Estimates of the Effective Population Size and Genetic Structure of the Critically Endangered Ship Sturgeon (*Acipenser nudiventris*) in the Chinese Section of the Ili River

Guo Hu 1, Feng Chen 1, Ying Zhang 1, Peixian Luan 1, Zhiyuan Luo 1, Jiangong Niu 2, Peng Zheng 1, Sai Wang 1, Tao Zhang 2, Yongjun Shu 3 and Feng Ji 1,*

1 Key Laboratory of Freshwater Aquatic Biotechnology and Breeding, Ministry of Agriculture and Rural Affairs, Heilongjiang River Fisheries Research Institute of Chinese Academy of Fishery Sciences, Harbin 150070, China; huguoh@hrfri.ac.cn (G.H.); fengcayh@163.com (F.C.); zhangying@hrfri.ac.cn (Y.Z.); luanpeixian@hrfri.ac.cn (P.L.); 15299255831@163.com (Z.L.); pengzheng420@163.com (P.Z.); saiwang320@163.com (S.W.)
2 Fisheries Science Research Institute of Xinjiang Uygur Autonomous Region, Urumqi 830000, China; njg20008@163.com (J.N.); 15099581796@163.com (T.Z.)
3 Key Laboratory of Molecular Cytogenetics and Genetic Breeding of Heilongjiang Province, College of Life Science and Technology, Harbin Normal University, Harbin 150025, China; syjun2003@126.com

*Correspondence: jifeng@hrfri.ac.cn; Tel.: +86-451-84861434

Abstract: Ship sturgeon (*Acipenser nudiventris*) is a critically endangered fish that is listed on the International Union for Conservation of Nature’s Red List of Threatened Species. Sixteen individuals from the Chinese section of the Ili River were genotyped using genome re-sequencing technology. By applying a genomic relatedness estimation with 1,527,694 genome-wide SNP markers, we found that the coancestry coefficients showed a high level of relatedness between individuals. The effective population sizes over 500 generations were estimated, and this showed that the effective population sizes began to dramatically decline from about 14,840 to 171 individuals when going back four generations from the current population. Artificial reproduction techniques guided by genomic relatedness may be a valuable approach to the conservation of this critically endangered fish species.

Keywords: ship sturgeon; genome re-sequencing; effective population size; Ili River; China

Key Contribution: As a critically endangered species, ship sturgeon is a second-class nationally protected wild animal in China and a first-class protected wild animal in the Xinjiang Uygur Autonomous Region. In this study of genomic relatedness, it was found that the individuals in the Ili River population were closely related to each other, and the estimated effective population size was small, with a high risk of extinction.

1. Introduction

Ship sturgeon (*Acipenser nudiventris*) is a large sturgeon and an ancient species with a very high ecological and economic value; it can be called a “living fossil” in water [1]. Ship sturgeons used to be widely distributed in the Black Sea and the Sea of Azov in Europe, the Caspian Sea and Aral Sea in Central Asia, and some rivers in these areas; it is of great value for maintaining the ecological balance of the waters of its habitat [2]. The caviar produced by ship sturgeon is of good quality and high price. It has always been a traditional and precious food in Europe. After years of overfishing, environmental changes, and fording projects, the sturgeon was recognized by the International Union for Conservation of Nature (IUCN) as critically endangered on its Red List of Endangered Species (www.iucnredlist.org).

