Development and application of diagnostic SNP marker resources for Northern (*Micropterus salmoides salmoides*) and Florida (*Micropterus salmoides floridanus*) largemouth bass.

By
Spencer W. Gowan

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
August, 1st 2015

Keywords: Single Nucleotide Polymorphism, hybridization, *Micropterus*, bass

Approved by

Eric Peatman, Associate Professor; School of Fisheries Aquaculture and Aquatic Sciences
Terry Hanson, Professor; School of Fisheries Aquaculture and Aquatic Sciences
Steven Sammons, Research Fellow; School of Fisheries Aquaculture and Aquatic Sciences
Benjamin Beck, Research Leader; USDA-ARS, Auburn AL

Abstract

Efforts to improve recreational fisheries have included widespread stocking of *Micropterus salmoides floridanus* outside its native range of peninsular Florida. Hybridization of Florida bass (*M. salmoides floridanus*) with Northern largemouth bass (*Micropterus salmoides salmoides*) has now dramatically expanded beyond a naturally occurring intergrade zone in the southeast U.S. In recent years, there has been growing interest in protecting the genetic integrity of native basses and assessing the impact and nature of *M.s. salmoides/M.s. floridanus* introgression from the standpoint of hatchery and sport-fishery managers, fish biologists, ecologists and evolutionary biologists. Here, RNA-seq-based sequencing of the transcriptomes of *M.s. salmoides, M.s floridanus* and their F1 hybrid was conducted and a set of 3674 SNP markers with fixed-allelic differences from 2112 unique genes were identified. A subset of 61 of these markers were then developed into a set of diagnostic multiplex assays and their capacity for assessing integrity and hybridization in hatchery and wild populations of Northern largemouth and Florida bass was evaluated.

Use of these markers for population comparisons and hybridization rate evaluations were demonstrated in populations spanning the state of Alabama. Geographic isolation by natural barriers (fall line and drainage basins) were found to lead to variation in introgression level, indicating limited effect of stocking efforts in some locations, while other populations appeared to have had successful introduction of FLMB alleles.

An attempt was also made to use these markers to assess the effect of FLMB allele introgression on trophy bass populations in Lake Guntersville, AL. Correlation between genotype and size was observed in fish sampled from tournaments. The heavier fish had more FLMB influence and higher heterozygosity.

Electrofishing surveys were also conducted to collect size at age data for the Lake Guntersville population and supplement the tournament samples. While growth differences were not apparent between genotype variants within this electrofishing sample, size and genotype differences were observed between fish caught by tournament anglers and fish caught by electrofishing surveys. Some individuals sampled from the tournament bass had higher observed weights than those found within the electrofishing sample. Also significantly higher mean FLMB allele frequency was observed in the tournament samples when compared to electrofishing samples.

The availability of this resource, high-quality transcriptomes and a large set of gene-linked SNPs, should continue to greatly facilitate functional and population genomics studies in these key species and allow the identification of traits and processes under selection during introgressive hybridization, as well as facilitate more efficient genetic management of hatchery and stocking programs aimed at enhancing or conserving various populations of largemouth bass.

Acknowledgments

I would like to thank my faculty advisor, Dr. Eric Peatman, for inviting me to work in his lab and for his role as a mentor during my time at Auburn University. I would also like to thank my committee members, Dr. Terry Hanson, Dr. Steve Sammons, and Dr. Benjamin Beck, for their time and guidance on this project.

I would like to thank Dr. Chao Li, Dr. Huseyin Kucuktas, and Dr. Milla Kaltenboeck, for their large contributions of time and effort into this project. They played an important role in sample collection, RNA extraction, executing assays on the Sequenom MassARRAY, and analysis. I also must acknowledge the rest of my lab mates who in many different ways contributed to the this project, Ammu Anil, Wilawan Thongda, Dongdong Zhang, and Honggang Zhao.

Thank you to Dave Belkoski of Dr. Sammons lab for being a second reader for otoliths. Thanks to the students at the Ireland Center, (Nick Feltz, Lee Grove, and Chris Kemp) for being willing to share and collect extra samples for genetic analysis. Many thanks to Keith Floyd, Phil Ekema, Damon Abernethy, Keith Henderson, Mike Holley, Brian Rinehard and the rest of the ALDCNR biologists and technicians that helped sample Lake Guntersville, as well as collected many of the samples from around Alabama. Thank you as well to GADNR biologists who also contributed samples. ALDCNR and GADNR were responsible for funding the projects I was involved in as a student which allowed me to have this opportunity in the Peatman Lab, and for that I am thankful.

A special thank you to my parents Karen and Samuel Gowan and my wife Kelly Gowan for supporting me in my decision to return to school to complete a graduate degree.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

DNA	Deoxyribonucleic acid
AU	Auburn University
ALDCNR	Alabama Department of Conservation and Natural Resources
FLMB	Florida Largemouth Bass *(Micropterus salmoides floridanus)*
F_{st}	Fixation index
GDNR	Georgia Department of Natural Resources
LMB	Largemouth bass *(Micropterus salmoides)*
mtDNA	Mitochondrial DNA
NLMB	Northern Largemouth Bass *(Micropterus salmoides salmoides)*
PCR	Polymerase Chain Reaction
SNP	Single Nucleotide Polymorphism
TL	Total length
W_r	Relative weight
Wt	Weight

Introduction

Black basses (*Micropterus spp.*) are ecologically and economically important members of a diverse array of ecosystems across North America, from small ponds and streams to large rivers and lakes (DeVries et al. 2014). With their reputation for aggressive feeding behavior and environmental adaptability, black basses have solidified their place as the most popular sport-fish in the United States; making bass fishing synonymous to American sport-fishing. In 2011 recreational fishing expenditures by anglers in the U.S. (including saltwater) reached $41.8 billion and, of the 33.1 million people participating in recreational angling nationwide, one-third (~11.2 million anglers) were primarily targeting black basses; far exceeding the number of anglers targeting any other species group (USDI 2011). Not only are these fish economically and ecologically valuable sport fish, but they are a growing segment of the US and Chinese aquaculture industry (Bai et al. 2008). Within the genus *Micropterus,* the American Fisheries Society recognizes eight species: shoal bass (*Micropterus cataractae*), redeye bass (*Micropterus coosae*), smallmouth bass (*Micropterus dolomieu*), Alabama bass (*Micropterus henshalli*), Suwannee bass (*Micropterus notius*), spotted bass (*Micropterus punctulatus*), largemouth bass (*Micropterus salmoides*), and Guadeloupe bass (*Micropterus treculii*) (Page et al. 2013). All of these species can be targeted sport-fish, however the majority of the focus is on the largemouth bass (LMB) (*Micropterus salmoides*). Popularity of this species relative to other black basses is due to their increased adaptability to a wide range of habitats and their ability to regularly attain sizes in the 5-20lb range. Their popularity is evident in the extreme anthropogenic expansion of the species' range. Originally inhabiting most of the Northern Gulf of Mexico drainages including the Mississippi River drainage and the Atlantic drainages south of Virginia, LMB are now found in most of North America as well as parts of Africa, South America, Europe, and Asia (e.g.

MacCrimmon & Robbins 1975; Azuma and Motomura 1999; Gratwicke & Marshall 2001).

Within its native range, the LMB taxa is composed of two subspecies: the Northern largemouth bass (NLMB) (*Micropterus salmoides salmoides*) and the Florida largemouth bass (FLMB) (*Micropterus salmoides floridanus*). The native ranges of these species were first described by Bailey and Hubbs (1949) based on scale counts. FLMB were shown to be restricted to peninsular Florida, while NLMB dominated all other regions in the LMB range. Bailey and Hubbs (1949) also identified an intergrade zone or hybrid zone occurring in the Gulf drainages from the Suwannee River west to the Choctawhatchee River, and Atlantic drainages from the St. Mary's River and north to the Savannah River; an area that roughly covers much of the Florida panhandle, Southeast Alabama, most of Georgia, and a small portion of southern South Carolina. They observed that, meristic traits were often unreliable in distinguishing the subspecies within the intergrade zone. Since their recognition, these subspecies have been intensely studied in an attempt to define their evolutionary history and their growth potential in a variety of natural and man-made aquatic ecosystems.

Much of the interest and research relating to these subspecies has been fueled by sport-fishing and this is reflected in the types of research conducted. Researchers have often focused on identifying variations in performance between the two subspecies (e.g. growth, angling susceptibility, and environmental tolerances). Conventional wisdom of anglers dictates that NLMB are aggressive feeders, exhibit fast initial growth, and are highly catchable; while FLMB offer the potential for larger maximum sizes, but are more elusive to the angler, and exhibit slower growth. Despite having little in the way of conclusive evidence of superiority or suitability to new environments, many state agencies and private individuals began extensive stockings of FLMB well outside of their native range throughout the second half of the 20th century, thus effectively

47 expanding the number of intergrade populations (Philipp et al. 1983; Phillip & Ridgeway 2002
48 Barthel 2010). Early efforts to substantiate these claims and justify FLMB stocking programs
49 resulted in contradictory and inconclusive results.

50 In head-to-head evaluations in small ponds and impoundments, early researchers did not
51 find any conclusive evidence of growth difference between the subspecies (Clugston 1964;
52 Addison & Spencer 1972). However, superior growth was observed in F1 progeny of the two
53 subspecies (Inman et al. 1977). Experimental conditions were often considered too variable to
54 make conclusions. But failure to find a large growth advantage in FLMB, led to the suggestion
55 that the apparent larger maximum size of the FLMB in the wild was due to better growing
56 conditions in Florida waterways rather than genetics. It was also suggested that growth may be the
57 same, but the FLMB may still get larger by living longer, citing that the FLMB appear to be heartier
58 and more resistant to handling mortality associated with angling (Miller 1965; Inman et al. 1977;
59 Bottroff & Lembeck 1978), or by being more elusive to anglers altogether (Addison & Spencer
60 1972). This was supported by research conducted in small farm ponds in southern Alabama, where
61 lower catch rates per unit effort for FLMB were observed (Zolczynski & Davies 1976). This also
62 corroborated similar results in California reservoirs (Sasaki 1961). However, in other studies, no
63 difference in angling susceptibility was observed between the two subspecies (Inman et al. 1977).
64 Ultimately, it was suggested that the lack of conclusive results in early studies may have also been
65 the result of using stocks that were not validated for purity (Philipp et al. 1983; Fields et al. 1987).
66 Since the meristic traits normally used to identify FLMB and NLMB are unreliable in these hybrid
67 populations, molecular methods were needed to validate population status when selecting stocks
68 for experimentation.

69 In order to develop a molecular method to validate the subspecies status of individuals,

Philipp et al. (1983) conducted a study involving 1800 largemouth bass collected from 90 populations across the United States. Enzymes extracted from tissue samples (white muscle and liver) were run through vertical starch gel electrophoresis to test for phenotypic variation of enzymes. Out of the 28 loci examined, only two loci, isocitrate dehydrase (Idh-B) and aspartate aminotransferase (Aat-B), were truly fixed between the subspecies. Another two loci, malate dehydrogenase (Mdh-B) superoxide dismutase (Sod-A), had variants that only occurred in one population even though the alternative allele occurred in all populations. At the conclusion of the study Philipp et al. (1983) used their marker data to redefine the intergrade zone to include much of the southeast (Northern Florida, Georgia, South and North Carolina, Virginia, Alabama, Mississippi) and parts of Maryland, Louisiana, Arkansas, Texas, and California. Another similar, but smaller study was conducted by Williamson et al. (1986), which validated the markers from Philipp et al. (1983) and identified five more polymorphic markers useful for distinguishing LMB subspecies. They included adosine deaminase (ADA), galactose-1-phosphate uridyltransferase (GALT), mannosephosphate isomerase (MPI), Peptidase-3 (PEP3), and triosephosphate isomerase-2 (TPI-2). Even though these markers, known as allozymes, are useful tools for determining purity of a whole population, the authors warn that the limited number of alleles and the non-fixed status of many of them, rendered them inappropriate for assigning status to an individual, citing that in a hybrid population some individuals will score as a "pure" NLMB largemouth and some will score as a pure FLMB. This means that large sample sizes are needed for validation of populations.

Using allozyme markers, researchers in the 1980's and 1990's were able to validate stocks in head-to-head evaluations as well as in artificially and naturally admixed populations. Using these markers, NLMB were found to exhibit faster growth in the first year (Isely et al. 1987;

93 Williamson & Carmichael 1990; Phillip & Whitt 1991), and FLMB were shown to have increased
94 growth and fecundity beyond age three, which may explain the larger maximum size achieved by
95 FLMB (Maceina et al. 1988). But others still found no growth differences between allozyme-
96 validated stocks (Leitner et al. 2002). Genotype-by-environment (GxE) interactions are thought to
97 ultimately govern the performance of the fish rather than exclusively genotype or exclusively
98 environment.

99 Since LMB are ectothermic, temperature plays a large role in their performance. As would
100 be expected, FLMB bass exhibit higher thermal tolerances (Fields et al. 1987), and NLMB exhibit
101 higher cold tolerance (Cichra et al. 1982; Williamson & Carmichael 1990). They also exhibit
102 variations in spawning periods and durations. NLMB have been shown to spawn slightly earlier
103 and over a shorter duration than FLMB (Rogers et al. 2006). Interestingly, when NLMB and FLMB
104 individuals were moved to new latitudes, they exhibited altered spawning times relative to their
105 native populations, but retained the relationship to one another demonstrating that genotype and
106 environmental cues play an important role in the temporal spawning habits of these fish (Rogers
107 et al. 2006). It has been suggested that rather than truly exhibiting superior growth, the earlier
108 spawn, swim up, and cold tolerance of NLMB may be misinterpreted as superior growth in the
109 first year, and may simply be the result of a size advantage when the water reaches the optimal
110 growth temperatures. This would be a greater advantage in northern regions which have a shorter
111 growing season (Isely et al. 1987). Despite the shift in spawning times, the two subspecies do not
112 fully exhibit assortative mating. Natural hybridization between "pure" NLMB and "pure" FLMB
113 is a common occurrence when the two subspecies cohabitate a waterbody (Isely et al. 1987; Philipp
114 & Ridgeway 2002; Rogers et al. 2006). The rate and success of this hybridization appears to be
115 the result of strikingly low selection (in the form of intrinsic genetic incompatibilities) against

hybrids within Centrarchidae when compared with other taxonomic groups (Bolnick & Near 2005; Seyoum et al. 2013). The performance of these hybrids is of great interest to FLMB stocking proponents and detractors alike.

