Project Report

Identifying Target Spawning Populations of Silver Carp *Hypophthalmichthys molitrix* Through Genomic Analysis for Directed Management

Submitted to Mississippi River Basin Panel on Aquatic Invasive Species

By

Guoqing Lu, Ph.D.

Department of Biology

University of Nebraska at Omaha

3/14/2025

Project Summary

Invasive silver carp pose a serious threat to native biodiversity and long-term ecosystem stability within the Mississippi River Basin (MRB). This study examined genetic diversity and variation in silver carp across the basin using high-throughput genomic sequencing. Analysis of populations from 16 locations revealed moderate to high overall genetic diversity, with significant genetic differentiation observed in specific invasion front populations, notably those from MOBG and AUB. Phylogenetic analyses further indicated that these populations form distinct lineages, suggesting divergence from core MRB populations. This divergence may be driven by factors such as multiple invasion events, genetic introgression, interspecific hybridization with bighead carp, and/or rapid local adaptation, which requires further investigation. The findings from this project have important implications for management, particularly in informing targeted control strategies for invasive populations.

1. Introduction

Silver carp (*Hypophthalmichthys molitrix*), a large filter-feeding fish native to eastern Asia, poses a significant threat to ecosystems in the Mississippi River Basin (Conover et al., 2007; Kolar et al., 2007). Imported to the United States for aquaculture, this invasive species has become a major ecological threat. Their voracious feeding habits disrupt food webs by consuming massive quantities of plankton, a critical food source for many native fish populations. Enormous efforts have been made by the U.S. government to prevent the introduction and establishment of invasive carp including silver carp in the Great Lakes (Chapman et al., 2023). Effective management of invasive species requires a deep understanding of their biology, including their genetic makeup (Lu et al., 2020). This project focuses on utilizing high-throughput next-generation sequencing to investigate the genetic diversity and differentiation of silver carp populations in the Mississippi River Basin. Understanding their genetic makeup is crucial for developing targeted control measures.

Studies suggest that silver carp and bighead carp (hereafter bigheaded carps) possess moderate to high heterozygosity when compared to other species (Farrington et al., 2017; Wang et al., 2019), although North American silver carp populations exhibit lower genetic diversity compared to their native Asian counterparts (Farrington et al., 2017; Lamer et al., 2014; Li et al., 2011). This could be attributed to the founder effect, where a small number of individuals establish a new population, leading to a loss of genetic variation (Lu et al., 2020). Additionally, limited population structure has been observed within the invasive range, possibly a consequence of the rapid spread (Farrington et al., 2017). However, one study indicates potential genetic differences within the invaded area, in particular between invasion fronts (Stepien et al., 2019) .

This project employs Restriction site-associated DNA sequencing (RAD-Seq), a high-throughput genomic technique, to assess genetic variation in silver carp populations across the Mississippi River Basin (MRB). The specific aims were (1) to investigate population genetic diversity and structure to better understand the genetic landscape of silver carp in the MRB and (2) to identify key populations that may serve as strategic targets for management interventions. By harnessing the power of population genomics, this study seeks to generate actionable insights for controlling invasive silver carp and safeguarding the ecological integrity of the MRB.

2. Materials and Methods

2.1 Specimen Collection

Silver carp samples were obtained from field scientists of state and federal agencies in the MRB region and our university collaborators (Lamer et al., 2014; Fig. 1). Sampling sites hereafter are labeled: ARLR, Arkansas River near Little Rock; ATCH, Atchafalaya River; AUB (Red River – below Lake Texoma, including RED1, RED2, RED3, and KIAM); CUCL, Cumberland River- Cheatham Lake; ILAG, Illinois River LaGrange Reach; IMAR, Illinois River Marseilles; MKY, Mississippi River (near Laketon, KY); MOBG, Missouri River below Gavins Point dam; MOO, Missouri River (near Omaha, NE); MSAG, Lower Mississippi River from the Arkansas River to the Gulf of Mexico; OHBC, Ohio River; PL20, Mississippi River at Pool 20; Pool 26, Mississippi River (near Alton, IL); P8, Upper Mississippi River at Pool 8; SC, Tennessee River - Kentucky Lake; WRNE, White River at the north end.

