Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A next-generation genotyping process to balance demographic goals and genetic health for the Paddlefish (Polyodon spathula)

Leah K. Berkman

berkmank@umsystem.edu

University of Missouri

Alexandra Clussman

Missouri Department of Conservation

Nathan Storts

Missouri Department of Conservation

Skyler Kramer

University of Missouri

Madison Feeler

University of Missouri

Edward J. Heist

Southern Illinois University Carbondale

Michael E. Byrne

University of Missouri

Research Article

Keywords: Genetic health, Paddlefish, genome duplication, conservation, fisheries, supplementation

Posted Date: November 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11434343/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Captive rearing and reintroduction programs are essential for many wild populations, but they are often associated with negative genetic outcomes, including the loss of diversity and adaptation to captivity. Optimizing practices to prevent negative impacts to wild populations requires accurate and cost-effective genetic monitoring. We developed and tested a microsatellite panel for Paddlefish (*Polyodon spathula*) designed for next generation sequencing. We then used the panel to evaluate the genetic diversity of paddlefish stocked in 2022 and compared results to a similar investigation from 2017, when the hatchery released paddlefish at a later stage in development. The new panel successfully assigned parentage to the 2022 broodstock parents (n = 29) and a sample of their offspring (n = 250). Fifty markers satisfied our error rate threshold (< 0.25) and Hardy-Weinberg Equilibrium expectations in two wild populations. Comparison of 2022 results to results from 2017 suggest that earlier stocking increases the effective to census population size ratio (N_e/N) of the captive population and likely reduces adaptation to captivity. Our genotyping method circumvented complex bioinformatics tasks typically required to filter out paralogous loci in species with recent genome duplications, like Paddlefish. Further, the use of existing bioinformatics tools and maximum likelihood assignment methods that allow genotyping error made genotyping and assignment easy to implement quickly for conservation and management investigations.

Introduction

Effectively preserving and bolstering wild fish populations using captive rearing requires careful attention to the genetic composition of both groups (Fraser 2008; George et al. 2009). Captive reproduction and rearing, though important conservation and management tools, have been associated with negative outcomes for wild populations including the loss of genetic diversity, reduced fitness, and adaptation to captivity (Ryman and Laikre 1991; Lynch and O'Hely 2001; Miller et al. 2004; Araki and Schmid 2010; Christie et al. 2012; McMillan et al. 2023). It is possible to mitigate these negative impacts and help maintain the genetic integrity and evolutionary potential of wild populations using knowledge gained from genetic studies that evaluate hatchery practices (Sloss et al. 2009; Fisch et al. 2015; LaCava et al. 2015; Osborne et al. 2020; Brooks et al. 2025). Thus, the development and optimization of genetic tools to understand how hatchery practices influence the genetic composition of stocked fish can help managers improve effective population size and overall success.

Recently, genetic marker analysis by high-throughput amplicon sequencing has become as cost effective as more traditional electrophoresis-based techniques while offering a cheaper and more user-friendly alternative to reduced representation approaches like RADSeq and Genotype-by-Sequencing (Meek and Larson 2019). In the case of fish, recent genome duplications can present challenges for population analyses that rely on filtering paralogous loci from reduced representation sequencing (McKinney et al. 2017). Circumventing the bioinformatics tasks required to effectively filter out paralogous loci by instead developing highly vetted genetic markers can save resources and produce population-level analyses for management quickly. Such approaches have been applied to economically important aquatic species

aiding their conservation management including Walleye (*Sander vitreus*; Bootsma et al. 2020), Kokanee Salmon (*Oncorhynchus nerka*; Chang et al. 2021), Norway Lobster (*Nephrops norvegicus*; Gallagher et al. 2022) and Lake Trout (*Salvelinus namaycush*; Marcy-Quay et al. 2023).

For Paddlefish (*Polyodon spathula*) further development of molecular tools for conservation is timely due to ongoing challenges. Slow maturation, infrequent spawning, high commercial value, and low tolerance for disturbance can lead to over harvest and extinction risk when Paddlefish habitat is modified and degraded (Birstein et al. 1997; Billard and Lecointre 2001). Their closest relative, the Chinese Paddlefish (*Psephurus gladius*), declined due to habitat loss and overharvest and was recently declared extinct (Scarnecchia 2023).

Commercial and recreational harvest of Paddlefish are permitted across the large river systems of eastern North America (Rider et al. 2019) where individuals frequently move long distances, often crossing interjurisdictional boundaries (Pracheil et al. 2012; Tripp et al. 2019). Natural reproduction and recruitment are rarely studied but they have been documented across the Paddlefish range (Jennings and Zigler 2009; Scarnecchia et al. 2019; Schwinghamer et al. 2019). In response to the loss and inaccessibility of spawning habitat, hatchery supplementation programs have been implemented to help restore native populations (Argent et al. 2009) and create harvest opportunities (Onders et al. 2001). In the latter case, hatchery-reared Paddlefish are stocked in reservoirs, traditionally without careful regard for their genetic composition, but these individuals intermix with Paddlefish spawned upstream (Schwinghamer et al. 2019) and may leave the reservoirs and mix with the population downstream in the free-flowing portions of rivers (Stancill et al. 2002; Hupfeld et al. 2018).