When performance of validated hybrids have been evaluated, researchers have found no evidence of heterosis in F1 or F2 reciprocal crosses (Isely et al. 1987; Maceina & Murphy 1988; Williamson & Carmichael 1990; Philipp & Whitt 1991; Philipp 1991). The only reported exception has been an observed higher chronic thermal maximum in the F1 of the FLMB♀ x NLMB♂, but not in the reciprocal cross (Fields et al. 1987). It has even been suggested that deleterious breakdown of co-adapted gene complexes and detrimental physiological consequences from outbreeding depression occurs in these intergrade stocks (Philipp & Ridgeway 2002; Cooke & Philipp 2006). This suggestion is based on observations in Illinois, where FLMB alleles may have deleterious GxE interactions and may not hold true for intergrade populations in the southern United States.

The rate of introgression following FLMB stocking into previously NLMB or intergrade populations, and the effect those allele combinations have on performance, is likely the key to understanding and developing best genetic management practices for hatcheries. Using allozyme markers, researchers in Texas and Alabama observed a decrease of NLMB alleles after stocking FLMB into already established populations (Maceina et al. 1988; Dunham et al. 1992; Mitchell et al. 1993; Brown & Murphy 1994). The altered F_x genotype and allele frequencies were shown to persist after annual stockings had ceased (Brown & Murphy 1994). However in Illinois, an immediate increase in NLMB alleles was observed after the conclusion of FLMB introductions (Philipp 1991). This supports the idea that GxE interactions have a large influence on performance and that one-size-fits-all approaches to stocking are inappropriate. To develop a better

understanding of GxE interactions in relation to FLMB and NLMB hybridization, higher marker resolutions are necessary.

A shift toward DNA-based marker technology was seen in the 1990's as molecular technology improved. DNA-based markers are advantageous because they eliminate the need for liver or muscle samples, relying instead on non-lethal fin clips, making sample collection and pre-extraction preservation much simpler. The first work in LMB to use DNA-based technology relied on mitochondrial DNA (mtDNA).

Polymorphisms in mtDNA were used to evaluate LMB population genetic distance and evolutionary history (Nedbal & Philipp 1994; Bremer et al. 1998; Williams et al. 1998; Kassler et al.2002; Near et al. 2003). Relying on a variety of restriction enzyme and polymerase chain reaction (PCR) techniques and later mtDNA sequencing, these studies were able to establish assays for distinguishing geographic populations and propose the timing and mechanisms for the speciation and evolutionary divergence among *Micropterus* spp. Based on mtDNA sequence polymorphisms, Kassler et al. (2002) made the case that NLMB and FLMB are in fact separate species rather than subspecies; citing that the sequence divergence between the two was 3.89%, which is significantly higher than the sequence divergence between *M. punctulatus* and *M. dolomieu* (1.20%); two long-recognized species. The use of mtDNA data is excellent for geographic population and evolutionary relationship studies. However, mtDNA is only maternally inherited, making it inappropriate for studying rates of hybridization between FLMB and NLMB in the intergrade zone. A nuclear DNA-based marker system would be preferable for this purpose.

Microsatellites are the tandem-repeat sequences in the non-coding regions of genomic DNA. Polymorphism in the number of repeats are common because mutations are not inhibited by functionality, and unlike mtDNA, these sequences are co-dominant, making them ideal

molecular markers for evaluating rates of introgression between species and subspecies. Lutz-Carrillo et al. optimized 11 (2006) and 52 (2008) microsatellite loci for distinguishing NLMB and FLMB. By more than doubling the number of diagnostic markers available, more reliable genotypes of individual fish were now a possibility. When these markers were used to genotype trophy bass in Texas (Lutz-Carrillo et al. 2006) and Arkansas (Lamothe & Johnson 2013), the results supported earlier findings that indicated a lack of heterosis in F1 fish, as well as high levels of Florida alleles in trophy bass. Unfortunately, these panels were optimized for technologies pre-dating capillary gel electrophoresis (e.g. ABI) and are unsuitable for higher-level multiplexing due to differing cycling and annealing temperature conditions. Barthel et al. (2010) and Seyoum et al. (2013), optimized an additional 18 microsatellite markers that can be used across the genus *Micropterus.* As advantageous as these markers are, this panel has no fixed-allelic differences between species at any locus and requires running of 3-4 ABI multiplexes, increasing the time and cost associated with performing the assay (Seyoum et al. 2013). In contrast, next generation sequencing technology has recently drastically reduced the cost and time involved in generation of needed molecular resources for non-model species (e.g. Wang et al. 2012), effectively opening a flood-gate of new marker technology.

The dramatic decline in sequencing costs associated with next generation sequencing has increased the accessibility of single nucleotide polymorphism (SNP) markers for population genetics and genomics in non-model organisms (Hohenlohe et al. 2011; Rice et al. 2011). SNP markers are valued for their genome wide distribution, abundance, ease of multiplexing and low genotyping error rate for high-throughput analyses (Slate et al. 2009; Pritchard et al. 2012). They are distributed across coding and noncoding regions of the genome, making SNPs particularly useful in studies examining traits and processes under selection during introgressive hybridization

(e.g. Fitzpatrick et al. 2009; Shen et al. 2012). Among teleost fish species, salmonid researchers have been pioneering in efforts to develop diagnostic SNP assays useful in assessing and managing genetic integrity and measuring impacts of hybridization with non-native, introduced populations (e.g. Stephens et al. 2009; Hohenlohe et al. 2011; Kalinowski et al. 2011; Lamaze et al. 2012; Pritchard et al. 2012, 2013; Lamer et al. 2014). SNP development and application in these studies have generally taken one of two approaches: RNA-seq on pooled samples followed by validation in greater numbers of individual samples (Lamaze et al. 2012) or reduced-representation sequencing of individual samples using RAD-seq or GBS approaches (e.g. Hohenlohe et al. 2011; Li et al. 2014). In this project the former approach was used.

The Alabama Department of Conservation and Natural Resources (ADCNR) has invested considerable time and capital into attempts to improve fishing opportunities by stocking FLMB into public reservoirs and state lakes. In order to understand and evaluate the success and impact of these introductions, statewide analysis of populations in major reservoirs have been completed in the past, often relying heavily on allozyme markers (Norgren et al.1986; Maceina & DiCenzo 1995). The goals for this project were to utilize RNA-seq to develop comprehensive transcriptomes for both NLMB and FLMB, identify a set of gene-based SNPs with fixed-allelic differences between the two subspecies, develop them into a robust marker panel for SNP genotyping, validate the panels on known populations, demonstrate their use by reevaluating introgression rates of populations previously assayed using allozymes (Norgren et al. 1986; Maceina & DiCenzo 1995), and to evaluate genetic contributions to growth and trophy LMB potential.

Materials and methods

Marker Development

Sample collection for RNA-seq

LMB were collected from genotyped stocks held by American Sport Fish Hatchery (Montgomery, AL, USA). Genotyping was conducted based on a subset of microsatellite markers from Lutz-Carrillo et al. (2006) and Seyoum et al. (2013). Sixty LMB were collected, these included 20 FLMB, 20 NLMB and 20 F1 hybrids (NLMB ♀ x FLMB ♂), with 10 males and 10 females selected from each group. Tissues collected from each fish, included brain, liver, skin, spleen, intestine, gonad, muscle and kidneys and were immediately stored in 5 mL RNA laterTM (Ambion, Austin, TX, USA) in separate tubes. Following an overnight incubation at 4 °C, the samples were stored at -80 °C until RNA extraction. Prior to RNA extraction, equal amounts of each tissue from the 20 fish within a group were homogenized into a master pool with mortar and pestle in the presence of liquid nitrogen. Total RNA was extracted using the RNeasy Universal Tissue Kit (Qiagen, Valencia, CA, USA). The three resulting master pools (FLMB, NLMB, and F1) were carried forward for library construction.

Library construction and RNA-seq

Sequencing libraries were prepared with 2.14–3.25 µg of starting total RNA and processed using the Illumina TruSeq RNA Sample Preparation Kit, as dictated by the TruSeq protocol. The libraries were amplified with 15 cycles of PCR and contained TruSeq barcode indices, identifying each of the three groups, within the Illumina adapters. Amplified library yields were 30 µL of 19.8–21.4 ng/µL with an average length of ~270 bp, indicating a concentration of 110–140 nM. After KAPA quantitation and dilution, based on included DNA standards (1–6), the libraries were

sequenced in a single lane on an Illumina HiSeq 2000 instrument with 100 bp paired-end (PE) reads at HudsonAlpha Genomic Services Lab (Huntsville, AL, USA). The image analysis, base calling, and quality score calibration were processed using ILLUMINA PIPELINE SOFTWARE v1.5. FASTQ files containing the raw sequencing reads, quality scores, and paired reads information were exported for the following trimming and assembly process.

De novo assembly and annotation of sequencing reads

Raw reads were processed for initial trimming by CLC Genomics Workbench (version 5.5.2; CLC Bio, Aarhus, Denmark). Before assembly, raw reads were trimmed by removing adapter sequences and ambiguous nucleotides. Reads with quality scores <20 and length below 30 bp were removed. The resulting high-quality sequences were used in the assembly. Assembly methodologies closely followed those described by Luo et al. (2014) and An et al. (2014), Li et al. (2014). Briefly, high-quality reads from the three barcoded pools (NLMB, FLMB and F1) were used to perform the de novo assembly using the Trinity assembler (v. 2014-04-13; Grabherr et al. 2011). This composite assembly was subsequently used for read mapping and SNP identification (below). The reads of each group were also assembled separately using Trinity, following the methodology of Luo et al. (2014), and subsequently annotated. The final assembled contigs from NLMB, FLMB, and their F1 were used as queries against the NCBI non-redundant (NR) protein database and the UniProtKB/SwissProt (Uniprot) database using BLASTX by setting the cut-off Expect value (E-value, the likelihood that the matching sequence is obtained by chance) of 1e-20 and score $\geq$ 100.

SNP and microsatellite marker identification

The SNP detection module (CLC Genomics Workbench) and composite reference assembly were used to identify SNPs. The composite reference assembly was used to identify SNPs utilizing the SNP detection module included in CLC Genomics Workbench (CLC Bio). Mapping of reads from each pooled sample to the composite reference assembly sequence was performed with mismatch cost of 2, deletion cost of 3 and insertion cost of 3. The highest scoring matches that shared ≥95% similarity with the reference sequence across ≥90% of their length were included in the alignment. A minimum coverage (read depth) ≥10 was set for each group to assess the quality of reads at positions for SNP detection. Only biallelic SNPs were allowed. Given the use of pooled samples, the identification of SNPs with fixed-allelic differences between NLMB and FLMB was the focus (e.g. homozygous 'A' in NLMB, homozygous 'T' in FLMB and heterozygous 'A/T' in their F1). SNPs which showed the consensus base (100% allele frequency) in one species, and the alternative allele in the other species, with both alleles present in the F1 hybrid read file (minor allele frequency ≥10%, minimum coverage ≥ 10), were carried forward as putative fixed-allele diagnostic SNPs.

Although not the focus of the current study, microsatellite markers were additionally mined from the NLMB and FLMB transcriptomes using MSATFINDER version 2.0.9 (Thurston & Field 2005), with a repeat threshold of eight dinucleotide repeats or five tri-, tetra-, penta-, or hexa-nucleotide repeats. The SSR loci with at least 50-bp sequence on both sides of the microsatellite repeats were considered sufficient for primer design and captured from the candidate marker list. The higher allelic richness of microsatellites makes them superior for some applications in structure and parentage analysis (Lapegue et al. 2014).

278 *Validation of fixed-allele interspecific SNPs*

279 A total of 119 samples from 5 populations (hatchery and wild) were used to validate and
280 markers and develop multi-plex panels. Subsets of each population were previously genotyped
281 using diagnostic microsatellite markers (Lutz-Carrillo et al. 2006; Seyoum et al. 2013). Samples
282 included 53 individuals from the Florida Bass Conservation Center, Webster, FL (directly and
283 indirectly through American Sport Fish and Alabama Department of Conservation and Natural
284 Resources), 37 individuals from American Sport Fish hatchery (Montgomery, AL) originally
285 sourced from an unknown Illinois lake, 20 individuals from Sugar Lake, MN, and 9 F1 individuals
286 from American Sport Fish hatchery.

287 DNA was extracted from both blood samples and fin clips, with the source depending on
288 scenarios in which different samples were collected (electrofishing/hatchery, etc.). Briefly,
289 approximately 20 mg of fin clip samples or 200 µL of blood were isolated using the Qiagen
290 DNeasy kit following the manufacturer's specifications (Qiagen). DNA concentration and purity
291 were estimated using a NanoDrop ND-2000 UV-VIS Spectrophotometer as well as by
292 electrophoresis on a 1.5% agarose gel.

293 The Sequenom MassARRAY platform (Sequenom, San Diego, CA, USA) was employed
294 to validate a subset of identified SNPs. Sequenom assays were designed using the MASSARRAY
295 ASSAY DESIGN Software with the goal to maximize multiplexing of 40 SNPs per well. Only
296 SNPs with at least a 100 bp flanking region on either side of the polymorphic site were selected
297 for the assay design. Amplification and extension reactions were performed using 20 ng (2 µL of
298 10 ng DNA) of DNA per sample and utilizing the iPLEX Gold Reagent Kit according to the
299 manufacturer's protocols. SNP genotypes were called using the SEQUENOM SYSTEM TYPER
300 4.0 Analysis software. This software uses a three parameter model to calculate the significance of

each genotype. A final genotype was called and assigned a particular name (e.g. conservative, moderate, aggressive, user call) based on the relative significance. No calls also were noted (e.g. low probability, bad spectrum). Individuals with lower than 90% call rates were removed or rerun. Based on initial screening, a subset of validated SNPs were merged into a single 25-plex multiplex panel through redesign of their extension primers. Another 38-plex marker panel from this dataset composed of "handpicked" SNPs based on sequence homology to important biological functions was also developed. The final multiplexes were run on both multi-plex panels for final validation.

