Platte River Pallid Sturgeon Genetics Project

Progress Report to Platte River Recovery Implementation Program

Ed Heist and Junman Huang

February 14, 2023

We Need Better Sturgeon Genetic ID

- Current ID based on 19 microsatellite markers
- Modeling in Jordan et al. (2019)*
  - 2% of pure Pallid Sturgeon identified as hybrids.
  - 9% of F_1 x Pallid back-crosses identified as pure Pallid Sturgeon.
  - 52% of F_1 x Shovelnose back-crosses identified as pure Shovelnose Sturgeon.
- We lack certainty about species ID of broodstock.
- We don't know what fraction of wild fish are hybrids.

*Jordan, G., E. J. Heist, B. R. Kuhajda, G. R. Moyer, P. Hartfield and M. Piteo (2019). "Morphological Identification Over-Estimates The Number Of Pallid Sturgeon (*Scaphirhynchus albus) In The Lower Mississippi River Due To Extensive Introgressive Hybridization."* *Transactions of the American Fisheries Society* ***148: 1004-1023.***

Genomic Sequencing of DNA to Develop Effective Single-nucleotide Polymorphism (SNP) Markers for the Endangered Pallid Sturgeon 2019-2022

- Funded through USGS Columbia Environmental Research Center
- Collaboration with Texas A&M Marine Genomics Laboratory
- Richard Flamio Jr. (Ph.D. 2022).

Received: 1 June 2022 | Revised: 2 November 2022 | Accepted: 22 November 2022

DOI: 10.1111/1755-0998.13742

MOLECULAR ECOLOGY RESOURCES WILEY

RESOURCE ARTICLE

Haploid gynogens facilitate disomic marker development in paleotetraploid sturgeons

Richard Flamio Jr.[1] | Dominic G. Swift[2] | David S. Portnoy[2] | Kimberly A. Chojnacki[3] | Aaron J. DeLonay[3] | Jeffrey Powell[4] | Patrick J. Braaten[3] | Edward J. Heist[1,5]