2.2 DNA Extraction, Library Preparation, and Sequencing

Genomic DNA was extracted from fin clips using the Qiagen DNeasy Blood and Tissue Kit (Germantown, MD). DNA concentration was quantified on a Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA). For RAD-Seq library preparation, 150 ng of DNA from each sample was used. Restriction enzyme digestion employed the enzyme EcoRI (FastDigest Enzymes, Thermo Fisher Scientific) followed by a 10-minute heat denaturation step at 80°C. Digested products were purified using MagBio HighPrep paramagnetic beads (Gaithersburg, MD) at a 1.8x ratio. Indexed TruSeq-compatible sequencing adapters (Illumina, San Diego, CA) were ligated onto the fragments using T4 DNA ligase (Invitrogen, Carlsbad, CA) followed by a bead-based size selection step. Ligation products were then amplified using KAPA HiFi Amplification Mix (Agilent, Santa Clara, CA) for 18 cycles. Another size selection step was performed using beads at specific ratios to achieve the desired fragment size range. Individual libraries were quantified using a Qubit, and fragment size distribution was assessed using an Agilent Fragment Analyzer. Finally, libraries were pooled in equimolar amounts for sequencing.

Initial sequencing was performed on an Illumina NextSeq 550 platform using a 2 x 150 bp paired-end read format at the UNMC Genomics Core Facility. However, some samples displayed low quality, and these samples were re-sequenced on an Element AVITI platform using a 2 x 75 bp paired-end read format. In 2024, we expanded our sequencing efforts, generating 2 x 75 bp paired-end reads on the Element AVITI platform at the UC Davis Genomics Core, which achieved broader geographical coverage of additional samples.

2.3 Bioinformatics Analysis

The quality of raw sequencing data was assessed using FastQC (Andrews, 2010). MultiQC was then employed to consolidate results from FASTQC into a single report

(Ewels et al., 2016). Reads were demultiplexed based on sample-specific barcodes. Low-quality reads with a Q-score below 20 were removed, and the remaining reads were trimmed to 50 bp using trimmomatic v0.39 (Bolger et al., 2014). Following this initial quality control step, the de novo assembly and SNP discovery pipeline within denovo_map.pl was used, including running ustas, cstacks, sstacks, tsv2bam, and gstacks programs in Stacks2 (Catchen et al., 2013). Population structure analysis was performed using the Populations program in Stacks2, with a minimum of 70% individuals across populations required to process a locus. De novo mapping and genetic structure analysis were conducted on the usegalaxy.eu server.

An alternative approach for de novo locus calling and genotyping software utilized (Rojas, 2018). Raw I7 reads (corresponding to the Eco RI restriction fragment) were used as input, and default parameters were applied with the exception of RE-AATTC and minDepth-5 settings within AftrRAD.pl, and MinReads-5 in Genotype.pl. Only loci successfully genotyped in at least 95% of the samples (allowing for up to 5% missing data per locus) were retained for further analysis (pctScored-95 in FilterSNPs.pl). Allele frequencies were used for Principal Component clustering analysis.

The CalculateHeterozygosity.pl script within AftrRAD was used to estimate genome-wide heterozygosity values based on the 75-bp reads. Genotype data were then formatted as appropriate and used as input to Genepop to calculate Fst values between population pairs, as input to Adegenet for clustering analysis based on discriminant analysis of principle components, and as input to Structure, with K=1-5, burn-in of 100000, and 500000 MCMC iterations to assess whether signals of population structure are present in the data (Pritchard et al., 2000). Additionally, the Genepop input file was restructured to treat all samples as coming from one hypothetical population, and Hardy-Weinberg tests were performed on each locus. The loci with the highest individual pairwise Fst values based on the Genepop Fst analysis were identified, and their corresponding aligned R2 sequences were obtained from the AftrRAD output.

3 Results

3.1 Sequencing Data

A total of 238 samples from 16 locations were sequenced using the Element AVITI platform. The NextSeq system generated approximately 951.3 million raw reads, of which 660.1 million high-quality reads remained after removing duplicates and low-quality sequences (Table 1). The resulting sequencing data have been made publicly available through the NCBI database under BioProject accession number

PRJNA1063311 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA1063311), providing a valuable resource for the research community.

3.2 Genetic Diversity and Variation

De novo mapping identified nucleotide positions ranging from 14,978,218 to 24,673,136 across samples from 16 locations, yielding a total of 317,010 variant sites and an average of 0.66% polymorphic loci (Table 2). Among the 16 sampling sites, private nucleotide positions ranged from 673 to 54,682 (mean ± SD: 8352 ± 13673). The highest private positions were found in AUB (Red River – below Lake Texoma), MOBG (Missouri River below Gavins Point dam), and P8 (Upper Mississippi River at Pool 8). Observed heterozygosity ranged from 0.00204 to 0.00279 (0.00236 ± 0.00022), while expected heterozygosity ranged from 0.00171 to 0.00278 (0.00204 ± 0.00026; Fig 2). Nucleotide diversity (π) values ranged from 0.00213 to 0.00319 (0.00247 ± 0.00028). Collectively, these results indicate moderate to high levels of genetic variation within invasive silver carp populations (Table 2).

