

Article

A comparison of non-destructive visceral swab and tissue biopsy sampling methods for genotyping-by-sequencing in the freshwater mussel *Fusconaia askewi*.

Matthew Harrison ¹, V. Alex Sotola ^{2*}, Alexander Zalmat ¹, Kyle T. Sullivan ³, Bradley M. Littrell ³, Timothy Bonner ¹ and Noland H. Martin ^{1*}

¹ Texas State University, Department of Biology, 601 University Drive, San Marcos, TX 78666

² State University of New York at Oneonta, Biology Department, 108 Ravine Parkway, Oneonta NY, 13820

³ BIO-WEST, Inc. 1405 United Drive, Suite 111, San Marcos, Texas 78666-2834

* Correspondence: noland.martin@txstate.edu (NHM), vasotola@oneonta.edu (VAS)

Abstract: Limiting harm to organisms via genetic sampling is an important consideration for rare species. Nondestructive sampling techniques have been developed to address this issue in freshwater mussels. Two methods, visceral swabbing and tissue biopsies, have proven to be effective for DNA sampling, though it is unclear as to which method is preferable for genotyping-by-sequencing (GBS). Tissue biopsies may cause undue stress and damage to organisms, while visceral swabbing potentially reduces the chance of such harm. Our study compared the efficacy of these two DNA sampling methods for generating GBS data for the Unionid freshwater mussel, Texas Pigtoe (*Fusconaia askewi*). Our results find both methods generate quality sequence data, though some considerations are in order. Tissue biopsies produced significantly higher DNA concentrations and larger numbers of reads when compared to swabs, though there was no significant association between starting DNA concentration and number of reads generated. Swabbing produced greater sequence depth (more reads per sequence) while tissue biopsies revealed greater coverage across the genome (at lower sequence depth). Patterns of genomic variation as characterized in principal component analyses were similar regardless of the sampling method, suggesting that the less invasive swabbing is a viable option for producing quality GBS data in these organisms.

Keywords: non-destructive DNA sampling; DNA collection methods; Louisiana Pigtoe; visceral swabbing; freshwater mussels; *Fusconaia askewi*; genotyping-by-sequencing; population genetic structure; genomic coverage; sequencing depth

1. Introduction

Studies examining the population genetic structure - or lack thereof - of rare and threatened species are critical tools for enabling biologists to recommend management strategies of these organisms [1-5]. Large sample sizes are often needed to obtain robust estimates of population genetic parameters, and it is important to minimize the degree to which genetic sampling results in the disruption or death of individual organisms, especially for rare species. A number of nondestructive sampling techniques (e.g., saliva, hair, swabs, etc.) have thus been developed to reduce the potentially harmful impact of genetic sampling [6-10].

In North America, Family Unionidae is a speciose group of freshwater mussels consisting of both widely distributed and locally endemic species, many of which exhibit varying conservation statuses [11-13]. A diverse array of DNA collection methods has been proposed for this group that reduce the mortality of sampled and re-released individuals. Collection methods include clipping of small pieces of mantle or foot tissue (i.e. "tissue biopsies"), swabbing of the viscera, foot scraping, and haemolymph extraction via hypodermic needles [14-16]. These four methods were specifically evaluated for the Freshwater Pearl Mussel (*Margaritifera margaritifera*) with respect to survivorship, growth

characteristics, and DNA quantity and quality of the sampled individuals [16]. All four methods reliably produced DNA, though yields from haemolymph and foot scraping were significantly lower than those produced from tissue biopsies and swabbing. Furthermore, survivorship (after 128 days) was 100% for all methods, leading the authors to conclude that visceral swabbing was the most appropriate and least invasive method for tissue collection in *M. margaritifera*. A separate investigation in the Pimpleback mussel *Quadrula pustulosa* further determined that tissue swabbing was preferred for DNA collection over mantle clips because it produced quality samples with less disruption to the organisms [15].