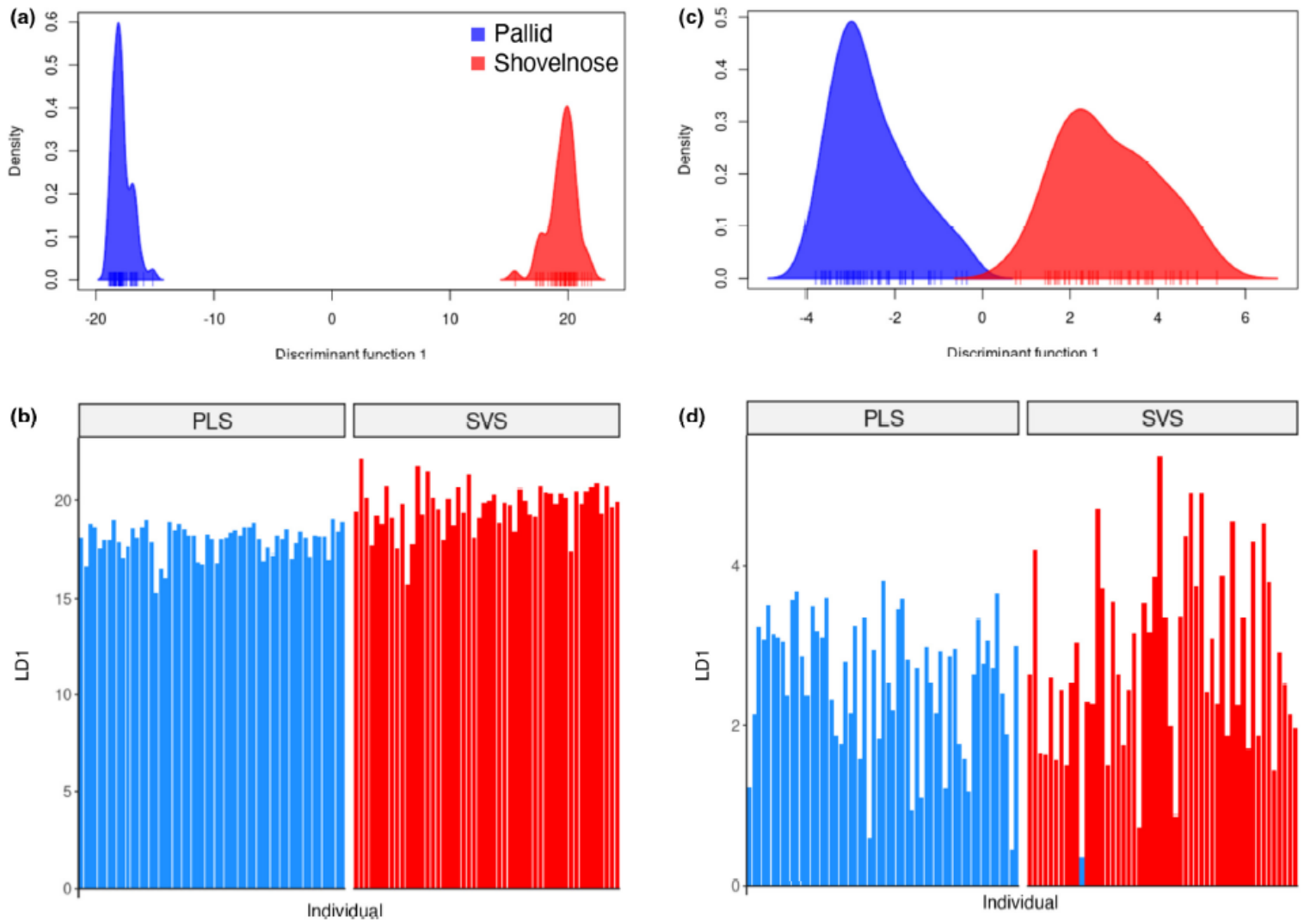


FIGURE 2 Discriminant analysis of principal components (DAPC) with k-means clusters = 2 and 15 principal components of 109 adult sturgeon using (a) 11,082 haplotyped loci and (c) 19 microsatellite loci. (b and d) Show assignment of field-identified pallid sturgeon individuals (*Scaphirhynchus albus*; PLS) and field-identified shovelnose sturgeon individuals (*S. platorynchus*; SVS) to genetic clusters based on DAPC of (a) and (c), respectively. (b) There was 100% correct assignment to species using 11,082 haplotyped loci. (d) One shovelnose sturgeon was misclassified as a pallid sturgeon using 19 microsatellite loci.