Sturgeons were moved from the Syr River to the Sino–Soviet cross-border river system of Ili River–Balkhash Lake by the former Soviet Union during the period from 1933 to 1934,
and a large fishing population was formed 16 years later [3]. The Ili River is the largest inland river in China and the most abundant river in Xinjiang, with an annual runoff of 16.5 billion cubic meters [4]. Due to the existence of suitable spawning grounds in the upper tributaries of the Ili River, mature parent fish often migrate from Kazakhstan to the upper reaches of the Ili River in Xinjiang, China. In the 1970s, stable economic catches were formed in the Xinjiang reach of the Ili River, with an annual yield of about 40 tons [5]. In the 1990s, around the collapse of the Soviet Union, the fishery management order of Kazakhstan was in chaos [6]. Affected by water conservancy projects, environmental pollution, illegal fishing, and other factors, wild ship sturgeon was on the verge of extinction in the Ili River–Balkhash Lake system of Kazakhstan. Since then, due to the lack of downstream population supplementation, the population size in the Ili River of Xinjiang was small and was only occasionally found, and it was unable to form a fishing output. This fish is listed as a first-class protected wild animal in Xinjiang Uygur Autonomous Region and as a second-class protected wild animal in China [3]. Over the years, people from all walks of life in Xinjiang have paid more and more attention to the protection of sturgeon species. Therefore, the rescue and protection of wild sturgeons has high social value.

In the Xinjiang section of the Ili River, males generally reach sexual maturity at 6–9 years old and most females do so at 12–14 years old. According to fishery resource surveys and monitoring records, few live sturgeons have been caught in the branches and tributaries of the Xinjiang section of the Ili River, and those that were caught were mainly young individuals. By collecting wild sturgeons for many years and domesticating them in an artificial environment, we achieved successful artificial reproduction in 2021 [3].

In this study, captive wild ship sturgeons collected in the Ili River in Xinjiang were re-sequenced to obtain SNP markers with high coverage across the whole genome in order to evaluate the effective population size and genetic structure and provide valuable information for the conservation and restoration of ship sturgeon species in the river system of Ili River–Balkhash Lake.

2. Materials and Methods

2.1. Genome Re-Sequencing and SNP Calling

Our team scientifically monitored ship sturgeons caught in the Chinese section of the river system of Ili River–Balkhash Lake (the green area in Figure S1; all distribution areas in Figure S1 came from the IUCN). The wild fish in this study were caught with permission provided by a scientific fishing license issued by the Agriculture and Rural Affairs Department of Xinjiang Uygur Autonomous Region. Fishing permits must be renewed and reapproved annually. In this study, there were 16 individuals—5 adults, 7 subadults, and 4 juveniles—and their genders were not identified. The compound 2-phenoxyethanol (C₈H₁₀O₂) was used as an anesthetic to prevent injury to the fish during handling and cutting of some of their caudal fins. Then, the fin samples were stored in absolute ethanol in a refrigerator at −20 °C until the extraction of genomic DNA. The genomic DNA was extracted by using the standard phenol–chloroform protocol, and the DNA samples were sequenced on a DNBseq platform (PE150 model) from the BGI company (Shenzhen, China), according to the manufacturer’s protocols. The raw sequence data were submitted to the Genome Sequence Archive [7] of the National Genomics Data Center [8] (Beijing Institute of Genomics (BIG), Chinese Academy of Sciences).

All sequencing reads were mapped to the genome of *A. ruthenus* (sterlet sturgeon) (https://www.ncbi.nlm.nih.gov/assembly/GCF_010645085.2) (Accessed on 18 December 2022) [9] by using the BWA software with the default parameters [10]. The alignments of the SAM files were converted into BAM files; by using SAMtools, they were sorted, and the individual IDs were inserted according to the file names [11]. SNP calling was executed for each alignment for individual fish samples by using the commands of “mpileup” and “call” in BCFtools with the default parameters [11]. The SNP sites were filtered by using VCFtools with the following settings, “--max-alleles 2 --min-alleles 2 --minDP 5 --max-missing 0.95 --remove-indels” [12].
2.2. Analysis of the Genomic Structure of the Population and Individual Kinship

The ROH (run of homozygosity) was estimated for each individual by using PLINK v. 1.9 [13,14]. Although no linkage disequilibrium (LD)-based pruning was performed, the minimum ROH length was set to 1 Mb to exclude short and common ROHs that appeared across the genome due to LD. The following PLINK parameters and thresholds were applied to define an ROH: (i) a sliding window of 50 SNPs across the genome; (ii) the proportion of homozygous overlapping windows was 0.05; (iii) the minimum number of consecutive SNPs included in an ROH was 50; (iv) the minimum length of an ROH was set to 1 Mb; (v) the maximum gap between consecutive homozygous SNPs was 1000 kb; (vi) a density of one SNP per 50 kb; and (vii) up to two heterozygous genotypes were allowed in an ROH. All ROHs were counted, and the average length of the ROHs was estimated. Additionally, the pairwise linkage disequilibrium (LD) was measured as a genotypic correlation coefficient ($r^2$), and the LD decay among the ship sturgeon individuals across the whole genome was generated by using the PopLDdecay software [15].