Visceral swabbing methods have produced quality DNA appropriate for use in traditional Sanger sequencing methods that include generating ND-1 mitochondrial sequences [15] and microsatellite loci [16]. However, such methods do not require large amounts of targeted DNA be extracted, and contamination with other non-target organisms is largely not problematic because primers are oftentimes designed to be species-specific. The advent of genotyping-by-sequencing (GBS) methods, on the other hand, has allowed for the production of single nucleotide polymorphism (SNP) data that can routinely produce orders of magnitude more genetic markers at much lower cost compared to traditional Sanger sequencing methods [17]. Massault et al. [18] evaluated the effectiveness of swabbing methods for GBS sequencing in the Silver-Lipped Pearl Oyster (*Pinctada maxima*, Family: Margaritiferidae) and found such methods to be appropriate for parentage analysis and pedigree reconstruction in these marine bivalves. Swab collections produced DNA of sufficient quantity and quality for genotyping, and SNP counts were highly correlated among tissue and swab collection methods [18]. However, it remains unclear whether swab methods are similarly appropriate for GBS sequencing of freshwater Unionid mussels, and further investigation is needed to address this question.

In the current study, we evaluated the degree to which visceral swab and tissue biopsies generate high quality GBS / SNP genotype data for the Unionoid freshwater mussel Texas Pigtoe (*Fusconaia askewi*). This was achieved by collecting and processing swab and tissue biopsy samples from the same individuals and generating separate SNP datasets for both types of sampling methods, with swab samples being collected prior to performing tissue biopsies. Our primary goal was to determine whether the less-disruptive swab sampling method was sufficient to undertake large-scale population genetic studies using GBS methods for freshwater mussels. Therefore, we sought to 1) compare DNA concentrations and number of raw reads produced from visceral swab samples and tissue biopsies collected across 14 different individuals, 2) determine whether and how SNP coverage and numbers differed across the two sampling methods and 3) determine whether the patterns of genetic variation among individuals are consistent across both datasets when assembled independently from one another, as well as when assembled to a larger tissue-based *de novo* reference assembly.

2. Materials and Methods

2.1. Sample Collection

Fourteen mussels were collected via snorkeling and tactile searches at two separate locations along the Sabine River in Texas, USA at 32.62986° North 32.62986° West (n = 6), and at 32.462220° North and 94.845864° West (n = 8). The mussels were identified in the field as *F. askewi* on the basis of morphology and were subsequently sampled using Isohelix SK-1S buccal swabs by carefully rubbing both sides of the swab along the mussel's foot tissue for 30 seconds. Swab heads were immediately placed in individual containment tubes and kept on dry ice until storage at -20° C. After swabbing, a ~ 0.5 cm³ tissue plug was collected from the same individual using a nasal biopsy tool, and the mussel was then returned to its original location. Tissue samples were stored separately in 95% ethanol on dry ice before transferring to a -20°C freezer for longer-term storage.

2.2. DNA Extraction and Quantification, Library Prep, and *de novo* Assemblies

DNA was extracted from the 14 swab samples and the replicated 14 foot-tissue plugs using Qiagen DNeasy blood and tissue kits in 96-well format. These samples were part of a larger group of 378 mussel tissue samples that were used in a separate population genetic study [19]. No attempts were made to equalize DNA concentrations prior to DNA library construction. However, DNA concentrations of replicated swab and biopsied tissue samples were quantified post-hoc to investigate whether initial concentrations affected the number of reads per individual in the resulting libraries. DNA concentrations were determined using the Qubit® dsDNA HS Assay (Invitrogen, USA) kit following standard protocols described in the Qubit manual. The DNA concentration of each of the 14 swab and tissue replicates was measured from a single aliquot three separate times, and a mean concentration was calculated for each sample.

Extracted DNA from the broader collection of mussels ($n = 378$ [19]) was used to create a reduced-complexity genomic library. This library included the 14 tissue and 14 swab replicates utilized here and was generated following modified protocols commonly used in our laboratory group across a wide diversity of species [5,19-22]. Restriction enzymes EcoRI and MseI were used to digest extracted DNA; EcoRI adapters (i.e., 10-20 base-pair multiplex identifier sequences: MIDs) were ligated onto the resulting fragments, and the 14 swab and 14 tissue replicates were each assigned different unique barcodes. Labeled products were amplified through two rounds of PCR using Illumina primers. PCR products were pooled into a single library and sent to the University of Texas Genomic Sequencing and Analysis Facility (Austin, TX) for sequencing on an Illumina Novaseq SR100 platform.