3.3 Population Differentiation

The majority of pairwise F_{ST} values were below 0.05, indicating low genetic differentiation among most populations, with the exception of AUB (Red River – below Lake Texoma) and MOBG (Missouri River – below Gavins Point Dam). AUB exhibited pairwise F_{ST} values ranging from 0.041 to 0.370 (0.201 ± 0.081), while MOBG exhibited values ranging from 0.071 to 0.155 (0.118 ± 0.027) (Table 3). According to the conventional interpretation, FST values less than 0.05 indicate low genetic differentiation, values between 0.05 and 0.15 suggest moderate differentiation and values greater than 0.15 reflect high genetic differentiation. These results suggest high genetic differentiation between AUB and other populations, as well as moderate differentiation between MOBG and other sites. Notably, both AUB and MOBG are considered invasion-front populations among the sampling sites (Fig. 1).

3.4 Population Genetics Structure

Population genetic structure analysis, utilizing the Evanno method (ΔK), indicated that K=3 best explained the observed genetic variation, suggesting three distinct genetic clusters within the studied populations. Examination of STRUCTURE results for K = 2 through K = 4 revealed that MOBG (Fig. 3) exhibited a unique genetic background, whereas AUB displayed higher levels of heterozygosity. Consistent with this, the Neighbor-joining tree revealed that MOBG and AUB formed distinct lineages (Table 4, Fig. 4). The remaining populations clustered into a separate clade. Within this clade, CUCL was positioned basally, while P8 was positioned at a tip, indicating a more recent divergence within the group. These findings suggest that the invasion front populations

(MOBG, AUB, CUCL, and P8) have significantly diverged from the core Mississippi River Basin populations. The observed phylogenetic pattern, with distinct lineages and varying levels of divergence among invasion front populations (MOBG, AUB, CUCL, and P8), may reflect a combination of factors, including multiple invasion events from different tributaries and across different years, potential genetic introgression and interspecific hybridization, and/or rapid evolutionary adaptation in the invasion fronts.

4 Summary and Future Directions

This project successfully generated a comprehensive dataset of next-generation sequencing reads from silver carp populations sampled across 16 locations within the Mississippi River Basin (MRB). These data have been submitted to NCBI and are publicly available for further research and comparative analyses (NCBI BioProject PRJNA1063311).

Our analysis revealed moderate to high levels of genetic diversity across populations, consistent with previous findings. A relatively high genetic diversity may enhance the adaptive potential of invasive species, contributing to their invasion success. However, further research is needed to identify specific genes or gene networks involved in invasion dynamics.

Population structure and phylogenetic analyses uncovered significant genetic differentiation in invasion front populations, particularly in AUB and MOBG, suggesting distinct evolutionary trajectories. These findings build upon and extend earlier work (e.g., Stepien et al., 2019), offering new insight into the genetic mechanisms underlying invasion spread.

For future research, the availability of well-annotated, complete reference genomes of invasive carp will be essential. Such a resource will enable the identification of loci associated with key traits, particularly those linked to invasion success. This information could prove critical for developing targeted and effective control strategies. Notably, interspecific hybridization between silver carp and bighead carp is a well-documented phenomenon. These hybrids may exhibit heterosis—enhanced fitness resulting from the combination of parental genomes—which could further increase their invasive potential. Therefore, future studies should explore the landscape genomics of invasive carp hybrids to assess the emergence and spread of potentially more problematic hybrid genotypes.

5 Acknowledgements

We acknowledge funding from the Mississippi River Basin Panel on Aquatic Nuisance Species. We thank panel members Eugene Braig, Duane Chapman, Patrick Kroboth, and Chris Steffen for their advice and suggestions. We also thank Greg

Conover and Elizabeth Brown for program support. Jim Lamer provided silver carp samples. Mike Sovic of Freedom Genomics LLC provided expertise in protocol optimization, library preparation, and data analysis. Stephen F. Spear's discussions enhanced the project's direction. We appreciate the field scientists from federal and state agencies who shared samples.

References

Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114-2120. https://doi.org/10.1093/bioinformatics/btu170

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124-3140.