The mixing of many hatchery-reared Paddlefish with few wild-spawned Paddlefish can degrade genetic integrity, decrease the diversity of the species, increase extinction risk, and hamper their evolutionary potential (Schwemm et al. 2019). More recently, attention has been given to analyzing and improving hatchery processes to reduce negative genetic impacts (e.g., Sloss et al. 2009). A study using microsatellites to assign parentage to 25- to 30-cm Paddlefish reared at the Blind Pony Hatchery (BPH) in Sweet Springs, Missouri USA revealed only 16% of the 2017 broodstock captured for spawning were represented in the stocked offspring, with a 50% reduction due to variable survival in rearing ponds (Heist and Gocker 2018). After this study, BPH began stocking hatchery-reared paddlefish at an earlier stage in development (12–20 cm), prior to experiencing the typical mortality that occurs in July and August when hatchery water temperatures in outdoor ponds can reach 29–32° C. Our goal was to repeat the parentage analysis at BPH with updated and improved molecular techniques aided by the recent publication of the Paddlefish genome (Cheng et al. 2021) and determine whether earlier stocking was associated with improved genetic diversity. We developed and tested a high-throughput microsatellite panel that can be used in future monitoring efforts and assessed whether earlier stocking of Paddlefish may reduce hatchery-related genetic diversity loss.

Materials and Methods

Locus Selection and Primer Design

Perfect simple sequence repeats (SSRs) were extracted from the published Paddlefish genome (Cheng et al. 2021) and primers were designed using Krait v1.4 (Du et al. 2018). Chromosomes 1 through 8 in Paddlefish have broad regions of homology among un-paired chromosomes, likely as a result of a Paddlefish-specific whole genome duplication event (Cheng et al. 2021) and were excluded to prevent selection of tetrasomic loci. We thus limited SSRs to chromosomes 9 through 60 and to di and tetra base pair combinations with a minimum number of seven and four repeats, respectively. Primers were designed for a product 100–300 base pairs in length. We then made a stratified random selection of one tetra SSR and 3 di SSRs per chromosome and eliminated loci situated within approximately 20,000 base pairs of a second locus (*n* = 5). Published primers for 27 paddlefish SSRs (Heist et al. 2002; Heist and Krampe 2011; Schwemm et al. 2014) were also included to produce a set of 230 loci for screening.

PCR and Sequencing

We screened primers with tissue samples chosen to represent genetic variation across the state of Missouri using two rounds of sequencing, similar to the strategies of Bootsma et al. (2020) and Marcy-Quay et al. (2023). Samples included for screening were obtained from the confluence of the Missouri and Osage rivers (*n* = 14) and from Table Rock Lake (*n* = 9) in March 2013 and from the Black River in October 2022 (*n* = 18).

We extracted DNA using the Qiagen DNeasy Blood and Tissue Kit with a modified protocol. All steps followed the manufacturer's recommendations except for the final elution step, where 60 µL of Buffer AE was incubated for 5 minutes at room temperature before elution. We quantified genomic DNA (gDNA) with a Qubit 2 fluorometer (ThermoFisher) and standardized to 20 ng/µL for the first and second round of screening but we did not normalize the gDNA concentration for the final parentage analysis.

Amplicons for sequencing were produced using two successive PCRs, the first to adhere to the genomic DNA template and apply an adaptor and the second to apply a molecular ID tag and adaptor for Illumina sequencing. We used PrimerPooler (Brown et al. 2017) to divide loci into two cocktails, one with 117 loci and one with 113 loci and kept them in separate locus pools through all PCRs and sequencing. For the second round of screening, we made two new cocktails (98 and 96 loci) and performed PCR separately but combined amplified products for the second PCR and sequencing. For the final round of sequencing, we amplified all loci using a single PCR cocktail.

The first PCR reaction consisted of 15 µL with 7.5 µL 2X Platinum Master Mix (Applied Biosystems), 1 µL gDNA, and a final primer concentration of 0.15 µM each. Reactions were subjected to pre-denaturation at 95°C for 5 minutes, 35 cycles of denaturation at 95° C for 30 seconds, annealing at 63°C for 1.5 minutes, and extension at 72°C for 1.5 minutes. PCR products were visualized on a 3% agarose gel, treated with ExoSAP-IT (Applied Biosystems), quantified by fluorometry, and then standardized to 1 ng/µL for the second PCR conducted in the University of Missouri's Genomics Technology Core (RRID:SCR_017778).

Products for sequencing were constructed by amplification of unique primers flanked by Illumina standard adapter sequences. The forward and reverse primer, in addition to the universal primer sequences, contain the full-length Illumina adapter sequences with dual 10-base indexes. PCR reactions contained 100 ng of DNA, 0.2 µM each of forward and reverse primers, 200 µM each of dNTPs, and 1 U Phusion High-Fidelity DNA Polymerase (ThermoScientific). PCR amplification was performed as follows: 98°C for 1 minute; then 6 cycles of 98° C for 15 seconds, 50° C for 30 seconds, 72° C for 45 seconds; followed by 72°C for 7 minutes. Amplified product (10 µl) from each PCR reaction was combined and thoroughly mixed to make a single pool. Pooled amplicons were purified by addition using AxyPrep MagPCR Clean-up beads (Axygen). The final amplicon pool was evaluated using the Advanced Analytical Fragment Analyzer automated electrophoresis system, quantified with the Qubit flourometer and diluted according to Illumina's standard protocol for sequencing on a MiSeq PE150 run.

