Title: Conservation genomics reveals low genetic diversity and multiple parentage in the threatened freshwater mussel, *Margaritifera hembeli*

Nicole L. Garrison[1]

Paul D. Johnson[2]

Nathan V. Whelan[1,3]

1: School of Fisheries, Aquaculture, and Aquatic Sciences, College of Agriculture, Auburn University, 203 Swingle Hall, Auburn, Alabama 36849, USA

2: Alabama Aquatic Biodiversity Center, Alabama Department of Conservation and Natural Resources, 2200 Highway 175, Marion, Alabama, USA

3: Southeast Conservation Genetics Lab, Warm Springs Fish Technology Center, US Fish and Wildlife Service, 203 Swingle Hall, Auburn, Alabama, 36849

Abstract

Margaritifera hembeli is a federally threatened freshwater mussel species restricted to three central Louisiana drainages. Currently, management efforts are being formulated without an understanding of population-level genetic patterns, which could result in sub-optimal conservation outcomes. In particular, information about riverscape genetic patterns is needed to design effective propagation and reintroduction plans. We apply a genomic approach (RADseq) to assess genetic diversity and structure among four wild populations sampled from across the species range. We also assess the genetic diversity of a captively reared cohort produced from a single female. We recovered population differentiation between individuals sampled to the north and south of the Red River. All sites had similarly low levels of heterogeneity and other measures of genetic diversity. The captive cohort displayed higher levels of genetic diversity than expected and likely represents a case of multiple paternity. Future propagation efforts will likely be able to produce genetically diverse cohorts from a small number of wild-caught females, and we recommend future reintroduction efforts utilize brooders within the sub-drainage closest to the reintroduction effort.

Keywords: Margaritiferidae, RADseq, Louisiana Pearlshell, endangered species act, population genomics.

Introduction

In a world of ever-increasing anthropogenic encroachment and climatic change, effective conservation decisions for species of concern must be made quickly. Regardless of taxon, detailed genetic assessments clarify the pattern and process of diversity across a landscape, diagnose specific conservation challenges, and answer crucial questions managers face during the decision making process (DeSalle and Amato 2004; Segelbacher et al. 2010; Richardson et al. 2016). Genetic information is invaluable to captive propagation programs, where data play a role in the selection of broodstock and maintenance of genetic diversity over time (Witzenberger and Hochkirch 2011). In cases where propagation methods are actively being developed and reintroduction efforts planned, genetic analysis can assist production and brood stock selection, better insure successful reintroduction outcomes, and enable continued monitoring of established populations (Jones et al. 2006; Schwartz et al. 2007; McMurray and Roe 2017). Despite the potential contribution of genetic information to successful conservation efforts, many species have little or no genetic data available; this is particularly true of freshwater mussels(Haag and Williams 2014; Strayer et al. 2019).

Conservation genetic information can now be inferred using thousands of markers from across a genome, providing a high resolution alternative to historic standbys like single gene or microsatellite analyses (Luikart et al. 2003; Davey and Blaxter 2010). Importantly, high throughput methods can be employed in non-model systems with no pre-existing genomic resources. Population genomic analyses can also directly aid in the development of markers for continued monitoring of wild and captive populations (Schwartz et al. 2007; Karlsson et al. 2011; Hendricks et al. 2018) and they represent a powerful tool for conservation managers and stakeholders overseeing long-term recovery programs (Witzenberger and Hochkirch 2011).

Leveraging the power and accessibility of current conservation genomic methods to examine understudied groups such as freshwater mollusks should be a major focus of conservation research as anthropogenic pressures increase.

Freshwater mussels are among the most critically imperiled aquatic organisms. A history of overexploitation and environmental impacts including habitat fragmentation, channel alterations, agricultural inputs, and pollutants has led to global endangerment of mussels. In the last 100 years, an estimated 28 species have gone extinct in North America alone. Of the remaining North American species, approximately 65% are currently at risk of extinction (Williams et al. 1993; Haag and Williams 2014). The importance of mussels to freshwater ecosystems globally cannot be overstated. Mussels are long-lived benthic invertebrates that actively filter suspended particles from water, providing valuable ecosystem services (Vaughn and Hakenkamp 2001). Furthermore, freshwater mussels are intimately linked to water quality, and their shells can provide a record of environmental changes, at both recent (Pfister et al. 2011) and historic scales (Fritts et al. 2017). As part of their complex life cycle, mussels produce parasitic larvae (glochidia) that must attach and mature on the gills of a host fish. This relationship can sometimes be very specific (i.e. a single host species) or more general (a particular family of fishes), and is required for the maturation and dispersal of juvenile mussels (Wächtler et al. 2001). A history of overexploitation and environmental impacts including habitat fragmentation, channel alteration, agricultural inputs, and pollution has led to global endangerment of mussels (Lopes-Lima et al. 2018b). Environmental degradation undoubtedly impacts mussels directly through physiological stress (Strayer et al. 2004), but may also impair recruitment via the reduction of host fish population densities (Bogan 1993; Haag 2012) and their ability to migrate freely (Watters 1996).

90 *Margaritifera hembeli*, Louisiana Pearlshell (Fig. 1a), is federally threatened under the
91 U.S. Endangered Species Act and is restricted to three tributary drainages of the Red River in
92 central Louisiana (Smith 1988)(Fig. 2). Belonging to the family Margaritiferidae, *M. hembeli* is a
93 morphologically and phylogenetically distinct group of mussels with only five extant North
94 American representatives, much fewer than the number of mussel species representing the family
95 Unionidae (Bogan 2008; Lopes-Lima et al. 2018a). *Margaritifera hembeli* individuals are long
96 lived (~45-75 years) and inhabit shallow, high velocity stream reaches with relatively stable
97 substrates (Johnson and Brown 1998; Johnson and Brown 2000). Healthy occurrences of this
98 species are characterized by large, dense beds (Fig. 1b), sometimes exceeding 300 individuals/m^2
99 (Johnson and Brown 1998). In the wild, *M. hembeli* glochidia have been found on *Noturus*
100 *phaeus* (Brown Madtom), *Luxilus chrysocephalus* (Striped Shiner), *Lythrurus umbratilis* (Redfin
101 Shiner), and *Notemigonus crysoleucas* (Golden Shiner), but these may be spurious records and
102 overall host suitability has not been confirmed (Hill 1986; Johnson and Brown 1998). In
103 particular, host fish specificity in the wild remains unclear given that *M. hembeli* glochidia have
104 only been found in small numbers on implicated fish species, yet they transform particularly well
105 on *Esox* spp. species in captivity (see below). Many negative factors influencing mussels on a
106 global scale are also affecting *M. hembeli*. The 1988 United States Fish and Wildlife (USFWS)
107 status report for the species noted that *M. hembeli* populations face threats from poor land
108 management practices (silviculture, gravel mining), reservoir construction, and pollution runoff.
109 During this initial USFWS assessment of *M. hembeli*, it was known only from 11 headwater
110 streams in Bayou Boeuf (Rapides Parish, Louisiana), which contributed to the designation of
111 endangered status under the U.S. Endangered Species Act (USFWS 1988). Upon the discovery

of additional populations in Grant Parish, it was later down listed to threatened (USFWS 1993). Nevertheless, *M. hembeli* remains in urgent need of conservation attention.

For *M. hembeli* and other mussels, there is considerable interest in captive propagation as a management tool (Haag and Williams 2014). Currently, most mussel propagation programs involve the capture of one or a few wild, gravid females followed by inoculation of glochidia on the appropriate host fish in captivity. Progress has been made in technical aspects of rearing certain species such as determining temperature thresholds (Steingraeber et al. 2007), appropriate nutrition (Gatenby et al. 1996), sediment requirements (Jones et al. 2005), host fish identification (Keller and Ruessler 1997; Hart et al. 2018), and creating capacity for large scale production (Barnhart 2006). However, genetic impacts of propagation are less clear. Though many propagation operations exist at federal and state government facilities across the southeast, genetic consequences of in-use propagation protocols for freshwater mussels is an area of research in its infancy. Recommendations exist (Jones et al. 2006), but few studies have explicitly examined the genetic consequences of specific propagation protocols for freshwater mussels.

An initial transformation and culture trial for *M. hembeli* producing individuals utilized in this study began in 2016. Transformation of *M. hembeli* juveniles was completed on a single *Esox americanus americanus* (Redfin Pickerel) by USFWS (Natchitoches National Fish Hatchery in Louisiana and the Ecological Services Office in Louisiana). Infection rate was robust and the effort resulted in nearly 9,000 juvenile *M. hembeli* transformed from the single *E. americanus americanus*. Once transformation had been completed, culture protocols were initiated at the Alabama Aquatic Biodiversity Center (AABC) in Marion, Alabama to evaluate three different methods for rearing *M. hembeli* juveniles. As this initial effort was focused

primarily on determining basic culture methods for *M. hembeli*, transformation of juveniles from a single gravid *M. hembeli* from Black Creek (Grant Parish, Louisiana) represented an opportunity to examine genetic variation of progeny found in a single female. Past research has indicated females of other *Margaritifera* species can mate with multiple males in the wild (Wacker et al. 2018) and multiple paternity has been noted in several Unionid mussel species (Christian et al. 2007; Bai et al. 2012; Ferguson et al. 2013). No study has examined this potential mating system in *M. hembeli*. Given conservation challenges facing *M. hembeli*, the prospect of reintroducing large numbers of low diversity individuals should be carefully considered. Understanding of both the genetic structure of wild populations and the opportunity to contrast against the diversity of a single-female brood could prove invaluable for informing future conservation and management recommendations for *M. hembeli* reintroduction efforts.

Despite genetic concerns for the management and survival of *M. hemebli,* modern genetic data has not been generated to inform management decisions. Previous genetic work on *M. hembeli* found extreme monomorphism ($H_o = 0$) across 25 allozyme markers (Curole et al. 2004). This observed near-absence of heterozygosity was hypothesized to result from the common (and likely natural) stochastic extirpation of *M. hembeli* beds, coupled with high re-colonization rates. A more recent study utilizing microsatellites (Roe 2009) found low heterozygosity at all sampled populations and some genetic structure among *M. hembeli* populations, but conclusions were drawn from only five loci. In total, previously generated genetic data offer scant information for use in designing management efforts. Modern, genomic-level data are necessary to assess population connectivity, measure genome-wide genetic diversity, and provide managers with useful information that can be used to inform propagation and reintroduction efforts.