The pairwise genomic coancestry coefficient was calculated between all samples in NGSRelate V2.0 by using the SNP data across the whole genome [16]. The first two principal components (PC1 and PC2) of a principal component analysis (PCA) were calculated to evaluate the population stratification in the PLINK software (v1.90) [13,14]. All of the figures presented for data analysis were drawn with the ggplot2 package in R.

2.3. Estimation of Ancestral and Contemporary Effective Population Sizes

Estimates of the effective population size (Ne) were implemented in the GONE software by using the SNP genotypes [17], and the GONE program was run with the following parameters: “PHASE = 2, cMMb = 1, DIST = 2”; other parameters were set to their default values. This included an unknown phase, an average rate of recombination of 0.01 cm/Mb, and genetic distance correction based on Haldane’s function.

3. Results

3.1. Statistics and Evaluation of the Sequencing Data

A total of 16 samples were sent to BGI for genomic DNA extraction, and a fragment library was established. The DNBseq platform was used for sequencing, and terminal PE150 sequences were paired. A total of 24.85 million reads were generated from 16 samples, with an average of 155 million reads per sample. The average sequencing depth was about $25 \times$, as shown in Table 1. The re-sequencing data were mapped to the genome of A. ruthenus with the BWA software. A total of 33,788,217 SNP sites were identified; low-quality sites were filtered, insertion and deletion sites were inserted, and 1,527,694 SNPs were retained for genetic analysis. These loci were evenly distributed throughout the genome, as shown in Figure S2.

Table 1. Sample information of ship sturgeon for genome re-sequencing.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Clean Reads</th>
<th>Clean Base</th>
<th>Q20 (%)</th>
<th>GC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS01</td>
<td>160,042,004</td>
<td>48,012,601,200</td>
<td>97.13</td>
<td>39.86</td>
</tr>
<tr>
<td>SS02</td>
<td>160,082,176</td>
<td>48,024,652,800</td>
<td>97.23</td>
<td>39.92</td>
</tr>
<tr>
<td>SS03</td>
<td>160,132,305</td>
<td>48,039,691,500</td>
<td>97.08</td>
<td>39.84</td>
</tr>
<tr>
<td>SS04</td>
<td>160,073,512</td>
<td>48,022,053,600</td>
<td>97.4</td>
<td>40.03</td>
</tr>
<tr>
<td>SS05</td>
<td>160,007,653</td>
<td>48,002,295,900</td>
<td>97.27</td>
<td>40.01</td>
</tr>
<tr>
<td>SS06</td>
<td>160,157,713</td>
<td>48,047,313,900</td>
<td>97.45</td>
<td>39.98</td>
</tr>
<tr>
<td>SS07</td>
<td>160,105,711</td>
<td>48,031,713,300</td>
<td>97.7</td>
<td>40.12</td>
</tr>
<tr>
<td>SS08</td>
<td>160,105,071</td>
<td>48,031,521,300</td>
<td>97.61</td>
<td>40.16</td>
</tr>
<tr>
<td>SS09</td>
<td>160,044,445</td>
<td>48,013,333,500</td>
<td>97.43</td>
<td>39.81</td>
</tr>
<tr>
<td>SS10</td>
<td>128,969,960</td>
<td>38,690,988,000</td>
<td>97.17</td>
<td>40.1</td>
</tr>
</tbody>
</table>
Table 1. Cont.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Clean Reads</th>
<th>Clean Base</th>
<th>Q20 (%)</th>
<th>GC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS11</td>
<td>159,114,361</td>
<td>47,734,308,300</td>
<td>97.33</td>
<td>39.59</td>
</tr>
<tr>
<td>SS12</td>
<td>160,037,844</td>
<td>48,011,353,200</td>
<td>97.29</td>
<td>39.45</td>
</tr>
<tr>
<td>SS13</td>
<td>145,365,836</td>
<td>43,6097,50,800</td>
<td>97.43</td>
<td>39.77</td>
</tr>
<tr>
<td>SS14</td>
<td>160,061,705</td>
<td>48,018,511,500</td>
<td>97.56</td>
<td>39.62</td>
</tr>
<tr>
<td>SS15</td>
<td>151,665,763</td>
<td>45,499,728,900</td>
<td>97.45</td>
<td>39.59</td>
</tr>
<tr>
<td>SS16</td>
<td>139,409,107</td>
<td>41,822,732,100</td>
<td>97.38</td>
<td>39.84</td>
</tr>
</tbody>
</table>