Bioinformatics and Marker Screening

Sequence data were made available by the University of Missouri Genomics Technology Core in FASTQ format and checked for quality using FASTQC (Andrews 2010). We processed all subsequent scripts with high power computing resources provided by the University of Missouri (Hellbender HPC, https://doi.org/10.32469/10355/97710/). For the first round of screening, we analyzed sequence output with the ampicon.py workflow (Supplemental S1 adapted from https://github.com/cornell-bioinformatics/amplicon), examined the distribution of reads per locus (i.e., all reads containing the locus-specific priming sequence), and removed over-amplifying loci based on a plot of the log of relative reads (Bootsma et al. 2020). A primer sequence file with no flanking regions was provided as input to MEGASAT (Zhan et al. 2017) which then demultiplexed loci and individuals. The maximum number of mismatches allowed was two and minimum depth threshold was 20. After the second round, we removed loci with fewer than an average of 30 reads per sample. We then calculated the error rate, defined as all genotype mismatches divided by the number of genotype comparisons using custom R (R Core Team 2025) scripts (Supplemental S1), and discarded loci with an error rate > 0.25. We remade the primer cocktail with a reduced set of loci (n = 104) for the third and final round of PCR.

Using data from the second round of PCR, we ran MEGASAT's associated plotting functions in R to examine the number and length of sequences discarded as they resemble electropherogram plots of microsatellites and exhibit the characteristic stutter (polymerase slippage) patterns. For the final parentage analysis from the third PCR, we discarded loci with plots in which the true allele (highest peak) was too difficult to ascertain from stutter or other fragments. We also identified loci that had more than two true alleles for some individuals and assumed them to be tetrasomic. Summary statistics (number of alleles and inbreeding coefficients averaged over the four sample groups) for this final set of markers were generated with GenAlEx (Peakall and Smouse 2006, 2012). We assessed whether markers were in Hardy-Weinberg Equilibrium (HWE) with a sample of wild-caught, presumably unrelated individuals: the BPH 2022 broodstock and Black River sample groups. We used the R package *adegenet* (Jombart 2008) with the method of Guo and Thompson (1992) and a sequential Bonferroni correction (α = 0.05).

Parentage Analysis

We determined genotypes for parents (20 males, 14 females) that were spawned at BPH and a sample of their offspring (*n* = 50 for each of 5 pools with fry from 7–11 females stocked in each pool) using the screened set of markers. We used the full likelihood method in Colony (Jones and Wang 2010), to find the most likely family clusters, assign offspring to parents and provide a second determination of error rates. We provided Colony with the error rate calculated above from repeated genotyping as both the drop out and other error rate. If we found zero errors from repeat genotyping, we set the dropout rate as 0.01 and false allele rate as 0.001 for the locus.

Results were compared to a study of BPH Paddlefish production in 2017 that used electrophoretic techniques for microsatellite genotyping (Heist and Gocker 2018). For this study, genomic DNA was extracted as above and parents (13 females and 16 males) and a sample of 144 offspring from five pools (20–34 per pool) were genotyped at 11 microsatellite loci including *Psp-18*, *-21*, *-26*, *-28* (Heist and Mustapha 2008) and *Psp-302*, *-402*, *-611*, *-612*, *-626*, *-640* and *– 678* (Heist and Krampe 2011). Fragments were resolved on an Applied Biosystems 3130xl DNA analyzer and analyzed using GeneMapper software. Colony was used to reconstruct a genotype for one female for which data could not be obtained and the program Cervus (Marshall et al. 1998) was used for assignment to parents.

To examine whether greater diversity was retained with the changes associated with stocking paddlefish at an earlier developmental stage we compared metrics relevant to the effective population size (N_e) of fish stocked from BPH for production years 2017 and 2022. Genotyping more individuals from each pool for the 2022 cohort may have increased the detection of parents in the sample. Thus, we used rarefaction with the R package *vegan* to estimate the number of moms and dads identified from each pool in 2022 and compared the mean across pools with the Wilcoxon rank-sum test.

We used the full dataset to calculate the effective population size with the PwoP method of Waples and Waples (2011) which is relatively insensitive to differences in sample size and is designed for situations when the offspring of some parents may go undetected. The method assumes separate sexes with no selfing and estimates N_e using the number of samples (S) and the number of progeny k assigned to parent i:

$$N_e = \frac{2S - 2}{\sum (k_i^2)/2S - 1}$$

Results

Marker Screening

For the first round of screening, we removed over-amplifying loci (i.e., loci that encompassed > 3.5% of all sorted reads) but retained one legacy locus that over-amplified. After the second round of genotyping, we examined the plots of 104 loci that passed our amplification and error rate thresholds and used this

set in our third and last PCR. For the final parentage analysis, we removed 28 loci that had unreadable plots (e.g., stutter and true peaks indistinguishable or many extraneous peaks). With the readable plots, we diagnosed 16 loci as tetrasomic (i.e., clearly expressing more than two alleles in some individuals).

