Range reduction of the Oblong Rocksnail, *Leptoxis compacta*, shapes riverscape genetic patterns

Aaliyah D. Wright[1], Nicole L. Garrison[2], Ashantye' S. Williams[3], Paul D. Johnson[4], Nathan V. Whelan[2,3]

[1]Tuskegee University, Tuskegee, Alabama, USA

[2]School of Fisheries, Aquaculture, and Aquatic Sciences, College of Agriculture, Auburn University, Auburn, Alabama USA

[3]Southeast Conservation Genetics Lab, Warm Springs Fish Technology Center, United States Fish and Wildlife Service, Auburn, Alabama, USA

[4]Alabama Aquatic Biodiversity Center, Alabama Department of Conservation and Natural Resources, Marion, Alabama, USA.

Corresponding Authors:
Aaliyah D. Wright[1]
Tuskegee University, Tuskegee, Alabama, USA
Email address: awright1863@tuskegee.edu

Nathan V. Whelan[2,3]
203 Swingle Hall, Auburn, AL 36849, USA
Email address: nathan_whelan@fws.gov

Keywords: Pleuroceridae, conservation genomics, Cahaba River, Alabama, narrow range endemic, water quality, habitat suitability.

Abstract

Many freshwater gastropod species face extinction, including 79% of species in the family Pleuroceridae. The Oblong Rocksnail, *Leptoxis compacta*, is a narrow range endemic pleurocerid from the Cahaba River basin in central Alabama that has seen rapid range contraction in the last 100 years. Such a decline is expected to negatively affect genetic diversity in the species. However, precise patterns of genetic variation and gene flow across the restricted range of *L. compacta* are unknown. This lack of information limits our understanding of human-altered freshwater systems. Here, we report a range extension for *L. compacta*, and show that the species has likely seen a species-wide decline in genetic diversity. Yet, remaining populations have relatively high genetic diversity. We also demonstrate that *L. compacta* does not display an isolation by distance pattern, contrasting patterns seen in most riverine taxa. Our findings indicate that range contraction has eliminated expected genetic patterns as the small distribution of *L. compacta* allows for relatively unrestricted gene flow across its remaining range despite limited dispersal abilities. Two collection sites had higher genetic diversity than others, and broodstock sites for future captive propagation and reintroduction efforts should utilize sites identified here as having the highest genetic diversity. Broadly, our results support the hypothesis that range contraction will result in the elimination of species-wide genetic diversity, and common riverscape genetic patterns cannot be assumed to be present in species facing extinction risk.

Introduction

Freshwater gastropods of the United States suffer one of the highest imperilment rates of any taxonomic group in North America (Johnson et al., 2013). Despite being critical components of many freshwater ecosystems, freshwater gastropods are grossly understudied compared to freshwater fish, mussels, and crayfish (Covich, Palmer & Crowl, 1999; Huryn, Benke & Ward, 1995; Strong et al., 2008). This creates a situation where desperately needed conservation efforts are hindered by a lack of information (Johnson et al., 2013). For example, data on the current range of many freshwater gastropods is lacking (Lydeard et al., 2004), but conservation assessments and effective management plans require detailed historical and contemporary range data (Potter & Thomas, 1983; NMFS, 2018). Population genetics data on freshwater gastropods are also needed to inform management efforts and provide basic understanding of freshwater ecosystems (Lysne et al., 2008).

The freshwater gastropod family Pleuroceridae is one group that suffers from a high imperilment rate (79%) and little research attention (Brown, Lang & Perez, 2008; Johnson et al., 2013; Perez & Minton, 2008). Pleurocerids are found east of the Rocky Mountains in North America, with most of their diversity concentrated in the southeastern United States (Lydeard & Mayden, 1995; Strong & Köhler, 2009). Pleurocerids lack a highly vagile veliger larval stage seen in many aquatic gastropod groups, and they are thought to move large distances only when washed downstream (Whelan et al., 2019; Whelan, Johnson & Harris, 2015). Only one study has been published on landscape and conservation genetics of pleurocerids, and that study focused exclusively on a single species, *Leptoxis ampla* (Whelan et al., 2019). Many freshwater species, including *L. ampla*, display common riverscape genetic patterns such as increased genetic diversity in downstream populations and isolation by distance (Hughes, Schmidt & Finn, 2009; Paz-Vinas et al., 2015). However, few studies have tested for such patterns in riverine species that have undergone drastic range reduction, and no such study has been done for a range restricted pleurocerid. Thus, more research is needed.

One pleurocerid in desperate need of more research is the Oblong Rocksnail, *Leptoxis compacta* (Figs. 1,2). This species is a narrow range endemic known historically from the middle Cahaba River and a single tributary in central Alabama, USA (Fig. 2; Goodrich, 1922). Until recently, *Leptoxis compacta* was considered extinct as it had not been collected, or at least identified correctly, from 1935 to 2011 (Goodrich, 1941; Johnson et al., 2006; Whelan, Johnson & Harris, 2012). As early as 1941, the decline of *L. compacta* was documented (Goodrich, 1941), and the species now occupies less than 5% of its historical range (Fig. 3; Whelan, Johnson & Harris, 2012). As a narrow range endemic with few historic collections, little is known about the species aside from recent survey efforts and limited life history data (Whelan, Johnson & Harris, 2012). Yet, the rediscovery of *L. compacta* in 2011 resulted in an emergency petition to list the species under the US Endangered Species Act (Kurth, 2017). For management agencies to assess the status of *L. compacta* and design effective conservation plans, detailed survey work and population genetics research are required. Modern population genomic tools such as restriction site associated DNA sequencing (RAD-seq) can provide data that will enhance *L. compacta* management options (Andrews et al., 2016). As a result of having a narrow range along a single river path, an effective recovery strategy for *L. compacta* will likely require reintroduction efforts to previously occupied habitat(s). Such a management effort will require detailed population genetics data to inform broodstock selection for maximizing heterogeneity of captively reared offspring.

In this study, we used 2bRAD-seq (Wang et al., 2012) to answer questions about population and riverscape genetics of *L. compacta* with thousands of single nucleotide polymorphisms (SNPs). Given the drastic range decline suffered by *L. compacta*, we set out to test the following hypotheses: 1) *Leptoxis compacta* has undergone a severe genetic bottleneck and 2) genetic diversity of *L. compacta* is considerably lower than *L. ampla*, a sympatric and wider ranging species. We also examined how genetic diversity of *L. compacta* varies across its current range, specifically assessing whether broad patterns seen in many other riverine taxa like isolation by distance and strong genetic structure are seen in *L. compacta*.