3.2. Analysis of the Genomic Structure of the Population and Individual Kinship

The PCA results showed that the 16 individuals were clustered into two subgroups; SS01, 02, 05, 06, 13, 14, 15, and 16 were in one subgroup, and the others were grouped in the other subgroup (Figure 1). However, due to the small number of samples collected, the PCA structure was not clear enough.

![Figure 1](image-url)

Figure 1. Principal component analysis (PCA) of all 16 re-sequenced individuals of the ship sturgeon population in the Chinese section of the Ili River.

By using the PLINK software, 981 ROHs with an average length of 1.40 Mb were identified from the ship sturgeon population. The longest ROHs were found in the SS01 genome, with a length of about 8.98 Mb. At the same time, the results showed that the numbers of ROHs in SS14–16 significantly decreased to 38, 28, and 30, respectively, whereas the average number of ROHs in the other individuals was 67, as shown in Figure 2. These results suggested that the sturgeon population could maintain heterozygosity and might avoid the tendency of genetic degradation of the population.

The largest ROH with a length of 8.98 Mb in individual SS01 is not shown in the figure.

We also calculated the coancestry coefficients between all pairs of individuals and found a high level of relatedness between individuals (Figure 3, Table S1). In addition, to understand the relationship between the SNP pairs in a linked group, we conducted a linked disequilibrium (LD) delay analysis. The average LD decay distance was about 300 kb ($r^2 = 0.1$) when all of the SNP markers were used for the LD analysis (Figure S3).
3.3. Estimation of Ancestral and Contemporary Effective Population Sizes

The effective population sizes over 500 generations were estimated, and this showed that the effective population size began to decline 500 to 450 generations back from the current population, then remained stable until it rapidly increased to a historical peak of about 19,241 individuals at 100 generations before the current population. However, the population size began to sharply decline again, especially in the last four generations, from about 14,840 to 171 individuals (Figure 4).
A. ruthenus. These results suggested a high similarity between segmented re-diploidation [9].

... similarity between was more than 99.47% and overall was about 99.50% (Table S2). These results suggested a high similarity between sturgeon genomes to genome sequences of A. ruthenus. Before using the genome sequences of A. ruthenus, genome mapping of reads from sturgeon genomes to genome sequences of A. ruthenus similar. Before using the genome sequences of A. ruthenus as reference genomes of ship sturgeons, genome mapping of reads from sturgeon genomes to genome sequences of A. ruthenus. were performed using the mapping tool “bwa mem”, and the results showed that the mapping rate for each sample was more than 99.47% and overall was about 99.50% (Table S2). These results suggested a high similarity between A. ruthenus and A. nudiventris not only morphologically but also in genomic sequence. Therefore, the A. ruthenus genome can be used as a reference genome for re-sequencing research on ship sturgeons.