A number of data processing steps were used to ensure high-quality sequencing data. First, Bowtie v.3 was used to identify PhiX control sequences, and reads that assembled to the PhiX genome were removed [23]. Custom Perl scripts were used to match sample IDs with unique barcode identifiers and remove MseI adapters and barcodes from sequence reads. The resulting sequence data were then organized into different assemblies for separate analyses. Since no reference genome is available for *F. askewi*, three separate de novo assemblies were built using different sets of individuals and/or collection methods. The first reference genome was created using a much larger dataset of tissue-sampled *F. askewi* individuals ($n=96$, [5]). We used this reference to align reads for both tissue and swab replicates, enabling a direct comparison of resulting SNP datasets (i.e., to directly assess coverage, depth and genetic variation of “tissue” and “swab” samples of the same 14 individuals). Two additional reference assemblies were also separately created; one exclusively using tissue samples (tissue-only assembly), and another for the replicated swab samples (swab-only assembly). Each of the three separate reference assemblies (i.e., control reference, tissue-only, and swab-only) followed the same processes using part of the dDocent variant calling pipeline [24]. First, unique reads were identified for each individual, and reads with less than four copies and shared across fewer than four individuals were filtered out. Subsequently, the remaining sequences (i.e., those that had more than four copies and were shared across more than four individuals) were assembled to form each reference using CD-hit by utilizing a similarity threshold of 80% [25,26].

2.3. Number of SNPs and Coverage

Sequence reads were assembled to reference scaffolds using the *aln* and *samse* algorithms from BWA (version 0.7.13-r1126; [27]). BCFtools (version 1.9) was used to both identify SNPs as well as to calculate Bayesian genotype likelihoods for each SNP [28]. Biallelic SNPs must have been represented by at least 80% of the individuals in a given assembly and have had a mean sequence depth of $\geq 2Xn$ (where n = number of individuals) in order to be included in the respective dataset. A custom Perl script was used to filter out potentially paralogous loci with exceptionally high sequence depth (mean sequence depth $>$ assembly-wide mean $+ 2*sd$). This script also filters loci based on mapping quality (minimum score of 30), as well as the difference in base and mapping quality between the reference and alternative alleles using Mann-Whitney U tests (z -score cutoff =

1.96). Genotype likelihoods were assigned to each SNP for each individual and used to estimate allele frequencies; SNP's with a minor allele frequency of <0.05 were not included in the final datasets. If there was more than one SNP identified on a contig, only a single SNP was randomly chosen to include for analyses in order to reduce effects of linkage disequilibrium.

2.4. Estimation of Genotype Probabilities

Rather than “calling” SNPs, we utilize genotype probabilities which account for uncertainty associated with sequencing. Genotype probabilities were estimated using Entropy, a hierarchical model that estimates Bayesian-based admixture proportions and genotype probabilities for each individual for a predetermined number of populations (k) [5,29]. Calculation of posterior distributions for $k = 2$ was done by combining two separate Markov Chain Monte Carlo (MCMC) runs including 50,000 total iterations with a burn-in of 5,000 and sampling every 10th iteration. Genotype probabilities were averaged across both runs after checking for chain convergence with Gelman-Rubin diagnostic statistics.

2.5. Data Analysis

We performed a paired t-test to determine whether there was a significant difference between DNA concentrations of tissue and swab samples. We then used an ANCOVA to assess the relationship between DNA concentrations and the total number of raw reads generated, while accounting for sample type. This ANCOVA was also used to determine whether there was a significant difference in the numbers of reads that each sampling method produced.

We ran Principal Components Analysis (PCA) on genotype probabilities to determine whether SNP alignments to either the sample-type reference or the tissue control reference influenced resulting patterns of genetic variation. In total, we ran four PCAs with: 1) swab data aligned to the swab de novo assembly, 2) tissue data aligned to the tissue de novo assembly, 3) swab data aligned to the control tissue de novo assembly, and 4) tissue data aligned to the control tissue de novo assembly. Rather than focusing on exact genetic relationships among individual samples, analyses were primarily focused on how relationships among individuals changed given the sampling type and the references to which they were aligned. To examine this, we ran Procrustes analysis from the vegan package in Program R and report the Procrustes R, a measure of correlation between two multivariate plots [30]. The first Procrustes analysis was run comparing the two PCAs which were sample-specific; swab and tissue data aligned to their specific de novo assemblies. The second Procrustes analysis compared the two sampling types when aligned to the control reference. We posited that if swab samples produced a SNP dataset with similar genetic patterns to the tissue samples, even when aligned to separate de novo assemblies, the PCAs would be highly (> 0.95) correlated.