Materials & Methods

Sample Collection

Leptoxis compacta was collected during two trips to the Cahaba River in June 2018 and June 2019. Specimens were collected from four sites (Fig. 3). These sites span the known range of *L. compacta*, including three sites where Whelan et al. (2012) failed to collect the species. One newly identified site was below the 2011 rediscovery site and two sites were above the 2011 rediscovery site. At each location, individuals were collected by hand and identified in the field. Despite being a narrow range endemic that has undergone distributional decline, *L. compacta* was locally abundant where found, and we are confident that lethal take was not detrimental to species survival. Twenty specimens from each site were transported live to the lab, sacrificed following Fukuda et al. (2008), and placed in 96-100% ethanol until tissue clips could be taken. Specimens were collected under an Alabama Department of Conservation and Natural Resources Educational Scientific Collections Permit (License #2019100990068680) or as an agent of the state (P.D. Johnson). All shells have been cataloged separately to correspond to associated molecular data and deposited at the Auburn University Museum of Natural History (AUMNH #####-#### [to be provided upon manuscript acceptance]; Table 1).

Molecular data generation

Tissue clips from 20 individuals per collection site were taken and subjected to a standard proteinase K digestion. DNA was extracted with the Qiagen DNeasy Plant Mini Kit with a minor modification to allow for tissue digested with proteinase K. We used a plant kit because it works well on freshwater gastropods that produce large amounts of mucus polysaccharides (Whelan et al., 2019). The integrity of whole genomic DNA was checked on a 1% agarose gel and quantified with a Qubit DNA assay. Extracted DNA was standardized to 120ng/µL for 2bRAD library prep.

A reduced representation genomic library was generated for genotyping using the AlfI enzyme and the 2bRAD library prep protocol of Wang et al., (2012). Following Whelan et al. (2019), we did a $1/16^{th}$ genomic reduction. All samples were dual-indexed for pooling. Sequencing occurred in multiple batches. The first batch had 48 *L. compacta* samples pooled in equimolar concentrations and sequenced on a single lane. The other individuals were pooled in equimolar concentrations with samples from other projects, and 87 individuals were sequenced per HiSeq 4000 lane. Although batch effects in RADseq data have been recently noted in studies that used different read lengths among sequencing runs (Leigh, Lischer & Keller, 2018) and in species introgression studies (Lambert et al., 2019), such issues were not relevant to our sequencing design or study objectives. Nevertheless, we took steps to limit potential batch effects by implementing strict filtering parameters during dataset assembly (see below). Pooled

libraries were sequenced on an Illumina HiSeq 4000 with 1 X 75bp chemistry at University of Oregon Genomics and Cell Characterization Core Facility.

Raw Illumina reads were demultiplexed with the STACKS 1.48 (Catchen et al., 2013) module *process_radtags*, allowing for one mismatch per barcode. Demultiplexed reads were quality filtered with the script QualFilterFastq.pl (http://github.com/Eli-Meyer/sequence_processing) for any read that had five or more base pairs with Phred quality scores less than 20. Loci were processed with scripts from SHRiMP 2.23 (Rumble et al., 2009) and subsequently trimmed to AlfI restriction enzyme recognition sites with the script AlfIExtract.pl (http://github.com/Eli-Meyer/2bRAD_utilities), resulting in removal of adaptor sequences and non-target sequences. Loci were assembled with the STACKS 1.48 pipeline *denovo_map.pl* as no reference genome is available for *L. compacta*. For *denovo_map.pl* parameters, we set minimum stack depth to five (-m 5), distance allowed between stacks to three (-M 3), and distance between catalog loci to two (-n 2). These parameters were determined to be most appropriate for our data following Paris et al. (2017). All other *denovo_map.pl* parameters were set to defaults.

After assembly, loci were filtered for missing data using the STACKS program *populations*. In order to pass filtering steps, a locus had to be present in 75% of individuals from any given collection site and also present at three collection sites. Loci that had a minimum minor allele frequency of less than 2.5% or heterozygosity higher than 50% were removed to limit the influence of paralogy and misassembly on final datasets. After filtering, a dataset that included all SNPs per locus and a dataset with only one random SNP per locus were generated as most downstream analyses assume unlinked loci. Analyses employed the one SNP per locus dataset unless otherwise noted.

Population genetics analyses

Average observed heterozygosity (H_o), expected heterozygosity (H_e), and nucleotide diversity (Π) at each collection site were calculated by *populations*. The number of private alleles at each site was also reported by *populations*. Average allelic richness (A_r) was calculated with the R (R Core Developent Team, 2020) package diveRsity (Keenan et al., 2013). An analysis of molecular variance (AMOVA; Excoffier, Smouse & Quattro, 1992) was done with the R packages poppr (Kamvar, Tabima & Grünwald, 2014) to test genetic structure among collection sites. AMOVA was implemented with the function "poppr.amova" using the ade4 method (Dray & Dufour, 2007) and 10,000 permutations.

We also tested for a pattern of isolation by distance by measuring the correlation between pairwise F_{ST} values and geographical distance between collection sites. Pairwise F_{ST} values were calculated using the Weir and Cockerham (1984) method with the R package hierfstat (Goudet, 2005). Stream distance was measured in Google Earth by tracing paths between collection sites along the Cahaba River (see Table 2). River distance was used, rather than straight line distances, because migration over land is impossible for gill breathing pleurocerids. A Mantel test was performed with the R package ade4 (Dray & Dufour, 2007), and significance was tested with 1,000 permutations. However, Mantel tests have been criticized as a method for testing isolation by distance (Legendre, Fortin & Borcard, 2015; Meirmans, 2015), so we also performed a multiple regression on distance matrices with 1,000 permutations using the R package ecodist and its MRM function (Goslee & Urban, 2007). In addition to a pattern of isolation by distance, past studies have shown that many freshwater organisms, including pleurocerids, display a pattern of increased genetic diversity in more downstream populations (Paz-Vinas et al., 2015;

Whelan et al., 2019). Therefore, to better assess riverscape genetic patterns of *L. compacta*, we performed linear regression of distance from the most downstream site against H_o, H_e, A_r, and Π. Linear regressions were done in R.

We inferred genomic admixture of *L. compacta* individuals with ADMIXTURE 1.3 (Shringarpure et al., 2016). To determine the best-fit number of clusters (K) for our data, K values from 1 to 5 were assessed with 20% cross-validation. Twenty replicates of ADMIXTURE were run at each K, and the best-fit K was determined as the value that had the lowest average CV score across replicates. ADMIXTURE results were visualized with Clumpak (Kopelman et al., 2015).

Genomic co-ancestry among individuals was also assessed with fineRADstructure (Malinsky et al., 2018). Unlike ADMIXTURE, fineRADstructure can use linked SNPs and provides additional information on individual genomic background. Thus, the multiple SNPs per locus dataset was used for fineRADstructure analyses. First, a co-ancestry matrix was inferred with the script RADpainter. Subsequently, clustering was done with the Markov chain Monte Carlo method of fineRADstructure, running for 500,000 generations and sampling every 1,000 generations; the first 200,000 generations were discarded as burn-in (non-default parameters: -x 200000 -y 300000 -z 1000). We also inferred a tree for visualization with fineRADstructure using the tree-building algorithm of Lawson et al. (2012) with 10,000 attempts (non-default parameters: -m T -x 10000). fineRADstructure results were plotted with R scripts included in the fineRADstructure package.