Statewide population evaluation

Sample collection and genotyping

To test performance of these markers and to revisit previous genetic evaluations of stocks in Alabama (Norgren et al.1986; Maceina & DiCenzo 1995) a total of 1,736 wild fish from 29 reservoirs or small rivers and 9 river drainages throughout Alabama. Samples, in the form of fin clips preserved in 95% ethanol, were collected by ADCNR, Georgia Department of Natural Resources (GDNR), or by other fisheries lab-groups at Auburn University (Ireland Center and Sammons Lab). DNA was extracted from fin clips and assayed with both multiplexes (61 markers total) using the Sequenom MassARRAY, as previously described. All individuals had >90% call rate (≥55/61 SNPs).

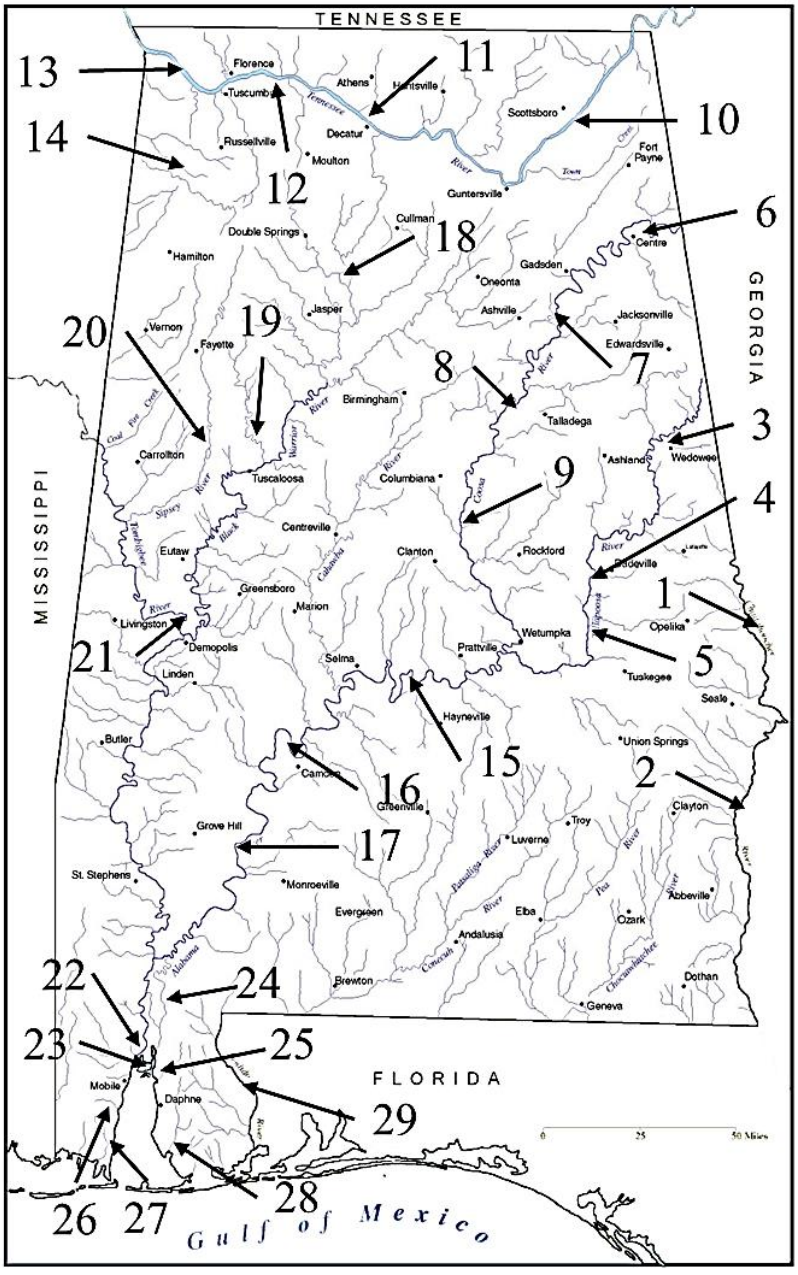
For each population (location and river drainage, including validation samples) NLMB allele frequency, FLMB allele frequency, mean heterozygosity and Q-value were quantified (Table 7). Only NLMB allele frequency, mean heterozygosity, and Q-value were used for analysis because FLMB allele frequency is the reciprocal of NLMB allele frequency and would be redundant.

The freshwater ecosystems of Alabama are a series of impoundments interconnected by a vast network of river systems. In order to look at the over-all genetic trends of LMB in Alabama, samples were first grouped into larger communities by river system for analysis, followed by a more specific within river system analysis.

The samples were group into 9 river systems for initial analysis. Six of the rivers systems identified in this analysis (the Tallapoosa River, the Coosa River, the Alabama River, the Black Warrior River and the Tombigbee River), all ultimately converge at various points and flow into the Gulf of Mexico (GOM) via the Mobile-Tensaw Delta, which was also grouped as a river system

for analysis. The other two major river systems (the Chattahoochee River and the Tennessee River) either flow directly into the GOM (Chattahoochee), or in the case of the Tennessee, flow into the GOM via the Ohio River and then the Mississippi River, or through a man-made canal (completed in 1984) that connects the Tennessee River at Pickwick Reservoir to the Tombigbee River. The final grouping of samples, labeled "Other GOM Drainages", is not a congruent river system, but represents three small rivers flowing directly into Mobile Bay (Fish, Fowl, and Dog Rivers), and one flowing into Perdido Bay (Styx River). These were grouped together because they are unique river systems from the Mobile-Tensaw system, but are fairly small compared to the other systems being considered.

Figure 1 Sample location map. **Chattahoochee River System:** [1]Lake Harding [2]Lake Eufaula; **Tallapoosa River System:** [3]Harris Reservoir [4]Lake Martin [5]Yates Reservoir; **Coosa River System:** [6]Weiss Reservoir [7]Neely Henry [8]Logan Martin Reservoir [9]Lay Lake; **Tennessee River System:** [10]Lake Guntersville [11]Wheeler Reservoir [12]Wilson Reservoir [13]Pickwick Reservoir [14]Bear Creek Reservoir; **Alabama River System:** [15]Jones Bluff [16]Miller's Ferry [17]Claiborne; **Black Warrior River System:**[18]Lewis Smith Reservoir [19]Lake Tuscaloosa; **Tombigbee River System:** [20]Sipsey River [21]Demopolis; **Mobile-Tensaw River System:** [22]Big Bayou Canot [23]Crab Creek [24]Tensaw Lake [25]D'Olive Bay; **Other GOM River Systems:** [26]Dog River [27]Fowl River [28]Fish River [29]Styx River



Statistical Analysis

STRUCTURE (version 2.2; Pritchard et al. 2000) clustering analysis was carried out on all samples (k=2, burnin=100,000, Markov Chain Monte Carlo (MCMC) reps after burnin = 200,000). From this analysis, Q-values were assigned to each individual. The Q-value indicates the proportion of times that individual was assigned to the NLMB cluster during the MCMC runs; the reciprocal indicates the proportion of times assigned to the FLMB cluster. In other words a reported Q-value of 1 indicates a bass is likely a NLMB and a Q-value of 0 indicates that a bass is likely a FLMB, with numbers near 0.5 indicating a putative F1. GENEPOP 4.0 (Rousset 2008) was used to calculate pairwise F_{st} values between populations (locations, and river systems). Fst value is the fixation index, and can be interpreted as the higher the value the more genetic distance between the populations being compared. The NLMB allele percentages, mean heterozygosity, and Q-value of each population (location and river systems) were calculated and compared using Kruskal-Wallis rank sum test and pairwise Mann-Whitney tests with a Bonferroni correction in RStudio (V. 0.98.1102).

Lake Guntersville evaluation

Lake Guntersville sample collection, aging and genotyping

Guntersville genotype analysis

Lake Guntersville, impounded in 1939, with a worldwide reputation for trophy LMB fishing opportunities, is the largest impoundment in the state of Alabama. In 2012, Lake Guntersville was ranked by Bassmaster Magazine as the third best bass fishing lake in the United States. This reputation and performance of the LMB bass fishery in the reservoir has a significant positive direct and indirect economic impact on the surrounding cities through tournament angling and recreational angling (Snellings 2015). Based on this economic significance of this population as well as the considerable investment of ADCNR has committed to hatchery and LMB stocking programs, understanding the genotypic influence on trophy largemouth bass in the reservoir is highly important.

The goal for this section of the study was to utilize the diagnostic SNP markers to identify which genotypes are contributing to the larger fish within the Lake Guntersville population. The initial strategy was to look mainly at the largest fish in the population by sampling tournaments. Fin clips from a total of 42 LMB were collected from 3 tournaments, including the Bassmaster Classic held in February 2014.

Genetic samples with corresponding length and Wt data were also collected by ALDCNR (N=15) as well as by other AU lab groups in 2014 (N=54) via electrofishing. Since tournament fish were not able to be sacrificed for aging, and age data was not available for the other samples, a total of 364 LMB were collected from 18 sites by ALDCNR electrofishing boats on March 24th and 25th of 2015. Sample sites are shown in figure 2. Fin clips were taken and stored in 95% ethanol and sagittal otoliths were removed, cleaned with water and stored dry. Total lengths (TL)

393 and weights (Wt) were recorded to the nearest mm, and 50 g respectively. Sex was observed and
394 recorded for each fish.

395 Relative weight (*Wr*) was calculated from the formula by Wege & Anderson (1978) for all
396 LMB with TL and Wt data:

397 $$Wr = (W/W_s)*100$$

398 Where W is the weight in grams and W_s is the standard weight from the LMB standard weight
399 curve (Henson 1991).

400 Ages were determined using otoliths with two independent readers. Because fish were
401 collected in the spring, before the April-July period that annuli formation occurs (Taubert &
402 Tranquilli 1982; Crawford et al. 1989), the age of the fish was considered to be the number of
403 observed annuli plus one. Otoliths were initially read whole-view, but, because of reported
404 inaccuracy in whole-view reads in older fish (Hoyer et al. 1985), all otoliths determined to have
405 two or more annuli, and those which had disagreement between readers, were sectioned and read
406 following the method described by Maceina et al.(1988). Disagreement between readers of
407 sectioned otoliths were revisited by readers together and a consensus was reached.

408 Genomic DNA from each sample was extracted at the Auburn University Aquatic Genetics
409 and Genomics Laboratory using the Puregene Tissue DNA Extraction Kit (QIAGEN, USA),
410 following the manufacturers protocol. The quality and concentration was quantified using a
411 NanoDrop 2000 UV-Vis Spectrophotometer (Thermo Scientific, DE, USA). Samples were then
412 genotyped on the Sequenom MassARRAY platform (Sequenom, San Diego, CA, USA) using the
413 2 previously described multiplex panels designed for NLMB and FLMB fixed allelic differences.

414 Differences in genetic composition between year classes, between Lake Guntersville
415 sampling gear, and electrofishing sample times (Spring 2015 or not) were tested by comparing Q-

value, mean heterozygosity, and NLMB allele frequency. Comparisons were tested for significance by ANOVA and Kruskal-Wallis rank sum test when normality assumptions were violated. Q-value and mean heterozygosity were compared to values of TL, Wt, and *Wr* in pooled samples (all age classes) by fitting simple linear regressions. To eliminate any age-based bias (larger fish actually being only older fish) TL and Wt were evaluated as a function of Q-value and mean heterozygosity by age classes 1 through 4. Age-based analysis was limited to age class 1 through 4 because of the limited number of samples available for each class beyond four years. Each comparison was fitted with a linear regression and slopes were tested for significant deviation from zero.