3. Results

3.1. DNA concentrations

Mean DNA concentrations significantly differed among tissue and swab samples (two-tailed paired t-test, $t = 2.171$, $P = 0.049$, $df = 13$) with tissue DNA concentrations (mean ± 1 SE; $36.2 \text{ ng/ml} \pm 14.6$) having 38% greater DNA concentrations than those of swab samples ($27.8 \text{ ng/ml} \pm 12.5$; Table S1). However, after omitting one swab sample that revealed exceptionally low amounts of DNA ($0.02 \text{ ng/ml} \pm 0.03 \text{ ng/ml}$), DNA concentrations did not significantly differ (two-tailed paired t-test, $p = 0.0996$, $t = 1.785$, $df = 12$).

3.2. SNP numbers, Coverage, and Correlations

The mean number of sequencing reads (± 1 S.D.) generated per individual from tissue samples ($1,608,863 \pm 307,130$) was 2.1X more than the number of reads generated from swab samples ($518,985 \pm 238,072$), and this was a significant difference (Table S2; $t = 9.693$,

$p < 0.001$). DNA concentration was not predictive of the number of raw reads that were generated regardless of the DNA sampling technique (ANCOVA; $t = 1.2003$, $p = 0.284$).

When both datasets were aligned to the control-reference, tissue samples produced more SNPs for each individual than swab samples, but at lower coverage. For the tissue dataset, there were 36,764 SNPs identified, each with a mean coverage of 3.64 ± 0.74 reads per SNP, whereas for the swab dataset, an order of magnitude fewer SNPs ($n = 3,138$) were identified with an average coverage of 4.29 ± 1.94 reads per SNP (Table S3). For the tissue-only assembly (i.e., tissue samples aligned to tissue-only reference), mean coverage of the 8,111 SNPs was 6.79 ± 1.48 , whereas for the swab-only assembly (i.e., swab samples aligned to swab-only reference), mean coverage for the 565 SNPs was dramatically higher at 22.64 ± 10.5 (Table S4).

3.3. Comparing PCA Plots

None of the four PCA plots revealed genetic structure among individuals sampled from the two separate locations along the Sabine River (Supplemental Figure 1). When each sample type was aligned to their respective assembly (i.e., tissue data aligned to swab-only reference and swab samples aligned to tissue-only reference), there was a high correlation between the two PCA plots (Procrustes $R = 0.982$ Figure 1A). Similarly, when tissue and swab data were aligned to the control reference, the PCAs are also highly correlated (Procrustes $R = 0.975$; Figure 1B).

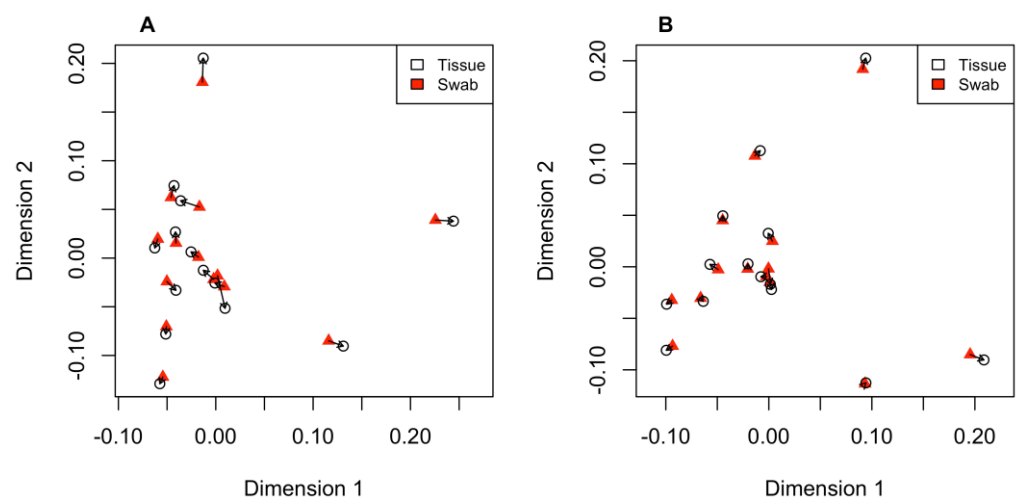


Figure 1. Procrustes plots showing differences between the two independent PCAs when samples were aligned to their own reference assemblies (A) and when they were aligned to the control reference (B). The open black circles are individual tissue samples, red triangles are individual swab samples, and arrows point to the relative change of positioning between the two PCAs.