Code and data availability

All bash and R scripts used for processing and analyzing data are available at github.com/nathanwhelan. Demultiplexed raw Illumina reads have been uploaded to NCBI under BioProject PRJNA631794. Assembled datasets in various file formats (e.g., vcf, genepop) and the 2bRAD library prep protocol are available on FigShare.

Results

Sample Collection

During survey work, we collected *L. compacta* from the Cahaba River at old Marvel slab upstream to Cahaba River at Booth's Ford (Fig. 3). This represents a 1.83 km downstream range extension and a 4.76 km upstream extension compared to the previously documented range of *L. compacta* (Whelan, Johnson & Harris, 2012). While this study was ongoing, 3 putative *L. compacta* individuals were collected at Cahaba River at Belle Ellen Shoals (Fig. 3) during a general mollusk survey (Johnson, 2019). However, species identification was uncertain and *L. compacta* appeared exceedingly rare. Therefore, individuals from Cahaba River at Belle Ellen Shoals were not included in our analyses, and we consider this record unconfirmed without additional positive survey results. At all four collection sites (Fig. 3), *L. compacta* could be found with relative ease, and we collected 20 individuals at each site. Based upon ease of collection, *L. compacta* appeared most abundant at Cahaba River above Shades Creek.

Molecular data and population genetics

DNA yields for two individuals were too low for library preparation so only 19 individuals were sequenced from Cahaba River at old Marvel Slab and Cahaba River at Booth's Ford. The number of demultiplexed raw reads per individuals varied from 930,062-10,146,649 (mean = 4,836,812). Much of the variation in raw reads can be attributed to whether the

individual was sequenced on a HiSeq 4000 lane with 48 or 87 samples. Aside from raw read number, we saw no evidence of batch affects. After initial raw-read filtering, the number of reads that passed quality filtering steps ranged from 865,314-9,838,187 (mean = 4,632,510). After assembly with STACKS and filtering with *populations*, the dataset consisted of 9,016 loci. Of those loci, 4,009 were invariant among all taxa, resulting in a final dataset of 5,007 loci with at least one SNP.

The number of private alleles at each site ranged from 28-262 (Table 1). Average observed heterozygosity at each collection site ranged from 0.0963-0.1568, and average expected heterozygosity ranged from 0.0980 to 0.1801 (Table 1). At each site, observed heterozygosity was lower than expected heterozygosity, except at Cahaba River at Boat Launch where observed heterozygosity was 0.001 greater than expected heterozygosity (Table 1). Allelic richness and nucleotide diversity ranged from 1.4511-1.8241 and 0.1010-0.1829, respectively (Table 1). Overall, genetic diversity was greatest at the most upstream site, Cahaba River at Booth's Ford, and lowest at the most downstream site, Cahaba River at old Marvel slab. All linear regressions of diversity statistics vs distance from the most downstream site were non-significant ($p \geq 0.169$). Nevertheless, we observed a slight trend of less genetic diversity in more downstream sites.

Pairwise F_{ST} values among sites ranged from 0.0-0.055 (Table 2). We found no evidence of an isolation by distance pattern among sites (Mantel test, $p = 0.843$; multiple regression, $p = 0.428$). According to the AMOVA, significant genetic structure was present among collection sites ($p = 0.004$), but only 4.16% of variation was explained by collection site. In contrast, 81.8% of genetic variation was explained by within individual variation, further indicating high amounts of gene flow among collection sites.

ADMIXTURE analyses indicated that genetic diversity from two ancestral populations were present in our data ($K = 2$). Most individuals from across the range of *L. compacta* had a genomic admixture profile that was dominated by a genomic background from a single ancestral population, likely indicating that overall genomic diversity has been lost across the range of *L. compacta*. Nevertheless, 14 individuals had varying levels of admixture with a second ancestral population (Fig. 3). fineRADstructure analyses generally corroborated ADMIXTURE analyses as two semi-distinct groupings were recovered (Fig. 4). fineRADstructure groupings did not correspond to collection site or any other obvious variable, indicating gene flow among collection sites. Notably, six individuals with comparably high co-ancestry proportions (upper right of co-ancestry matrix in Fig. 4) correspond to individuals in ADMIXTURE analyses with a large proportion of genetic background from the less common ancestral population (represented by orange in Fig. 3). Given the well-documented decline of *L. compacta*, a small number of individuals with a less common genomic background suggests that the species has lost genetic diversity through bottleneck and drift.

Discussion

Our findings provide reasons to be optimistic about the survival of *L. compacta*. We found *L. compacta* more widespread than previously known, and the species retains a relatively high amount of genetic diversity. Across its range, *L. compacta* had similar levels of observed heterozygosity and nucleotide diversity to *L. ampla* (Whelan et al., 2019) with the lowest *L. compacta* values being greater than the lowest values determined for *L. ampla*. This observation rejects one of our main hypotheses that *L. compacta* would have lower genetic diversity than the more widespread *L. ampla*. Nevertheless, *L. compacta* is restricted to a 9.2 km stretch of river,

and *L. compacta* has likely lost range-wide genetic diversity. This probable loss of evolutionary potential could be detrimental to the long-term survival of the species.

Observed *L. compacta* genetic patterns often conflicted with predictions made by broad-scale hypotheses about riverscape genetics. For example, we did not see an isolation by distance pattern, which is common among freshwater taxa (Hughes, Schmidt & Finn, 2009) and was documented in *L. ampla* (Whelan et al., 2019). We also did not uncover a pattern of increased genetic diversity in downstream populations, despite such a pattern being present in numerous plants and animals (Paz-Vinas et al., 2015), including *L. ampla* (Whelan et al., 2019). Patterns determined for *L. compacta* are likely explained by a drastic range reduction and the limited scale at which we performed the current study. That is, gene flow across the 9.2 km contemporary range of *L. compacta*, likely explains observed patterns of riverscape genetic diversity. Other factors such as small-scale habitat suitability differences may also help explain differences in genetic diversity among sites separated by less than 10 km.