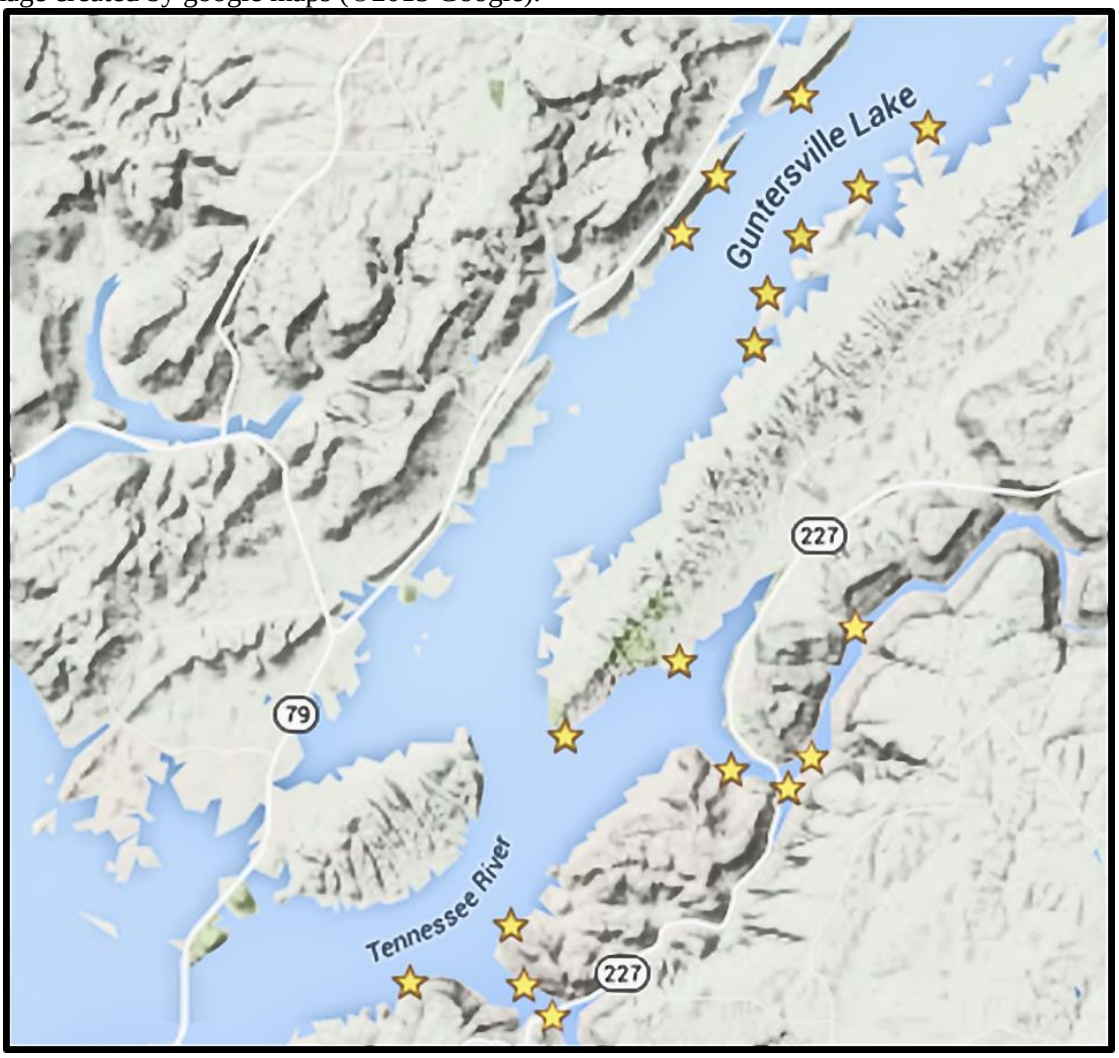
Von Bertalanffy growth curves for two relative categories of Q-value (high >0.787> low) were calculated based on length at age data using the von Bertalanffy growth formula using R-Studio (version: 0.98.1102).

$$L_{age}=L_{\infty}\,(1-e^{-k\,(age-t0)})+error$$ (Rafail 1973)

In this formula L_{age} is the length at age, L_{∞} is the maximum length, k is the growth coefficient, and t_0 is the estimated time that length was equal to zero. The k, L_{∞}, and t_0 are estimated parameters and were used to compare curves.

Finally allele usage analysis was conducted on Lake Guntersville samples with GENEPOP (Rousset 2008) by pairwise exact G tests (dememorisation=10000, batches=100, iterations per batch= 5000) and Fst estimations for each locus between trophy sized LMB (>7 lbs.), memorable size LMB (> 5 lbs.) and all others.

Figure 2. Lake Guntersville sampling map from March 24th and 25th of 2015. Yellow stars indicate sample sites. Image created by google maps (©2015 Google).



Results and Discussion

Marker Development

Transcriptome sequencing, assembly and annotation

Illumina sequencing on pooled, barcoded multi-tissue RNA samples from NLMB, FLMB and their F1 hybrid generated over 273 million 100 bp reads with >84 million reads from each sample pool (Table 1). Raw reads are archived at NCBI Sequence Read Archive (SRA) under Accession SRP042097.

Table 1 Sequencing statistics from *Micropterus* sp. RNA-seq samples.

	FLMB	NLMB	F1	Total
Number of reads	92,432,310	96,180,872	84,476,928	273,090,110
Avg read length (bp)	100	100	100	
Number of reads after trimming	88,102,383	91,029,685	80,279,632	259,411,700
Percentage kept after trimming	95.32%	94.64%	95.00%	
Avg read length after trimming (bp)	91.8	91.6	91.4	

To generate a comprehensive reference transcriptome for SNP detection, the reads from NLMB, FLMB and their F1 were pooled together to generate a composite assembly using Trinity. A total of 343,632 contigs were generated with average contig size 788.9 bp and N50 size of 1182 bp for the composite assembly (Table 1). Simultaneously, species-specific assemblies were generated using Trinity for NLMB, FLMB and their F1 hybrid. Reads were assembled into 166,934 FLMB contigs, 227,220 NLMB contigs and 123,503 F1 contigs. Average contig sizes and N50 were 984.4 and 2,096 bp, respectively, for FLMB, 1,556 and 914.4 bp, respectively, for NLMB, and 2,176 bp and 1,017.1 bp for the F1, respectively (Table 2).

463 **Table 2** Summary of Trinity *de novo* assembly results of Illumina RNA-seq data from NLMB， FLMB,
464 F1 and the composite assembly. LMB sourced from American Sportfish Hatchery.

	FLMB	NLMB	F1	Composite assembly
Contigs	166,934	227,220	123,053	343,632
Largest contig (bp)	17,360	14,346	15,275	31,075
Large contigs (≥1000bp)	47,272	67,467	36,394	69,664
Large contigs (≥500bp)	73,801	117,243	56,057	156,229
N50 (bp)	2,096	1,556	2,176	1,182
Average contig length (bp)	984.4	914.4	1,017.1	788.9

465

466 Transcriptome assemblies have been deposited to NCBI's Transcriptome Shotgun
467 Assembly (TSA) under Accessions GBFM00000000 (FLMB), GBGA00000000 (NLMB),
468 GBFO00000000 (F1). Annotation was carried out by BLAST against the Uni-Prot and NR
469 databases for NLMB, FLMB and their F1. Using the stringent criteria (E-value ≤ 1 e-20, score ≥
470 100), similar results were obtained from all three groups, with between 17,258 and 19,053
471 annotated unigene matches against UniProt and between 23,468 and 27,244 annotated unigene
472 matches against NR (Table 3). Previous work generated a transcriptome from NLMB from the
473 liver, gonad and brain tissues using 454 sequencing. This previous effort captured 7,395 annotated
474 genes, which, along with un-annotated features, were used to develop a toxicology-focused Agilent
475 microarray (Garcia-Reyero et al. 2008; Mehinto et al. 2014; Richter et al. 2014). Contig sequences
476 from this project are not publicly available and were short in length. Our results provide a more
477 comprehensive transcriptome from the two bass species, encompassing many more genes and
478 benefitting from longer contig lengths. As LMB are an important model for aquatic toxicology
479 (Denslow et al. 2007), this resource should aid future QPCR, microarray and RNA-seq studies in
480 this field as well as others.

481

Table 3 Summary of gene identification and annotation of assembled FLMB and NLMB and F1 hybrid contigs based on BLAST homology searches against various protein databases (UniProt and NR) as well as statistics of fixed interspecific SNPs identified between FLMB and NLMB sourced from American Sport Fish Hatchery.

Transcriptome and SNP coverage	FLMB	NLMB	**F1**
Unigene matches (UniProt)	19,053	22,412	17,258
Unigene matches (NR)	23,709	27,244	23,468
SNPs with coverage 20X	160	140	718
SNPs with coverage 50X	1,055	1,878	1,400
SNPs with coverage 100X	1,148	1,114	841
SNPs with coverage 500X	1,262	519	692
SNPs with coverage >500X	49	23	23
Average coverage	109	68	70
SNP annotation			
Total number of SNPs		3,674	
Annotated SNPs (NR)		3,445	
Annotated SNPs from unique genes		2,112	

Microsatellite marker identification in FLMB and NLMB transcriptomes

In FLMB, from a total of 13,354 microsatellites identified by MSATFINDER, 51.71% (6,905) had sufficient flanking regions to allow design of primers. The microsatellite-bearing contigs had 4,376 putative gene matches to the NR database from 2,576 unique genes. Similarly, in NLMB, from a total of 13,099 microsatellites identified by MSATFINDER, 59.60% (7,807) had sufficient flanking regions to allow design of primers. The microsatellite-bearing contigs had 4,964 putative gene matches to the NR database from 2,249 unique genes.

Table 4 Statistics of simple sequence repeats (SSRs) identified from FLMB and NLMB transcriptomes.

SSR mining	FLMB	NLMB
Total number of sequences examined	166,934	227,220
Total size of examined sequences(base pairs [bp])	164,323,248	207,758,528
Total number of identified SSRs	13,354	13,099
Total number of SSRs with primers	6,905	7,807
Contigs containing SSRs with primers	6,101	7,102
SSRs with primers associated with gene matches in nr	4,376	4,964
SSRs with primers associated with unique gene matches	2,576	2,249
Distribution of SSRs in different repeat types		
Dinucleotide	6,070	5,403
Trinucleotide	6,592	7,120
Tetranucleotide	633	527
Pentanucleotide	37	27

SNP identification in FLMB and NLMB

Given the complexity of determining genotypes from pooled populations, we focused on the identification of SNPs with fixed-allelic differences between species (i.e. homozygous 'A' in NLMB, homozygous 'T' in FLMB and heterozygous 'A/T' in F1) similar to the approach of Lamaze et al. (2012). We detected a set of 3,674 SNPs with fixed-allelic differences using the parameters and cut-off values described in the Materials and Methods section. These SNP contigs had 3,445 putative gene matches to the NR database from 2,112 unique genes (Table 1). Average read coverage in FLMB, NLMB and F1 was 109 reads/SNP, 68 reads/SNP and 70 reads/SNP, respectively.

Validation of SNPs by Sequenom MassARRAY

To determine the accuracy and usefulness of this resource for the study of genetic integrity and introgression of NLMB and FLMB, a subset of the fixed-allelic SNPs was tested on 119 individual bass samples in multiplex panels on the Sequenom MassARRAY. Failing SNPs, although amplifying, showed allelic patterns deviating from those expected by RNA-seq, likely

512 representing either rare alleles previously uncaptured or SNPs within duplicated genes. Ultimately
513 63 were amenable to remultiplexing through redesign of mass-specific extension primers. The
514 final 25-plex, and 38-plex SNP panels were genotyped across the 119 individuals from four
515 populations sourced from hatchery and wild populations. Details of the 25 SNP markers and 38
516 SNP markers are provided in Tables 5 and 6, including contig ID, species-specific genotypes and
517 coverage based on RNA-seq, and gene annotation. Putative functions of the encoding genes are
518 also given. Future studies examining phenotypic differences between the two bass species and
519 selective pressures on allele usage in hybrid populations may benefit from use of these markers
520 (Redenbach & Taylor 2003; Fitzpatrick et al. 2009). Two SNP markers (Contig25196 and
521 Contig11367-1) from the 25-plex were ultimately omitted from future analysis because of high
522 failure rate and were not included in any genotyping. Multiplex primer information are provided
523 in Appendix 1.

Table 5 Details of 25 SNP multiplex, with genotype, reads coverage in FLMB, NLMB and F1 largemouth bass with the gene annotation.

Contig	Position	FLMB Genotype	NLMB Genotype	F1 Genotype	Gene Name	Function
Contig2930	875	T(232)	A(304)	T/A(171)	Carboxypeptidase D	Immune (1)
Contig26936	5048	T(191)	G(134)	T/G(120)	Splicing factor, proline-and glutamine-rich	Immune/stress (2)
Contig25677	809	C(179)	T(158)	C/T(195)	Vacuolar protein-sorting-protein 25	Endocytosis (3)
Contig17385	4548	A(298)	G(269)	G/A(461)	Clustered mitochondria protein homolog	Mitochondrial
Contig8751	2930	G(212)	A(234)	A/G(236)	Mitochondrial glutamate carrier 1	Glucose homeostasis (4)
Contig34438	153	T(167)	A(66)	A/T(62)	Protein kinase C and casein kinase substrate	Endocytosis (5)
Contig4716	1720	G(430)	A(285)	A/G(406)	CpG-binding protein-like	Expression regulation (6)
Contig10770	6042	T(119)	A(258)	T/A(87)	Protein VPRBP-like	Immune (7)
Contig25196	1374	A(99)	G(71)	A/G(129)	Putative transferase CAF17 homolog, mitochondrial	Heme biosynthesis (8)
Contig5903	2988	T(652)	A(326)	T/A(480)	Calcium/calmodulin-dependent protein kinase type II	Neural function (9)
Contig33105	2477	C(99)	T(196)	T/C(81)	Acyl-CoA dehydrogenase, very long chain	Fatty acid metabolism (10)
Contig15421	896	T(226)	A(252)	A/T(168)	Serine incorporator 1	Lipid biosynthesis (11)
Contig35139	851	T(400)	C(292)	C/T(207)	Repressor of RNA polymerase III transcription MAF1	Nutrient-dependent growth (12)
Contig2993	1992	T(458)	A(259)	T/A(335)	DNAJ homolog subfamily C member 7	Steroid receptor chaperone (13)
Contig11367	889	A(154)	G(164)	G/A(369)	Kinesin-like protein KIF22	Neural function (14)
Contig11367	748	G(96)	A(120)	A/G(289)	Kinesin-like protein KIF22	Neural function (14)
Contig35112	957	A(277)	C(362)	A/C(103)	Calreticulin	Chaperone (15)
Contig33087	1916	A(466)	G(345)	A/G(355)	Heat shock protein 60 kDa, mitochondrial	Chaperone (16)
Contig6106	1199	T(111)	C(168)	C/T(232)	Fanconi anemia group F protein	DNA repair (17)
Contig20911	1554	G(149)	A(215)	A/G(308)	Mitochondrial import receptor subunit TOM40	Mitochondrial (18)
Contig19092	233	G(627)	T(269)	G/T(237)	Nonspecific cytotoxic cell receptor protein-1	Immune (19)
Contig31992	101	C(1686)	T(72)	C/T(383)	Choriogenin L	Reproductive (20)
Contig5885	2325	G(485)	A(149)	G/A(506)	Spermatogenesis associated 2-like	Reproductive (21)
Contig31857	326	C(831)	T(337)	C/T(481)		
Contig32455	234	C(1226)	A(514)	A/C(585)		

Note: 1.Hadkar & Skidgel 2001 2.Imamura et al. 2014 3.Yorikawa et al. 2005 4.Casimir et al. 2009 5.Goh et al. 2012 6.Ansari et al. 2008 7.Kassmeier et al. 2012 8.Mandilaras & Missirlis 2012 9.Rodrigues et al. 2004 10.Tucci et al. 2010 11.Inuzuka et al. 2005 12.Rideout et al. 2012 13.Moffatt et al. 2008 14.Blaker-Lee et al. 2012 15.Wang et al. 2012 16.An et al. 2014 17.Zhao et al. 2014 18.Bender et al. 2013 19.Cai et al. 2013 20.Bugel et al. 2014 21.Onisto et al. 2001

524

Table 6 Details of 38 SNP multiplex, with genotype, reads coverage in FLMB, NLMB and F1 largemouth bass with the gene annotation.