Chapman, D., Amberg, J., Calfee, R., Hlavacek, E., Hortness, J., Jackson, P. R., Kazyak, D. C., Knights, B., & Roberts, J. (2023). *US Geological Survey invasive carp strategic framework, 2023–27*. US Geological Survey.

Conover, G., Simmonds, R., Whalen, M., & editors. (2007). *Management and control plan for bighead, black, grass, and silver carps in the United States.* . Aquatic Nuisance Species Task Force, Washington, D.C

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047-3048.

Farrington, H. L., Edwards, C. E., Bartron, M., & Lance, R. F. (2017). Phylogeography and population genetics of introduced Silver Carp (Hypophthalmichthys molitrix) and Bighead Carp (H-nobilis) in North America. *Biological Invasions*, *19*(10), 2789-2811. https://doi.org/10.1007/s10530-017-1484-3

Kolar, C. S., Chapman, D. C., Courtenay, W. R., Housel, C. M., Williams, J. D., & Jennings, D. P. (2007). *Bigheaded carps: a biological synopsis and risk assessment*. American Fisheries Society.

Lamer, J. T., Sass, G. G., Boone, J. Q., Arbieva, Z. H., Green, S. J., & Epifanio, J. M. (2014). Restriction site-associated DNA sequencing generates high-quality single nucleotide polymorphisms for assessing hybridization between bighead and silver carp in the United States and China. *Mol Ecol Resour*, *14*(1), 79-86.

Li, S., Xu, J., Yang, Q., Wang, C., Chapman, D., & Lu, G. (2011). Significant genetic differentiation between native and introduced silver carp (Hypophthalmichthys molitrix) inferred from mtDNA analysis [Article]. *Environmental Biology of Fishes*, *92*(4), 503-511. https://doi.org/10.1007/s10641-011-9870-7

Lu, G., Wang, C., Zhao, J., Liao, X., Wang, J., Luo, M., Zhu, L., Bernatzhez, L., & Li, S. (2020). Evolution and genetics of bighead and silver carps: Native population

conservation versus invasive species control. *Evolutionary Applications*, *13*(6), 1351-1362.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945-959.

Rojas, K. A. V. (2018). Estructura genética mediante SNPs en bancos de almeja generosa (Panopea globosa; Dall 1898) del Alto Golfo de California.

Stepien, C. A., Snyder, M. R., & Elz, A. E. (2019). Invasion genetics of the silver carp Hypophthalmichthys molitrix across North America: Differentiation of fronts, introgression, and eDNA metabarcode detection. *PLoS One*, *14*(3). https://doi.org/ARTN e020301210.1371/journal.pone.0203012

Wang, J., Gaughan, S., Lamer, J. T., Deng, C., Hu, W. T., Wachholtz, M., Qin, S. S., Nie, H., Liao, X. L., Ling, Q. F., Li, W. T., Zhu, L. F., Bernatchez, L., Wang, C. H., & Lu, G. Q. (2019). Resolving the genetic paradox of invasions: Preadapted genomes and postintroduction hybridization of bigheaded carps in the Mississippi River Basin. *Evolutionary Applications*. https://doi.org/10.1111/eva.12863

Tables (See attached file)

Table 1. Sequencing Summary of Silver Carp Populations from 16 Sampling Locations in the Mississippi River Basin. ARLR, Arkansas River near Little Rock; ATCH, Atchafalaya River; AUB (Red River – below Lake Texoma, including RED1, RED2, RED3, and KIAM); CUCL, Cumberland River- Cheatham Lake; ILAG, Illinois River LaGrange Reach; IMAR, Illinois River Marseilles; MKY, Mississippi River (near Laketon, KY); MOBG, Missouri River below Gavins Point dam; MOO, Missouri River (near Omaha, NE); MSAG, Lower Mississippi River from the Arkansas River to the Gulf of Mexico; OHBC, Ohio River; PL20, Mississippi River at Pool 20, Pool 26, Mississippi River (near Alton, IL); P8, Upper Mississippi River at Pool 8; SC, Tennessee River - Kentucky Lake; WRNE, White River at the north end.