Genetic diversity across a small landscape

The two most distant populations in this study were separated by a smaller distance (9.2 km) than all but two sites sampled for *L. ampla* in a previous study (Whelan et al., 2019). Therefore, it is difficult to make direct comparisons between genetic patterns across the two *Leptoxis* species. However, we can leverage differences in geographical scale between the two studies to make inferences about fine-scale versus long-distance genetic patterns in pleurocerids. F_{ST} values among *L. compacta* collection sites (Table 2) were much lower than values determined for populations of *L. ampla* (F_{ST} 0.377-0.773; Whelan et al. 2019). Furthermore, even though AMOVA indicated significant genetic structure among *L. compacta* collection sites, the small amount of genetic variation that is explained by collection site casts doubt on its biological relevance. Overall, these data indicate that pleurocerid riverscape genetic patterns across small distances will not always follow common patterns such as isolation by distance and increased genetic diversity at more downstream collection sites. This is likely attributable to gene flow and random drift that prevent the establishment of genetic patterns that are typically seen across more geographically separated collection sites. From a historical standpoint, we hypothesize that *L. compacta* previously displayed an isolation by distance pattern across its range, similar to the patterns determined for *L. ampla* (Whelan et al., 2019). We think this scenario is likely given limited dispersal abilities of pleurocerids and patterns established for *L. ampla*, a species that retains a much larger portion of its historic range in the Cahaba River drainage than *L. compacta*. Whether or not there was a similar historical pattern of increased genetic diversity in downstream populations of *L. compacta* is more difficult to infer, as such a pattern may not be influenced solely by dispersal ability. Nevertheless, historical *L. compacta* population sizes were likely higher at downstream sites as the physical size of the Cahaba River, and therefore available habitat, is larger as the river flows downstream. If so, larger downstream populations would have likely contained greater genetic diversity compared to upstream populations.

ADMXITURE analyses indicated two ancestral populations, but the two sites with individuals that had admixture profiles containing a signature from the less common ancestral population were not adjacent (Fig. 3). A similar pattern was resolved with co-ancestry proportions inferred by fineRADstructure as groupings were not driven by geography (Fig. 4). Although individuals with some admixture from the uncommon ancestral population may be present in unsampled individuals at Cahaba River at old Marvel slab and Cahaba River at canoe

launch, they would be uncommon. Recent migration as the primary driver of observed co-ancestry profiles would indicate that a sizeable population of *L. compacta* exists elsewhere in the Cahaba River. If an unsampled population of *L. compacta* exists, migration would likely be a rare event, even if *L. compacta* has higher dispersal abilities than *L. ampla*. Therefore, the most likely hypothesis for explaining observed co-ancestry profiles (Figs. 3, 4) is a genetic bottleneck resulting from species decline in the 20th century. In this scenario, *L. compacta* was genetically diverse across its historical range prior to decline, but range contraction caused a considerable loss of genetic diversity. In turn, genetic drift resulted in the observed admixture pattern of one ancestral population being more common in extant individuals (Fig. 3). Similarly, the high co-ancestry inferred by fineRADstructure for a small number of individuals likely results from only a few individuals maintaining a genomic signature that has been mostly lost to a genetic bottleneck and drift. The bottleneck and drift hypothesis is corroborated by observed heterozygosity being considerably lower than expected heterozygosity at the two most genetically diverse sites (Table 1). Furthermore, nearly identical observed and expected heterozygosity values at the other two sites may be a result of recent movement into previously unsuitable habitat (Table 1).

For a narrow range endemic that has seen a range reduction of over 95%, differences in habitat suitability across small distances would not be surprising. For instance, small pockets of ideal habitat may have served as a refugia for *L. compacta* during range contraction in the 20th century. Strong anthropogenic influence on the Cahaba River over the last 50 years caused poor water quality and habitat degradation in many parts of the upper Cahaba River drainage (Onorato, Angus & Marion, 2000; Pitt, 2000; Shepard et al., 1994; Tolley-Jordan, Huryn & Bogan, 2015). This begs the question: how did *L. compacta* survive within its current range, while it became extirpated in other locations? We speculate that the position of Cahaba River above Shades Creek provided a refugia where pollution may have been lowest across the historical range of *L. compacta*. In this scenario, the refugia persisted because the Cahaba River above Shades Creek site is far downstream from pollution sources in the Birmingham, Alabama metropolitan area but above Shades Creek, which was likely a conduit of pollution into the Cahaba River and is still considered to have poor water quality (ADEM, 2007; ADEM, 2012). Recent efforts have been made to improve the water quality in the Shades Creek drainage (Friends of Shades Creek, 2020) and monitoring by the Alabama Department of Environmental Management reported slightly improved water quality between 2007 and 2012 (ADEM, 2007; ADEM, 2012). These water quality improvements potentially resulted in newly suitable habitat for *L. compacta* in the Cahaba River below the Shades Creek confluence, including the sampled site Cahaba River at old Marvel slab. Future study on habitat preference and fine-scale habitat characterization of the Cahaba River drainage will be necessary to fully understand factors influencing the historical and current range of *L. compacta*.

Conservation of Leptoxis compacta

Leptoxis compacta suffered a massive decline during the 20th century, a period of intense mining, forestry, and urban development in the Cahaba River drainage. The decline was so drastic that *L. compacta* was considered extinct less than a decade ago. Fortunately, *L. compacta* is not extinct. However, conservation efforts are needed to ensure the long-term survival of *L. compacta* as the species is at risk from both chronic habitat degradation and one-time catastrophic events. Two potential management strategies for *L. compacta* are habitat restoration and reintroduction with captively reared individuals.

In this study, we report an 8.26 km known range extension for *L. compacta*. One site, Cahaba River at old Marvel Slab, was previously the focus of intense habitat restoration through the removal of a low-level dam (Johnson et al. 2013). The site may have also benefited from improved water quality in Shades Creek as the site is just below its confluence with the Cahaba River. Since removal of the low-level dam, increases in fish abundance and diversity have been reported (Bennett et al., 2015). Considering *L. compacta* was not found at this site by Whelan, Johnson & Harris (2012), we think habitat either improved from a point where *L. compacta* could not survive or from a point of considerably lower carrying capacity. As the only undammed, major river in the southeastern United States, the Cahaba River is much less modified than most other systems in the southeast. Our findings suggest that imperiled gastropods will benefit from water quality and habitat improvements even in relatively "pristine" river systems. Improving habitat, or identifying suitable habitat, will be a necessary starting point for *L. compacta* reintroduction efforts.

In addition to having a small range, *L. compacta* only exists along a single river path. This means that one catastrophic event such as a massive point source pollution event above Cahaba River at Booth's Ford could spell doom for *L. compacta*. Such an event is not merely a hypothetical. In 2016, a gasoline pipeline spill came perilously close to the Cahaba River (Pillion, 2016). To mitigate the risks of a single catastrophic event, reintroduction efforts should emphasize range expansion outside the mainstem Cahaba River. Of course, reintroduction efforts also must be limited by the historical range of any given species. Thus, lower Buck Creek is potentially an ideal reintroduction site if habitat quality is sufficient for the persistence of *L. compacta*. Once a suitable reintroduction site is chosen, managers will need to choose a broodstock site. This decision should be informed with genetic data. The absence of an isolation by distance effect across the current range of *L. compacta* indicates that managers do not need to prioritize potential broodstock sites based on whether they are geographically proximate to reintroduction sites. Rather, sites with high genetic diversity and ease of access should be prioritized for broodstock choice. Therefore, the Cahaba River above Shades Creek is likely an ideal broodstock choice, given its comparably high genetic diversity. Moreover, *L. compacta* is easy to sample and relatively easy to distinguish from other sympatric species at Cahaba River above Shades Creek, making it ideal from both a genetic and sampling standpoint.