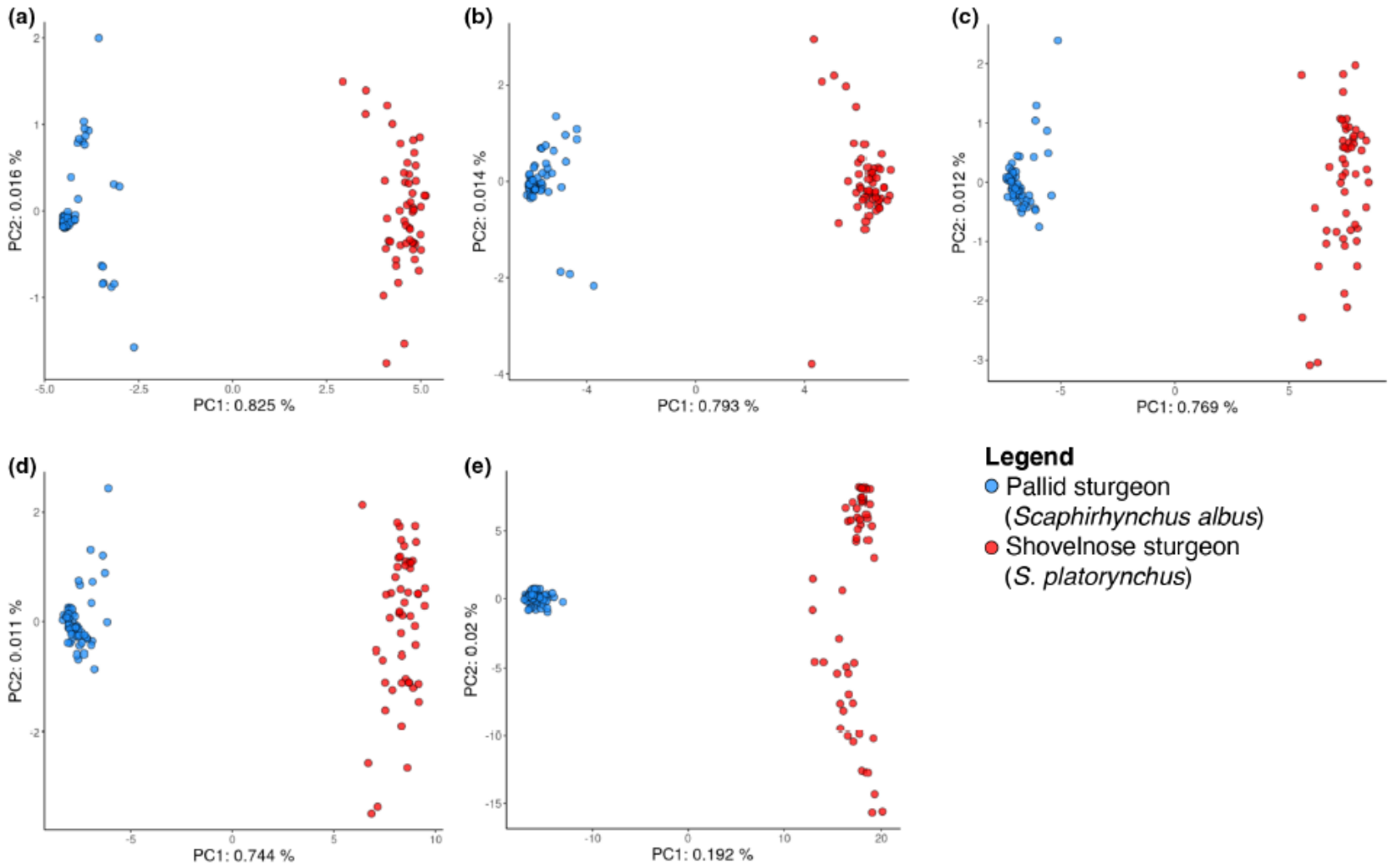


FIGURE 3 Plots of the top two principal components (PC1 and PC2) of principal component analysis by species of the (a) top 50 most informative loci, (b) top 100 most informative loci, (c) top 150 most informative loci, (d) top 200 most informative loci, and (e) all 11,082 haplotyped loci.

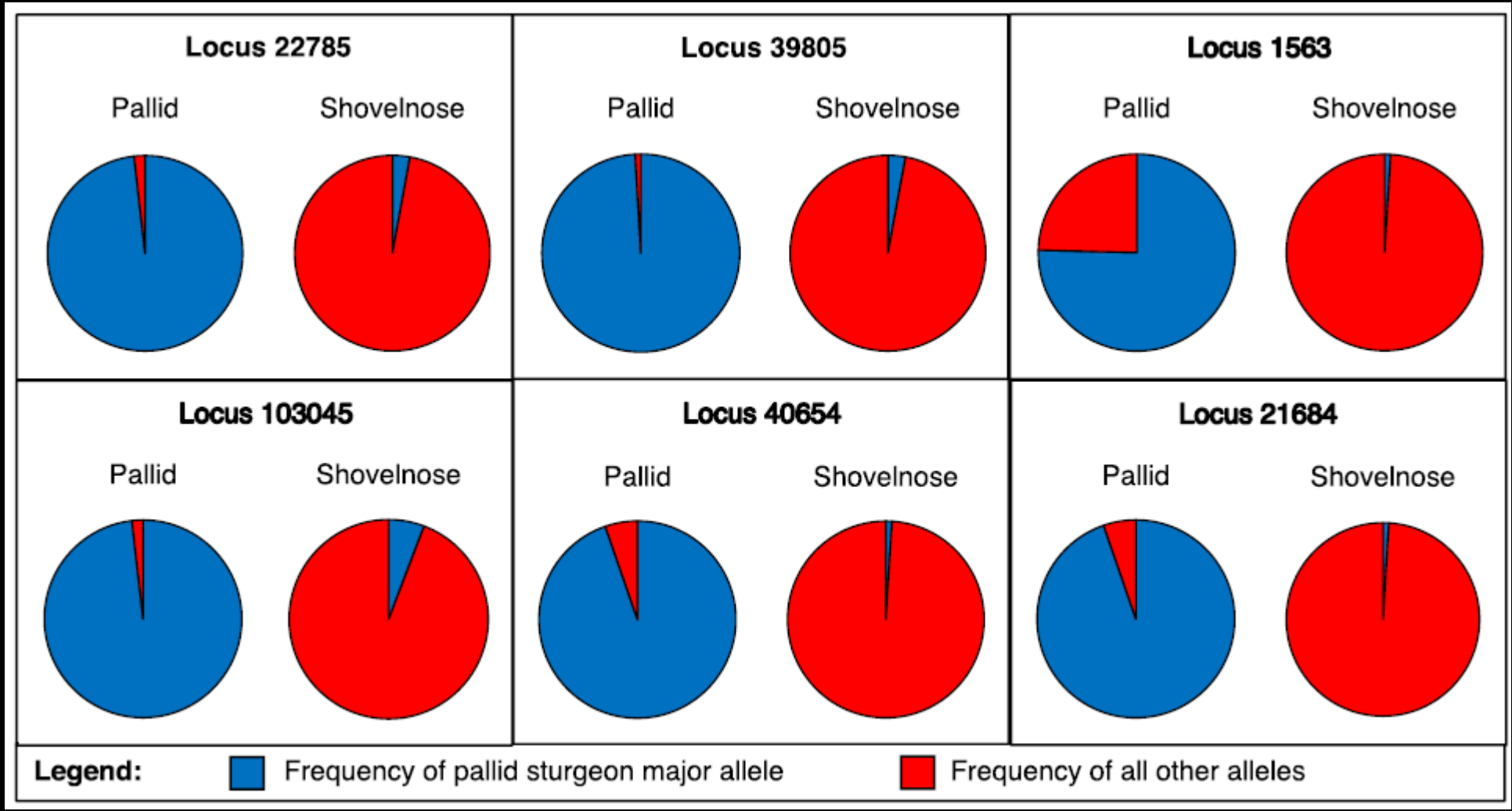
GTseq – Platte River Recovery Implementation Program