157 The goals of this study were to 1) assess the genetic diversity of the 2016 AABC captive
158 *M. hembeli* cohort 2) determine whether previous reports of low genetic diversity of wild *M.*
159 *hembeli* populations using allozymes is supported by genomic markers and 3) and characterize
160 patterns of gene flow or potential barriers to aid in reintroduction of future culture efforts. To
161 accomplish these objectives, a restriction enzyme associated high-throughput sequencing method
162 was employed (RADseq) to generate a dataset of thousands genomic loci. RADseq is capable of
163 utilizing genomic DNA samples collected non-lethally to generate millions of sequence reads for
164 each individual. These reads can be processed with existing tools to identify thousands of single
165 nucleotide polymorphisms (SNPs) from across the genome, providing high resolution data for
166 determining demographic and evolutionary dynamics within a species. Ultimately, these data
167 will inform recovery efforts of *M. hembeli* and provide the first genetic profile of a propagated
168 cohort for this threatened freshwater mussel.
169

170 **Materials and Methods**

171 *Sample Collection and Sequencing*

172 We sampled 20 individuals each from four sites in the Red River drainage in Louisiana,
173 and 20 captively reared individuals produced from a single wild caught female from Black
174 Creek. Two sampling localities were from Rapides Parish (Brown Creek – Bayou Rapides and
175 Loving Creek – Bayou Boeuf) and two were from Grant Parish (Jordan Creek and Black Creek –
176 Bayou Rigolette). Sampling sites were selected for ease of access, abundance of *M. hembeli*, and
177 because they encompass the major drainages from which *M. hembeli* is known to occur (Fig. 2).
178 Mussels were collected by hand and effort was taken to select a broad size range likely
179 representing individuals from multiple age classes. A sterile, individually wrapped buccal swab

180 was used to collect genetic material from the foot of each mussel and immediately placed in
181 swab stabilization buffer from the Buccal-PrepPlus DNA isolation kit (Isohelix™). Mussels were
182 photographed on site and returned immediately to their bed after swabbing. The captively reared
183 individuals were lethally sampled at the AABC and a clip of the foot was placed in 96-100%
184 ethanol until DNA extractions could be completed. The mother of the captive cohort was not
185 available for sampling and a tissue voucher/buccal swab was never taken. Shells of captively
186 reared individuals were deposited at the Auburn Museum of Natural History (AUM 45578-
187 45595).

188 DNA extractions of buccal swabs and tissue were completed with Isohelix Xtreme DNA
189 isolation kit and the Qiagen DNeasy blood and tissue kit respectively, following the
190 manufacturers' instructions. After DNA extraction, each sample was treated with RNase A at a
191 final concentration of 100 µg/ml and incubated at 37°C for 15 minutes to remove any co-purified
192 RNA. Each extraction was quantified with a Qubit Fluorometer and checked for integrity of high
193 molecular weight DNA through standard gel electrophoresis. Samples were standardized to a
194 concentration of 20 ng/µL and 50 µL of standardized DNA was sent to Floragenex Inc.
195 (Portland, OR) for RADseq library preparation using the SbfI restriction enzyme following Baird
196 et al. (2008). Samples were tagged with unique barcode identifiers, pooled, and sequenced in
197 three replicate lanes on the Illumina HiSeq 4000 platform using 100bp paired-end chemistry.

198 *Sequence Processing and Variant Identification*

199 Reads from all three sequenced lanes were combined for each individual and processed
200 with the STACKS v2.1 pipeline for population genomic analysis (Rochette et al. 2019).
201 STACKS demultiplexes raw sequencing data, aligns reads to form stacks of loci, identifies
202 variants (i.e. SNPs), and generates descriptive population statistics. Paired-end reads were

203 demultiplexed with the *process_radtags* command using default settings that allow for barcode
204 sequences to be rescued if the barcode varies by only one nucleotide from the expected sequence.
205 Input file preparation details and *process_radtags* settings can be found online
206 (https://github.com/nlgarrison/ConservationGenomics). The denovo_map.pl script was used to
207 automate the STACKS pipeline, as a reference genome is not available for *M. hembeli* or any
208 closely related species. Stack assembly required a minimum of five reads per locus (-m 5),
209 allowed for three mismatches within stacks of the same individual initially (-M 3) and two
210 mismatches between stacks of different individuals (-r 2). All other parameters were set to
211 default. To identify SNPs in stacked loci, a catalog of potential sites must be formed; in our
212 pipeline only individuals from the wild populations were used to generate the catalog. To do this,
213 at the *cstacks* step in the denovo_map.pl pipeline, a population map including only wild-caught
214 individuals was provided to the catalog building command. This approach was chosen to reduce
215 bias in called variants that might be introduced by uneven population sampling, and it allowed
216 for a true test of population assignment for the captive individuals.

217 We ran the *populations* command on two sets of samples: wild-caught individuals only
218 (NoCaptive) and all sampled individuals (WithCaptive). The captive cohort was excluded in
219 initial runs of *populations* to eliminate bias when estimating the baseline population parameters
220 for wild individuals. The NoCaptive dataset was analyzed with a minimum minor allele
221 frequency setting of 0.025, maximum heterozygosity setting of 0.50, and a requirement that a
222 variant be present in at least three sampling sites and at least 50% of individuals within each
223 sampling site; these settings were also used in the analysis of the WithCaptive dataset. For both
224 the NoCaptive and WithCaptive datasets, one *populations* analysis was generated allowing only
225 one random SNP per locus (denoted “S”) and a second was produced allowing multiple SNPs

226 per locus (denoted "M"), resulting in a total of four datasets; NoCaptiveS, NoCaptiveM,
227 WithCaptiveS, WithCaptiveM. This was done because some downstream analyses assume
228 unlinked loci, whereas others can use multiple SNPs originating from the same locus. For
229 downstream analyses requiring subsets of individuals as input, the program VCFtools (Danecek
230 et al. 2011) was used to generate reduced datasets as needed from the WithCaptiveS/M vcf files.
231

232 *Population Genomic Analyses*

233 For each dataset, average heterozygosity, nucleotide diversity, pairwise F_{ST} among each
234 sampling site, and F_{IS} at each sampling site was reported by *populations*. The *basicStats* function
235 of the R (R CORE Team 2019) package diveRsity (Keenan et al. 2016) was used to calculate
236 allelic richness. We assessed population structure among sampling sites with an Analysis of
237 Molecular Variance (AMOVA; Excoffier, Smouse, and Quattro 1992) and a series of clustering
238 methods. AMOVA was performed using the *poppr.amova* command in the R package adegenet
239 (Kamvar et al. 2014; Jombart et al. 2018). Individuals were stratified by sample site and whether
240 the site was located north or south of the Red River to assess whether the Red River serves as a
241 barrier to gene flow or if genetic structure is better explained by local landscape characteristics.
242 Significance was tested with a 500 permutation randomization test.

243 Discriminant analysis of principal components (DAPC) was implemented using the R
244 package adegenet following Jombart and Collins (2015), using the NoCaptiveM and
245 WithCaptiveM datasets. The best-fit number of clusters (K) was assessed using *k*-means
246 clustering with Bayesian information criteria. The *snmf* function of the R package LEA (Frichot
247 and François 2015) was used to identify clusters of individuals and determine admixture
248 proportions, and the best-fit K was determined using the cross-entropy criterion. For these

249 analyses, only the unlinked SNPs (NoCaptiveS, WithCaptiveS) were used. Though similar in
250 function to STRUCTURE (Pritchard et al. 2000), LEA can be more accurate than STRUCTURE
251 in the face of inbreeding (Frichot et al. 2014) and handles genomic data more efficiently.

252 We also used the model-based method fineRADstructure (Malinsky et al. 2018) to
253 generate a summary of haplotype coancestry for all individuals. A major advancement provided
254 by fineRADstructure is its ability to utilize linkage and polymorphism in genomic data without a
255 reference genome, allowing fine scale patterns of relatedness among individuals to be examined.
256 This analysis was conducted with the NoCaptiveM, WithCaptiveM, and the captive cohort in
257 isolation. The captive cohort was examined independently in order to make more accurate
258 inferences about potential multiple paternity.

259 We tested for a signature of isolation-by-distance using a Mantel test of correlation
260 between geographical distance and pairwise F_{ST} values for the wild populations. Geographic
261 distances among sites were measured by plotting sample collection points in QGIS (QGIS
262 Development Team 2014) and hand tracing stream distance between sampling sites in a pairwise
263 fashion. As river connections were sometimes difficult to assess and to account for minor
264 idiosyncrasies associated with tracing river path, hand tracing was repeated three times for each
265 pair and an average of distances in meters was taken as the geographic stream distance. F_{ST}
266 values used for the Mantel test were those reported by STACKS (Table 1). The Mantel test was
267 done using the R package 'ade4' (Dray, Dufour, et. al 2007); significance was evaluated with
268 1,000 random permutations. As Mantel tests have received criticism for use as a measure of
269 isolation by distance (Legendre et al. 2015; Meirmans 2015), we also performed a multiple
270 regression on distance matrices with the MRM function of the Ecodist R package (Goslee and
271 Urban 2007); significance was tested with 10,000 permutations.

272 The programs diveRsity and Migrate-n version 4.2.14 (Beerli and Palczewski 2010) were
273 used to evaluate connectivity between wild populations of *M. hembeli*. The *divMigrate* function
274 of the R package diveRsity uses differences in allele frequencies to model asymmetric, relative
275 rates of migration between populations (Sundqvist et al. 2016). Though it was designed for use
276 with microsatellite data, *divMigrate* is capable of handling genomic SNP datasets and has been
277 shown to reflect biologically realistic scenarios of population connectivity in recent studies (e.g.
278 Woodings et al. 2018; Manuzzi et al. 2019). An analysis of the WithCaptiveS dataset was done
279 with the *divMigrateOnline* implementation (Keenan 2012) to calculate relative rates of migration
280 between wild populations using a cutoff value of 40, an alpha of 0.05, and the G_{ST} statistic (Nei
281 1973). Support for the asymmetry of migration rates was evaluated with 1000 bootstrap
282 replicates. Although the captive population was included in our initial *divMigrate* analysis, due
283 to its high similarity (G_{ST} = 1.0) with Black Creek individuals it was masked using the "exclude
284 population" option without recalibrating rates.

285 The *divMigrateOnline* analysis was complemented by the Bayesian population genetics
286 program migrate-n. Given computational demands of migrate-n, a subset of 100 polymorphic
287 loci were randomly selected from the NoCaptiveM dataset. Migrate-n analyses were done with
288 loci in their entirety, rather than individual SNPs, as the SNP model has not been thoroughly
289 tested (see migrate-n manual). Loci that appeared more than once in the random subset were
290 filtered out, leaving 95 polymorphic loci for inclusion in the migrate-n analysis. Geographic
291 distances between sites were calculated as described previously. Five migration models were
292 investigated as indicated by geography and other population genomic analyses 1) full migration
293 2) northern panmixia with unidirectional gene flow from Loving Creek to Brown Creek 3)
294 northern panmixia with bidirectional gene flow between Loving Creek and Brown Creek 4)

295 panmixia (Fig. 3). For each migration model, Bayesian inference was performed using the DNA
296 sequence model (sampling of 20,000,000 total steps, 10,000 steps discarded as burn-in, default
297 priors) and combined with thermodynamic integration (four parallel heated chains) at
298 temperatures 1, 1.5, 3.0, and $1x10^6$. Log marginal likelihood values were calculated with Bezier
299 approximation within migrate-n and log Bayes factors were used to rank the models following
300 Beerli & Palczewski (2010).