Conclusions

Even though *L. compacta* was considered extinct less than a decade ago, we now know more about this species than most other freshwater gastropods. This bodes well for the conservation of *L. compacta* as the biggest barrier to effective conservation strategies for most freshwater gastropods is a lack of data. Future research efforts should focus on differences in dispersal dynamics among pleurocerids and causes of differences in riverscape genetic patterns seen between *L. ampla* and *L. compacta*. As more population genomic data becomes available for pleurocerids, we will be better suited to develop strategies to conserve these critically important components of many North American riverine ecosystems.

Acknowledgements

We thank Lauren Allred, Jecca Shumante, and Living River: A Retreat on the Cahaba for providing access to field sites. Haley Dutton (Auburn University) assisted in the field. Jeffrey Garner (Alabama Department of Conservation and Natural Resources), Matthew Galaska (University of Washington), Kenneth Halanych (Auburn University) and Katherine Bockrath

(U.S Fish and Wildlife Service) provided advice on multiple components of this work. 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 was work funded by a reverted Section 6 grant from Alabama Department of Conservation and Natural Resources and by United States Fish and Wildlife Service. This study was also supported in part by NSF grant DBI-1658694, the Alabama Agricultural Experiment Station, and the Hatch program of the National Institute of Food and Agriculture, U.S. Department of Agriculture. This work was also made possible in part by a grant of high-performance computing resources and technical support from the Alabama Supercomputer Authority.

References

Alabama Department of Environmental Management. 2007. 2007 rivers and streams monitoring summary report: Shades Creek at Jefferson County Road 6. Available at http://adem.alabama.gov/programs/water/wqsurvey/table/2007/2007ShadesCk.pdf (accessed 15 April 2020)

Alabama Department of Environmental Management. 2012. 2012 rivers and streams monitoring summary report: Shades Creek at Parkwood Road in Jefferson County. Available at http://adem.alabama.gov/programs/water/wqsurvey/table/2012/2012ShadesCk.pdf (accessed 15 April 2020)

Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics* 17:81-92. DOI: i:10.1038/nrg.2015.28.

Bennett MG, Howell JH, Kuhajda BR, Freeman PL. 2015. New upstream records for fishes following dam removal in the Cahaba River, Alabama. *Southeastern Fishes Council Proceedings* 55:50-61.

Brown KM, Lang B, Perez KE. 2008. The conservation ecology of North American pleurocerid and hydrobiid gastropods. *Journal of the North American Benthological Society* 27:484-495. DOI: 10.1899/07-062.1.

Catchen JM, Hohenlohe P, Bassham S, Amores A, Cresko W. 2013. Stacks: an analysis tool set for population genomics. *Molecular Ecology* 22:3124-3140. DOI: 10.1111/mec.12354.

Covich AP, Palmer MA, Crowl TA. 1999. The role of benthic invertebrate species in freshwater ecosystems: zoobenthic species influence energy flows and nutrient cycling. *Bioscience* 49:119-127. DOI: 10.2307/1313537.

Dray S, Dufour AB. 2007. The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software* 22:1-20. DOI: 10.18637/jss.v022.i04.

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

Friends of Shades Creek. 2020. Shades Creek Watershed Management Plant. Available at https://www.gmcplanning.com/ShadesCreek (accessed 15 April 2020).

Fukuda H, Haga T, Tatara Y. 2008. *Niku-nuki*: a useful method for anatomical and DNA studies on shell-bearing molluscs. *Zoosymposia* 1:15-38. DOI: 10.11646/zoosymposia.1.1.5.

Goodrich C. 1922. The *Anculosae* of the Alabama River drainage. *University of Michigan Museum of Zoology Miscellaneous Publication No 7*:63.

Goodrich C. 1941. Distribution of the gastropods of the Cahaba River, Alabama. *Occasional Papers of the Museum of Zoology* 426:1-30.

Goslee SC, Urban DL. 2007. The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software* 22:1-17. DOI: 10.18637/jss.v022.i07.

Goudet J. 2005. HIERFSTAT, a package for R to compute and test hierarchical *F*-statistics. *Molecular Ecology Resources* 5:184-186. DOI: 10.1111/j.1471-8286.2004.00828.x.

Hughes JM, Schmidt DJ, Finn DS. 2009. Genes in streams: using DNA to understand the movement of freshwater fauna and their riverine habitat. *Bioscience* 59:573-583. DOI: 10.1525/bio.2009.59.7.8.

Huryn AD, Benke AC, Ward GM. 1995. Direct and indirect effects of geology on the distribution, biomass, and production of the freshwater snail Elimia. *Journal of the North American Benthological Society* 14:519-534. DOI: 10.2307/1467538.

Johnson PD. 2019. Results of AABC freshwater mollusk survery Cahaba River historic shoal at Belle Ellen Bibb County, AL-7 May 2019. Report submitted to Alabama Department of Conservatio and Natural Resources. Marion, Alabama. p 6.

Johnson PD, Bogan AE, Brown KM, Burkhead NM, Cordeiro JR, Garner JT, Hartfield PD, Lepitzki DAW, Mackie GL, Pip E, Tarpley TA, Tiemann JS, Whelan NV, Strong EE. 2013. Conservation status of freshwater gastropods of Canada and the United States. *Fisheries* 38:247-282. DOI: 10.1080/03632415.2013.785396.

Johnson PD, Sides JD, McGregor SW, Ahlstedt SA, Garner JT, Butler RS, Novak SF. 2006. Inventory of freshwater mollusks in the Cahaba River Basin, Alabama. Submitted to the Alabama Department of Conservation and Natural Resources. p 153.

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. DOI: 10.7717/peerj.281.

Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl PA. 2013. diveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution* 48:782-788. DOI: 10.1111/2041-210X.12067.

Kopelman NM, Mayzel J, Jakobsson M, Rosebnberg NA, Mayrose I. 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* 15:1179-1191. DOI: 10.1111/1755-0998.12387.

Kurth JW. 2017. Endangered and threatened wildlife and plants; 90-day findings for five species. *Federal Register* 82:60362-60366.

Lambert SM, Streicher JW, Fisher-Reid MC, Méndez de la Cruz FR, Martínez-Méndez N, García-Vázquez UO, de Oca ANM, Wiens JJ. 2019. Inferring introgression using RADseq and D_{FOIL}: Power and pitfalls revealed in a case study of spiny lizards (*Sceloporus*). *Molecular Ecology Resources* 19:818-837. DOI: 10.1111/1755-0998.12972.

Lawson DJ, Hellenthal G, Myers S, Falush D. 2012. Inference of population structure using dense haplotype data. *PLoS Genetics* 8:e1002453. DOI: 10.1371/journal.pgen.1002453.

Legendre P, Fortin M-J, Borcard D. 2015. Should the Mantel test be used in spatial analysis. *Methods in Ecology and Evolution* 6:1239-1247. DOI: 10.1111/2041-210X.12425.