4. Discussion

Identifying DNA collection methods that minimize mortality and disruption to organisms is important when taxa being studied are rare or of high conservation priority [6-10]. Our results demonstrated that both buccal swabs and tissue biopsy methods produce quality DNA in the Unionid mussel *F. askerwi*. Of the 14 individuals sampled, there was one notable exception where the DNA concentration for a single swabbed sample was near-zero (two of three DNA quantification measurements on the qubit system were “zero” and one measurement was at the lowest possible reading of $0.1 \text{ ng}/\mu\text{L}$). As we did not re-extract the sample, it is unknown whether this was a failure of the swab collection method itself or whether this was simply a failed DNA extraction. It is important to note that, excepting the potentially failed extraction, should the quantity of DNA collected be important for the study being conducted, swabs produce about 20 percent less DNA than tissue biopsies (though this was not significantly less in the current study [$p = 0.097$]), and

this may need to be taken under advisement should other experimental protocols require higher concentrations of DNA.

Both methods of DNA collection studied here have proven to be effective at producing high-quality DNA for Sanger sequencing in other species of freshwater mussels [14–15]. In the current study, GBS sequencing of tissue samples generated twice as many raw reads per individual than swab samples. Tissue samples also generated more SNPs per individual regardless of whether the reads were aligned to the reference assembly or the tissue-only de novo assembly, but the mean read depth of each tissue-based SNP was lower than that of swab samples. Regardless of the sampling method, however, the Procrustes analysis revealed nearly identical relationships among all of the included samples. Thus, the less invasive swab method of DNA collection is likely to be an acceptable sampling method for performing population genomic work in Unionid mussels. However, it should be kept in mind that these results should be interpreted within the context of their sampling distribution. The genetic diversity of the individuals collected from these two closely-located sampling locales (separated by ~97 river km) within the Sabine River is likely lower than what might be expected in individuals collected from a larger sample size across a wider sampling distribution. It is therefore possible that the extra genomic data provided from tissue biopsy samples may provide a substantial benefit when it comes to identifying the pattern and extent of structure of more diverse collections.

We found that visceral swabbing and tissue biopsies reveal similar patterns of genetic variation among individual samples of *F. askewi* when performing GBS. Such methods are potentially transferable to other mussel systems as well [16]. Future researchers should consider the potential advantages and disadvantages of both methods. Here, tissue samples provided substantially more *genomic coverage* (i.e. more total SNPs) than viscera swabbing, yet the mean sequence coverage per SNP (and the concomitant increase in accuracy in genotype probabilities or genotype “calls” at each of these fewer SNPs) is higher for the viscera swabbing method. Researchers should not only keep in mind these potential tradeoffs, but also take into consideration the degree to which these different methods have the potential to disrupt growth and survivorship of these species.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Figure S1: PCA plots for all four datasets: A) swab aligned to swab reference, B) tissue aligned to tissue reference, C) swab aligned to control reference, D) tissue aligned to control reference; Table S1: DNA concentrations (ng/ml) as measured with a Qubit system for each sample. Each individual was measured three times (Measure 1, Measure 2, and Measure 3). Also provided are the mean and one standard deviation (SD) for each individual; Table S2: Number of raw reads for each sampling method from each individual. Also provided are the mean and one standard deviation (SD) for each individual; Table S3: Median, mean, and one standard deviation (SD) of read depth for each sample type when aligned to the tissue-based control reference; Table S4: Median, mean, and one standard deviation (SD) of read depth for each sample type when aligned to their respective de novo references (e.g., swab samples aligned to swab-only reference).