- Genotyping-in-thousands (Campbell et al. 2015).
- Efficient genotyping of SNP loci.
- Requires next generation sequencing.
  - Illumina MiSeq
- Genotype thousands of *Scaphirhynchus* with new GTseq markers
  - Ph.D. student Junman Huang

Two types of GTseq loci

- **S-loci (species)**
  - Markers with high F_{ST} values between species
  - Species discrimination
- **P-loci (polymorphic)**
  - Markers with high heterozygosity in pallid sturgeon
  - Also polymorphic in shovelnose sturgeon
  - Used for individual and population-level analysis
    - Allele frequency differences among management units
    - Effective population size
    - Parentage and relatedness
- Both markers will be scored simultaneously in all specimens

Examples of S-loci allele frequencies (Flamio et al. 2022).



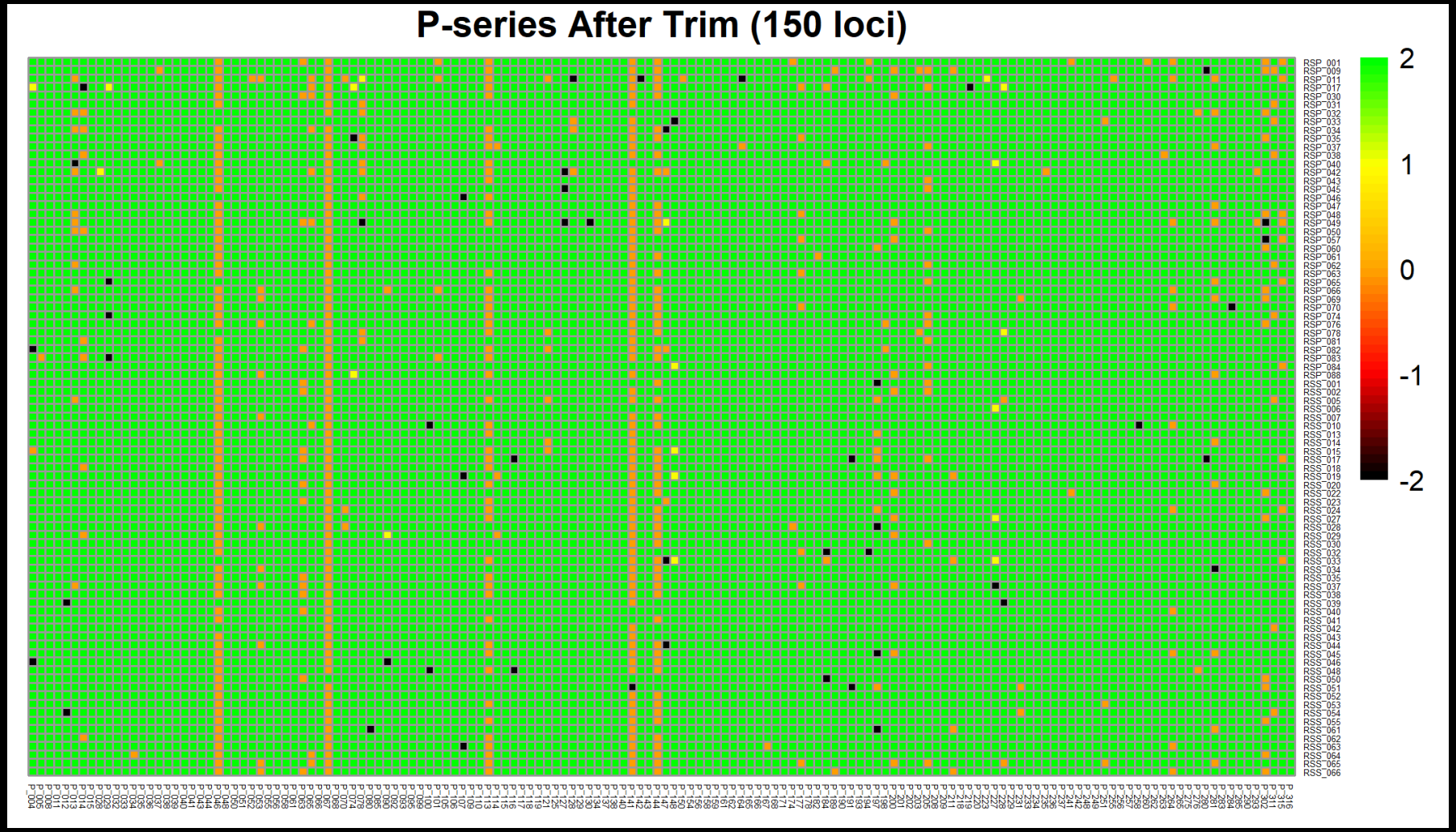
Validation

- Compare genotypes between original ddRAD study and GTseq
- 93 individuals
  - 38 pallid sturgeon
  - 55 shovelnose sturgeon
  - Mix of CLMU and GPMU fish for each species
- Loci with high disparity between studies eliminated
  - Some loci disomic in ddRAD and tetrasomic in GTseq
- Agreement between very different methods provides high confidence of genotyping accuracy.