Table 2. Genetic Diversity Metrics of Silver Carp Populations across 16 Locations in the Mississippi River Basin. Sampling sites are labeled: labeled: ARLR, Arkansas River near Little Rock; ATCH, Atchafalaya River; AUB (Red River – below Lake Texoma, including RED1, RED2, RED3, and KIAM); CUCL, Cumberland River-Cheatham Lake; ILAG, Illinois River LaGrange Reach; IMAR, Illinois River Marseilles; MKY, Mississippi River (near Laketon, KY); MOBG, Missouri River below Gavins Point dam; MOO, Missouri River (near Omaha, NE); MSAG, Lower Mississippi River from the Arkansas River to the Gulf of Mexico; OHBC, Ohio River; PL20, Mississippi River at Pool 20; Pool 26, Mississippi River (near Alton, IL); P8, Upper Mississippi River at Pool 8; SC, Tennessee River - Kentucky Lake; WRNE, White River at the north end.

Table 3. Genetic Differentiation (Fst) among Silver Carp Populations from 16 Sampling Locations in the Mississippi River Basin. Sampling sites are labeled: labeled: ARLR, Arkansas River near Little Rock; ATCH, Atchafalaya River; AUB (Red River – below Lake Texoma, including RED1, RED2, RED3, and KIAM); CUCL, Cumberland River- Cheatham Lake; ILAG, Illinois River LaGrange Reach; IMAR, Illinois River Marseilles; MKY, Mississippi River (near Laketon, KY); MOBG, Missouri River below Gavins Point dam; MOO, Missouri River (near Omaha, NE); MSAG, Lower Mississippi River from the Arkansas River to the Gulf of Mexico; OHBC, Ohio River; PL20, Mississippi River at Pool 20; Pool 26, Mississippi River (near Alton, IL); P8, Upper Mississippi River at Pool 8; SC, Tennessee River - Kentucky Lake; WRNE, White River at the north end.

Table 4 Pairwise Genetic Distances among 16 Silver Carp Populations in the Mississippi River Basin. Sampling sites are labeled: labeled: ARLR, Arkansas River

near Little Rock; ATCH, Atchafalaya River; AUB (Red River – below Lake Texoma, including RED1, RED2, RED3, and KIAM); CUCL, Cumberland River- Cheatham Lake; ILAG, Illinois River LaGrange Reach; IMAR, Illinois River Marseilles; MKY, Mississippi River (near Laketon, KY); MOBG, Missouri River below Gavins Point dam; MOO, Missouri River (near Omaha, NE); MSAG, Lower Mississippi River from the Arkansas River to the Gulf of Mexico; OHBC, Ohio River; PL20, Mississippi River at Pool 20; Pool 26, Mississippi River (near Alton, IL); P8, Upper Mississippi River at Pool 8; SC, Tennessee River - Kentucky Lake; WRNE, White River at the north end.