Contig	Position	FLMB Genotype	NLMB Genotype	F1 Genotype	Gene Name	Function
Contig12358	4347	G(21)	A(178)	G/A(23)	Carbonic anhydrase 5B, mitochondrial precursor	Mitochondrial Ion Transport
Contig12388	222	T(111)	A(58)	T/A(69)	Interferon regulatory factor 2-binding protein 2-A	Expression Regulation
Contig1240	2826	T(67)	C(62)	T/C(13)	Tumor necrosis factor receptor 1	Immune
Contig13020	1550	G(212)	A(146)	A/G(144)	Protein kinase, cAMP-dependent, regulatory	Kinase
Contig15950	3646	T(52)	C(38)	T/C(29)	V-type proton ATPase 116 kDa subunit a isoform 2	Ion Transport
Contig16665	2088	A(13)	C(58)	C/A(16)	Aquaporin 7	Membrane
Contig17151	814	A(236)	T(267)	A/T(52)	Proto-oncogene protein c-Fos	Expression Regulation
Contig18101	3609	C(39)	A(31)	A/C(57)	Pepatic lipase	Fatty Acid Metabolism
Contig1826	318	T(192)	A(46)	A/T(112)	Growth arrest-specific 8	Reproduction
Contig18667	326	G(126)	A(69)	G/A(100)	Peptidyl-prolyl cis-trans isomerase, mitochondrial'	Protein Folding
Contig19961	828	T(255)	C(126)	C/T(161)	Interleukin enhancer binding factor 2	Immune
Contig20908	6427	A(201)	G(59)	A/G(34)	SPARC related modular calcium binding 1	Ocular Function
Contig21621	5166	G(26)	A(80)	G/A(27)	Angiotensin-converting enzyme	Metal Binding
Contig21676	856	G(1149)	A(35)	G/A(306)	CD9 antigen	Membrane
Contig21917	2492	C(314)	T(122)	T/C(224)	Insulin-induced gene 2 protein-like	
Contig2242	1101	T(113)	C(32)	C/T(47)	Trimethyllysine dioxygenase, mitochondrial precursor	Biosynthesis
Contig22709	1569	T(151)	G(56)	G/T(88)	Hepatoma-derived growth factor-related protein 2	Cellular regulation
Contig22803	556	T(244)	G(126)	G/T(130)	Interferon regulatory factor 9	Expression Regulation
Contig23008	878	G(224)	T(51)	G/T(73)	Follistatin-related protein 1 precursor	Ion Binding
Contig23633	1578	C(83)	T(52)	T/C(71)	V-ATPase subunit A	ATP Metabolism
Contig2635	166	A(218)	G(51)	G/A(133)	Growth arrest-specific 7	Neural Development
Contig28601	2907	A(135)	G(131)	A/G(75)	Interleukin-1 receptor type 1-like	Cellular regulation
Contig2880	197	A(101)	T(56)	A/T(30)	C-X-C chemokine receptor type 3A isoform 1	Chemotaxis
Contig31979	1197	T(240)	A(129)	A/T(136)	Tumor suppressor protein p53	Expression Regulation
Contig3296	1848	T(80)	G(23)	T/G(27)	Sodium/potassium-transporting ATPase subunit	Ion Transport
Contig3379	1277	G(130)	A(43)	G/A(55)	Hsp90 co-chaperone Cdc37-like 1	Protein Folding

Table 6 continued Details of 38 SNP multiplex, with genotype, reads coverage in FLMB, NLMB and F1 largemouth bass with the gene annotation.

Contig	Position	FLMB Genotype	NLMB Genotype	F1 Genotype	Gene Name	Function
Contig3616	1054	A(212)	G(74)	A/G(56)	Sperm-associated antigen 1	Cellular metabolism
Contig36172	1764	T(363)	C(233)	T/C(162)	Succinate dehydrogenase iron-sulfur subunit	Mitochondrial
Contig4773	3667	C(136)	T(54)	T/C(54)	Growth arrest-specific protein 6-like	Ion Binding
Contig4919	267	T(263)	G(141)	G/T(127)	V-type proton ATPase subunit S1-like	ATP Metabolism
Contig4936	206	C(275)	T(73)	C/T(171)	N-acetylglutamate synthase, mitochondrial-like	Mitochondrial
Contig5713	1304	A(147)	G(60)	A/G(55)	Calcium-binding mitochondrial carrier protein SCaMC	Ion Transport
Contig6127	1788	C(49)	G(55)	G/C(11)	Na-K-2Cl cotransporter	Ion Transport
Contig6920	1282	C(109)	T(70)	T/C(102)	carnitine O-palmitoyltransferase 2, mitochondrial-like	Mitochondrial metabolism
Contig8717	243	T(86)	C(255)	T/C(57)	Pyruvate dehydrogenase E1 component subunit beta	Mitochondrial metabolism
Contig9758	2937	G(319)	T(110)	T/G(180)	peroxisomal 3-ketoacyl-CoA thiolase A	Mitochondrial metabolism
Contig9870	2987	G(38)	A(21)	A/G(18)	suppressor of cytokine signaling 5	Intercellular Signaling

525

A total of 20 individuals were run on multiple plates with 99.6% of genotypes matching among technical replicates (data not shown). As summarized in Table 7, 'pure' Florida bass (n = 53), based on previous microsatellite genotyping, had the FLMB allele in 100% of genotypes, while 'pure' largemouth bass (n = 57) had the NLMB allele in 99% of genotypes on average. Known F1 hybrids (n = 9) were heterozygous (50% FLMB, 50% NLMB allele frequencies) at all 61 loci. To avoid ascertainment bias, in all cases, we genotyped additional 'pure' individuals from populations not present in the original RNA-seq pools. In our attempt to avoid natural or anthropogenic 'intergrade' individuals, diagnostic markers were validated on fish from the more isolated edges of largemouth bass and Florida bass ranges. It is likely therefore that additional genetic variation will be revealed at some of the loci through future genotyping of fish closer to the still disputed intergrade zone (Bailey & Hubbs 1949; Philipp et al. 1983). In many cases, however, widespread stocking of FLMB in the southeastern US presents a significant obstacle to distinguishing between natural intraspecific variation in NLMB and historical signatures of hybridization with introduced Florida bass.

Table 7 Summary of populations genotyped with 61 fixed diagnostic SNP markers for distinguishing NLMB and FLMB. *Q-values from STRUCTURE cluster analysis where near 1.0 represents NLMB, near 0.0 represents FLMB, and near 0.5 represents putative F1 hybrid. **Stocking is represented as number on record of FLMB ever stocked.

	Population	N	NLMB Allele frequency	*S.E.*	FLMB Allele Frequency	*S.E.*	Mean Heterozygosity	*S.E.*	Q-Value*	*S.E.*	Stocking**
Validation	Florida Bass Conservation Center, FL	53	0.00	0.00	1.00	0.00	0.01	0.00	0.00	0.00	NA
	American Sport Fish (Illinois)	37	0.99	0.00	0.01	0.00	0.01	0.01	0.99	0.00	NA
	Sugar Lake, MN	20	0.98	0.00	0.02	0.00	0.01	0.00	1.00	0.00	NA
	American Sport Fish (F1 Tiger Bass)	9	0.50	0.00	0.50	0.00	1.00	0.00	0.56	0.00	NA
	Total	119									
Chattahoochee River	Lake Harding	51	0.42	0.01	0.58	0.01	0.41	0.01	0.49	0.02	NA
	Lake Eufaula	38	0.32	0.01	0.68	0.01	0.36	0.01	0.37	0.01	13,800
	Total/Average	89	0.38	0.01	0.63	0.01	0.39	0.01	0.44	0.01	
Tallapoosa River	Harris Reservoir	42	0.53	0.01	0.47	0.01	0.43	0.01	0.64	0.01	123,939
	Lake Martin	37	0.53	0.01	0.47	0.01	0.40	0.01	0.64	0.01	1,361,607
	Yates Reservoir	30	0.49	0.01	0.51	0.01	0.36	0.02	0.60	0.02	7,920
	Total/Average	109	0.52	0.01	0.48	0.01	0.40	0.01	0.63	0.01	
Coosa River	Weiss Reservoir	49	0.67	0.01	0.33	0.01	0.33	0.01	0.83	0.01	499,210
	Logan Martin	47	0.63	0.01	0.37	0.01	0.33	0.01	0.78	0.02	1,489,847
	Neely Henry	50	0.67	0.01	0.33	0.01	0.32	0.01	0.83	0.01	231,043
	Lay	70	0.51	0.02	0.49	0.02	0.37	0.02	0.64	0.03	1,372,912
	Total/Average	216	0.61	0.01	0.39	0.01	0.34	0.01	0.76	0.01	

Table 7 continued Summary of populations genotyped with 61 fixed diagnostic SNP markers for distinguishing NLMB and FLMB. *Q-values from STRUCTURE cluster analysis where near 1.0 represents NLMB, near 0.0 represents FLMB, and near 0.5 represents putative F1 hybrid. **Stocking is represented as number on record of FLMB ever stocked.

	Population	N	NLMB Allele frequency	*S.E.*	FLMB Allele Frequency	*S.E.*	Mean Heterozygosity	*S.E.*	Q-Value*	*S.E.*	Stocking**
Tennessee River	Guntersville	491	0.67	0.00	0.32	0.00	0.39	0.00	0.78	0.00	571,119
	Wheeler	104	0.84	0.01	0.16	0.01	0.20	0.01	0.94	0.00	799,970
	Wilson	50	0.83	0.01	0.17	0.01	0.24	0.01	0.94	0.00	223,046
	Pickwick	46	0.84	0.01	0.16	0.01	0.22	0.02	0.95	0.01	97,544
	Bear Creek	50	0.71	0.01	0.29	0.01	0.31	0.02	0.86	0.02	283,810
	Total/Average	741	0.72	0.00	0.27	0.00	0.34	0.00	0.83	0.00	
Alabama River	Jones Bluff	36	0.70	0.01	0.30	0.01	0.18	0.01	0.93	0.01	192,551
	Miller's Ferry	49	0.72	0.00	0.28	0.00	0.15	0.01	0.97	0.00	338,898
	Claiborne	20	0.73	0.01	0.27	0.01	0.13	0.01	0.97	0.01	53,236
	Total/Average	105	0.72	0.00	0.28	0.00	0.16	0.01	0.96	0.01	
Black Warrior River	Lewis Smith	50	0.72	0.01	0.28	0.01	0.19	0.01	0.95	0.01	1,917,753
	Tuscaloosa	20	0.71	0.01	0.29	0.01	0.23	0.02	0.91	0.01	378,709
	Total/Average	70	0.72	0.01	0.28	0.01	0.21	0.01	0.94	0.01	
Tombigbee River	Sipsey River	88	0.73	0.00	0.27	0.00	0.16	0.00	0.97	0.00	NA
	Demopolis	139	0.72	0.01	0.28	0.01	0.20	0.01	0.93	0.01	1,021,470
	Total/Average	227	0.72	0.00	0.28	0.01	0.18	0.01	0.95	0.01	

Table 7 continued Summary of populations genotyped with 61 fixed diagnostic SNP markers for distinguishing NLMB and FLMB. *Q-values from STRUCTURE cluster analysis where near 1.0 represents NLMB, near 0.0 represents FLMB, and near 0.5 represents putative F1 hybrid. **Stocking is represented as number on record of FLMB ever stocked. *** Low number of individuals sampled, so results may be unreliable.

	Population	N	NLMB Allele frequency	*S.E.*	FLMB Allele Frequency	*S.E.*	Mean Heterozygosity	*S.E.*	Q-Value*	*S.E.*	Stocking**
Mobile-Tensaw Delta	Big Bayou Canot	24	0.73	0.01	0.27	0.01	0.13	0.01	0.99	0.00	NA
	Crab Creek	25	0.72	0.01	0.28	0.01	0.12	0.01	0.99	0.00	NA
	D'Olive Bay	24	0.71	0.02	0.29	0.02	0.14	0.01	0.95	0.02	NA
	Tensaw Lake	29	0.72	0.01	0.28	0.01	0.12	0.01	0.97	0.01	NA
	Total/Average	102	0.72	0.00	0.28	0.00	0.13	0.01	0.97	0.01	
Other GOM Drainages	Fish River***	5	0.59	0.07	0.41	0.07	0.25	0.06	0.78	0.11	NA
	Dog River	22	0.68	0.03	0.31	0.03	0.15	0.01	0.93	0.04	NA
	Fowl River	30	0.61	0.02	0.39	0.02	0.25	0.02	0.81	0.02	NA
	Styx River	20	0.62	0.02	0.38	0.02	0.24	0.03	0.80	0.02	NA
	Total/Average	77	0.63	0.01	0.37	0.01	0.22	0.01	0.84	0.02	

541

Statewide population evaluation

Analysis by River System

Overall higher FLMB influence and heterozygosity was observed in the rivers in the eastern part of Alabama relative to the rest of the state, which is not surprising considering the proximity to the originally identified intergrade zone (Bailey and Hubbs 1949). Likewise, NLMB influence increased in an east to west pattern across the state, which is similar to the results found by others with allozyme markers (Norgren et al. 1986 and Maceina & Dicenzo 1995).