Author Contributions: Conceptualization, M.H., V.A.S., K.T.S., B.M.L., T.B., AND N.H.M.; methodology, M.H., V.A.S., A.Z., K.T.S., B.M.L., T.B., and N.H.M.; formal analysis, M.H., V.A.S., N.H.M.; investigation, M.H., V.A.S., A.Z., K.T.S., B.M.L., T.B., and N.H.M.; data curation, M.H. and N.H.M.; writing—original draft preparation, M.H., V.A.S., A.Z., and N.H.M.; writing—review and editing, M.H., V.A.S., A.Z., K.T.S., B.M.L., T.B., and N.H.M.; visualization, M.H., V.A.S., N.H.M.; supervision, N.H.M., T.B.; project administration, N.H.M. T.B.; funding acquisition, N.H.M., T.B.; All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Sabine River Authority of Texas and by the Texas Comptroller of Public Accounts.

Data Availability Statement: Raw reads data are provided at www.XXXXXXXXXX.XXXXXX [to be provided upon acceptance of publishing].

Acknowledgments: We thank several individuals involved with the study concept and with assistance in the field including Bill Kirby (Sabine River Authority of Texas), Cody Craig (Arrowhead

Ecology), and Colin McDonald (Texas Comptroller of Public Accounts). Members of the Fuess/Martin/Nice/Ott discussion group provided valuable feedback to an early version of this manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Galbraith, H.S.; Zanatta, D.T.; Wilson, C.C. Comparative analysis of riverscape genetic structure in rare, threatened and common freshwater mussels. *Conservation Genetics* **2015**, *16*, 845–857.
- Kelly, M.W.; Rhymer, J.M. 2005. Population genetic structure of a rare unionid (*Lampsilis cariosa*) in a recently glaciated landscape. *Conservation Genetics* **2005**, *6*, 789–802.
- Lucas, L.K.; Fries, J.N.; Gabor, C.R.; Nice, C.C. Genetic variation and structure in *Eurycea nana*, a federally threatened salamander endemic to the San Marcos springs. *Journal of Herpetology* **2009**, *43*(2), 220–227.
- Nice, C.C.; Fordyce, J.A.; Sotola, V.A.; Crow, J.; Diaz, P.H. 2021. Geographic patterns of genomic variation in the threatened Salado salamander, *Eurycea chisholmensis*. *Conservation Genetics*, **2021**, *22*, 811–821.
- Sotola, V.A.; Ruppel, D.S.; Bonner, T.H.; Nice, C.C.; Martin, N.H. Asymmetric introgression between fishes in the Red River basin of Texas is associated with variation in water quality. *Ecology and Evolution* **2019**, *9*(4), 2083–2095.
- Banks, S.C.; Piggott, M.P. Non-invasive genetic sampling is one of our most powerful and ethical tools for threatened species population monitoring: a reply to Lavery et al. *Biodiversity and Conservation* **2022**, *31*(2), 723–728.
- Baus, I.; Miño, C.I.; Monge, O. Current trends in the application of non-invasive genetic sampling to study Neotropical birds: Uses, goals, and conservation potential. *Avian. Biol. Res.* **2019**, *12*.
- Carroll, E.L.; Bruford, M.W.; DeWoody, J.A.; Leroy, G.; Strand, A.; Waits, L.; Wang, J. Genetic and genomic monitoring with minimally invasive sampling methods. *Evol. Appl.* **2018**, *11*, 1094–1119.
- Storer, C.; Daniels, J.; Xiao, L.; Rossetti, K. Using noninvasive genetic sampling to survey rare butterfly populations. *Insects* **2019**, *10*(10), 311.
- Waits, L.P.; Paetkau, D. Noninvasive genetic sampling tools for wildlife biologists: A review of applications and recommendations for accurate data collection. *J. Wildl. Manag.* **2005**, *69*, 1419–1433.
- Graf, D. L.; Cummings, K.S. A 'big data' approach to global freshwater mussel diversity (Bivalvia: Unionoda), with an updated checklist of genera and species. *Journal of Molluscan Studies* **2021**, *87*(34) eyaa034. <https://doi.org/10.1093/mollus/eyaa034>