The coancestry coefficient estimates the actual pairwise relatedness levels and may be helpful in species conservation, including captive breeding and reintroduction programs. Both the PCA results and the analysis of the genomic coancestry coefficient measures suggested that these individuals probably came from a small population, and the parents of these juveniles came from the adults in our study. The genetic base of the population is already very narrow now, and very few individuals have been found while monitoring the
Ili River. These sporadic individuals should be bred in artificial conditions, as it is difficult for them to meet in the wild and successfully naturally reproduce. Reproductive strategies guided by genomic relatedness may be a valuable approach to the conservation of this rare fish species. Individuals involved in mating can be directly controlled by setting a genomic coancestry coefficient index for wild fish individuals without a pedigree record [20].

The habitat of ship sturgeons was originally extensive, from the Black and Azov Seas to the inland Caspian and Aral Seas, as well as the rivers that connected them. With the collapse of the Soviet Union, the management of fisheries in many of these areas was thrown into disarray, and, worse, state-led efforts to increase and release ship sturgeons came to a halt. Under various pressures, a large number of native populations rapidly shrank and were even depleted. The populations introduced into the Ili River–Balkhash Lake system face similar challenges. By the time the management of fisheries was restored, the stocks of species had reached extremely endangered levels. As this is a transboundary river–lake system, the overall population of the individual fish is not very clear, and the effective population size is difficult to predict based on traditional approaches. In this study, we reconstructed the historical effective population sizes of this population with genome-wide SNP markers. This population has clear historical records from populations in the Aral Sea–Syr River system, and the results suggested that the effective population size of this population decreased after it left the original population and colonized the Ili River–Balkhash Lake system about 10 generations ago; the effective population size was relatively stable in the following four generations and then declined in a collapsing manner. As this population migration event occurred only about 100 years ago, it was clearly documented. By contrast, the results of these population genomics analyses were consistent with historical records. Due to the rarity of the global population size, the results of effective population size evaluations of other geographic populations have not been published. The sample size was also very small and could only be used as a very loose reference in this study. However, the results also suggested that the ship sturgeon population appears to have been caught in an extinction vortex in the Ili River.

5. Conclusions

In summary, the evidence obtained in this study indicated that the effective population size of wild ship sturgeons in the Chinese section of the Ili River is very small, and the genomic kinships among individuals in the population are also very close. Captive breeding and aquaculture in farmed ponds should be considered to sustainably maintain this sturgeon species in the Ili River.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/fishes8070354/s1, Table S1. Estimated genomic coancestry coefficients between each pair of 16 re-sequenced individuals. Table S2. Genome mapping of reads from ship sturgeon genome to genome sequences of A. ruthenus. Figure S1. Global habitat distribution map of the ship sturgeon according to information from the IUCN. Figure S2. Genome-wide distribution of the SNP markers identified in this study. Figure S3. Linkage disequilibrium decay analysis with the SNP markers identified in this study.

Author Contributions: Conceptualization, G.H. and F.J.; methodology, G.H. and Y.S.; investigation, F.C., Y.Z., P.L., J.N., P.Z., S.W., and T.Z.; data curation, G.H. and Y.S.; writing—original draft preparation, G.H. and F.J.; writing—review and editing, G.H. and F.J.; supervision, F.J.; funding acquisition, F.J. and G.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Key Research and Development Projects of Ili Kazak Autonomous Prefecture (YZ2021A001), the Central Public-Interest Scientific Institution Basal Research Fund (2020TD22), and the Special Project for Agricultural Financial Funding “Survey of fishery resources and environment in key waters in Northwest China” from the Ministry of Agriculture and Rural Affairs of China.

Institutional Review Board Statement: All procedures on the experimental animals were carried out in accordance with the guidelines for Experimental Animal Care and Use of Heilongjiang Fisheries
Research Institute, Chinese Academy of Fishery Sciences. The animal experiments were examined and approved by the Experimental Animal Welfare and Ethics Committee of Heilongjiang Fisheries Research Institute, Chinese Academy of Fishery Sciences (approval number: 2018-11-14-01).

**Data Availability Statement:** The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive of the National Genomics Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences (GSA: CRA011020).

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

1. Bekbergenova, V. *Current Biological Data on the Ship Sturgeon Acipenser Nudiventris Lovetsky;* 1828 (Review); Springer International Publishing: Cham, Switzerland, 2023.


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.