301 Effective population size (N_e) and the number of effective breeders (N_b) was estimated
302 for each sampling site, geographically proximate wild populations combined, and the captive
303 cohort in isolation with NeEstimator2 (Do et al. 2014) using the NoCaptiveS and WithCaptiveS
304 datasets as input. Given results indicating that the Red River represents a barrier to gene flow
305 (see below), we combined sites to assess N_e of "northern" (Black Creek and Jordan Creek) and
306 "southern" (Loving Creek and Brown Creek) populations. The molecular coanscestry method of
307 Nomura (2008) and linkage disequilibrium method (Hill 1981; Waples 2006) as implemented in
308 NeEstimator2 were used to evaluate N_e and N_b without a genomic map.

309 The program COLONY (Jones and Wang 2010) was used to validate suspected multiple
310 paternity in the captive cohort. A captive-only subset of SNPs filtered from the WithCaptiveM
311 dataset using vcftools was used as input. Three different filtering strategies were tested, allowing
312 sites with missing data proportions of 0%, 25%, and 50%. Files were converted to COLONY
313 input format using tidy_vcf and the write_colony function of the R package radiator (Gosselin et
314 al. 2020). Long, full likelihood runs were performed with 10 replications for each filtered
315 dataset; male and female polygamy settings were used and half sibship of all individuals with an
316 unknown mother was specified in the input file.
317

318 **Results**

319 Sequencing resulted in >460 million reads per lane and >1 million reads per individual
320 (n=94). Sequence reads filtered out due to barcode ambiguity, unclear restriction cut sites, and
321 potential Illumina adapter contamination comprised ~15% of the total raw reads; 1.4 billion
322 paired-end reads were included in the assembly process. We saw no evidence of contamination
323 in our samples and most loci present in Black Creek were also present in the Captive Cohort,
324 indicating that there was no bias associated with DNA sampling method in our study.
325 Furthermore, there were no detectable differences in the quality or quantity of sequences
326 obtained from the two sample collection methods (swab and tissue clips). The STACKS pipeline
327 identified 1,185,792 SNPs across 20,464 loci, further filtered by *populations* program constraints
328 to 2,563 putatively independent variant SNPs across the wild populations. When multiple SNPs
329 per locus were allowed, the number of variant SNPs recovered across all wild populations
330 increased to 7,601. After initial runs of the STACKS pipeline, four individuals having less than
331 half the identified SNPs found in the rest of the samples (Jordan18, Captive17, Loving20, and
332 Brown5) were removed, resulting in datasets generated for downstream analyses which included
333 90 individuals from the original set.

334 The vast majority of markers sampled were not variable across the four wild populations;
335 out of ~20,000 potential markers (i.e. SNPs), about 10%, met the population parameters
336 specified and contained informative variation. Average nucleotide diversity (π) ranged from
337 0.21-0.23. Observed heterozygosity was considerably lower than expected heterozygosity across
338 all sites (H_{obs} = 0.08-0.09, H_{exp} = 0.20-0.22), likely indicating genetic bottlenecks at all sites.
339 Allelic richness was also similar across all sampled groups (1.16-1.51). Analysis of the captive
340 population generated 2,416 variable SNPs (8,069 when multiple SNPs per locus were allowed)

with an average observed H_{obs} of 0.10 and an average nucleotide diversity of 0.25. Pairwise F_{ST} values were low, with the Brown Creek population presenting the highest differentiation from other populations (0.063-0.082; see Table 1 for all pairwise F_{ST} values); the captive cohort and the Black Creek population had the lowest pairwise F_{ST} (0.035). F_{IS} values for each collection site ranged from 0.44 to 0.48 and was 0.39 for the captive cohort.

Genetic structure was seen among collection sites to varying degrees depending on analytical method. AMOVA was significant at all hierarchical levels ($p = 0.001$), but only 3.89% of genetic variation was explained by whether the sampling site was north or south of the Red River. A further 6.7% of genetic variation was explained by collection sites within regions, indicating landscape barriers to gene flow that are more complex than one major river. The Mantel test and multiple regression for isolation-by-distance was significant ($p < 0.05$), suggesting isolation by distance patterns in the data. DAPC analysis of wild individuals indicated that a K of 4 most accurately captured the diversity of the samples, with individuals from Jordan and Black (Grant Parish) overlapping and Brown and Loving (Rapides Parish) each forming their own distinct cluster (Fig. 4a). When all populations were included, captive mussels overlapped entirely with the Black Creek cluster (Fig. 4b). Notably, the captive mussels showed similar spread in the DAPC analysis as wild individuals from Black Creek, indicating the captive cohort possesses variation in genetic diversity comparable to wild populations.

Analysis with LEA suggested data were best explained by two genetic clusters when only wild populations were included. LEA analysis with K = 2 showed that individuals from sites north of the Red River (Grant Parish) had more similar admixture profiles to each other than those from south of the Red River (Rapides Parish) and *vice versa* (Fig. 5). Hierarchical clustering analysis with fineRADstructure, mirrored DAPC and LEA analyses (Fig. 6). Brown

364 Creek and Loving Creek formed groupings, to the exclusion of most other individuals, with
365 particularly high coancestry values detected within the Brown Creek population. Individuals
366 from Jordan Creek, Black Creek, and the Captive cohort are nearly indistinguishable from each
367 other with only minor internal structuring comprising a subset of Black Creek and Captive
368 individuals (Fig. 6). When examined in isolation, the captive cohort displayed structured
369 relationships at varying levels of coancestry; at the top of the coancestry value range a cluster of
370 individuals (Cap3, Cap10, and Cap4) may represent a grouping of full siblings (Fig. 7).

371 The program *divMigrateOnline* detected relatively high migration rates between the two
372 northern sampling sites, Jordan Creek and Black Creek. To a lesser degree, the analysis showed
373 migration from the northern populations to Loving Creek and unidirectional movement from
374 Loving Creek to Brown Creek (Fig. 8). The asymmetry of migration rates inferred by
375 *divMigrateOnline* were not supported by 1000 bootstrap replicates, meaning that while gene
376 flow was detected between Jordan Creek and Black Creek, no strong directionality could be
377 inferred. Of the models assessed by migrate-n, the most likely was model 3 (Table 3) with
378 populations occurring north of the Red River being panmictic, bidirectional gene flow from the
379 panmictic northern population to Loving Creek, and bidirectional migration between Loving
380 Creek and Brown Creek.

381 Estimates of effective population size for some populations in isolation resulted in
382 negative or infinite values, suggesting that the values were driven by sampling error as a result of
383 insufficient sample size or marker informativeness (Marandel et al. 2019). Combining
384 individuals from the northern sites (Jordan Creek, Black Creek) and southern sites (Brown Creek
385 and Loving Creek) allowed for more realistic estimation of effective population sizes (Table 2).
386 The captive cohort generated the lowest estimates of N_e (86 -111 individuals) but an N_b of 5.

387 This high estimate for the number of effective breeders within a cohort known to originate from
388 a single female is a strong indication of multiple paternity.
389 COLONY also indicated multiple paternity was present in the captive cohort samples, but
390 the point estimates of male input varied with the level of missing data allowed into the analysis.
391 Using the most complete dataset (sites every individual shared), COLONY output indicated
392 several full-sibling clusters and 7 fathers. The most relaxed filtering strategy (50% site coverage)
393 indicated a different father for each individual in the sample (n = 19). Estimates of N_e produced
394 by COLONY (N_e = 4, 95% CI = 2-12) were stable across analyses and reflected both the
395 NeEstimator2 values for the captive samples (Table 2) and the fineRADstructure clustering
396 analysis (Fig. 7).
397
398 **Discussion**
399 Population genomic data revealed low levels of heterozygosity across *M. hembeli*
400 populations and complex patterns of gene flow among sites. Although low levels of genetic
401 diversity in wild populations is concerning from a conservation standpoint, our results are similar
402 to past studies on *M. hembeli* (Curole et al. 2004; Roe 2009). Moreover, low genetic diversity
403 and a signature of bottlenecks at every sampling site may reflect natural processes such as rapid
404 colonization after the loss of a habitat that results from stream meanders and cutoffs that are
405 common in low-elevation, flat terrains like those in central Louisiana. Despite low genetic
406 diversity, high gene flow among some sample sites was observed. Genetic structure generally
407 followed a pattern of isolation-by-distance, but the Red River also appeared to be a factor in
408 observed genetic structure. This suggests movement between populations on the same side of the
409 Red River may occur during major flood events, thereby facilitating migration with infected

410 fishes moving to adjacent stream channels and does not require host-fish passage through the
411 Red River channel.

412 Our data showed a captive cohort of *M. hembeli* produced from a single female had
413 comparable genetic diversity to the wild population from which the female was sampled.
414 Broadly, analyses indicate evidence of multiple paternity, which has considerable implications
415 for propagation efforts of *M. hembeli*. Although multiple paternity has been known to occur in at
416 least some freshwater mussels (Christian et al. 2007; Bai et al. 2012; Ferguson et al. 2013;
417 Wacker et al. 2018), more recent studies have empirically demonstrated wild female mussels can
418 mate with multiple males to produce genetically diverse captive cohorts (Wacker et al. 2019).
419 Multiple paternity had not been empirically observed in *M. hembeli* prior to this study. Results
420 suggest progeny generated from a few (randomly selected) wild-fertilized females annually
421 would mirror the heterogenity found in the wild host populaton, likely resulting in low detectable
422 differences between propagated and wild mussels with a recovery effort that spans multiple
423 years.