- Lydeard, C.; Cowie, R.H.; Ponder, W.F.; Bogan, A.E.; Bouchet, P.; Clark, S.A.; Cummings, K.S.; Frest, T.J.; Gargominy, O.; Herbert, D.G.; Hershler, R. The global decline of nonmarine mollusks. *BioScience* **2004**, *54*(4), 321–330.
- Williams, J.D.; Warren Jr., M.L.; Cummings, K.S.; Harris, J.L.; Neves, R.J. Conservation status of freshwater mussels of the United States and Canada. *Fisheries* **1993**, *18*, 6–22.
- Berg, D.J.; Haag, W.R.; Guttman, S.I.; Sickel, J.B. Mantle biopsy: a technique for nondestructive tissue-sampling of freshwater mussels. *Journal of the North American Benthological Society* **1995**, *14*(4), 577–581.
- Henley, W.F.; Grobler, P.J.; Neves, R.J. Non-invasive method to obtain DNA from freshwater mussels (Bivalvia: Unionidae). *Journal of Shellfish Research* **2006**, *25*(3), 975–977.
- Karlsson, S.; Larsen, B.M.; Eriksen, L.; Hagen, M. Four methods of nondestructive DNA sampling from freshwater pearl mussels *Margaritifera margaritifera* L. (Bivalvia: Unionoida). *Freshwater Science* **2013**, *32*(2), 525–530.
- Andrews, K.R.; Good, J.M.; Miller, M.R.; Luikart, G.; Hohenlohe P.A. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics* **2016**, *17*, 81–92.
- Massault, C.; Jeffrey, C.; Jones, D.B.; Barnard, R.; Strugnell, J.M.; Zenger, K.R.; Jerry, D.R. Non-invasive DNA collection for parentage analysis for bivalves: A case study from the silver-lipped pearl oyster (*Pinctada maxima*). *Aquaculture* **2022**, *552*, 738036.
- Harrison, M. Population Genetic Structure of Unionid Mussels Across Multiple Gulf Drainages. Master's Thesis, Texas State University **2022**.
- Gompert, Z.; Lucas, L.K.; Nice, C.C.; Fordyce, J.A.; Forister, M.L.; Buerkle, C.A. Genomic regions with a history of divergent selection affect fitness of hybrids between two butterfly species. *Evolution* **2012**, *66*, 2167–2181.
- Sung, C.J.; Bell, K.L.; Nice, C.C.; Martin, N.H. Integrating Bayesian genomic cline analyses and association mapping of morphological and ecological traits to dissect reproductive isolation and introgression in a Louisiana Iris hybrid zone. *Molecular Ecology* **2018**, *27*(4), 959–978.
- Zalmit, A.S.; Sotola, V.A.; Nice, C.C.; Martin, N.H. Genetic structure in Louisiana Iris species reveals patterns of recent and historical admixture. *American Journal of Botany* **2021**, *108*(11), 2257–2268.
- Langmead, B.; Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **2012**, *9*, 357–359. <https://doi.org/10.1038/nmeth.1923>.
- Puritz, J. B.; Hollenbeck, C. M.; Gold, J. R. dDocent: A RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ* **2014**, *2*, e431. <https://doi.org/10.7717/peerj.431>.
- Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics* **2012**, *28*, 3150–3152. <https://doi.org/10.1093/bioinformatics/bts565>.
- Li, W.; Godzik, A. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **2006**, *22*, 1658–1659. <https://doi.org/10.1093/bioinformatics/btl158>.
- Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*(14), 1754–1760.

-
28. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **2011**, 27, 2987–2993. <https://doi.org/10.1093/bioinformatics/btr509>.
 29. Gompert, Z.; Lucas, L.K.; Buerkle, C.A.; Forister, M.L.; Fordyce, J.A.; Nice, C.C. Admixture and the organization of genetic diversity in a butterfly species complex revealed through common and rare genetic variants. *Molecular Ecology* **2014**, 23(18), 4555–4573.
 30. Oksanen, J.; Blanchet, F.; Friendly, M.; Kindt, R.; Legendre, P.; McGlinn, D.; Minchin, P.; O'Hara R.; Simpson, G.; Solymos, P.; Stevens, M.; Szoecs, E.; Wagner, H.; vegan: community ecology package version 2.5-6. **2019**. <https://cran.r-project.org/web/packages/vegan/index.html>. **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.