Loci on x-axis
Individuals on y-axis

Concordance = green
Missing ddRAD = yellow
Missing GTseq = orange
Missing both = red
Discordance = black

Concordance = 99.58%

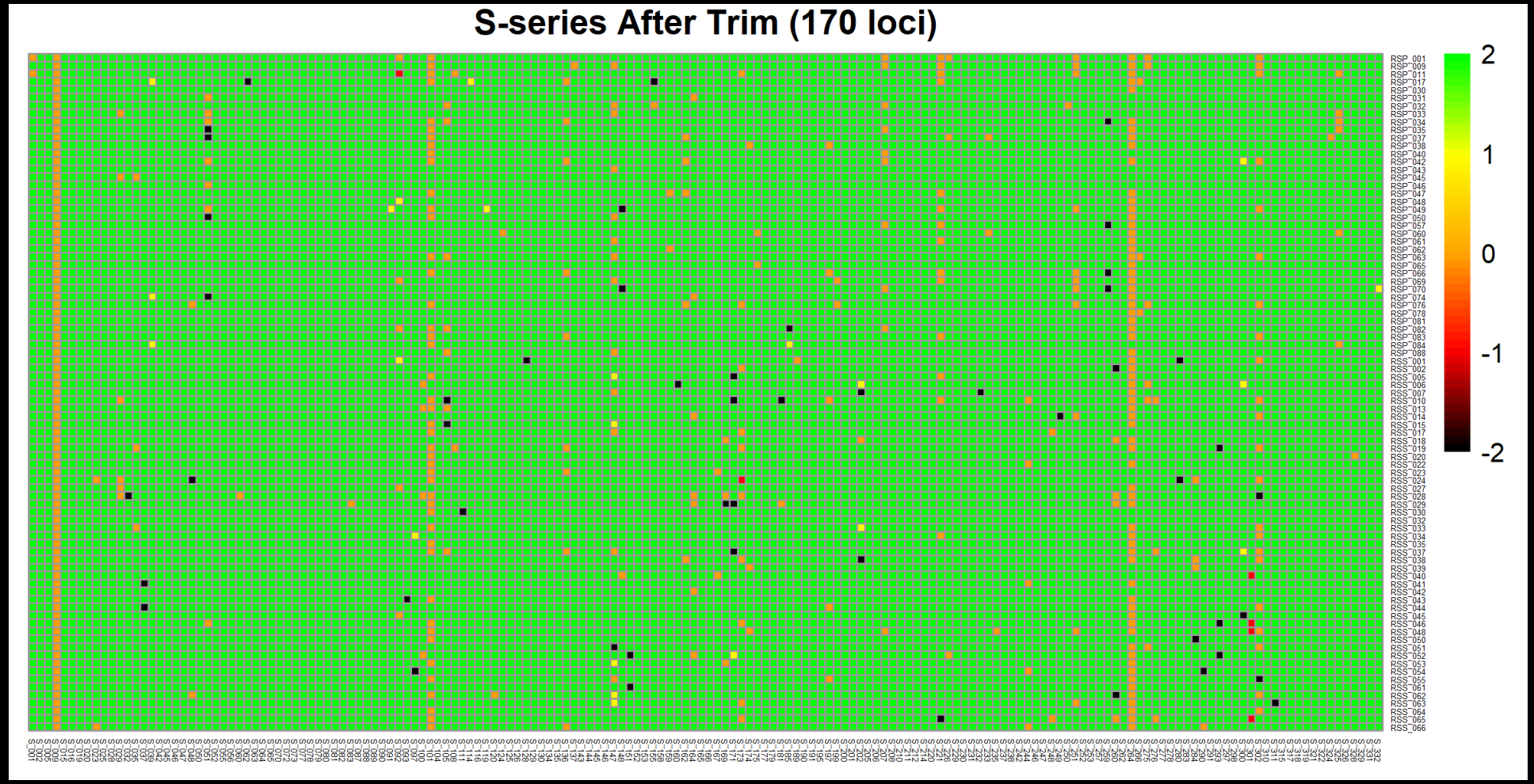


Note: loci with high percentage of missing data will be eliminated if they fail in future runs.

Loci on x-axis
Individuals on y-axis

Concordance = green
Missing ddRAD = yellow
Missing GTseq = orange
Missing both = red
Discordance = black

Concordance = 99.64%



Note: loci with high percentage of missing data will be eliminated if they fail in future runs.