424

425 *Genetic Diversity and Effective Population Size*

426 Low observed heterozygosity relative to expected heterozygosity (Table 2), suggests *M.*
427 *hembeli* populations are small and experiencing associated effects such as inbreeding and genetic
428 bottlenecks. This is also reflected by high F_{IS} values (Table 2). Given the threatened status of *M.*
429 *hembeli* under the U.S. Endangered Species Act, anthropogenic activities have clearly caused
430 severe population declines. However, life history, natural demographics, and stochastic habitat
431 disturbances probably also play a possibly an outsized role, in influencing low heterozygosity of
432 *M. hembeli*. For instance, a single beaver dam once led to the extirpation of an *M. hembeli* bed

433 (~1000 individuals) located on Forest Service land (Stewart 1990). Furthermore, Johnson and
434 Brown (2000) showed that channel stability of the study area on the time-scale of a single year
435 can be low, even though they noted that *M. hembeli* appeared to be associated with relatively
436 stable microhabitats. Assuming rapid colonization ability of *M. hembeli* (see Curole, Foltz, and
437 Brown 2004), low heterozygosity is likely the result of repeated bottlenecks caused by habitat
438 destruction followed by colonization of newly exposed suitable habitat. Although natural
439 destruction of mussel beds and rapid colonization of new habitats may be a normal part of *M.*
440 *hembeli* biology, habitat fragmentation caused by anthropogenic activity exacerbates
441 contemporary population decline by restricting recruitment across populations (Geist and
442 Auerswald 2007). The habitat fragmentation witnessed through modification of waterways or
443 changes in landuse are likely limiting re-colonization options for *M. hembeli*.

444 When considered in isolation, N_e estimates for all sampling sites except Black Creek
445 were inferred to be infinite, which is indicative of N_e estimates being driven by sampling error,
446 rather than large population sizes (Marandel et al. 2019). Pooling individuals from north and
447 south of the Red River, respectively, allowed for more precise estimates of N_e (Table 2).
448 Effective population size is difficult to accurately estimate even with genome-wide markers as
449 many methods make assumptions that are typically violated such as sampling of non-overlapping
450 generations (Waples et al. 2016; Wang et al. 2016; Marandel et al. 2019). The linkage
451 disequilibrium method applied here is known to be downwardly biased (as much as 30%) when
452 samples consist of different age classes (Waples et al. 2014). We anticipate that our estimates of
453 N_e for wild populations are much lower than reality, as multiple age classes were sampled. Thus,
454 N_e estimates reported here may be of value to resource managers, but they should be approached
455 with caution and not the sole genetically derived metric used for assessing populations of *M.*

456 *hembeli*. The temporal trajectory of N_e is more crucial to conservation managers than a point
457 estimate, and we recommend additional sampling of size/age class cohorts to reveal trends
458 through time.
459
460 *Population Connectivity*
461 In general, analyses determined genetic demarcation between populations occurring north
462 of the Red River in those occurring south of it. However, fine-scale relationships and migration
463 patterns among populations that our SNP-based approach illuminated are novel. Sampled sites
464 appeared to demonstrate an isolation-by-distance effect, which is a general pattern seen in many
465 freshwater organisms (Meffe and Vrijenhoek 1988; Whelan et al. 2019). Furthermore, AMOVA
466 indicated significant genetic structure with a genetic break occurring between collection sites
467 north and south of the Red River, with further significant genetic structure between populations
468 in the north and south. These patterns were recovered to varying degrees by clustering analyses,
469 with DAPC indicating at least some genetic distinctness among all four sites (Fig. 4).
470 Analyses that examined finer-scale gene-flow patterns indicated a high amount
471 connectivity between Grant Parish populations (Jordan Creek and Black Creek), with the best-fit
472 model as assessed by migrate-n having a panmictic Jordan Creek and Black Creek (Fig.
473 2).Whether the observed pattern is a result of active gene flow or a historical connection is
474 unknown. Black Creek has been putatively isolated from Jordan Creek for the past 60 years
475 behind Iatt Lake, but the long lifespan of *M. hembeli* may result in a longer time-period for
476 genomic signatures of isolation to be detectable. Documentation of Iatt Lake's management
477 history (Moses et al. 2016) indicates *Esox americanus americanus*, was not detected in fish
478 surveys but other *Esox* species (*Esox americanus vermiculitus*, *Esox niger*) have been collected

479 during Iatt Lake surveys. Iatt Lake is prone to flooding and has experienced several high water
480 events in recent history (Moses et al. 2016), which could facilitate fish passage between
481 tributaries. Assuming the presence of *Esox* spp. around Lake Iatt and periodic flooding it is
482 possible that a connection was recent and possibly intermittent between Black Creek and Jordan
483 Creek drainages.

484 Relative genetic homogeneity among northern sites contrasts with the relative isolation of
485 those sampled from south of the Red River. Loving Creek and Brown Creek shared similar
486 genomic backgrounds, as indicated by LEA (Fig. 4), but *divMigrate* and migrate-n analyses
487 indicated somewhat limited gene flow between the two southern sites. Several analyses indicated
488 that Brown Creek was the most isolated group of individuals sampled, which may be partially
489 explained by a higher stream distance from its Red River confluence compared to other sites.
490 Connectivity analyses appear to indicate that Loving Creek represents Brown Creek's only
491 connection to the rest of the species range (Fig. 6). Though there are sites not sampled during the
492 course of this study which occur in streams located between Loving and Brown Creek, our data
493 still indicates that movement of *M. hembeli* in the southern part of its range is relatively more
494 restricted than in the northern section of its range.

495 Further natural history work is needed to better understand the conditions required for
496 successful dispersal of *M. hembeli*. Although a detailed host fish trial has not been completed for
497 *M. hembeli*, formal trials have been completed for *M. marrianae*, the Alabama Pearlshell. These
498 trials indicate *Esox vermicularis vermicularis* is the primary host in Alabama. A minimal
499 transformation rate was also documented for *Noturus leptacanthus* (Speckled Madtom) but it's
500 not a primary host for *M. marrianae* (Fobian et al., *in prep.*). In contrast, *M. hembeli* readily
501 transforms transform on *Esox* spp. in captivity (Schmidt-Frater, *pers. comm.*) but a formal host

trial has not been completed. Additionally, several non-sampled populations of *M. hembeli* occur in headwater streams isolated behind reservoirs – future studies including those sites will likely provide further insights into the impact of such barriers on the dispersal of *M. hembeli* across its range.

Captive Propagation and Reintroduction

Our analysis of captively reared individuals revealed a surprising amount of genetic diversity given that the cohort was produced by a single female from Black Creek. Notably, genetic diversity estimates for this cohort were virtually identical to estimates from the wild population (Table 2). Clustering analyses provided evidence these data were more than sufficient to assign captive individuals to their population of origin, always grouping them with wild Black Creek individuals (Fig. 4b). Multiple paternity was also evident based on estimates of the number of effective breeders for the captive cohort (N_b =5). Coupled with inferences from COLONY and fineRADstructure, where multiple paternal genotypes (7-19) and multiple clusters of individuals with high (but not identical) co-ancestry were observed, our study supports the presence of a multiple paternity strategy within *M. hembeli*. Overall, this represents a best-case scenario for managers as a limited number of wild-caught *M. hembeli* females can be brought into captivity annually and produce a genetically diverse cohort for reintroduction efforts. However, given the difficulty of *Margaritifera* spp. production in a hatchery setting (Paul Johnson *pers. obs.*), any serious reintroduction effort would likely be a decades-long endeavor at minimum.

The ability to produce genetically diverse individuals from a small number of females should facilitate propagation programs; however, care must still be taken when choosing brood

525 stock and determining reintroduction sites. Managers should be guided by our findings of
526 population structure and isolation by distance. Broadly, brood stock should be as geographically
527 proximate to the chosen reintroduction site as possible, coming from one or a few sites within the
528 same drainage. At the very least, broodstock should come from the same side of the Red River as
529 the chosen reintroduction site. Furthermore, genetic diversity and number of effective breeders
530 for another Margaritiferid, *M. margaritifera*, was recently determined to be higher when females
531 were fertilized in the wild relative to those fertilized in captivity (Wacker et al. 2019). Our work
532 suggests multiple paternity is likely the case for *M. hembeli* as well. The best chance we have at
533 maintaining appropriate levels of diversity is to utilize wild-fertilized broodstock while it is still
534 available, rather than attempt to establish a captive breeding colony of *M. hembeli*.

535 Our data also give reasons to avoid augmentation (i.e., placing captively reared
536 individuals on top of a natural population) in favor of reintroductions (i.e., placing captively
537 reared individuals at a site from which *M. hembeli* has been extirpated). Each population
538 analyzed here was considered genetically distinct in at least one analysis, albeit with limited or
539 no genetic distinction between Black Creek and Jordan Creek in most analyses. That said,
540 consequences of outbreeding depression are impossible to predict at this time, and augmentation
541 violates recent recommendations for freshwater mussel propagation and release (Mobile River
542 Basin Mollusk Restoration Committee 2010; Cumberlandian Region Mollusk Restoration
543 Committee 2010; Strayer et al. 2019). Our data support such recommendations. If managers are
544 faced with no suitable sites for propagation and release other than sites with natural *M. hembeli*
545 population, then we argue that habitat restoration should be a higher priority than captive
546 propagation of *M. hembeli*. Good animals placed into poor habitat will not have the desired
547 outcome (Geist and Auerswald 2007).

548

549 **Conclusions**

550 This study provides information that can be used to facilitate successful propagation
551 efforts and a framework for studying the existing diversity in imperiled mussel species using
552 modern methods. We have demonstrated that genetically diverse cohorts of margaritiferids may
553 be produced from a small number of wild-caught, gravid females. Importantly, our findings also
554 indicate that the occurrence of multiple paternity in freshwater mussels may be more widespread
555 than the limited number of explicitly documented cases. Our findings can likely be generalized
556 to closely related species such as the federally endangered Alabama Pearlshell, *M. marrianae,*
557 the sister species to *M. hembeli*, which is currently the focus of intense propagation and
558 management efforts. More broadly, we have demonstrated the utility of RAD-seq approaches,
559 compared to older technologies, in providing fine-scale information for freshwater mussel
560 conservation. Although RAD-seq is now widely used for many conservation genetics studies of
561 non-model organisms, its use for freshwater mussel research is still rare.

562 In terms of Louisiana Pearlshell recovery, our recommendation is to continue propagation
563 efforts utilizing wild-fertilized females with a focus on habitat restoration and continued life
564 history research. Given the likelihood of multiple parentage, captive cohorts produced from a
565 single female will have more diversity than might have been previously expected. However,
566 efforts should be made to not re-use the same females over multiple propagation years, and the
567 use of multiple broodstock females per year is encouraged when possible. Broodstock selected
568 from populations north of the Red River should not be used to propagate individuals that are to
569 be released into localities south of the river. While our findings provide some hope for the
570 efficacy of propagating the Louisiana Pearlshell using existing populations as sources of diverse

brood stock, it also indicates that high levels of inbreeding and loss of population connectivity may be a looming problem for long-term survival of the species. Indeed given the genetic bottlenecks at multiple sites sampled in this study, further analyses might reveal hererogeneity of a reintroduction effort could be improved by mixing propagules from multiple adjacent populations, taking care to keep individual efforts within their respecitive subdrainages. More work is needed to ensure that reintroduced and existing populations of this threatened species form a connected, contemporarily recruiting network of individuals capable of sustaining itself if true recovery is to be achieved.

