, D.L., et al., Extinction of one of the world's largest freshwater fishes: Lessons for conserving the endangered Yangtze fauna, *Sci. Total Environ.*, 2020, vol. 710, Article 136242.
<https://doi.org/10.1016/j.scitotenv.2019.136242>