MO River Species Discrimination

- Female broodstock fish 4713130264

4713130264	Pallid	Hybrid	Shovelnose
Microsatellites	0.954	0.046	0.000
GTseq	>0.999	0.000	0.000

- Female broodstock Fish 6C00111938

6C00111938	Pallid	Hybrid	Shovelnose
Microsatellites	0.976	0.024	0.000
GTseq	>0.999	0.000	0.000

Neither fish would be used as broodstock based on microsatellite assignment

MO River Species Discrimination

- MO River larva 20-07947
- MO River pallid larvae used for feeding study

20-07947	Pallid	Hybrid	Shovelnose
Microsatellites	0.946	0.053	0.000
GTseq	0.000	>0.999	0.000

We ~~Need~~ Have Better Sturgeon Genetic ID

- ~~Current~~ ID based on 19 microsatellite markers plus 146 SNP loci.
- Modeling in Jordan et al. (2019)*
  - 2% of pure Pallid Sturgeon identified as hybrids.
    - Some former hybrids will come out as pure pallids.
  - 9% of F_1 x Pallid back-crosses identified as pure Pallid Sturgeon.
    - Some former pure pallid sturgeon will be identified as hybrids.
  - 52% of F_1 x Shovelnose back-crosses identified as pure Shovelnose Sturgeon.
    - Ability to better characterize the population makeup.
- We ~~lack~~ will have certainty about species ID of broodstock.
- We ~~don't~~ will know what fraction of wild fish are hybrids.

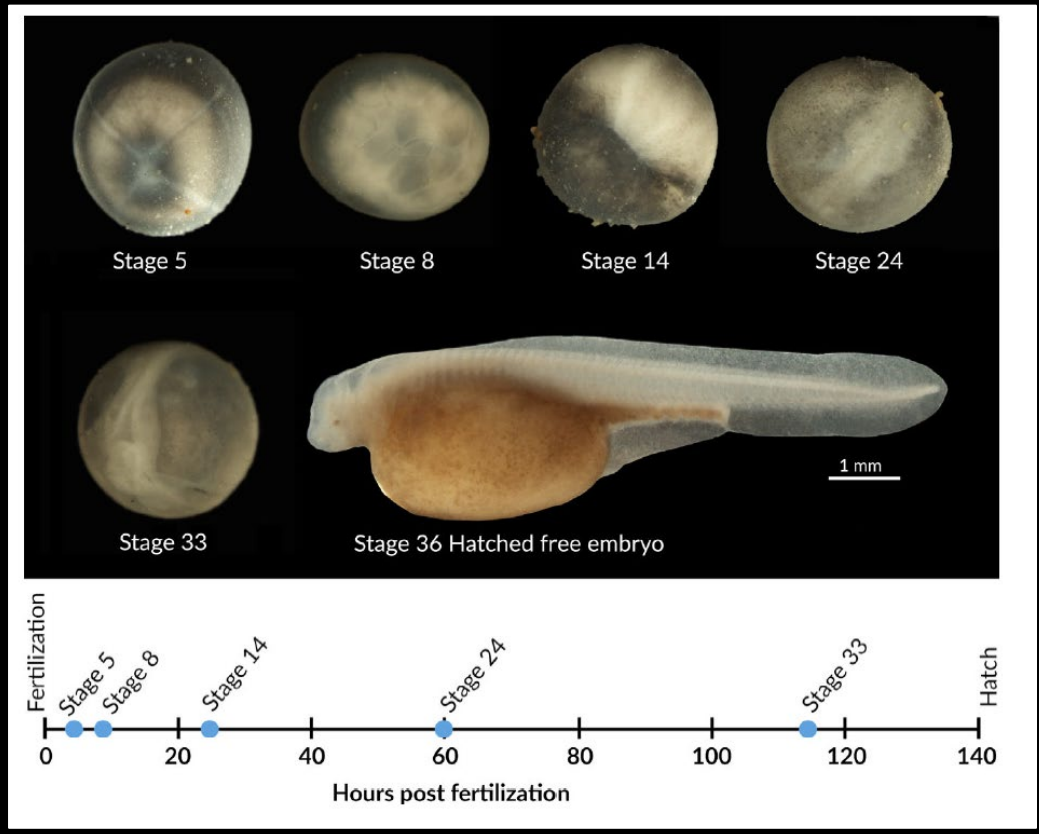
*Jordan, G., E. J. Heist, B. R. Kuhajda, G. R. Moyer, P. Hartfield and M. Piteo (2019). "Morphological Identification Over-Estimates The Number Of Pallid Sturgeon (*Scaphirhynchus albus) In The Lower Mississippi River Due To Extensive Introgressive Hybridization."* *Transactions of the American Fisheries Society* ***148: 1004-1023.***