Figures

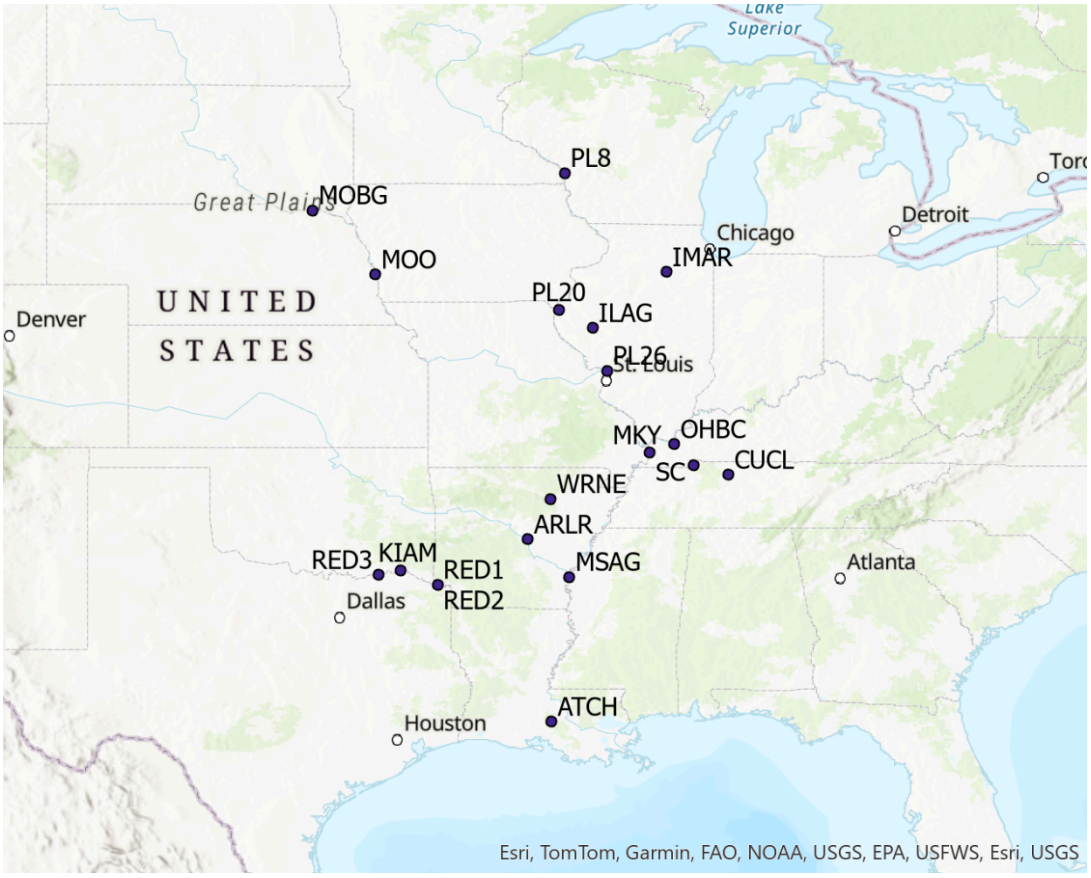

Fig 1. Sampling sites for genetic analysis: ARLR, Arkansas River near Little Rock; ATCH, Atchafalaya River; AUB (Red River – below Lake Texoma, including RED1, RED2, RED3, and KIAM); CUCL, Cumberland River- Cheatham Lake; ILAG, Illinois River LaGrange Reach; IMAR, Illinois River Marseilles; MKY, Mississippi River (near Laketon, KY); MOBG, Missouri River below Gavins Point dam; MOO, Missouri River (near Omaha, NE); MSAG, Lower Mississippi River from the Arkansas River to the Gulf of Mexico; OHBC, Ohio River; PL20, Mississippi River at Pool 20; Pool 26, Mississippi River (near Alton, IL); P8, Upper Mississippi River at Pool 8; SC, Tennessee River - Kentucky Lake; WRNE, White River at the north end.