The remaining marker set (n = 60) contained 1–15 alleles per locus (Supplemental S2) and one monomorphic locus (*Psp-611*) was dropped for analysis in Colony. Error rates from repeat genotyping were 0-0.22 with an average of 0.048. Mistyping estimated from parent-offspring mismatches in Colony was 0-0.29 with average of 0.045 for allelic dropout and 0-0.14 with average 0.020 for false alleles. Nine loci were out of HWE ($p < 0.0008$) in at least one of the two groups assessed. Of these nine, six had relatively high allelic dropout rates (> 0.1; Fig. 1, Supplemental S2). For the remaining 50 loci that did not deviate from HWE, average inbreeding coefficients (F) ranged from − 0.288–0.460.

Parentage Analysis

All parentage assignments made with Cervus achieved 95% confidence and all assignments to clusters in Colony were made with 1.0 probability. One female that contributed to the 2022 cohort was not genotyped, but Colony assigned her offspring to an inferred genotype. For 2017, we identified 2–3 moms and 1–2 dads in each pool and, for 2022, we identified 7–8 moms and 7–10 dads in each pool (Table 1). Average rarified number of moms was 2.7 in 2017 and 6.4 in 2022 and differed significantly by year (p = 0.012). Average rarified number of dads was 1.2 in 2017 and 6.7 in 2022 and also differed significantly by year (p = 0.011).

Table 1
Summary results for genetic parentage assignment of American Paddlefish (*Polyodon spathula*) produced at the Blind Pony Hatchery in Missouri, USA. The raw numbers of parents detected from offspring in each pool are given along with the rarefied estimate standardized to the smallest sample (n = 20).

Sample Source	*n*	Moms Detected	Rarefied Moms	Dads Detected	Rarefied Dads
2017 Pond 1	20	3	30	1	10
2017 Pond 2	39	2	15	2	15
2017 Pond 3	25	3	30	1	10
2017 Pond 4	34	3	30	2	16
2017 Pond 5	26	3	28	1	10
2022 Pond 1	50	8	61	8	61
2022 Pond 2	50	8	72	10	78
2022 Pond 3	50	7	61	7	61
2022 Pond 4	50	8	63	9	71
2022 Pond 5	50	7	63	8	65

Female reproductive variance and the number of moms contributing to the cohort were similar in both years but male reproductive variance and the number of dads contributing to the cohort were much higher in 2022. The resulting N_e and ratio of N_e to the number of broodstock collected were also higher in 2022 than 2017 (Table 2).

Table 2
Demographic and sampling metrics of propagated paddlefish (*Polyodon spathula*). Metrics were collected and estimated using genetic parentage assignment studies in 2017 and 2022 at the Blind Pony Hatchery in Missouri, USA.

Broodstock Collected	**2017**	**2022**
	13 females, 16 males	**15 females, 20 males**
Females Producing Fry	11	11
Offspring genotyped	144	250
# dads detected in stockers	3	11
# moms detected in stockers	7	11
Variance in male output	1839	417
Variance in female output	402	480
N_e	53	124
N_e/Broodstock Collected	011	035

Discussion

We successfully identified and screened a set of microsatellite markers for use in the Paddlefish and used it to document an improvement in the genetic diversity of stocked Paddlefish following a change in the supplementation process, specifically, stocking paddlefish at an earlier stage of development. Supplementation programs have served Paddlefish management and conservation for decades, often providing angling opportunities in impounded portions of the Paddlefish range, where natural reproduction is very low or nonexistent (Hupfeld et al. 2018). Given the resources devoted them and the advancement of technologies that impact Paddlefish fisheries (e.g., live sonar; Cooke et al. 2025) managers should continue to evaluate supplementation programs and the systems they may affect (Schwemm et al. 2019). For instance, increased reservoir supplementation to offset improved angler efficiency could result in a significant portion of hatchery-reared Paddlefish moving past dams and mixing with the self-sustaining population in un-impounded portions of the major rivers, thus reducing their fitness (Ryman and Laikre 1991). Vetted genetic markers and case studies such as ours provide tools for continual and careful evaluation of supplementation programs to ensure they balance competing priorities and do no harm to the species (George et al. 2009).

Despite a recent genome duplication in Paddlefish, we were able to identify markers that expressed disomic inheritance and, thus, will be valuable for population genetic studies and parentage assignment. For parentage analysis, errors and missing genotypes can impact results and reduce power (Wang 2010). Repeated genotyping can help overcome these errors but may become a drain on resources. Although several of our loci had many missing genotypes and error rates considered high (> 5%) for microsatellites (Pampanon et al. 2005; Karaket and Poompuang 2012), the strategy of genotyping many loci simultaneously and incorporating error rates into the maximum likelihood estimation (e.g., Karaket and Poompuang 2012; Christie et al. 2013) allowed us to use many, albeit imperfect, loci for successful assignment.

We further followed a genotyping process that takes advantage of next-generation sequencing. Our marker set can be genotyped with a relatively simple and ready-to-use bioinformatics pipeline rather than one that relies on whole-genome alignment and assembly which might be otherwise complicated by the paddlefish genome duplication. We anticipate this marker set will be useful for other researchers exploring management and conservation of paddlefish, who may not have access to extensive bioinformatics support.