Leigh DM, Lischer HEL, Keller LF. 2018. Batch effects in a multiyear sequencing study: false biological trends due to changes in read lengths. *Molecular Ecology Resources* 18:778-788. DOI: 10.1111/1755-0998.12779.

Lydeard C, Cowie RH, Ponder W, F., Bogan AE, Bouchet P, Clark SA, Cummings KS, Frest TJ, Gargominy O, Herbert DG, Hershler R, Perez KE, Roth B, Seddon M, Strong EE, Thompson FG. 2004. The global decline of nonmarine mollusks. *Bioscience* 54:321-330. DOI: 10.1641/0006-3568(2004)054[0321:TGDONM]2.0.CO;2.

Lydeard C, Mayden RL. 1995. A diverse and endangered ecosystem of the southeast United States. *Conservation Biology* 9:800-805. DOI: 10.1046/j.1523-1739.1995.09040800.x.

Lysne SJ, Perez KE, Brown KM, Minton RL, Sides JD. 2008. A review of freshwater gastropod conservation: challenges and opportunities. *Journal of the North American Benthological Society* 27:463-470. DOI: 10.1899/07-061.1.

Malinsky M, Trucchi E, Lawson DJ, Falush D. 2018. RADpainter and fineRADstructure: population inference from RADseq Data. *Molecular Biology and Evolution* 35:1284-1290. DOI: 10.1093/molbev/msy023.

Meirmans PG. 2015. Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology* 24:3223-3231. DOI: 10.1111/mec.13243.

National Marine Fisheries Service. 2018. Interim Endangered and Threatened Speices Recovery Planning Guidance Version 1.4. National Oceanic and Atmospheric Administration. Silver Spring, Maryland. p 122.

Onorato D, Angus RA, Marion KR. 2000. Historical changes in the ichtyofaunal assemblages of the upper Cahaba River in Alabama associated with extensive urban development in the watershed. *Journal of Freshwater Ecology* 15:47-63. DOI: 10.1080/02705060.2000.9663721.

Paris JR, Stevens JR, Catchen JM. 2017. Lost in parameter space: a road map for Stacks. *Methods in Ecology and Evolution* 8:1360-1373. DOI: 10.1111/2041-210X.12775.

Paz-Vinas I, Loot G, Stevens VM, Blanchet S. 2015. Evolutionary processes driving spatial patterns of intraspecific genetic diversity in river ecosystems. *Molecular Ecology* 24:4586-4604. DOI: 10.1111/mec.13345.

Perez KE, Minton RL. 2008. Practical applications for systematics and taxonomy in North American freshwater gastropod conservation. *Journal of the North American Benthological Society* 27:471-483. DOI: 10.1899/07-059.1.

Pillion D. 2016. Alabama pipeline leak: what we know so far about the spill, gas shortages and more. Available at https://www.al.com/news/birmingham/2016/09/how_alabama_pipeline_leak_led.html (accessed 19 March 2020)

Pitt RE. 2000. Water quality conditions in the Cahaba River and likely pollutant sources. Birmingham, Alabama: Department of Civil and Environmental Engineering, University of Alabama at Birmingham. Available at http://rpitt.eng.ua.edu/Publications/MonitoringandStormwater/Cahaba%20problems%20and%20sources.PDF (accessed 15 April 2020)

Potter JJ, Thomas WL. 1983. Endangered and threatened species listing and recovery priority guidelines. *Federal Register* 48:43098-43105.

R Core Development Team. 2020. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rumble SM, Lacroute P, Dalca AV, Fiume M, Sidow A, Brundo M. 2009. SHRiMP: accurate mapping of short color-space reads. *PLoS Computational Biology* 5:e1000386. DOI: 10.1371/journal.pcbi.1000386.

Shepard TE, O'Neil PE, McGregor SW, Harris SC. 1994. Water-quality and biomonitoring studies in the upper Cahaba River drainage of Alabama. *Geological Survery of Alabama Bulletin* 160:1-118.

Shringarpure SS, Bustamante CD, Lange K, Alexander DH. 2016. Efficient analysis of large datasets and sex bias with ADMIXTURE. *BMC Bioinformatics* 17:218. DOI: 10.1186/s12859-016-1082-x.

Strong EE, Gargominy O, Ponder W, F., Bouchet P. 2008. Global diversity of gastropods (Gastropoda; Mollusca) in freshwater. *Hydrobiologia* 595:149-166. DOI: 10.1007/S10750-007-9012-6.

Strong EE, Köhler F. 2009. Morphological and molecular analysis of '*Melania jacqueti*' Dautzenberg and Fischer, 1906: from anonymous orphan to critical basal offshoot of the Semisulcospiridae (Gastropoda: Cerithioidea). *Zoologica Scripta* 38:483-502. DOI: 10.1111/j.1463-6409.2008.00385.x.

Tolley-Jordan LR, Huryn AD, Bogan AE. 2015. Effects of land-use change on a diverse pleurocerid snail assemblage. *Aquatic Conservation: Marine and Freshwater Ecosystems* 25:235-249. DOI: 10.1002/aqc.2474.

Wang S, Meyer E, McKay JK, Matz MV. 2012. 2b-RAD: a simple and flexible method for genome-wide genotyping. *Nature Methods* 9:808-810. DOI: 10.1038/nmeth.2023.

Weir BS, Cockerham CC. 1984. Estimating *F-statistics* for the analysis of population structure. *Evolution* 38:1358-1370. DOI: 10.1111/j.1558-5646.1984.tb05657.x.

Whelan NV, Galaska MP, Sipley BN, Weber JM, Johnson PD, Halanych KM, Helms BS. 2019. Riverscape genetic variation, migration patterns, and morphological variation of the threatened Round Rocksnail, *Leptoxis ampla*. *Molecular Ecology* 28:1593-1610. DOI: 10.1111/mec.15032.

Whelan NV, Johnson PD, Harris PM. 2012. Rediscovery of *Leptoxis compacta* (Anthony, 1854) (Gastropoda: Cerithioidea: Pleuroceridae). *PLoS One* 7:e42499. DOI: 10.1371/journal.pone.0042499.

Whelan NV, Johnson PD, Harris PM. 2015. Life-history traits and shell morphology in the genus *Leptoxis* Rafinesque, 1819 (Gastropoda: Cerithioidea: Pleuroceridae). *Journal of Molluscan Studies* 81:85-95. DOI: 10.1093/mollus/eyu058.