PRRIP Objectives

- Coordinate with U. of Nebraska Lincoln field program
- Identification of sturgeon fin clips and larvae
  - GTseq for positive species ID
    - 7-day turnaround possible
    - Fewer samples cost more per sample
  - Microsatellites for parentage analysis
    - 2 business day turnaround for telemetry fish
  - Identification of embryos and larvae
    - 7-day turnaround possible
  - Parentage analysis of larvae (where possible)
    - GTseq or microsatellites to match larvae with telemetered fish
- Species/hybrid composition
    - Ongoing, final report June 2025

Eggs and Free Embryos

- Two eggs and one free embryo analyzed with SNPs/microsatellites.
- Unfertilized eggs are a single cell
  - Not enough nuclear DNA for SNP analysis*
- Fertilized eggs have enough DNA by stage 14
  - Around 24 hours post fertilization.
- Two eggs failed to amplify for nuclear SNPs.
  - mtDNA SNPs confirmed that they were sturgeon.
- Free embryo identified as a shovelnose sturgeon.



* Kashiwagi, T., A. J. DeLonay, P. J. Braaten, K. A. Chojnacki, R. M. Gocker, and E. J. Heist. 2020. Improved genetic identification of acipenseriform embryos with application to the endangered pallid sturgeon *Scaphirhynchus albus*. *Journal of Fish Biology 96(2):10.*

Parentage analyses microsatellites

26 unique fin clips provided.
25 pallid, one hybrid sturgeon.

24/25 pallid sturgeon of hatchery origin

- 2001 – 2019 year-classes

ID	Genetic.ID	Species	Origin	Mother	Father
UNL-007	16275	Pallid	Hatchery 2007	471A2C1013	412C3D4B11
UNL-009	16276	Pallid	Hatchery 2016	4627201358	48683A3B7D
UNL-010	16280	Pallid	Hatchery 2008	423373582F	432A191F35
UNL-013	16281	Pallid	Hatchery 2008	1F497F1801	1F4849755B
UNL-014	16278	Pallid	Hatchery 2016	4627201358	48683A3B7D
UNL-029	16298	Pallid	Hatchery 2013	460E52494D	43497E1577
UNL-043	16299	Pallid	Hatchery 2015	4718485A15	4626111802
UNL-080	16279	Pallid	Wild	NA	NA
UNL-103	16284	Pallid	Hatchery 2019	4626D2971	48685D1608
UNL-159	16303	Pallid	Hatchery 2001	411D262C1F	411D0E2C5F
UNL-189	16289	Pallid	Hatchery 2005	115676635A	1F50072169
UNL-191	16286	Pallid	Hatchery 2018	4626641923	47191F7F39
UNL-231	16300	Pallid	Hatchery 2015	4718485A15	4626111802
UNL-232	16301	Pallid	Hatchery 2013	412C34470E	4627061C6F
UNL-236	16304	Pallid	Hatchery 2002	116224546A	116167123A
UNL-243	16293	Pallid	Hatchery 2018	462704502D	4715591B05
UNL-250	16294	Pallid	Hatchery 2018	4715674971	434A582F17
UNL-311	16291	Pallid	Hatchery 2002	116224546A	1F477B3A65
UNL-313	16288	Pallid	Hatchery 2018	4626711111	4626773563
UNL-320	16282	Pallid	Hatchery 2008	423373582F	435F60206E
UNL-335	16296	Pallid	Hatchery 2018	46270E6C3C	47041F697D
UNL-340	16302	Pallid	Hatchery 2018	4626641923	47191F7F39
UNL-341	16297	Pallid	Hatchery 2009	412C20001A	486762580F
UNL-342	16283	Hybrid	Wild	NA	NA
UNL-346	16292	Pallid	Hatchery 2018	46270E6C3C	47041F697D
UNL-347	16277	Pallid	Hatchery 2001	220E345E09	7F7D3C5708