Even though the statewide patterns as measured with SNP markers appeared to be similar to the patterns as measured with allozymes in the past, the multiplex panels consistently indicate higher FLMB allele frequencies than the allozyme markers. One explanation for this discrepancy could be continued stocking efforts increasing the level of introgression of FLMB alleles in the 20+ years between studies. However it is more likely related to the increased sensitivity that comes with a 30-fold increase in fixed marker number.

Estimates of NLMB allele frequency, mean heterozygosity, and Q-value among river systems were tested for significance between river systems. Because of violations in normality assumptions, Non-parametric means were used. The comparisons included validation samples (pure FLMB and NLMB), so it was no surprise that Kruskal-Wallis rank sum tests were found to be significant for all parameters. Mann-Whitney non-parametric tests with Bonferroni corrections were conducted for pairwise comparisons for all parameters and are reported in the appendices, and comprehensive summary of the results can be found in table 8. A p-value cut-off of less than 0.05 were required to reject the null and infer significance. Pairwise Fst values were also calculated between river systems and are presented in the appendices.

Statewide pairwise analysis by river system (Table 8) indicates that the observed NLMB

allele percentages in the Chattahoochee, Tallapoosa, and Coosa rivers were each significantly different from all others, but were similar between the Coosa River and the small GOM drainages near Mobile Bay. The NLMB allele frequencies of the Alabama River (which is the sum of the Coosa and Tallapoosa Rivers below the fall line) was found to be similar to those in the Mobile-Tensaw Delta (downstream) and all river systems upstream on the western side of the state (Tombigbee and Black Warrior Rivers), and the Tennessee River system in the most northern part of the state.

Pairwise comparisons of Q-value estimates among river systems (table 8) indicated again that the Tombigbee, Black Warrior, and the Alabama River are all genetically similar. And that the Chattahoochee, Tallapoosa, and the Coosa River systems were all significantly different than all other river systems.

Pairwise F_{st} calculated between river-systems supported the patterns found in the pairwise Mann-Whitney tests. The overall synopsis indicated that the Alabama River LMB population (below the fall line) in the central part of the state shares genetic identity with LMB populations in river systems in the western part of the state (Black Warrior River, and Tombigbee River), while the LMB populations in river systems above the fall line (Coosa River, and Tallapoosa River) in the eastern part of the state, as well as those in the Chattahoochee River were shown to have a unique genetic identity, but some genetic identity was observed between geographic neighbors in these three populations. The Tennessee River system, the Mobile-Tensaw Delta, and the other GOM drainages group were shown to have a more complicated genetic relationship with the other systems, which would be expected for the Tennessee River, based on it being part of non-adjacent major river system (the Mississippi River), but is surprising for the other two groups because of their proximity and, in the case of the Mobile-Tensaw Delta, interconnectivity with many of the

other systems. These systems are also great examples that demonstrate the difference between Q-value estimates and the other metrics of mean heterozygosity and NLMB allele frequency.

In the most northwest corner of the state, the Tennessee River reservoirs of Pickwick, Wheeler and Wilson, showed the highest average NLMB allele frequencies (84%, 83%, and 84% respectively). This is not surprising, considering that the Tennessee River was the only river sampled (besides the NLMB used for validation) that is part of the Mississippi River drainage, which is the major region identified as the native range of the NLMB (Barthel et al. 2010). Interestingly, when the Q-value estimates from the Bayesian analysis are considered, the NLMB influence is exaggerated in all populations compared to the NLMB allele frequency, but not at a consistent proportion. For example the Q-value estimates for the lower Alabama River and Mobile-Tensaw Delta indicate a high level of NLMB influence (Q=0.97), even indicating that some are close to pure NLMB populations (Q=0.99 in Big Bayou Canot and Crab Creek), while they have a reported NLMB allele frequency of only ~72%). Likewise the previously mentioned Tennessee River populations which showed high NLMB allele frequencies appear more moderate (Q=~0.94) when Q-value is considered. This is because Q-value is not simply a proportion of alleles from each population, it is a probability of that individual being from the assigned population. So Q-value actually relies on both the number of NLMB alleles present as well as the number of heterozygous loci present, when the model is constructed and the estimates are given. This make Q-values much more informative for analysis.

The Q-value averages for the Mobile Tensaw Delta populations tracked closely to the results of allozyme studies that place isolated populations of NLMB in the Mobile River Delta (Phillip 1983, Norgren et al.1986, Maceina & DiCenzo 1995). Differences in mean heterozygous loci in this Delta population (table 8) was also found to be unique in relation to other populations

612 in the state, potentially indicating low levels of hybridization. These Delta populations that appear
613 to be an isolated NLMB population, often referred to as 'Delta Bass', have physiological and
614 ecological differences in comparison to other LMB (DeVries et al.2014) and may be a unique
615 genetic stock that needs to be considered during marker development, but is beyond the scope of
616 this thesis.

617

Table 8. Summary of mean NLMB allele frequency, heterozygosity, and Q-value for populations genotyped with 61 fixed diagnostic SNP markers for distinguishing NLMB and FLMB for each of the nine river systems or drainage groupings. Letters beside values indicate that p-value of pairwise Mann-Whitney test with Bonferroni correction indicated a lake of significant difference between populations that have the same letters (significance cut-off of p=0.05). *Indicates that mean value for population is significantly different from all other populations shown. **Q-values from STRUCTURE cluster analysis where near 1.0 represents NLMB, near 0.0 represents FLMB, and near 0.5 represents putative F1 hybrid.

Population	N	NLMB Allele frequency	Mean Heterozygosity	Q-Value**
Tennessee River	741	0.72 a	0.34 a	0.83 a
Chattahoochee River	89	0.38*	0.39 b	0.44*
Tallapoosa River	109	0.52*	0.40 b	0.63*
Coosa River	216	0.61 b	0.34 a	0.76*
Alabama River	105	0.72 a	0.16 c	0.96 b
Black Warrior River	70	0.72 a	0.21 d	0.94 b
Tombigbee River	227	0.72 a	0.18 d c	0.95 b
Mobile-Tensaw Delta	102	0.72 a	0.13*	0.97*
Other GOM drainages	77	0.63 b	0.22 d	0.84 a

618

Statewide analysis by reservoir or sample location

The river systems of Alabama, punctuated by a network of major and minor locks and dams, are often biologically discontinuous with drastic variations in habitat and ecosystems within the same river system. Therefore analyzing data grouped by whole river systems may give some perspective, but analysis by impoundments or sample locations is necessary for a thorough understanding of the systems.

As with the previous river system analysis, significant differences between NLMB allele frequency, mean heterozygosity, and Q-value between each sampling location were tested. Mann-Whitney non-parametric tests with Bonferroni corrections were conducted for pairwise comparisons for all parameters and are reported in the appendices. A p-value cut-off of less than 0.05 were required to reject the null hypothesis and infer significance. A summary of this information in presented in tables 9 through 16. Pairwise Fst values were also calculated between sampling locations and are presented in the appendices.

When individual populations were evaluated with Mann-Whitney pairwise comparisons, patterns within river systems emerged. These systems will be discussed in an east-to-west direction; starting with the Chattahoochee River system and ending with the Tennessee River system in the northern part of the state.

The Chattahoochee River system, sampled at Lake Harding and then downstream in Lake Eufaula, forms part of the eastern boarder of Georgia and Alabama as it flows into the Lake Seminole in Florida and then into the GOM via the Apalachicola River. Lake Harding LMB populations were found to have a NLMB allele frequency, mean heterozygosity and Q-value of 0.42, 0.41, and 0.49 respectively. LMB populations of Lake Eufaula, which sits below the fall line

with respect to Lake Harding, were found to have a NLMB allele frequency, mean heterozygosity and Q-value of 0.32, 0.36, and 0.37 respectively. NLMB allele frequency and, Q-value differences were found to be significant at p-values of $9.33E^{-06}$, and $1.59E^{-05}$ respectively. However mean heterozygosity differences were not significant and the pairwise F_{st} estimate (0.033) suggests that there is little difference between the populations. The results indicate that the Eufaula population has higher FLMB influence.

Table 9. Summary of mean NLMB allele frequency, heterozygosity, and Q-value for populations genotyped with 61 fixed diagnostic SNP markers for distinguishing NLMB and FLMB with-in the Chattahoochee River System. Letters beside values indicate that p-value of pairwise Mann-Whitney test with Bonferroni correction indicated a lake of significant difference between populations that have the same letters (significance cut-off of p=0.05). *Indicates that mean value for population is significantly different from all other populations shown. **Q-values from STRUCTURE cluster analysis where near 1.0 represents NLMB, near 0.0 represents FLMB, and near 0.5 represents putative F1 hybrid.

Population	N	NLMB Allele frequency	Mean Heterozygosity	Q-Value**
Lake Harding	51	0.42*	0.41 a	0.49*
Lake Eufaula	38	0.32*	0.36 a	0.37*

Three reservoirs were sampled on the Tallapoosa River system (Harris Reservoir, Lake Martin, and Yates Reservoir). Even though slight decreases were observed in NLMB allele frequency, mean heterozygosity, and Q-value between LMB populations in a downstream progression, none of the differences were found to be significant by pairwise comparisons. This was supported by pairwise F_{st} values that indicate only minor differences between the populations.

Table 10. Summary of mean NLMB allele frequency, heterozygosity, and Q-value for populations genotyped with 61 fixed diagnostic SNP markers for distinguishing NLMB and FLMB with-in the Tallapoosa River System. Letters beside values indicate lake of significant differences between populations that have the same letters. *Indicates that mean value for population is significantly different from all other populations shown. **Q-values from STRUCTURE cluster analysis where near 1.0 represents NLMB, near 0.0 represents FLMB, and near 0.5 represents putative F1 hybrid.

Population	N	NLMB Allele frequency	Mean Heterozygosity	Q-Value**
Harris Reservoir	42	0.53 a	0.43 a	0.64 a
Lake Martin	37	0.53 a	0.40 a	0.64 a
Yates Reservoir	30	0.49 a	0.36 a	0.60 a

The Coosa River system (Weiss Reservoir, Neely Henry Reservoir, Logan Martin Reservoir, and Lay Lake) populations of LMB were observed to have a downstream decrease of NLMB allele frequency and Q-values. All pairwise comparisons indicated similarity, excluding NLMB allele frequencies between Lay Lake and Neely Henry Reservoir (p-value = 0.00038), and between Lay Lake and Weiss Reservoir (p-value = $6.72E^{-05}$). The pairwise F_{st} values indicate that there is a high level of genetic similarity between these populations. The significantly higher FLMB influence found in Lay compared to the other reservoirs may be related to stocking. ALDCNR stocking reports indicate 1,372,912 FLMB over the duration of stocking efforts in Alabama. This is much higher than the stocking numbers reported for Weiss and Neely Henry, which received a modest 499,210 and 231,043 FLMB respectively. Logan Martin on the other hand was not observed to have a different level of FLMB influence when NLMB allele frequency was considered, which may have some relation to the 1,489,847 FLMB reportedly stocked. Another explanation could be sampling bias from the sampling method. On Lay Lake many of the samplers were tournament angler collected; an issue to be discussed later when looking at data from Lake Guntersville.

Table 11. Summary of mean NLMB allele frequency, heterozygosity, and Q-value for populations genotyped with 61 fixed diagnostic SNP markers for distinguishing NLMB and FLMB with-in the Coosa River System. Letters beside values indicate that p-value of pairwise Mann-Whitney test with Bonferroni correction indicated a lake of significant difference between populations that have the same letters (significance cut-off of p=0.05). *Indicates that mean value for population is significantly different from all other populations shown. **Q-values from STRUCTURE cluster analysis where near 1.0 represents NLMB, near 0.0 represents FLMB, and near 0.5 represents putative F1 hybrid.

Population	N	NLMB Allele frequency	Mean Heterozygosity	Q-Value**
Weiss Reservoir	49	0.67 a	0.33 a	0.83 a
Neely Henry	50	0.67 a	0.32 a	0.83 a
Logan Martin	47	0.63 a b	0.33 a	0.78 a
Lay	70	0.51 b	0.37 a	0.64 a

LMB at three locations were sampled on the Alabama River (Jones Bluff, Miller's Ferry, and Claiborne) which represents the confluence of the Tallapoosa and Coosa River systems below the fall line. No differences were observed between populations, excluding the Q-value difference between Jones Bluff (0.93), found in the upstream portion of the system, and Claiborne (0.97) found in the lower end of the system (p-value = 0.031). Pairwise F_{st} estimates indicated that there is little genetic distance between these populations; with a negative value being reported for the Miller's Ferry and Claiborne.

Table 12. Summary of mean NLMB allele frequency, heterozygosity, and Q-value for populations genotyped with 61 fixed diagnostic SNP markers for distinguishing NLMB and FLMB with-in the Alabama River System. Letters beside values indicate that p-value of pairwise Mann-Whitney test with Bonferroni correction indicated a lake of significant difference between populations that have the same letters (significance cut-off of p=0.05). *Indicates that mean value for population is significantly different from all other populations shown. **Q-values from STRUCTURE cluster analysis where near 1.0 represents NLMB, near 0.0 represents FLMB, and near 0.5 represents putative F1 hybrid.

Population	N	NLMB Allele frequency	Mean Heterozygosity	Q-Value**
Jones Bluff	36	0.70 a	0.18 a	0.93 a
Miller's Ferry	49	0.72 a	0.15 a	0.97 a b
Claiborne	20	0.73 a	0.13 a	0.97 b

The Black Warrior River system, a tributary of the Tombigbee River, was sampled at two locations (Lewis Smith Reservoir, and Lake Tuscaloosa). Pairwise comparisons indicated that there is no difference between these reservoir populations when NLMB allele frequency and mean heterozygosity were considered. However the reported Q values of 0.95 (Lewis Smith) and 0.91 (Tuscaloosa) are different (p-value = 0.01), indicating slightly higher NLMB influence in Lewis Smith Reservoir. The low pairwise F_{st} estimates indicate there is very little genetic difference between these two reservoirs.