Tables

Table 1: Summary statistics and AUMNH catalog numbers of *L. compacta* at each collection site

Collection Site	Private alleles	H_o (sd)	H_e (sd)	A_r (sd)	Π (sd)	AUMNH #
Cahaba River at Booth's Ford	43	0.1568 (0.1349)	0.1801 (0.1222)	1.8241 (0.2742)	0.1855 (0.1261)	###-###
Cahaba River at boat launch	32	0.1046 (0.1552)	0.1045 (0.1421)	1.4511 (0.4518)	0.1075 (0.1459)	###-###
Cahaba River above Shades Creek	262	0.1363 (0.1334)	0.1779 (0.1245)	1.8072 (0.2883)	0.1829 (0.1281)	###-###
Cahaba River at Old Marvel Slab	28	0.0963 (0.1400)	0.0980 (0.1343)	1.4606 (0.4387)	0.1010 (0.1382)	###-###

H_o: Observed heterozygosity, H_e: Expected heterozygosity, A_r: Allelic richness, Π: nucleotide diversity

Table 2: Pairwsie F_{st} and distances (km) between sites. F_{st} below diagonal and distances above diagonal.

	Booth's Ford	boat launch	above Shades Creek	old Marvel slab
Cahaba River at Booth's Ford	-	4.62	5.55	9.2
Cahaba River at boat launch	0.04	-	0.98	4.57
Cahaba River above Shades Creek	0	0.05	-	3.64
Cahaba River at old Marvel slab	0.03	0.04	0.03	-

Figures

Figure 1: Photograph of live *L. compacta*. Photo Credit: Thomas Tarpley, ADCNR.

Figure 2: Shells of representative individuals that we sequenced. A) Cahaba River at Canoe Launch, B) Cahaba River at Booth's Ford, C) Cahaba River above Shades Creek, D-F) Cahaba River at old Marvel slab. Scale bar = 1 cm

Figure 3: Map of known historical and current range of *L. compacta*, collection sites, and other landmarks. Lines from collection sites lead to ADMXITURE plots with K = 2 for each site. Each column is an individual with ADMIXTURE proportions of the two inferred ancestral populations.