Acknowledgements

We thank Jared Streeter (Louisiana Department of Wildlife and Fisheries), Ted Soileau, and Steve Shively (US Forest Service) for assistance in the field. Nathan Johnson (US Geological Survey) and Jürgen Geist (Technische Universität München) provided advice on swabbing and extraction methods. A special thanks to the staff of the Alabama Aquatic Biodiversity Center for maintaining captively reared *M. hembeli*, developing propagation techniques for the species, and allowing access to tissues. We also thank the landowners who provided access to sampling sites on privately owned land. The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the U.S. Fish and Wildlife Service.

Funding This work was funded by a 1311 system-wide funding grant from U.S. Fish and Wildlife Service.

Availability of data and material

Demultiplexed Illumina sequence data have been uploaded to NCBI SRA (ascession numbers to be provided upon manuscript acceptance). Processed datasets in various file formats and certain program input and output files (e.g., colony) are available on FigShare (private link for

596 reviewers: https://figshare.com/s/2707ff9ee3050d6b10f6; *DOI to be provided upon acceptance*).

597 Additional code related to the execution of pipelines used are available online at

598 https://github.com/nlgarrison/ConservationGenomics.

599

600 **References**

601 Bai Z, Luo M, Zhu W, et al (2012) Multiple paternity in the freshwater pearl mussel Hyriopsis
602 cumingii (Lea, 1852). J Molluscan Stud 78:142–146.
603 https://doi.org/10.1093/mollus/eyr049

604 Baird NA, Etter PD, Atwood TS, et al (2008) Rapid SNP Discovery and Genetic Mapping Using
605 Sequenced RAD Markers. PLoS ONE 3:e3376.
606 https://doi.org/10.1371/journal.pone.0003376

607 Barnhart MC (2006) Buckets of muckets: A compact system for rearing juvenile freshwater
608 mussels. Aquaculture 254:227–233. https://doi.org/10.1016/j.aquaculture.2005.08.028

609 Beerli P, Palczewski M (2010) Unified Framework to Evaluate Panmixia and Migration
610 Direction Among Multiple Sampling Locations. Genetics 185:313–326.
611 https://doi.org/10.1534/genetics.109.112532

612 Bogan AE (1993) Freshwater Bivalve Extinctions (Mollusca: Unionoida): A Search for Causes.
613 Am Zool 33:599–609. https://doi.org/10.1093/icb/33.6.599

614 Bogan AE (2008) Global diversity of freshwater mussels (Mollusca, Bivalvia) in freshwater. In:
615 Balian EV, Lévêque C, Segers H, Martens K (eds) Freshwater Animal Diversity
616 Assessment. Springer Netherlands, Dordrecht, pp 139–147

617 Christian AD, Monroe EM, Asher AM, et al (2007) Methods of DNA extraction and PCR
618 amplification for individual freshwater mussel (Bivalvia: Unionidae) glochidia, with the
619 first report of multiple paternity in these organisms: TECHNICAL ARTICLE. Mol Ecol
620 Notes 7:570–573. https://doi.org/10.1111/j.1471-8286.2007.01743.x

621 Cumberlandian Region Mollusk Restoration Committee (2010) Plan for the Population
622 Restoration and Conservation of Imperiled Freshwater Mollusks of the Cumberlandian
623 Region

624 Curole JP, Foltz DW, Brown KM (2004) Extensive allozyme monomorphism in a threatened
625 species of freshwater mussel, Margaritifera hembeli Conrad (Bivalvia: Margaritiferidae).
626 Conserv Genet 5:271–278. https://doi.org/10.1023/B:COGE.0000030005.14751.76

627 Danecek P, Auton A, Abecasis G, et al (2011) The variant call format and VCFtools.
628 Bioinformatics 27:2156–2158

Davey JW, Blaxter ML (2010) RADSeq: next-generation population genetics. Brief Funct Genomics 9:416–423. https://doi.org/10.1093/bfgp/elq031

DeSalle R, Amato G (2004) The expansion of conservation genetics. Nat Rev Genet 5:702. https://doi.org/10.1038/nrg1425

Do C, Waples RS, Peel D, et al (2014) NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Mol Ecol Resour 14:209–214

Dray S, Dufour A-B, others (2007) The ade4 package: implementing the duality diagram for ecologists. J Stat Softw 22:1–20

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Ferguson CD, Blum MJ, Raymer ML, et al (2013) Population structure, multiple paternity, and long-distance transport of spermatozoa in the freshwater mussel Lampsilis cardium (Bivalvia:Unionidae). Freshw Sci 32:267–282. https://doi.org/10.1899/12-028.1

Frichot E, François O (2015) LEA: an R package for landscape and ecological association studies. Methods Ecol Evol 6:925–929

Frichot E, Mathieu F, Trouillon T, et al (2014) Fast and Efficient Estimation of Individual Ancestry Coefficients. Genetics 196:973–983. https://doi.org/10.1534/genetics.113.160572

Fritts AK, Fritts MW, Haag WR, et al (2017) Freshwater mussel shells (Unionidae) chronicle changes in a North American river over the past 1000 years. Sci Total Environ 575:199–206. https://doi.org/10.1016/j.scitotenv.2016.09.225

Gatenby CM, Neves RJ, Parker BC (1996) Influence of Sediment and Algal Food on Cultured Juvenile Freshwater Mussels. J North Am Benthol Soc 15:597–609. https://doi.org/10.2307/1467810

Geist J, Auerswald K (2007) Physicochemical stream bed characteristics and recruitment of the freshwater pearl mussel (Margaritifera margaritifera). Freshw Biol 52:2299–2316

Goslee SC, Urban DL (2007) The ecodist package for dissimilarity-based analysis of ecological data. J Stat Softw 22:1–19

Gosselin T, Lamothe M, Grewe P (2020) radiator: RADseq Data Exploration, Manipulation and Visualization using R

Haag WR (2012) North American Freshwater Mussels: Natural History, Ecology, and Conservation. Cambridge University Press

Haag WR, Williams JD (2014) Biodiversity on the brink: an assessment of conservation strategies for North American freshwater mussels. Hydrobiologia 735:45–60. https://doi.org/10.1007/s10750-013-1524-7

Hart MA, Haag WR, Bringolf R, Stoeckel JA (2018) Novel technique to identify large river host fish for freshwater mussel propagation and conservation. Aquac Rep 9:10–17. https://doi.org/10.1016/j.aqrep.2017.11.002

Hendricks SA, Schweizer RM, Wayne RK (2018) Conservation genomics illuminates the adaptive uniqueness of North American gray wolves. Conserv Genet. https://doi.org/10.1007/s10592-018-1118-z

Hill A (1986) The Glochidia (larvae) of the Freshwater Mussel Margaritifera hembeli (Unionacea: Margaritiferidae) Fish Host Species, Morphology, and Period of Fish Host Infection

Hill WG (1981) Estimation of effective population size from data on linkage disequilibrium. Genet Res 38:209–216. https://doi.org/10.1017/S0016672300020553

Johnson PD, Brown KM (2000) The importance of microhabitat factors and habitat stability to the threatened Louisiana pearl shell, Margaritifera hembeli (Conrad). 78:7

Johnson PD, Brown KM (1998) Intraspecific life history variation in the threatened Louisiana pearlshell mussel, Margaritifera hembeli. Freshw Biol 40:317–329

Jombart T, Collins C (2015) A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0. 0. Imp Coll Lond-MRC Cent Outbreak Anal Model 43:

Jombart T, Kamvar ZN, Collins C, et al (2018) Package ‘adegenet’

Jones J, Mair R, Neves R (2005) Factors Affecting Survival and Growth of Juvenile Freshwater Mussels Cultured in Recirculating Aquaculture Systems. North Am J Aquac 67:210–220. https://doi.org/10.1577/A04-055.1

Jones JW, Hallerman EM, Neves RJ (2006) Genetic management guidelines for captive propagation of freshwater mussels (Unionoidea). J Shellfish Res 25:527–535. https://doi.org/10.2983/0730-8000(2006)25[527:GMGFCP]2.0.CO;2

Jones OR, Wang J (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. Mol Ecol Resour 10:551–555

Kamvar ZN, Tabima JF, Grünwald NJ (2014) Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2:e281. https://doi.org/10.7717/peerj.281

Karlsson S, Moen T, Lien S, et al (2011) Generic genetic differences between farmed and wild Atlantic salmon identified from a 7K SNP-chip. Mol Ecol Resour 11:247–253. https://doi.org/10.1111/j.1755-0998.2010.02959.x

Keenan K (2012) divMigrate-online: Visualise and test gene flow patterns among populations. https://popgen.shinyapps.io/divMigrate-online/

Keenan K, McGinnity P, Cross TF, et al (2016) diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. Methods Ecol Evol 782–788. https://doi.org/10.1111/2041-210X.12067@10.1111/(ISSN)2041-210X.ecologyandevolutionireland

Keller AE, Ruessler DS (1997) Determination or Verification of Host Fish for Nine Species of Unionid Mussels. Am Midl Nat 138:402–407. https://doi.org/10.2307/2426831

Legendre P, Fortin M-J, Borcard D (2015) Should the Mantel test be used in spatial analysis? Methods Ecol Evol 6:1239–1247

Lopes-Lima M, Bolotov IN, Do VT, et al (2018a) Expansion and systematics redefinition of the most threatened freshwater mussel family, the Margaritiferidae. Mol Phylogenet Evol 127:98–118. https://doi.org/10.1016/j.ympev.2018.04.041

Lopes-Lima M, Burlakova LE, Karatayev AY, et al (2018b) Conservation of freshwater bivalves at the global scale: diversity, threats and research needs. Hydrobiologia 810:1–14. https://doi.org/10.1007/s10750-017-3486-7

Luikart G, England PR, Tallmon D, et al (2003) The power and promise of population genomics: from genotyping to genome typing. Nat Rev Genet 4:981. https://doi.org/10.1038/nrg1226

Malinsky M, Trucchi E, Lawson DJ, Falush D (2018) RADpainter and fineRADstructure: population inference from RADseq data. Mol Biol Evol 35:1284–1290

Manuzzi A, Zane L, Muñoz-Merida A, et al (2019) Population genomics and phylogeography of a benthic coastal shark (Scyliorhinus canicula) using 2b-RAD single nucleotide polymorphisms. Biol J Linn Soc 126:289–303. https://doi.org/10.1093/biolinnean/bly185

Marandel F, Lorance P, Berthelé O, et al (2019) Estimating effective population size of large marine populations, is it feasible? Fish Fish 20:189–198. https://doi.org/10.1111/faf.12338

