Identifying Target Spawning Populations through Genomic Analysis for Directed Management

MRBP is interested in whether genetic population structure can be identified that corresponds to spawning stocks of sliver carp throughout the Mississippi River basin. There are two primary studies to date – Farrington et al 2017 documented little to no significant genetic structure throughout the basin with microsatellites, while Stepien et al 2019 (mtDNA and nuclear markers) documented significant structure between several pairwise group comparisons, though the magnitude of differentiation was consistently small (ThetaST < 0.01 in all comparisons) and comparable to that seen in the previous study. Both studies employed the conventional approach of genetic markers, which may limit the fine-scale resolution in revealing population structure.

Identification of structure could suggest target spawning populations for directed management activities. Based on the data available to date, any genetic structure that exists is likely to occur at low levels. As such, confidently detecting structure that could be useful for informing management activities will likely require relatively powerful datasets that include large sample sizes and/or high numbers of informative genetic markers. Genomic-scale genotyping-by-sequencing approaches such as RAD-seq currently offer the best options for generating such datasets. RAD-seq has become widely utilized over the past ~10 yrs to address a wide variety of questions including population structure in a large range of species, and is especially useful in non-model species, as generating data with this approach doesn't require any prior genetic information.

Additional questions were raised during the conversation – maybe the most pertinent of which is the issue of hybridization between silver and bighead carp. Such hybridization could be a confounding factor in detecting genetic structure within silver carp, but could probably be alleviated by screening samples for "pure" individuals prior to analysis. However, the genomic datasets generated could also be useful for addressing questions surrounding hybridization such as the frequency at which hybrids occur, and relationships between morphological data and hybrid status if morphological data become available for some or all of the samples. Focusing on the population structure aspects of the study, three potential phases were suggested:

1.) Development of a set of informative genomic markers: The number of informative genetic markers (SNPs) that are generated with the RADseq method needs to be optimized for a given study system. This involves identifying and selecting specific parameters of the wet-lab protocol such as the restriction enzyme(s) used and the size range of genomic fragments retained throughout the protocol. Such work generally does not require large sample sizes (less than 20 individuals total), and does not necessarily need to be performed on samples that will be included in subsequent analyses. Therefore, this phase could potentially begin prior to samples being collected in the field if archived samples could be identified and obtained.
2.) Preliminary test for population structure: Upon identification of a set of informative markers, the markers could be applied to a small number (2-3) of potential spawning

populations from geographically distant locations to give the best chance of identifying population structure between the target groups. Potential groups to target might include populations from the upper Missouri River, the mid to upper Wabash River, and potentially one additional river in the lower Mississippi basin.

3.) Analyze population structure for all populations of interest: If population structure is documented in phase 2, sampling and analysis could be expanded to larger numbers of potential spawning populations to more thoroughly assess population structure throughout the basin.