Understanding the composition of hatchery-reared fish is important, particularly for those that mingle with conspecific wild populations. Paddlefish supplementation programs present at least two potential genetic hazards: the loss of within population diversity and reduced fitness due to adaptation to captive conditions (Busack and Currens 1995). If a very small number of broodstock to contribute a substantial portion of offspring to the wild population, diversity and fitness of the wild population are negatively impacted (Ryman and Laikre 1991). This hazard is particularly relevant for Paddlefish for which relatively few individuals (< 25) are spawned (Sloss et al. 2009, Scarnecchia et al. 2026), due in part to invasive methods of egg collection (Mims 2001). Further, the parents initially selected for broodstock are rarely equally represented in the offspring that are ultimately stocked and choices throughout the hatchery process can exacerbate this loss of diversity (Sloss et al. 2009). We saw a marked increase in effective population size for the 2022 production year compared to 2017 which may be attributed to stocking at an earlier developmental stage suggesting that this choice better preserves the genetic contributions of broodstock and mitigates diversity loss (Waples et al. 2016).

In addition to diversity preservation, stocking at an earlier developmental stage may prevent adaptation to captivity and preserve fitness for the natural environment. Paddlefish experience the vast majority of mortality and considerable selection pressure within the earliest stages of growth but mortality rates can be greatly reduced in the hatchery (Mims 2001). Generally, paddlefish that can grow quickly have better survival (Scarnecchia et al. 2019), but the adaptations that allow paddlefish to feed optimally and grow quickly in the hatchery may not be optimal in the wild. Scarnecchia et al. (2026) provide evidence that suggests adaptation to hatchery feeding results in slower paddlefish growth and reproductive potential later in life, but many more domestication effects may occur that are difficult to document once paddlefish are released.

Stocking at an earlier developmental stage has its tradeoffs and uncertainties as well. This strategy will reduce the number of hatchery reared paddlefish that recruit to the population and, thus, the ability of the hatchery to meet its supplementation goals. Further, it is uncertain whether the benefit of greater diversity from this strategy will be retained as the cohort matures, but the paddlefish that are released will experience selection pressures in the natural environment at a critical point in their development and will ultimately be better fit for the natural system.

Our results point to a significant tradeoff between maximizing diversity and fitness of the stocked cohort and producing paddlefish likely to recruit into the population. Understanding the degree of intermixing among the supplemented reservoir and wild big river populations is important to properly evaluate this tradeoff and calibrate management practices accordingly (Waples et al. 2016). Many studies in Missouri have contributed to the knowledge of paddlefish movement (e.g., Stancill et al. 2002; Hupfeld et al. 2018) but not enough is known about reproduction and the effective size of wild river natural population (Scarnecchia et al. 2019). The useful, simple, and effective genetic marker set that we produced for such investigations will help inform and fine tune future paddlefish management and conservation practices.

Declarations

Author contributions

LKB and AC wrote the main manuscript and prepared tables and figures LKB, EJH, and NS conceived of the study and secured funding. NS collected samples and organized collection efforts. AC and MF performed lab work and provided lab support. KS and LKB wrote data analysis scripts. MEB provided administration and facilities. All authors reviewed and edited the manuscript.

Funding

The Missouri Department of Conservation provided funding for this project.

Data availability

Data are available in the supplemental materials. Any additional raw files will be made available upon request from the corresponding author.

Competing interests.

The authors declare no competing interests.

Acknowledgement

The Missouri Department of Conservation Management and Evaluation fund provided the resources for this project. We thank Tom Boersig, Kasey Whiteman, Charlayna Cammarata, Dave Herzog, and Jake Colehour for project development and manuscript review. Lyndon Coghill, Hong An, Nick Sievert, Grant

Zane, and Nathan Bivens provided technical support for marker sequencing and bioinformatics coding. This work was made possible by the Genomics and Bioinformatics Cores, High Performance Computing, and the School of Natural Resources at the University of Missouri