McMurray SE, Roe KJ (2017) Perspectives on the Controlled Propagation, Augmentation, and Reintroduction of Freshwater Mussels (Mollusca: Bivalvia: Unionoida). Freshw Mollusk Biol Conserv 20:1. https://doi.org/10.31931/fmbc.v20i1.2017.1-12

Meffe GK, Vrijenhoek RC (1988) Conservation genetics in the management of desert fishes. Conserv Biol 2:157–169

Meirmans PG (2015) Seven common mistakes in population genetics and how to avoid them. Mol Ecol 24:3223–3231

Mobile River Basin Mollusk Restoration Committee (2010) A Plan for the Population Restoration and Conservation of Freshwater Mollusks of the Mobile River Basin

Moses RD, Mathews L, Richard S (2016) Iatt Lake; Lake History and Management Issues

Nei M (1973) Analysis of Gene Diversity in Subdivided Populations. Proc Natl Acad Sci 70:3321–3323. https://doi.org/10.1073/pnas.70.12.3321

Nomura T (2008) Estimation of effective number of breeders from molecular coancestry of single cohort sample. Evol Appl 1:462–474

Pfister CA, McCoy SJ, Wootton JT, et al (2011) Rapid Environmental Change over the Past Decade Revealed by Isotopic Analysis of the California Mussel in the Northeast Pacific. PLOS ONE 6:e25766. https://doi.org/10.1371/journal.pone.0025766

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

QGIS Development Team (2014) QGIS Geographic Information System. Open Source Geospatial Foundation

Richardson JL, Brady SP, Wang IJ, Spear SF (2016) Navigating the pitfalls and promise of landscape genetics. Mol Ecol 25:849–863. https://doi.org/10.1111/mec.13527

Rochette NC, Rivera-Colón AG, Catchen JM (2019) Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. Mol Ecol 28:4737–4754. https://doi.org/10.1111/mec.15253

Roe K (2009) Conservation Genetics of the Freshwater Mussel Margaritifera hembeli (Bivalvia: Margaritiferidae). Final report to USFWS Natchitoches National Fish Hatchery. Natchitoches, LA.

Schwartz M, Luikart G, Waples R (2007) Genetic monitoring as a promising tool for conservation and management. Trends Ecol Evol 22:25–33. https://doi.org/10.1016/j.tree.2006.08.009

Segelbacher G, Cushman SA, Epperson BK, et al (2010) Applications of landscape genetics in conservation biology: concepts and challenges. Conserv Genet 11:375–385. https://doi.org/10.1007/s10592-009-0044-5

Smith DG (1988) Notes on the biology and morphology of Margaritifera hembeli(Conrad, 1838)(Unionacea: Margaritiferidae). Nautilus 102:159–163

Steingraeber MT, Bartsch MR, Kalas JE, Newton TJ (2007) Thermal Criteria for Early Life Stage Development of the Winged Mapleleaf Mussel (Quadrula Fragosa). Am Midl Nat 157:297–311. https://doi.org/10.1674/0003-0031(2007)157[297:TCFELS]2.0.CO;2

Stewart JH (1990) Louisiana Pearlshell (Margaritifera hembeli) Recovery Plan

Strayer DL, Downing JA, Haag WR, et al (2004) Changing Perspectives on Pearly Mussels, North America's Most Imperiled Animals. BioScience 54:429. https://doi.org/10.1641/0006-3568(2004)054[0429:CPOPMN]2.0.CO;2

Strayer DL, Geist J, Haag WR, et al (2019) Essay: Making the most of recent advances in freshwater mussel propagation and restoration. Conserv Sci Pract 1:e53

Sundqvist L, Keenan K, Zackrisson M, et al (2016) Directional genetic differentiation and relative migration. Ecol Evol 6:3461–3475. https://doi.org/10.1002/ece3.2096

Team RC (2019) R: a language and environment for statistical computing, version 3.0. 2. Vienna, Austria: R Foundation for Statistical Computing; 2013

U.S. Fish and Wildlife Service (USFWS) (1988) Endangered and threatened wildlife and plants; final endangered status for the Louisiana pearlshell *Margaritifera hembeli*. Federal Register 53(24): 3567-3569.

U.S. Fish and Wildlife Service (USFWS) (1993) Endangered and Threatened Wildlife and Plants: Determination to Reclassify the Louisiana Pearlshell (*Margaritifera hembeli*) from Endangered to Threatened. Federal Register 58(184): 49935-49937.

Vaughn CC, Hakenkamp CC (2001) The functional role of burrowing bivalves in freshwater ecosystems. Freshw Biol 46:1431–1446. https://doi.org/10.1046/j.1365-2427.2001.00771.x

Wächtler K, Dreher-Mansur MC, Richter T (2001) Larval Types and Early Postlarval Biology in Naiads (Unionoida). In: Bauer G, Wächtler K (eds) Ecology and Evolution of the Freshwater Mussels Unionoida. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 93–125

Wacker S, Larsen BM, Jakobsen P, Karlsson S (2018) High levels of multiple paternity in a spermcast mating freshwater mussel. Ecol Evol 8:8126–8134. https://doi.org/10.1002/ece3.4201

Wacker S, Larsen BM, Jakobsen P, Karlsson S (2019) Multiple paternity promotes genetic diversity in captive breeding of a freshwater mussel. Glob Ecol Conserv 17:e00564. https://doi.org/10.1016/j.gecco.2019.e00564

Wang J, Santiago E, Caballero A (2016) Prediction and estimation of effective population size. Heredity 117:193–206. https://doi.org/10.1038/hdy.2016.43

Waples RK, Larson WA, Waples RS (2016) Estimating contemporary effective population size in non-model species using linkage disequilibrium across thousands of loci. Heredity 117:233–240. https://doi.org/10.1038/hdy.2016.60

Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci*. Conserv Genet 7:167. https://doi.org/10.1007/s10592-005-9100-y

Waples RS, Antao T, Luikart G (2014) Effects of Overlapping Generations on Linkage Disequilibrium Estimates of Effective Population Size. Genetics 197:769–780. https://doi.org/10.1534/genetics.114.164822

Watters GT (1996) Small dams as barriers to freshwater mussels (Bivalvia, Unionoida) and their hosts. Biol Conserv 75:79–85. https://doi.org/10.1016/0006-3207(95)00034-8

Whelan NV, Galaska MP, Sipley BN, et al (2019) Riverscape genetic variation, migration patterns, and morphological variation of the threatened Round Rocksnail, Leptoxis ampla. Mol Ecol 28:1593–1610

Williams JD, Warren ML, Cummings KS, et al (1993) Conservation Status of Freshwater Mussels of the United States and Canada. 18:17

Witzenberger KA, Hochkirch A (2011) Ex situ conservation genetics: a review of molecular studies on the genetic consequences of captive breeding programmes for endangered animal species. Biodivers Conserv 20:1843–1861. https://doi.org/10.1007/s10531-011-0074-4

Woodings LN, Murphy NP, Doyle SR, et al (2018) Outlier SNPs detect weak regional structure against a background of genetic homogeneity in the Eastern Rock Lobster, Sagmariasus verreauxi. Mar Biol 165:185. https://doi.org/10.1007/s00227-018-3443-7

Table 1 Pairwise FST (top) and estimated in-stream distance in km (bottom) for all populations analyzed.

	Jordan	**Black**	**Loving**	**Brown**	**Captive**
Jordan	-	0.043	0.063	0.078	0.048
Black	36.5	-	0.064	0.082	0.035
Loving	77.1	91.4	-	0.063	0.071
Brown	87.8	102.5	41.2	-	0.086
Captive	36.5	0	91.4	102.5	-

888
889

Table 2 Population summary statistics (in part, full summary available online) from left to right; number of individuals (N), number of loci recovered (Loci), number of private alleles (PA), nucleotide diversity (Π) and standard deviation (SD), observed heterozygosity (H_o), expected heterozygosity (H_e), coefficient of inbreeding (F_{IS}), effective populaiton size (N_e) and 95% confidence interval (CI), number of effective breeders (N_b), and allelic richness (AR).

890

Population	N	Loci	PA	Π (SD)	H_o (SD)	H_e (SD)	F_{IS} (SD)	N_e (CI)	N_b	AR (CI)
Jordan	18	2134	91	0.25 (±0.04)	0.10 (±0.02)	0.24 (±0.03)	0.45 (±0.17)	INF	4.2	1.47 (1.41-1.51)
Black	20	2203	88	0.25 (±0.04)	0.10 (±0.02)	0.24 (±0.04)	0.43 (±0.16)	205 (188-225)	3.9	1.51 (1.46-1.55)
North	38	--	--	--	--	--	--	177 (170-183)	3.9	--
Loving	16	1672	40	0.27 (±0.04)	0.09 (±0.02)	0.25 (±0.04)	0.46 (±0.17)	INF	4.3	1.16 (1.12-1.19)
Brown	17	1941	142	0.27 (±0.04)	0.09 (±0.01)	0.26 (±0.03)	0.49 (±0.17)	INF	3.9	1.35 (1.30-1.38)
South	33	--	--	--	--	--	--	488 (429-566)	4.1	--
Captive	19	2024	41	0.25 (±0.04)	0.10 (±0.02)	0.24 (±0.04)	0.40 (±0.16)	70 (64-78)	5.0	1.37 (1.33-1.41)

891
892
893
894
895
896
897
898
899
900

Table 3 Description of models assessed by migrate-n, sorted by model rank; model numbers correspond to those in Fig. 2.

Number	Model Description	Log marginal likelihood value	Rank
3	**Northern panmixia, bidirectional South**	**-93595.56**	**1**
4	Northern panmixia, full migration	-93803.10	2
1	Full migration	-94306.55	3
5	Panmixia	-94030.00	4
2	Northern panmixia, unidirectional South	-94612.37	5

Fig. 1 Map depicting the extent of M. hembeli distribution in central Louisiana. Pictures illustrate mussel bed density at Black Creek site (upper) and individual shell (lower).