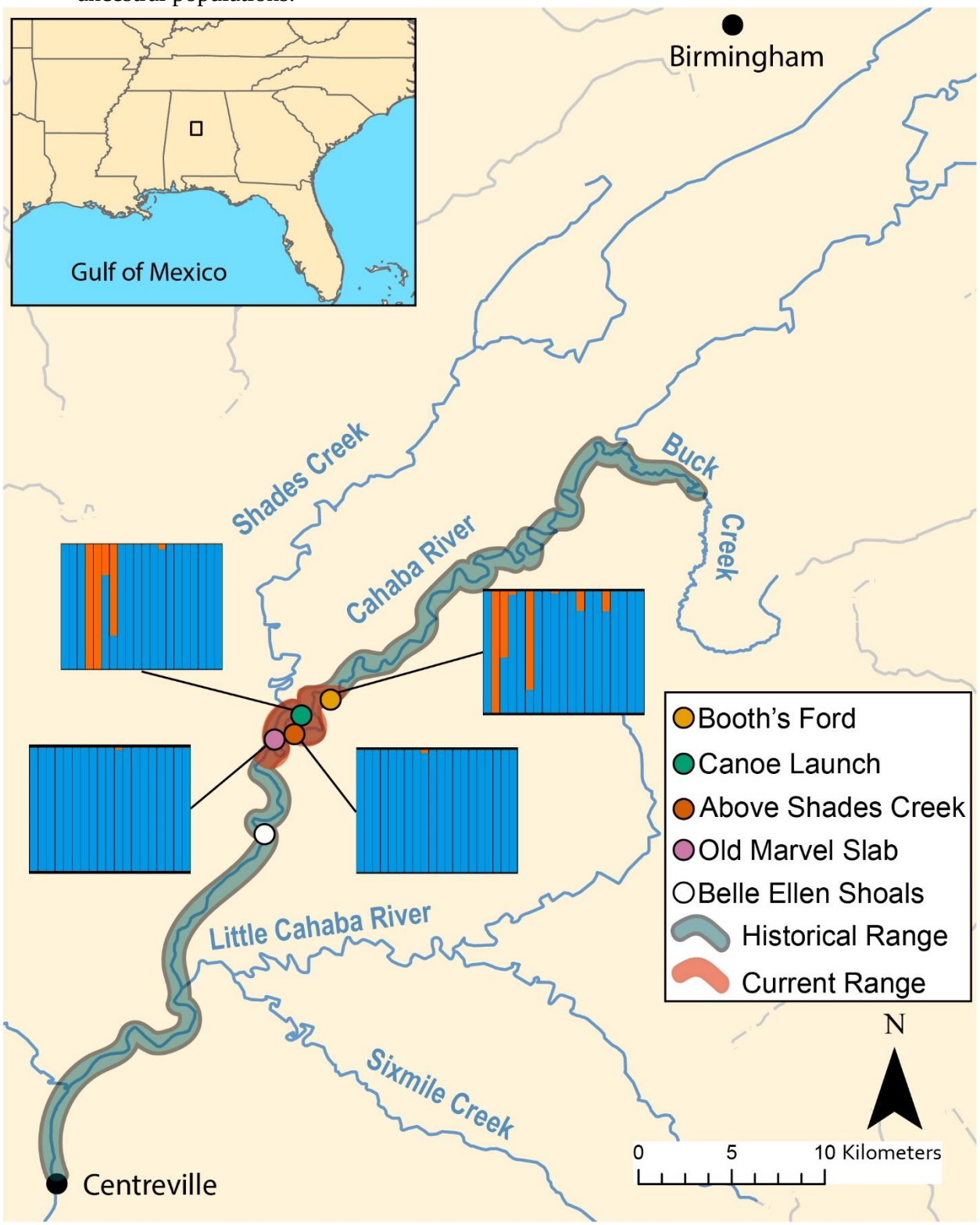
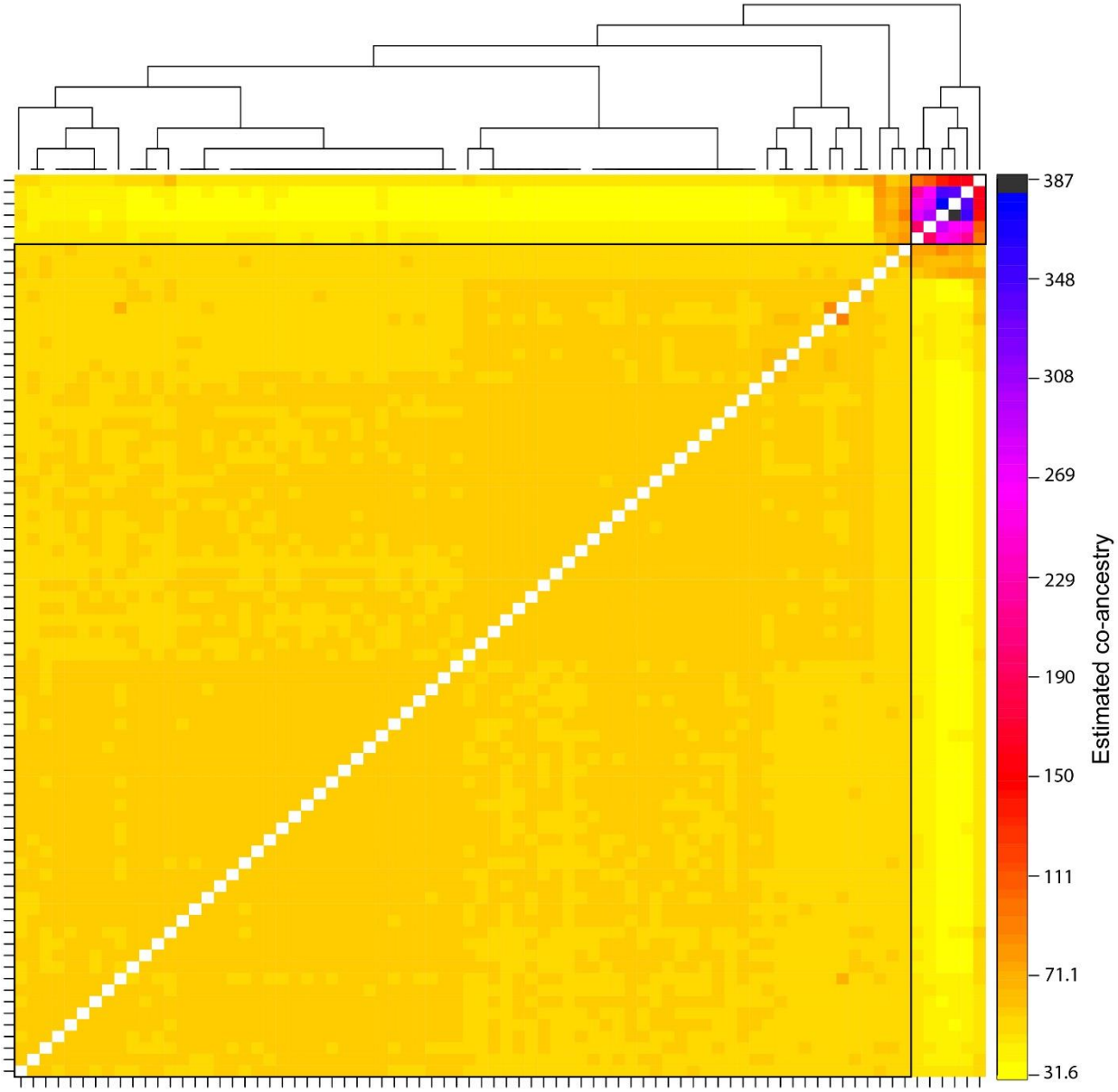
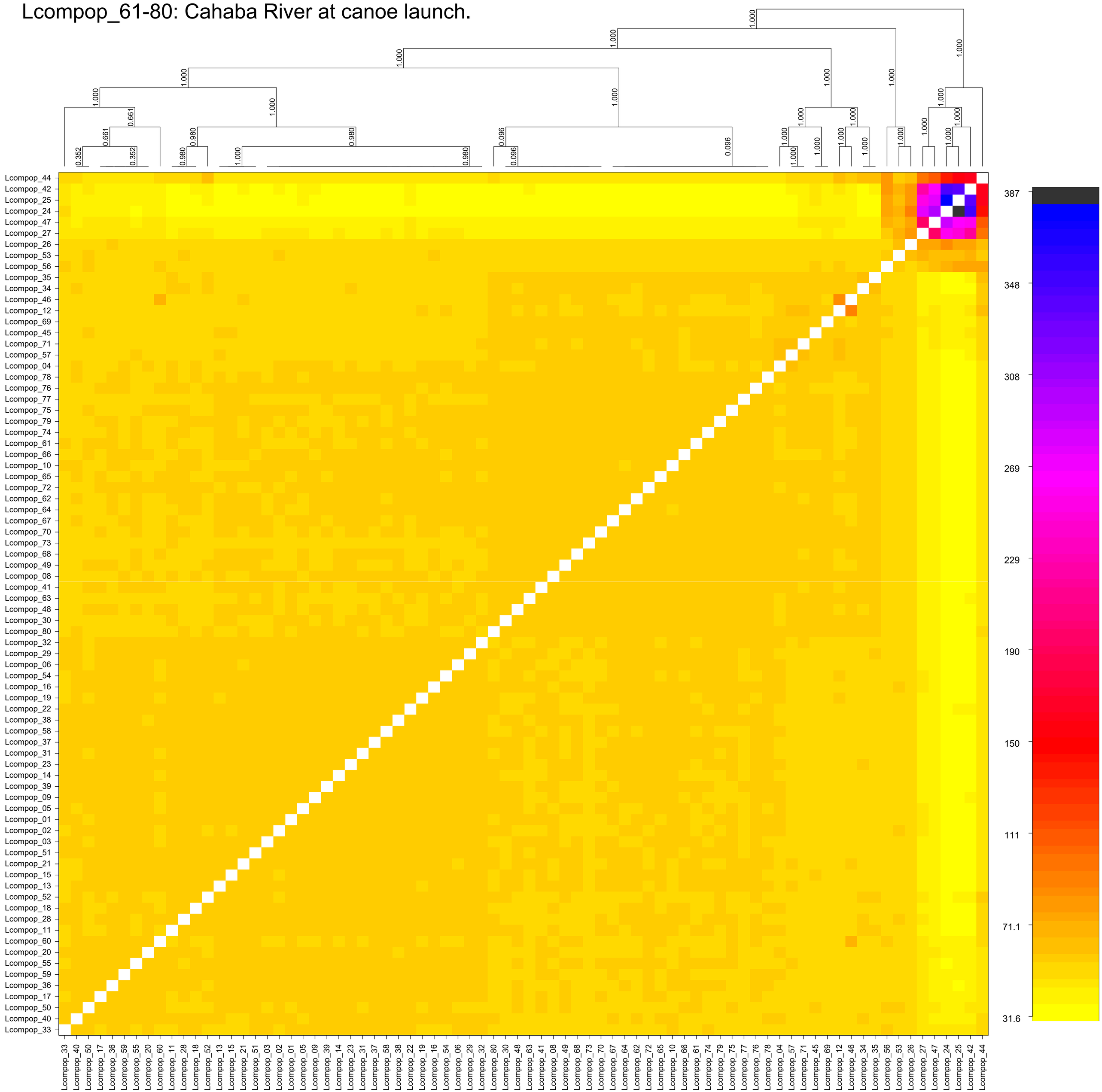


Figure 4: Pairwise co-ancestry matrix and simple tree inferred with fineRADstructure. Boxes surround the two main groupings. Tick marks represent individuals, but labels have been removed for visualization. For a figure with full taxon labels, see Supplementary Figure 1.



Supplementary Figure 1: Pairwise co-ancestry matrix and simple tree inferred with fineRADstructure. Boxes surround the two main groupings. Tick marks represent individuals. Lcompop_01-20: Cahaba River at old Marvel slab; Lcompop_21-40: Cahaba River above Shades Creek; Lcompop_41-60: Cahaba River at Booth's Ford; Lcompop_61-80: Cahaba River at canoe launch.