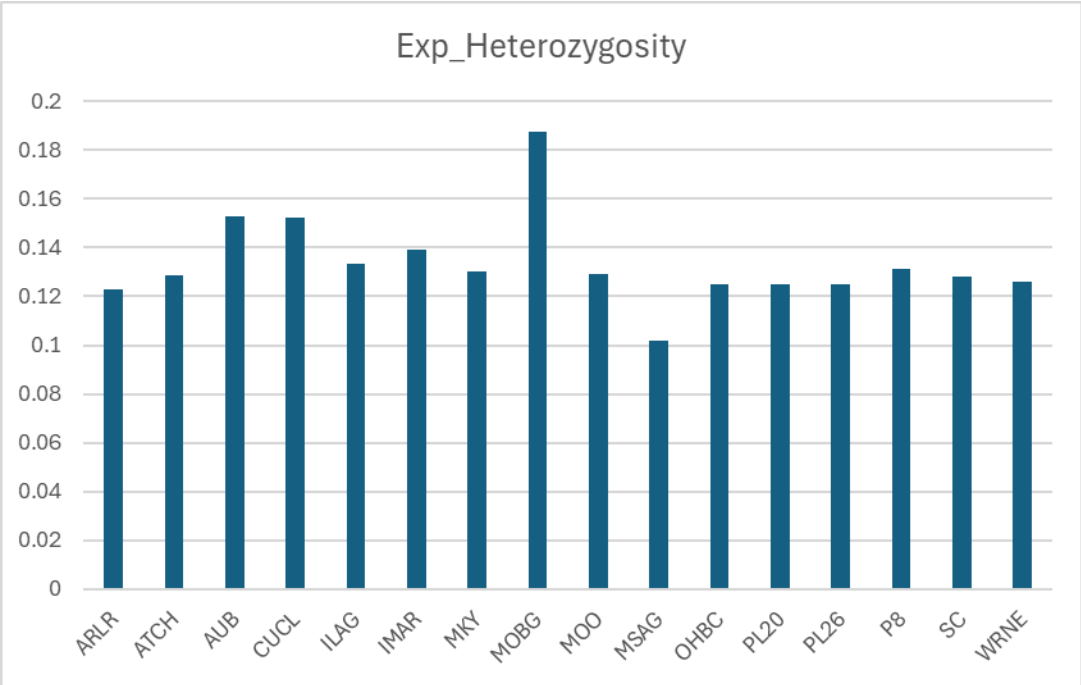

Fig 2. Expected heterozygosity (He) in Silver Carp populations at 16 sampling sites, labeled: ARLR, Arkansas River near Little Rock; ATCH, Atchafalaya River; AUB (Red River – below Lake Texoma, including RED1, RED2, RED3, and KIAM); CUCL, Cumberland River- Cheatham Lake; ILAG, Illinois River LaGrange Reach; IMAR, Illinois River Marseilles; MKY, Mississippi River (near Laketon, KY); MOBG, Missouri River below Gavins Point dam; MOO, Missouri River (near Omaha, NE); MSAG, Lower Mississippi River from the Arkansas River to the Gulf of Mexico; OHBC, Ohio River; PL20, Mississippi River at Pool 20; Pool 26, Mississippi River (near Alton, IL); P8, Upper Mississippi River at Pool 8; SC, Tennessee River - Kentucky Lake; WRNE, White River at the north end.