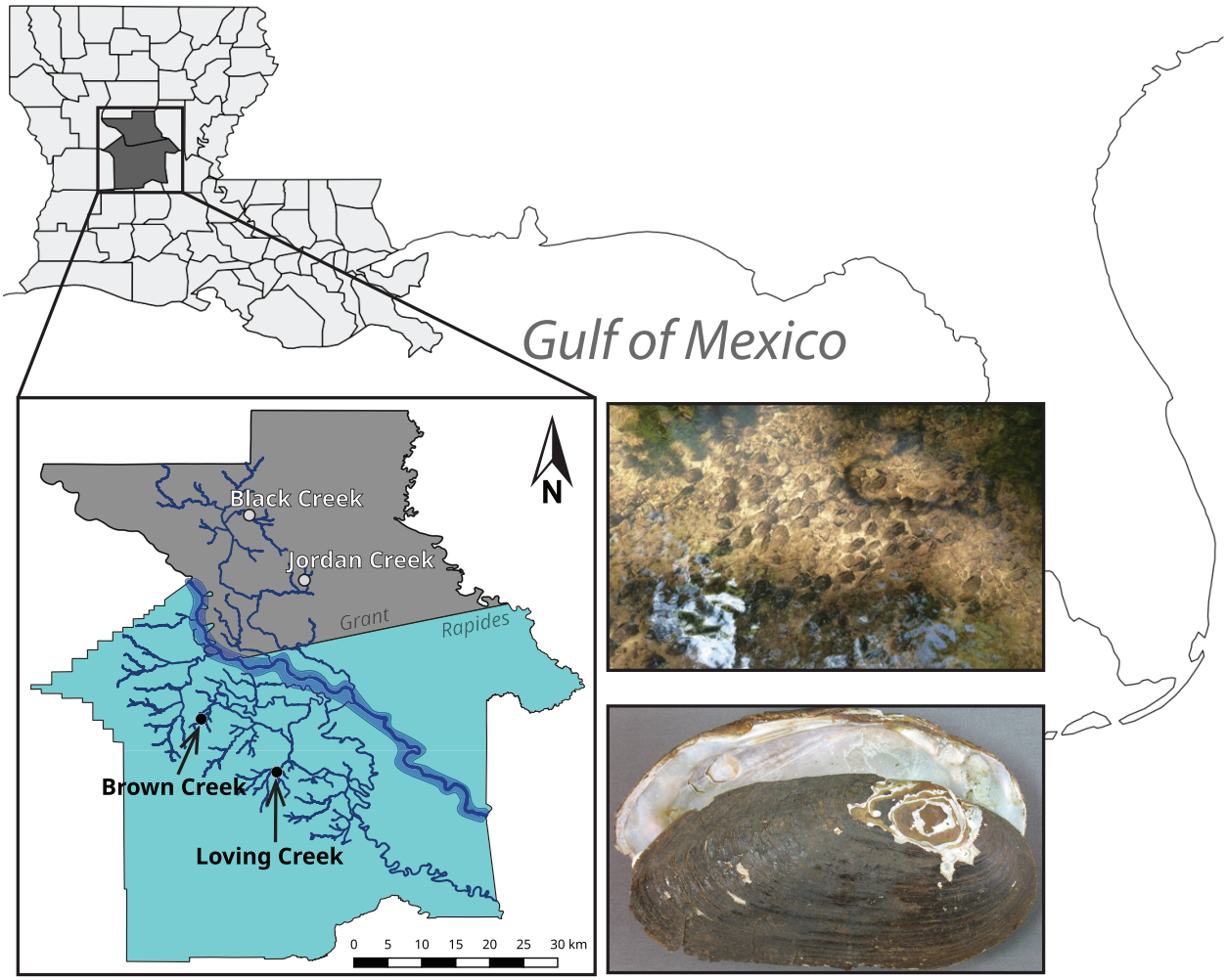

Fig. 2 Population migration models assessed using migrate-N.

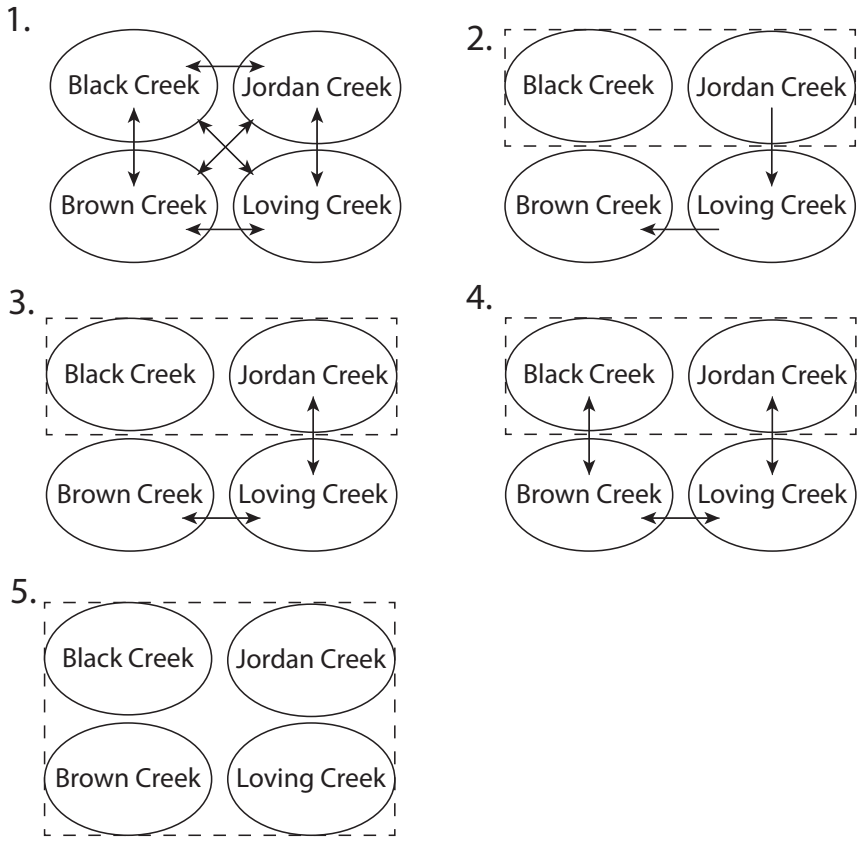

Fig. 3 Discriminant analysis of principal components output showing clustering of a) all wild populations and b) wild + captive individuals using the multiple SNP per locus datasets.

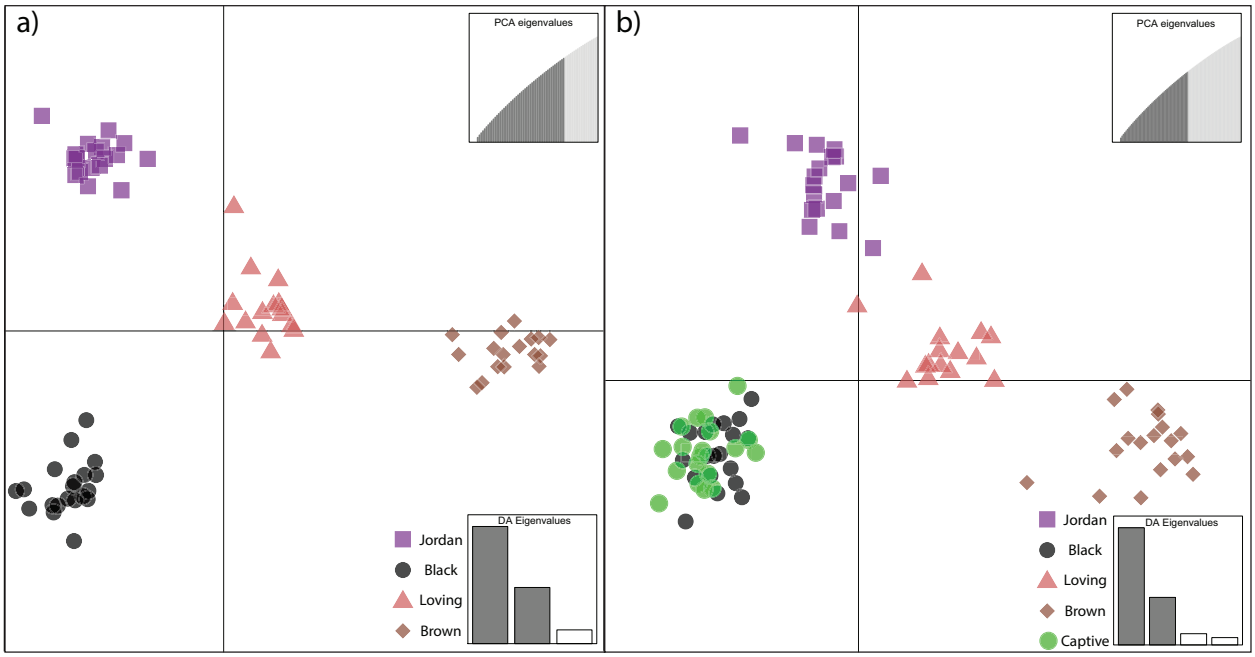

Fig. 4 Individual admixtures for all wild individuals as inferred by LEA analysis, using the best-fit K of 2. Colors correspond to parish designations depicted in Fig. 1.

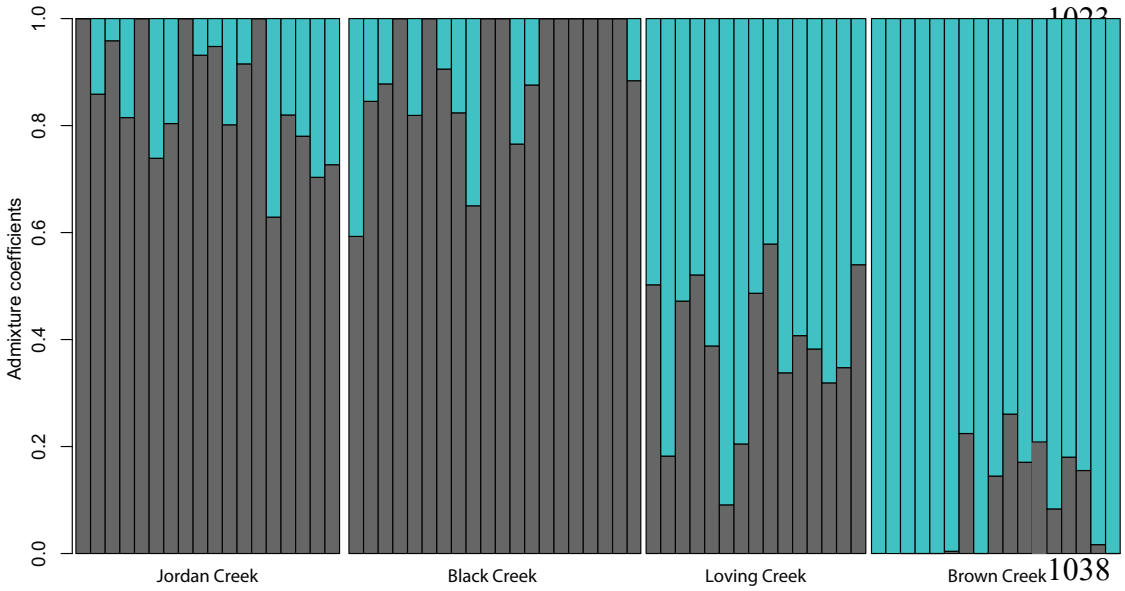

Fig. 5 Hierarchical heatmap generated by fineRADstructure, all individuals included. Letters correspond to population identifiers (A = Jordan Creek, B = Black Creek/Captive cohort, C = Brown Creek, D = Loving Creek). Boxes highlight patterns of population clustering, colors represent relative co-ancestry values averaged by population.