Microsatellite-based species ID of 2022 Platte River sturgeon

ID	mPAL	mSho	mF1	mF2	mBxP	mBxS
UNL-007	0.999	0.000	0.000	0.000	0.001	0.000
UNL-009	0.998	0.000	0.000	0.000	0.001	0.000
UNL-010	0.999	0.000	0.000	0.000	0.001	0.000
UNL-013	0.998	0.000	0.000	0.000	0.002	0.000
UNL-014	0.999	0.000	0.000	0.000	0.001	0.000
UNL-029	0.999	0.000	0.000	0.000	0.001	0.000
UNL-043	0.999	0.000	0.000	0.000	0.001	0.000
UNL-080	0.996	0.000	0.001	0.001	0.003	0.000
UNL-103	0.999	0.000	0.000	0.000	0.000	0.000
UNL-159	1.000	0.000	0.000	0.000	0.000	0.000
UNL-189	0.999	0.000	0.000	0.000	0.001	0.000
UNL-191	0.999	0.000	0.000	0.000	0.000	0.000
UNL-231	0.999	0.000	0.000	0.000	0.000	0.000
UNL-232	1.000	0.000	0.000	0.000	0.000	0.000
UNL-236	0.999	0.000	0.000	0.001	0.001	0.000
UNL-243	0.982	0.000	0.006	0.001	0.010	0.000
UNL-250	0.999	0.000	0.000	0.000	0.000	0.000
UNL-311	0.996	0.000	0.000	0.001	0.003	0.000
UNL-313	0.998	0.000	0.000	0.000	0.001	0.000
UNL-320	0.997	0.000	0.000	0.000	0.002	0.000
UNL-335	0.999	0.000	0.000	0.000	0.001	0.000
UNL-340	0.999	0.000	0.000	0.000	0.001	0.000
UNL-341	0.999	0.000	0.000	0.000	0.000	0.000
UNL-342	0.000	0.068	0.879	0.030	0.019	0.004
UNL-346	0.999	0.000	0.000	0.000	0.000	0.000
UNL-347	1.000	0.000	0.000	0.000	0.000	0.000
DLT-082	0.000	0.999	0.000	0.000	0.000	0.000

mPAL = Pallid
mSho = Shovelnose
mF1 = F_1 hybrid
mF2 = F_2 hybrid
mBxP = F_1 backcross to pallid
mBxS = F_1 backcross to shovelnose

Compare microsatellites with GTseq

ID	mPAL	mSho	mF1	mF2	mBxP	mBxS	gPal	gSho	gF1	gF2	gBxP	gBxS
UNL-007	0.999	0.000	0.000	0.000	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-009	0.998	0.000	0.000	0.000	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-010	0.999	0.000	0.000	0.000	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-013	0.998	0.000	0.000	0.000	0.002	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-014	0.999	0.000	0.000	0.000	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-043	0.999	0.000	0.000	0.000	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-080	0.996	0.000	0.001	0.001	0.003	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-103	0.999	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-159	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-189	0.999	0.000	0.000	0.000	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-191	0.999	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-231	0.999	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-232	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-236	0.999	0.000	0.000	0.001	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-243	0.982	0.000	0.006	0.001	0.010	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-250	0.999	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-311	0.996	0.000	0.000	0.001	0.003	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-313	0.998	0.000	0.000	0.000	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-320	0.997	0.000	0.000	0.000	0.002	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-335	0.999	0.000	0.000	0.000	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-340	0.999	0.000	0.000	0.000	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-341	0.999	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-342	0.000	0.068	0.879	0.030	0.019	0.004	0.000	0.000	1.000	0.000	0.000	0.000
UNL-346	0.999	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-347	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
DLT-082	0.000	0.999	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000

Compare microsatellites with GTseq

ID	mPAL	mSho	mF1	mF2	mBxP	mBxS	gPal	gSho	gF1	gF2	gBxP	gBxS
UNL-007	0.999	0.000	0.000	0.000	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-009	0.998	0.000	0.000	0.000	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-010	0.999	0.000	0.000	0.000	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-013	0.998	0.000	0.000	0.000	0.002	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-014	0.999	0.000	0.000	0.000	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-029	**0.999**	0.000	0.000	0.000	0.001	0.000	0.001	0.000	0.000	0.000	**0.999**	0.000
UNL-043	0.999	0.000	0.000	0.000	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-080	0.996	0.000	0.001	0.001	0.003	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-103	0.999	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-159	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-189	0.999	0.000	0.000	0.000	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-191	0.999	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-231	0.999	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-232	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-236	0.999	0.000	0.000	0.001	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-243	0.982	0.000	0.006	0.001	0.010	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-250	0.999	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-311	0.996	0.000	0.000	0.001	0.003	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-313	0.998	0.000	0.000	0.000	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-320	0.997	0.000	0.000	0.000	0.002	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-335	0.999	0.000	0.000	0.000	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-340	0.999	0.000	0.000	0.000	0.001	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-341	0.999	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-342	0.000	0.068	0.879	0.030	0.019	0.004	0.000	0.000	1.000	0.000	0.000	0.000
UNL-346	0.999	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
UNL-347	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
DLT-082	0.000	0.999	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000

Academic Benchmark Schedule – Junman Huang

- Admitted to Graduate School – August 2022.
- Signed committee form -- December 2022.
- Plan of Study (list of formal courses) – May 2023.
- Dissertation proposal –December 2023.
- Preliminary examination –May 2025.
- Dissertation defense/graduation – May 2027.

Research Objective Timeline

- Marker development/validation – December 2022
- Refined species ID and baselines – December 2023
- Population structure/Redefine management units – June 2024
- Population composition by species/hybrid – December 2024
- Demographics – pallid sturgeon N_e by population – June 2025
- Final report to PRRIP – June 2026