Table 13. Summary of mean NLMB allele frequency, heterozygosity, and Q-value for populations genotyped with 61 fixed diagnostic SNP markers for distinguishing NLMB and FLMB with-in the Black Warrior River System. Letters beside values indicate that p-value of pairwise Mann-Whitney test with Bonferroni correction indicated a lake of significant difference between populations that have the same letters (significance cut-off of p=0.05). ***Indicates that mean value for population is significantly different from all other populations shown. *Q-values from STRUCTURE cluster analysis where near 1.0 represents NLMB, near 0.0 represents FLMB, and near 0.5 represents putative F1 hybrid.

Population	N	NLMB Allele frequency	Mean Heterozygosity	Q-Value*
Lewis Smith	50	0.72 a	0.19 a	0.95*
Tuscaloosa	20	0.71 a	0.23 a	0.91*

687

688 The Tombigbee River system was also only sampled in two locations (the Sipsey River
689 and Demopolis Reservoir). Demopolis is an impoundment at the conjunction of the Black Warrior
690 and the Tombigbee River, and the Sipsey River is a small tributary of the Tombigbee thought to
691 contain some LMB that may represent the native Alabama LMB, meaning they are thought to be
692 unadulterated by FLMB stocking. Despite a reported 1,021,470 FLMB planted into Demopolis,
693 these two populations were not found to be different when NLMB allele frequencies, mean
694 heterozygosity, or Q-value were considered. A low pairwise F_{st} for this comparison was estimated,
695 confirming that these two populations within the Tombigbee system are genetically equivalent.

696

Table 14. Summary of mean NLMB allele frequency, heterozygosity, and Q-value for populations genotyped with 61 fixed diagnostic SNP markers for distinguishing NLMB and FLMB with-in the Tombigbee River System. Letters beside values indicate that p-value of pairwise Mann-Whitney test with Bonferroni correction indicated a lake of significant difference between populations that have the same letters (significance cut-off of p=0.05). *Indicates that mean value for population is significantly different from all other populations shown. **Q-values from STRUCTURE cluster analysis where near 1.0 represents NLMB, near 0.0 represents FLMB, and near 0.5 represents putative F1 hybrid.

Population	N	NLMB Allele frequency	Mean Heterozygosity	Q-Value**
Sipsey River	88	0.73 a	0.16 a	0.97 a
Demopolis	139	0.72 a	0.20 a	0.93 a

697

698

699 The Mobile-Tensaw Delta system is the river and delta system that extends from the
700 confluence of the Tombigbee and Alabama Rivers to the head of Mobile Bay. The sampling
701 locations (Tensaw Lake, Big Bayou Canot, Crab Creek, and D'Olive Bay) appeared to be a
702 homogenized population with no deviations from one another in observed NLMB allele frequency,
703 mean heterozygosity, or Q-value. This observation was supported by low pairwise F_{st} estimates.
704 Interestingly, differences were not found between observed values from these populations and the
705 Alabama River, Black Warrior, and the Tombigbee River, which corresponds to river system-wide
706 observations that linked all of the Mobile River drainage with exception of the Alabama River
707 drainage above the fall-line (the Tallapoosa and Coosa Rivers).
708 Samples were taken from four other small independent river systems that did not neatly fit
709 into any particular group for analysis. All of these rivers flow into saline water of the northern
710 GOM, so they were grouped together as 'other GOM drainages". These included one small river
711 that flows into Perdido Bay (Styx River) which marks the boarder of Florida and Alabama, as well
712 as three small rivers that flow into the saline water of Mobile Bay; one on the eastern shore (Fish
713 River) and two on the western shore (Dog and Fowl Rivers). The Fish River sample size was
714 limited (n=5), so the pairwise comparisons involving this population were often found to be
715 insignificant. Between all four populations, no differences were not found for NLMB allele
716 frequency, mean heterozygosity, and Q-value, with the exception of between Dog and Fowl Rivers
717 for all three parameters, and between Fowl River and Styx River when Q-value was considered
718 (p-value = 1.42E-08). These observations indicate that the Dog River population, which is closer
719 to the Mobile-Tensaw system, has higher NLMB influence than the populations of the other small
720 GOM drainages. The observed values for the Dog River population were not different from any
721 of the populations identified as the Mobile-Tensaw Delta system, while the observations for the

722 Fowl and Styx River populations were found to have higher FLMB influence.

Table 15. Summary of mean NLMB allele frequency, heterozygosity, and Q-value for populations genotyped with 61 fixed diagnostic SNP markers for distinguishing NLMB and FLMB with-in the Mobile-Tensaw Delta and other small GOM drainages grouping. Letters beside values indicate that p-value of pairwise Mann-Whitney test with Bonferroni correction indicated a lake of significant difference between populations that have the same letters (significance cut-off of p=0.05). *Indicates that mean value for population is significantly different from all other populations shown. **Q-values from STRUCTURE cluster analysis where near 1.0 represents NLMB, near 0.0 represents FLMB, and near 0.5 represents putative F1 hybrid.

Population	N	NLMB Allele frequency	Mean Heterozygosity	Q-Value**
Tensaw Lake	29	0.72 a	0.12 a	0.97 a
Big Bayou Canot	24	0.73 a	0.13 a	0.99 a
Crab Creek	25	0.72 a	0.12 a	0.99 a
D'Olive Bay	24	0.71 a	0.14 a	0.95 a
Dog River	22	0.68 a b c	0.15 a	0.93 a
Styx River	20	0.62 b c	0.24 a b	0.80 b
Fish River	5	0.59 a b c	0.25 a b	0.78 a b
Fowl River	30	0.61 b	0.25 b	0.81 b

723

724 The Tennessee River System flows south from Tennessee into Alabama, and then flows
725 from the northeastern corner of the state to the northwestern corner only to then flow north to join
726 the Ohio River which ultimately joins the Mississippi river where the Kentucky Illinois and
727 Missouri state boarders meet. LMB samples from all four Tennessee River reservoirs (Lake
728 Guntersville, Wheeler Reservoir, Wilson Reservoir, Pickwick Reservoir, and Bear Creek
729 Reservoir) were collected. The Lake Guntersville population (67% NLMB alleles, Q-value of
730 0.78) is different in relation to the other populations in this system (average of 81% NLMB alleles,
731 average Q-value of 0.92) (pairwise p-values range from 1.39E-18 to 2.20E-39). This high level of
732 FLMB alleles, more indicative of reservoirs in the Coosa River, is extremely curious. A similar
733 relationship was observed between the Lake Guntersville population and the other Tennessee
734 River populations when samples were assayed with allozymes in 1992-1995 (Maceina & DiCenzo
735 1995). The observed differences cannot be solely due to differences in stocking because FLMB
736 stocking rates over the past 30 years have been similar (~.3 fish/acre/year) for Guntersville,

737 Wheeler, and Wilson. And Pickwick has had considerably less stocking effort than the others, but
738 Bear Creek, a small reservoir on a tributary draining into Pickwick, has experienced intense FLMB
739 stocking (~13.7 fish/acre/year), but still retains a similar level of FLMB influence (71% NLMB
740 alleles, Q-value of 0.86) as does the Lake Guntersville population. Ancestral genetic makeup, river
741 drainage, latitude, lake characteristics, etc. are all likely impacting the success of introgressive
742 hybridization (Norgren et al. 1986). In this case, Lake Guntersville may have a different
743 environmental condition in relation to Wheeler and Wilson, (*e.g.* greater vegetative coverage)
744 perhaps making it a better environment for FLMB, and therefore aiding in the introgression of
745 FLMB alleles. The genetic composition of the founding stock in Lake Guntersville could also
746 have an effect on the current status, however allozyme results obtained prior to state-directed
747 stocking of Florida bass reported a 92.5% NLMB contribution in Lake Guntersville (Philipp et al.
748 1983). This apparent successful introgression of FLMB alleles, and its productive trophy LMB
749 fishery, makes Lake Guntersville an ideal location for assessing the effect of introgression on
750 trophy bass.

Table 16. Summary of mean NLMB allele frequency, heterozygosity, and Q-value for populations genotyped with 61 fixed diagnostic SNP markers for distinguishing NLMB and FLMB with-in the Tennessee River System. Letters beside values indicate that p-value of pairwise Mann-Whitney test with Bonferroni correction indicated a lake of significant difference between populations that have the same letters (significance cut-off of p=0.05). *Indicates that mean value for population is significantly different from all other populations shown. **Q-values from STRUCTURE cluster analysis where near 1.0 represents NLMB, near 0.0 represents FLMB, and near 0.5 represents putative F1 hybrid.

Population	N	NLMB Allele frequency	Mean Heterozygosity	Q-Value**
Guntersville	491	0.67 b	0.39*	0.78*
Wheeler	104	0.84 a	0.20 b	0.94 a
Wilson	50	0.83 a	0.24 a b	0.94 a b
Pickwick	46	0.84 a	0.22 a b	0.95 a
Bear Creek	50	0.71 b	0.31 a	0.86 b

751

Lake Guntersville evaluation

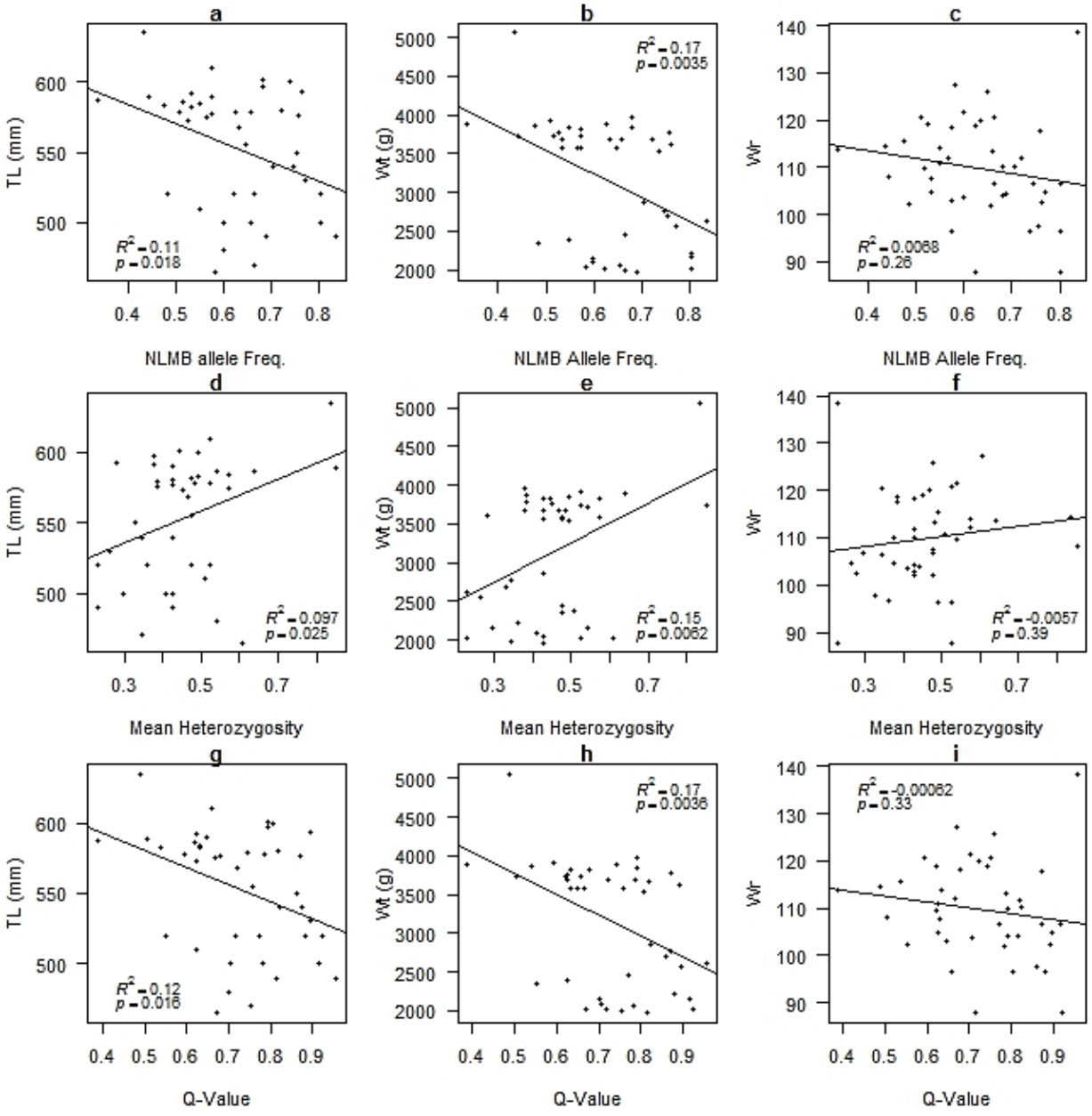
Tournament caught LMB

Analysis of the Q-value, mean heterozygosity, and NLMB allele frequencies relative to TL (mm), Wt (g), and *Wr*, were all conducted on the entire set of Lake Guntersville tournament bass samples with simple linear models. Scatterplots containing R^2 and p-values are reported in Figure 3.

These simple regressions indicate a positive correlation of TL, and Wt in relation to mean heterozygosity and a negative correlation in relation to NLMB allele frequency and Q-value. All, except the slope for *Wr* as compared to mean heterozygosity and Q-value, were significantly different than zero. However, the R^2 values were fairly small, indicating these models are weak, but suggested that there may be some relationship between FLMB and F1 genotypes and larger fish.

Because these fish were sampled from high-profile tournaments, sacrificing these fish for otolith collection and aging was not an option. Having age data with genotype and size data would help identify which genotypes were truly contributing to trophy bass. Without this information it would be difficult to ascertain whether the larger fish having higher levels of FLMB influence are larger because of better growth or simply they are older fish. So a sample of LMB from Lake Guntersville was collected in the spring of 2015 from which otoliths could be collected for aging.