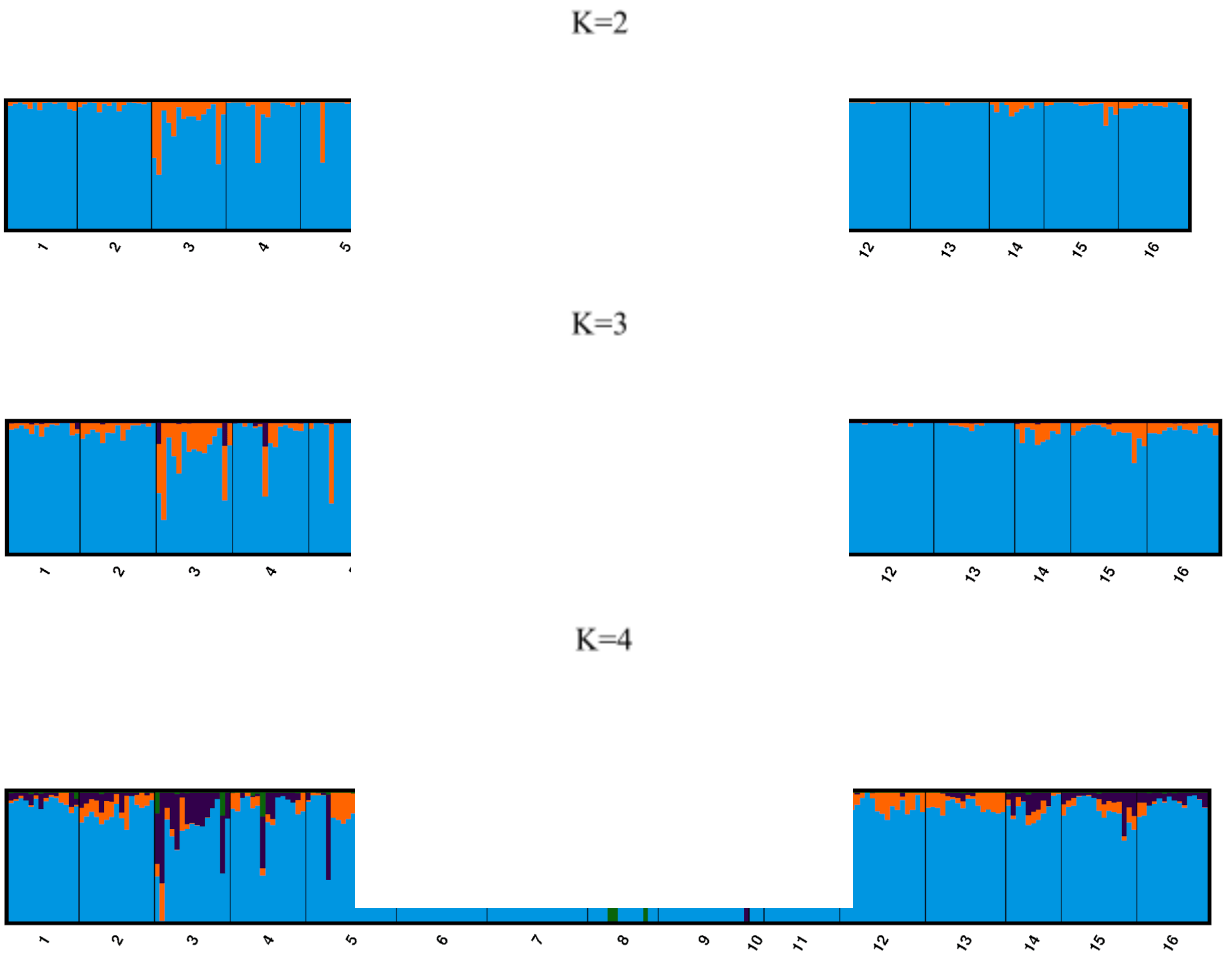

Fig. 3. Silver Carp population structure of 16 populations (ARLR-WRNE) as determined by STRUCTURE and STRUCTURESELECTOR, displayed for K=2, 3, and 4. Population labels: 1- Arkansas River near Little Rock, 2- Atchafalaya River, 3- Red River below Lake Texoma, 4- Cumberland River at Cheatham Lake, 5- Illinois River LaGrange Reach, 6- Illinois River Marseilles, 7- Mississippi River (near Laketon, KY), 8- Missouri River below Gavins Point dam, 9- Missouri River (near Omaha, NE), 10- Lower Mississippi River from the Arkansas River to the Gulf of Mexico, 11- Ohio River, 12- Mississippi River at Pool 20, 13- Mississippi River (near Alton, IL), 14- Upper Mississippi River at Pool 8, 15- Tennessee River - Kentucky Lake, 16- White River at the north end.

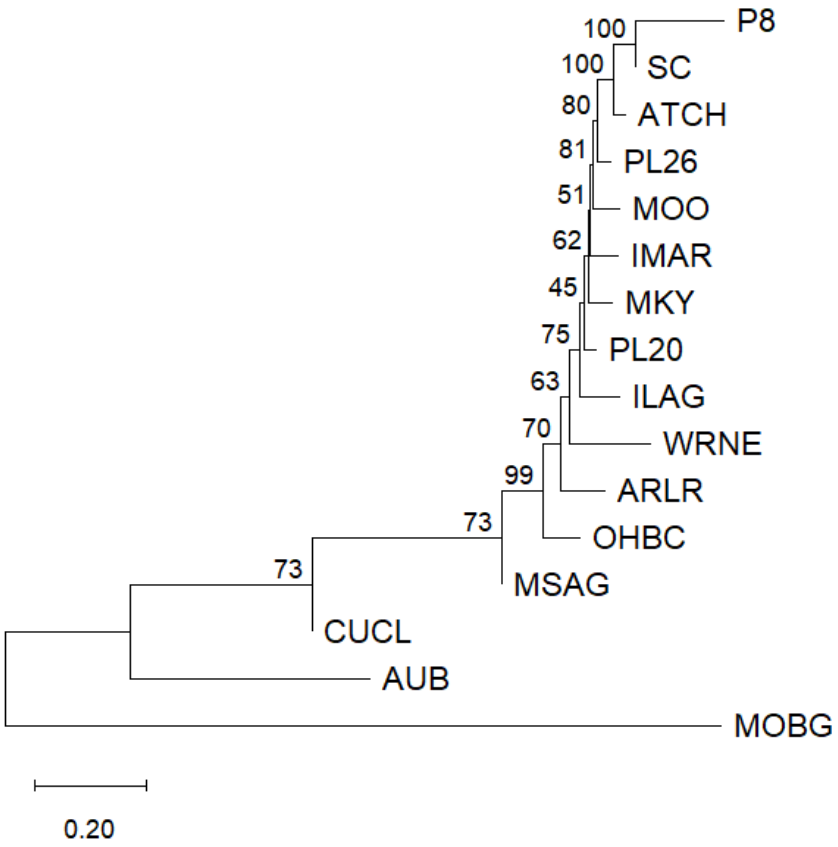

Fig 4. Phylogenetic relationships among 16 silver carp populations from the Mississippi River basin, depicted as a neighbor-joining tree. Branch lengths reflect nucleotide substitutions per site. Bootstrap support values (500 replicates) are indicated at the nodes. The scale bar represents 0.20 nucleotide substitutions per site. Sampling sites are labeled: labeled: ARLR, Arkansas River near Little Rock; ATCH, Atchafalaya River; AUB (Red River – below Lake Texoma, including RED1, RED2, RED3, and KIAM); CUCL, Cumberland River- Cheatham Lake; ILAG, Illinois River LaGrange Reach; IMAR, Illinois River Marseilles; MKY, Mississippi River (near Laketon, KY); MOBG, Missouri River below Gavins Point dam; MOO, Missouri River (near Omaha, NE); MSAG, Lower Mississippi River from the Arkansas River to the Gulf of Mexico; OHBC, Ohio River; PL20, Mississippi River at Pool 20; Pool 26, Mississippi River (near Alton, IL); P8, Upper Mississippi River at Pool 8; SC, Tennessee River - Kentucky Lake; WRNE, White River at the north end.