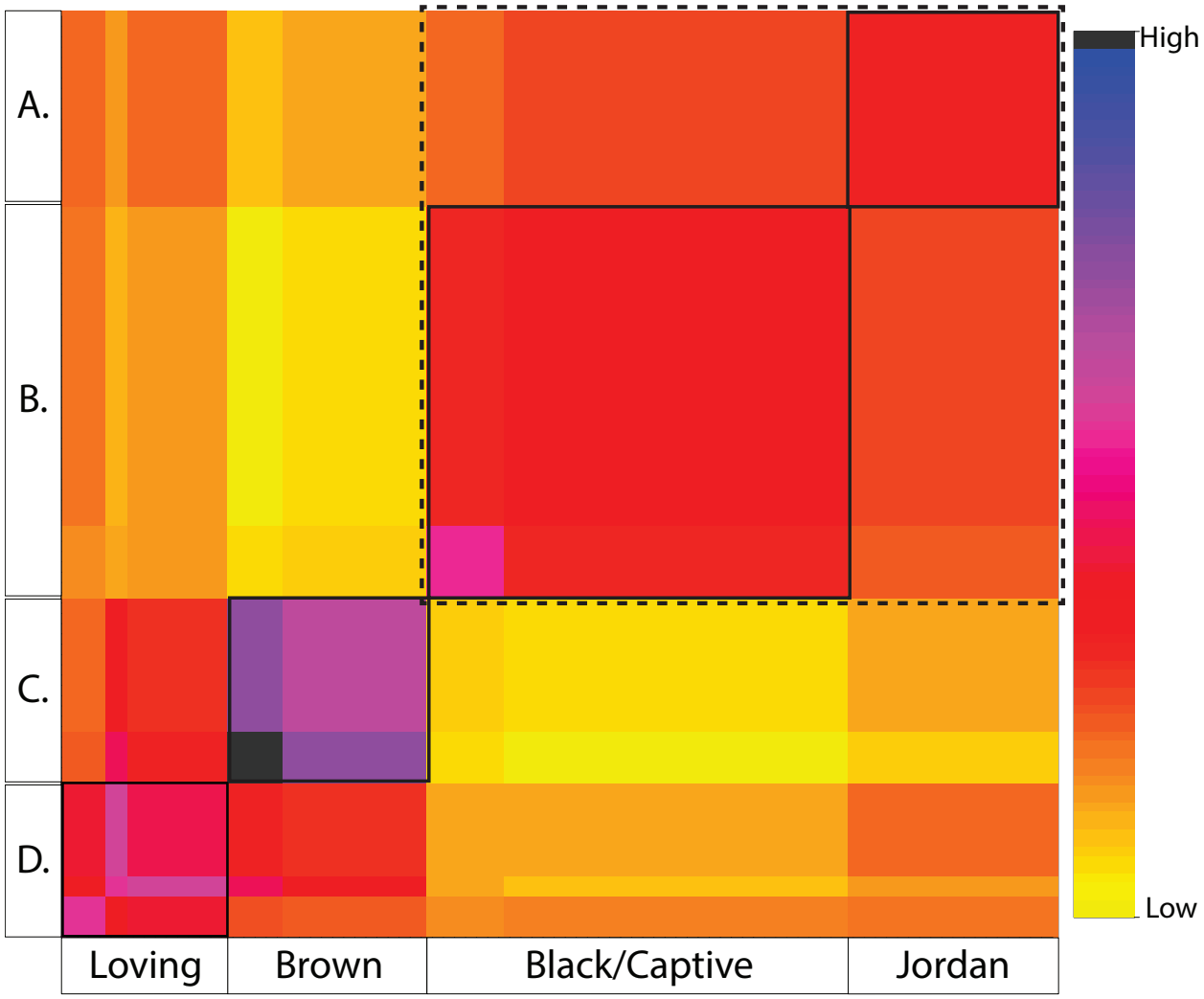

Fig. 6 Hierarchical heatmap generated by fineRADstructure, only captive individuals. Boxes highlight patterns of individual clustering, colors represent relative co-ancestry values averaged by cluster.

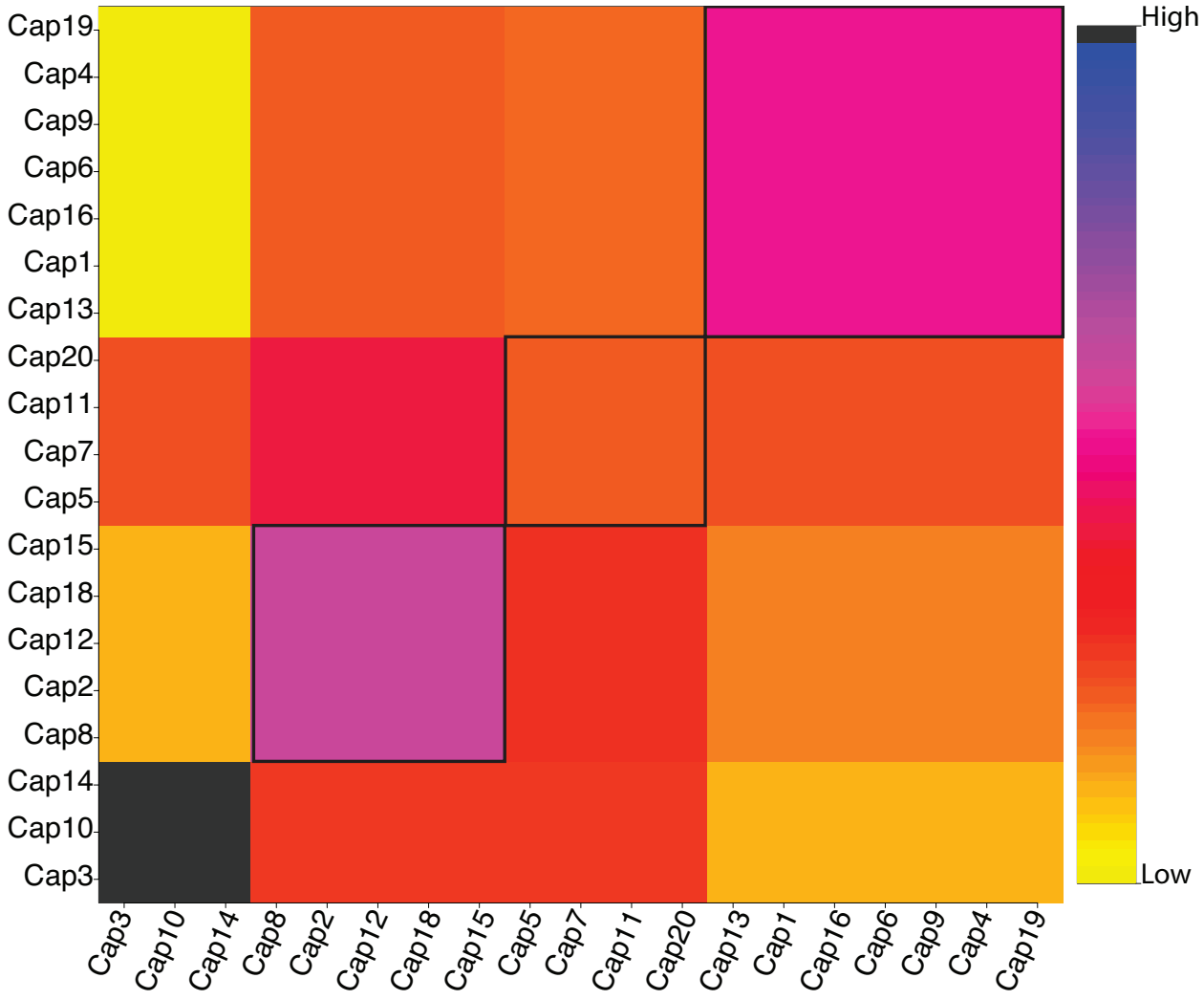
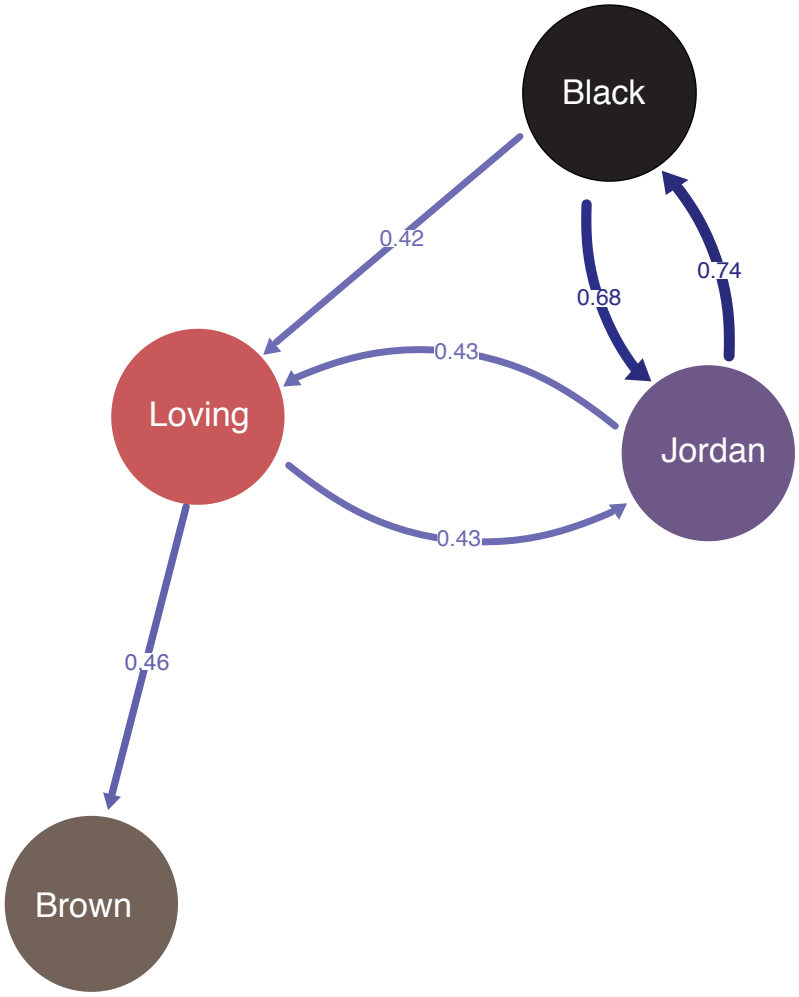
Fig. 7 Network generated by divMigrateOnline using the WithCaptiveS dataset; GST statistic, 0.40 filtering level, alpha 0.05. Colors correspond to those in Fig. 3.