References

1. Andrews S (2010) FastQC: A Quality Control Tool for High Throughput Sequence Data http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ Accessed 25 July 2024
2. Araki H, Schmid C (2010) Is hatchery stocking a help or harm? Evidence, limitations and future directions in ecological and genetic surveys. Aquaculture 308:S2–S11. https://doi.org/10.1016/j.aquaculture.2010.05.036
3. Argent DG, Kimmel WG, Lorson R et al (2009) Paddlefish restoration in the upper Ohio and Allegheny River systems. In: Paukert CP, Scholten GB (eds) Paddlefish management, propagation and conservation in the 21st century: building from 20 years of research and management. American Fisheries Society Symposium 66, Bethesda, Maryland, pp 2397–2409
4. Billard R, Lecointre G (2001) Biology and conservation of sturgeon and Paddlefish. Rev Fish Biol Fish 10:355–392. https://doi.org/10.1023/A:1012231526151
5. Birstein VJ, Bemis WE, Waldman JR (1997) The threatened status of acipenseriform species: a summary. Environ Biol Fishes 48:427–435. https://doi.org/10.1023/A:1007382724251
6. Bootsma ML, Gruenthal KM, McKinney GJ et al (2020) A GT-seq panel for walleye (*Sander vitreus*) provides important insights for efficient development and implementation of amplicon panels in non-model organisms. Mol Ecol Resour 20:1706–1722. https://doi.org/10.1111/1755-0998.13226
7. Brooks JA, Berkman LK, Zimmerscheid MM et al (2025) Genetic analysis of Missouri's Topeka Shiners with implications for the propagation of understudied small-bodied freshwater fishes. Trans Am Fish Soc 154:372–384. https://doi.org/10.1093/tafafs/vnaf014
8. Brown SS, Chen YW, Wang M et al (2017) PrimerPooler: automated primer pooling to prepare library for targeted sequencing. Biol Methods Protoc 2:bpx006. https://doi.org/10.1093/biomethods/bpx006
9. Busack CA, Currens KP (1995) Genetic risks and hazards in hatchery operations: fundamental concepts and issues. Am Fish Soc Symp 15:71–80
10. Chang SL, Ward HG, Russello MA (2021) Genotyping-in-Thousands by sequencing panel development and application to inform kokanee salmon (*Oncorhynchus nerka*) fisheries management at multiple scales. PLoS ONE 16:e0261966. https://doi.org/10.1371/journal.pone.0261966
11. Cheng P, Huang Y, Lv Y et al (2021) The American paddlefish genome provides novel insights into chromosomal evolution and bone mineralization in early vertebrates. Mol Biol Evol 38:1595–1607. https://doi.org/10.1093/molbev/msaa326

12. Christie MR, Marine ML, French RA, Blouin MS (2012) Genetic adaptation to captivity can occur in a single generation. Proc Natl Acad Sci USA 109:238–242. https://doi.org/10.1073/pnas.1111073109
13. Christie MR, Tennessen JA, Blouin MS (2013) Bayesian parentage analysis with systematic accountability of genotyping error, missing data and false matching. Bioinformatics 29:725–732. https://doi.org/10.1093/bioinformatics/btt039
14. Cooke SJ, Neely BC, Hasler CT et al (2025) Emerging live sonar technologies in freshwater recreational fisheries: Issues and opportunities. Fisheries 50:66–74. https://doi.org/10.1093/fshmag/vuae003
15. Du L, Zhang C, Liu Q, Zhang X, Yue B (2018) Krait: an ultrafast tool for genome-wide survey of microsatellites and primer design. Bioinformatics 34:681–683. https://doi.org/10.1093/bioinformatics/btx665
16. Fisch KM, Kozfkay CC, Ivy JA, Ryder OA, Waples RS (2015) Fish hatchery genetic management techniques: integrating theory with implementation. N Am J Aquacult 77:343–357. https://doi.org/10.1080/15222055.2014.999846
17. Fraser DJ (2008) How well can captive breeding programs conserve biodiversity? A review of salmonids. Evol Appl 1:535–586. https://doi.org/10.1111/j.1752-4571.2008.00036.x
18. Gallagher J, Lordan C, Hughes GM, Jonasson JP, Carlsson J (2022) Microsatellites obtained using high throughput sequencing and a novel microsatellite genotyping method reveals population genetic structure in Norway Lobster, Nephrops norvegicus. J Sea Res 179:102139. https://doi.org/10.1016/j.seares.2021.102139
19. George AL, Kuhajda BR, Williams JD et al (2009) Guidelines for propagation and translocation for freshwater fish conservation. Fisheries 34:529–545. https://doi.org/10.1577/1548-8446-34.11.529
20. Guo SW, Thompson EA (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. Biometrics 48:361–372. https://doi.org/10.2307/2532296
21. Heist EJ, Nicholson EH, Sipiorski JT, Keeney DB (2002) Microsatellite markers for the paddlefish (Polyodon spathula). Conserv Genet 3:205–207. https://doi.org/10.1023/A:1015272414957
22. Heist EJ, Krampe MS (2011) Development of additional disomic polymorphic microsatellite loci in paddlefish (Polyodon spathula). Department of Zoology, Carbondale
23. Heist EJ, Gocker R (2018) A genetic assessment of survival bias in hatchery-reared paddlefish. Final Report for Cooperative Agreement N. 411 Missouri Department of Conservation, Jefferson City
24. Hupfeld RN, Phelps QE, Tripp SJ, Herzog DP (2018) Quantitative evaluation of Paddlefish sport fisheries in Missouri's large reservoirs: implications for the management of trophy sport fisheries. N Am J Fish Manage 38:295–307. https://doi.org/10.1002/nafm.10018
25. Jennings CA, Zigler SJ (2009) Biology and life history of paddlefish in North America: an update. In: Paukert CP, Scholten GB (eds) Paddlefish management, propagation and conservation in the 21st century: building from 20 years of research and management. American Fisheries Society Symposium 66, Bethesda, Maryland, pp 1–22

26. Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405. https://doi.org/10.1093/bioinformatics/btn129
27. Jones OR, Wang J (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. Mol Ecol Resour 10:551–555. https://doi.org/10.1111/j.1755-0998.2009.02787.x
28. Karaket T, Poompuang S (2012) CERVUS vs. COLONY for successful parentage and sibship determinations in freshwater prawn Macrobrachium rosenbergii de Man. Aquaculture 324:307–311. https://doi.org/10.1016/j.aquaculture.2011.10.045
29. LaCava M, Fisch K, Nagel M et al (2015) Spawning behavior of cultured Delta Smelt in a conservation hatchery. N Am J Aquacult 77:255–266. https://doi.org/10.1080/15222055.2015.1007192
30. Lande R, Barrowclough GF (1987) Effective population size, genetic variation, and their use in population management. In: Soulé ME (ed) Viable populations for conservation. Cambridge University Press, Cambridge, pp 87–124
31. Lynch M, O'Hely M (2001) Captive breeding and the genetic fitness of natural populations. Conserv Genet 2:363–378. https://doi.org/10.1023/A:1012550620717
32. Marcy-Quay B, Wilson CC, Osborne CA, Marsden JE (2023) Optimization of an amplicon sequencing-based microsatellite panel and protocol for stock identification and kinship inference of lake trout (*Salvelinus namaycush*). Ecol Evol 13:e10020. https://doi.org/10.1002/ece3.10020
33. Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Mol Ecol 7:639–655. https://doi.org/10.1046/j.1365-294x.1998.00374.x
34. McKinney GJ, Waples RK, Seeb LW, Seeb JE (2017) Paralogs are revealed by proportion of heterozygotes and deviations in read ratios in genotyping-by-sequencing data from natural populations. Mol Ecol Resour 17:656–669. https://doi.org/10.1111/1755-0998.12613
35. McMillan JR, Morrison B, Chambers N et al (2023) A global synthesis of peer-reviewed research on the effects of hatchery salmonids on wild salmonids. Fish Manage Ecol 30:446–463. https://doi.org/10.1111/fme.12643
36. Meek MH, Larson WA (2019) The future is now: Amplicon sequencing and sequence capture usher in the conservation genomics era. Mol Ecol Resour 19:795–803. https://doi.org/10.1111/1755-0998.12998
37. Miller LM, Close T, Kapuscinski AR (2004) Lower fitness of hatchery and hybrid rainbow trout compared to naturalized populations in Lake Superior tributaries. Mol Ecol 13:3379–3388. https://doi.org/10.1111/j.1365-294X.2004.02347.x
38. Mims SD (2001) Aquaculture of paddlefish in the United States. Aquat Living Resour 14:391–398. https://doi.org/10.1016/S0990-7440(01)01137-8
39. Onders RJ, Mims SD, Wang C, Pearson WD (2001) Reservoir ranching of paddlefish. N Am J Aquacult 63:179–190. https://doi.org/10.1577/1548-8454(2001)063%3C0179:RROP%3E2.0.CO;2

40. Osborne MJ, Dowling TE, Scribner KT, Turner TF (2020) Wild at heart: Programs to diminish negative ecological and evolutionary effects of conservation hatcheries. Biol Conserv 251:108768. https://doi.org/10.1016/j.biocon.2020.108768
41. Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295. https://doi.org/10.1111/j.1471-8286.2005.01155.x
42. Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537–2539. https://doi.org/10.1093/bioinformatics/bts460
43. Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. Nat Rev Genet 6:847–859. https://doi.org/10.1038/nrg1707
44. Pracheil BM, Pegg MA, Powell LA, Mestl GE (2012) Swimways: protecting Paddlefish through movement-centered management. Fisheries 37:449–457. https://doi.org/10.1080/03632415.2012.722877
45. R Core Team (2025) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. r-project.org
46. Rider S, Riecke DK, Scarnecchia DL (2019) Proactive harvest management of commercial Paddlefish Fisheries. Am Fish Soc Symp 88:267–297
47. Ryman N, Laikre L (1991) Effects of supportive breeding on the genetically effective population size. Conserv Biol 5:325–329. https://www.jstor.org/stable/2385902
48. Scarnecchia DL (2023) The Extinction of the Chinese Paddlefish Psephurus gladius: Transnationalism, Technology Transfer, and Timescape. Rev Fish Sci Aquacult 31:396–419. https://doi.org/10.1080/23308249.2023.2201636
49. Scarnecchia DL, Schooley JD, McAdam SO et al (2019) Factors affecting recruitment of Paddlefish: hypotheses and comparisons with sturgeons. In: Schooley JD, Scarnecchia DL (eds) Paddlefish: ecological, aquacultural, and regulatory challenges of managing a global resource. American Fisheries Society, Symposium 88, Bethesda, Maryland, pp 103–126
50. Scarnecchia DL, Lim Y, Slominski A et al (2026) Comparative Long-Term, within-Cohort Growth of Wild and Hatchery-Reared and Released Paddlefish Polyodon spathula. Trans Kans Acad Sci 129:59–77. https://doi.org/10.1660/062.129.0107
51. Schwemm MR, Echelle AA, Van Den Bussche RA, Schooley JD (2014) Development of diploid microsatellite markers for the North American paddlefish (Polyodon spathula). Conserv Genet Resour 6:217–218. https://doi.org/10.1007/s12686-013-0060-0
52. Schwemm MR, Asher AM, Heist EJ, Turner TF, Echelle AA (2019) Genetic management of North American Paddlefish: case studies and recommendations. In: Schooley JD, Scarnecchia DL (eds) Paddlefish: ecological, aquacultural, and regulatory challenges of managing a global resource. American Fisheries Society, Symposium 88, Bethesda, Maryland, pp 29–48

53. Schwinghamer CW, Tripp S, Phelps QE (2019) Using ultrasonic telemetry to evaluate Paddlefish spawning behavior in Harry S. Truman Reservoir, Missouri. N Am J Fish Manage 39:231–239. https://doi.org/10.1002/nafm.10263
54. Sloss BL, Klumb RA, Heist EJ (2009) Genetic conservation and paddlefish propagation. In: Paukert CP, Scholten GB (eds) Paddlefish management, propagation and conservation in the 21st century: building from 20 years of research and management. American Fisheries Society Symposium 66, Bethesda, Maryland, pp 307–327
55. Stancill W, Jordan GR, Paukert CP (2002) Seasonal migration patterns and site fidelity of adult paddlefish in Lake Francis Case, Missouri River. N Am J Fish Manage 22:815–824. https://doi.org/10.1577/1548-8675(2002)022%3C0815:SMPASF%3E2.0.CO;2
56. Tripp SJ, Phelps QE, Hupfeld RN et al (2019) Sturgeon and Paddlefish migration: evidence to support the need for interjurisdictional management. Fisheries 44:183–193. https://doi.org/10.1002/fsh.10215
57. Wang J (2010) Effects of genotyping errors on parentage exclusion analysis. Mol Ecol 19:5061–5078. https://doi.org/10.1111/j.1365-294X.2010.04865.x
58. Waples RS, Waples RK (2011) Inbreeding effective population size and parentage analysis without parents. Mol Ecol Resour 11:162–171. https://doi.org/10.1111/j.1755-0998.2010.02942.x
59. Waples RS, Hindar K, Karlsson S, Hard JJ (2016) Evaluating the Ryman–Laikre effect for marine stock enhancement and aquaculture. Curr Zool 62:617–627. https://doi.org/10.1093/cz/zow060
60. Zhan L, Paterson IG, Fraser BA et al (2017) MEGASAT: automated inference of microsatellite genotypes from sequence data. Mol Ecol Resour 17:247–256. https://doi.org/10.1111/1755-0998.12561

Figures

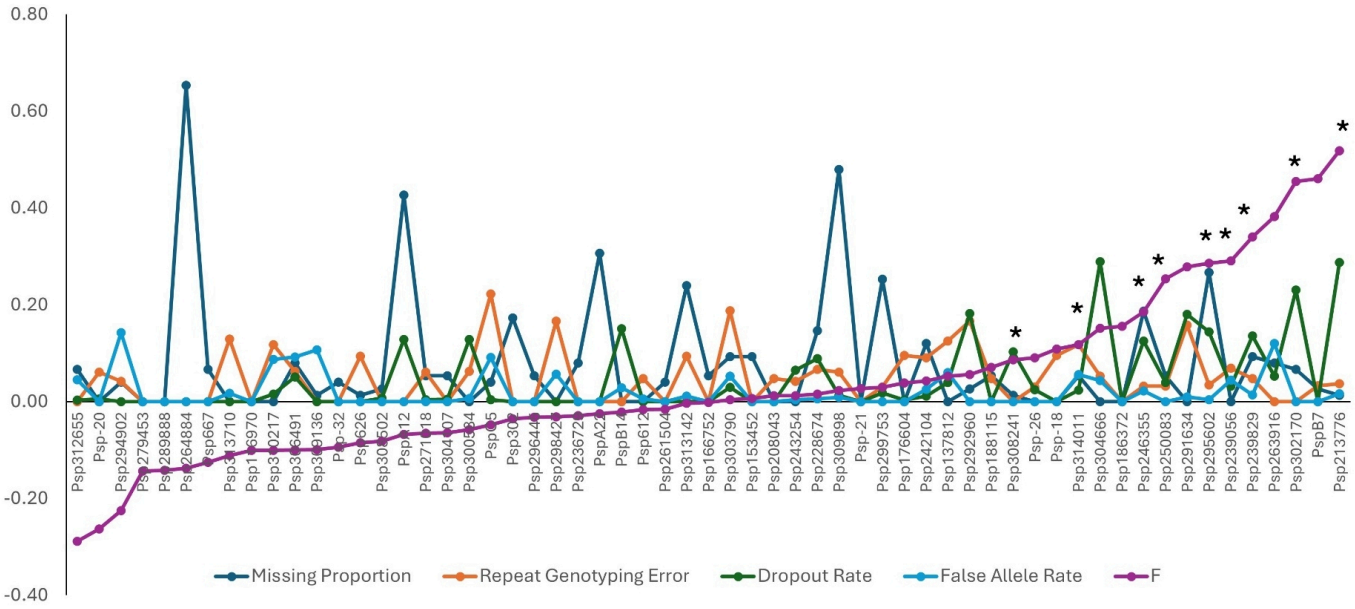


Figure 1

Attributes of 59 microsatellite markers designed for the American Paddlefish (*Polyodon spathula*) and assessed using samples from wild Missouri, USA, populations. The proportion missing, error rates (dropout rate and false allele rate from parent-offspring comparisons), and inbreeding coefficient (*F*) are presented with asterisks denoting markers that deviated significantly from Hardy-Weinberg Equilibrium.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplemental.zip