Figure 3 For all Lake Guntersville tournament caught fish: **a)** TL as a function of NLMB allele frequency **b)** Wt as a function of NLMB allele frequency **c)** *Wr* as a function of NLMB allele frequency **d)** TL as a function of mean heterozygosity **f)** Wt as a function of mean heterozygosity **g)** *Wr* as a function of mean heterozygosity **h)** TL as a function of Q-value* **i)** Wt as a function of Q-value* **j)** *Wr* as a function of Q-value*. Reported p-values of <0.05 indicate that the slope of the regression is significantly different from zero. *Q-value of 0 is pure FLMB and 1 is pure NLMB



Tournament caught bass vs electrofishing collected bass

When the observed values for samples collected by the tournament anglers were compared to electrofishing samples, the tournament samples included a fish that were beyond the maximum sizes collected by electrofishing (sample distributions by Wt are shown in Figure 4). This was somewhat expected given that the tournament fish are biased in that the anglers are targeting the largest fish, and only keeping the largest fish they catch. What was not expected, was genotypic differences between electrofishing samples and tournament angler samples.

Overall analysis by Kruskal-Wallis rank sum tests indicated differences in mean heterozygosity (46%, 39%), Q-value (0.82, 0.85), and NLMB allele frequencies (63%, 71%) between tournament bass and electro-fished bass, respectively (p-value<0.05). This could be interpreted to suggest that FLMB are more catchable by angling, going against the common perception that they are considerably more angler wary(Sasaki 1961; Addison & Spencer 1972; Zolczynski & Davies 1976) However, because the tournament bass are not a randomly selected population of angled bass, this would be a poor assumption. A more appropriate conclusion would be that the largest fish caught by tournament anglers may be more FLMB influenced and heterozygous than the overall population of Lake Guntersville.

Figure 4 Frequency of largemouth bass by age collected in March 2015 from Lake Guntersville via electrofishing.

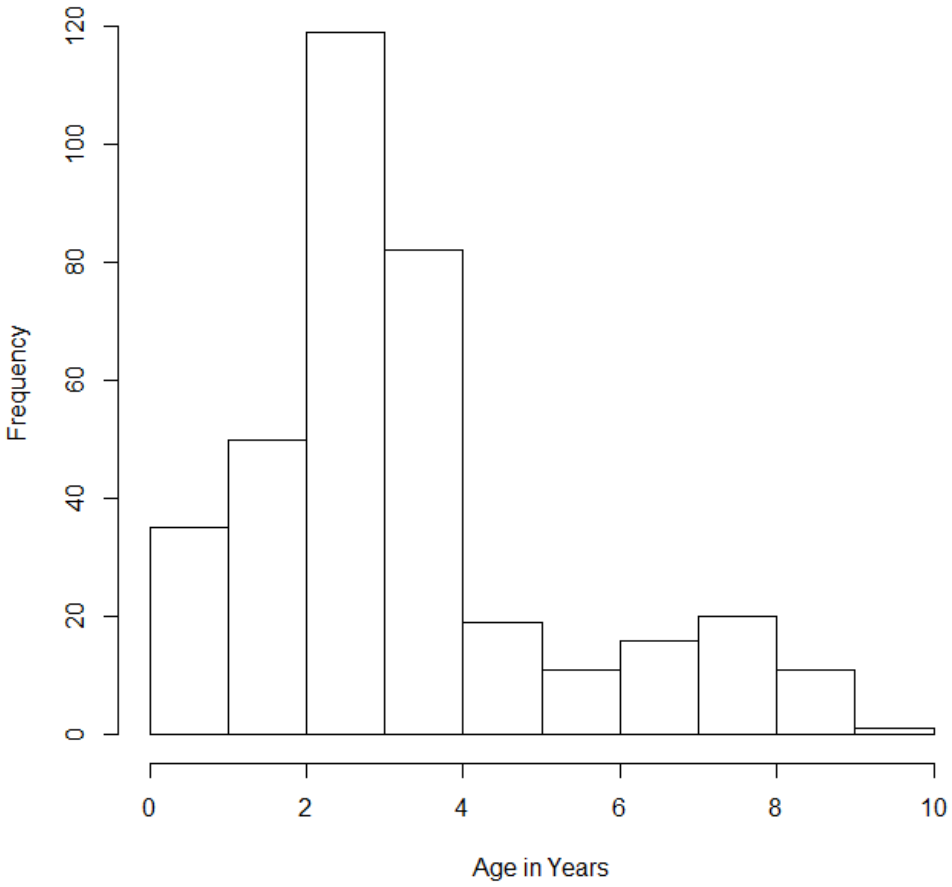
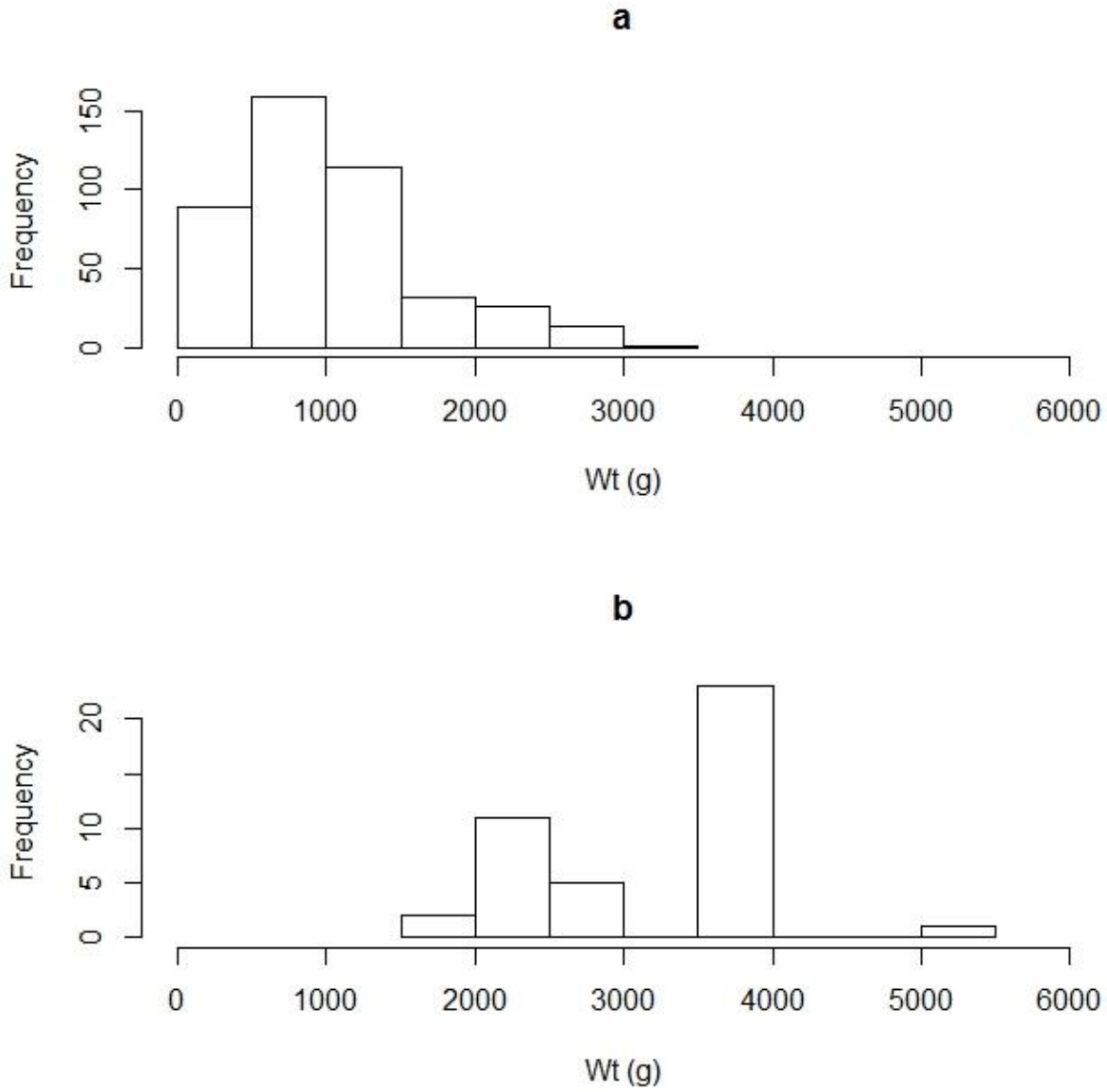


Figure 5 Frequency distribution by Wt of **a)** electrofishing samples and **b)** tournament angler samples

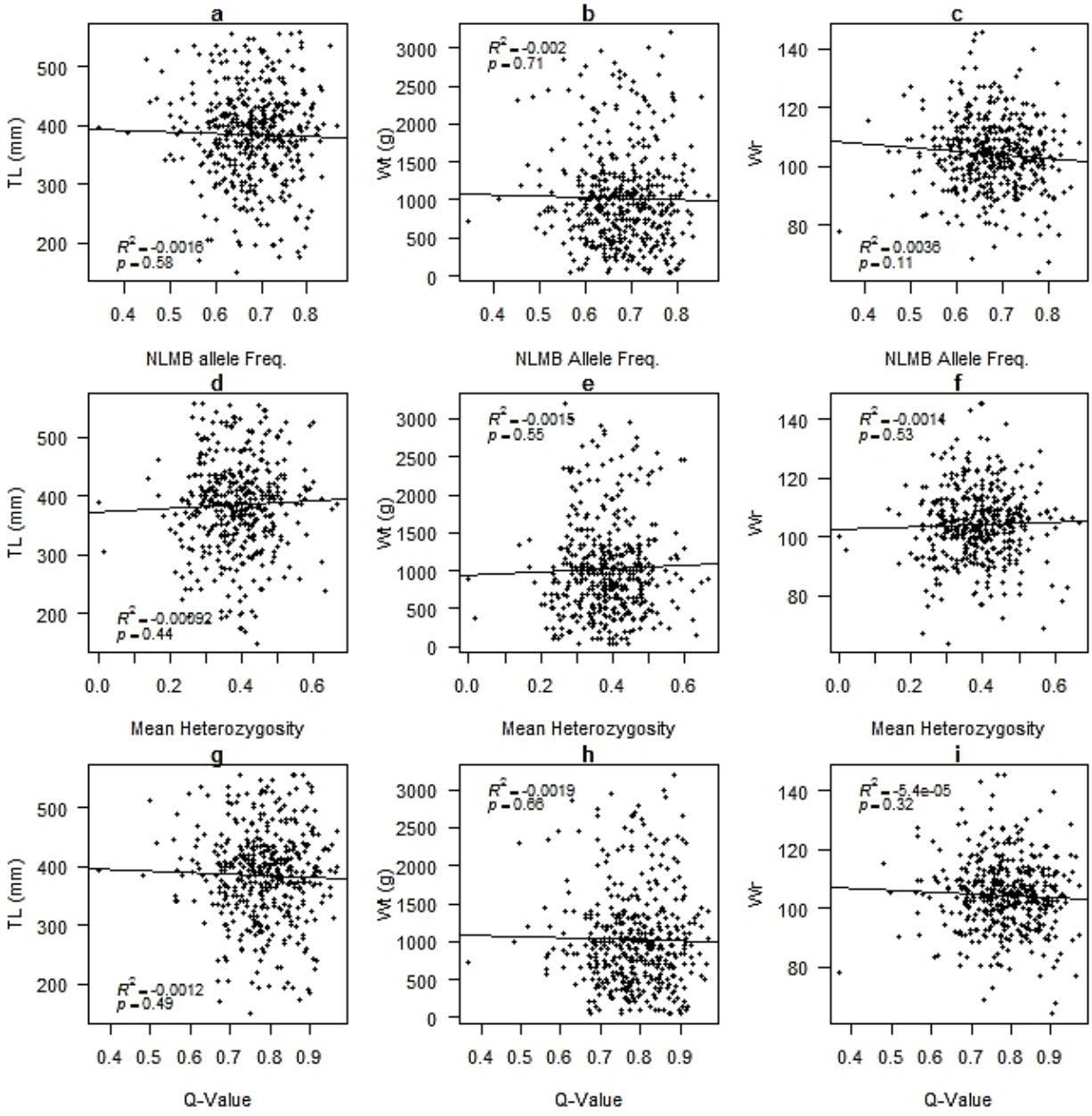


Analysis of electrofishing samples

Samples taken in March 2015 represented the majority (84%) of electrofishing samples and spawning timing differences between the subspecies could be a factor (Isely et al. 1987, Fields et al.1987, Rogers et al.2006), NLMB allele frequency, mean Heterozygosity, and Q-value between the spring 2015 samples and summer electrofishing samples taken in previous years needed to be tested. However, no genetic differences were found between these populations.

Analysis of the Q-value, mean heterozygosity, and NLMB allele frequencies relative to TL (mm), Wt (g), and *Wr*, was all conducted on the entire set of Lake Guntersville samples that were collected by electrofishing with simple linear models. Scatterplots containing R^2 and p-values are reported in Figure 6. Contrary to the tournament samples, no correlation of genotype to either size metric or *Wr* was found within electro-fished samples.

Figure 6 For all Lake Guntersville electrofishing samples: **a)** TL as a function of NLMB allele frequency **b)** Wt as a function of NLMB allele frequency **c)** *Wr* as a function of NLMB allele frequency **d)** TL as a function of mean heterozygosity **f)** Wt as a function of mean heterozygosity **g)** *Wr* as a function of mean heterozygosity **h)** TL as a function of Q-value **i)** Wt as a function of Q-value **j)** *Wr* as a function of Q-value. Reported p-values of <0.05 indicate that the slope of the regression is significantly different from zero.



Age based analysis

Size at age is an important factor in considering growth and size potential of an organism. In order to eliminate age related biases, ages were determined for all of the fish collected during the spring of 2015. A simple ANOVA, and Kruskal-Wallis rank sum test where necessary, was used to identify any differences in genotype (NLMB allele frequency, mean heterozygosity, and Q-value) between year classes. No differences between any ages (year class 1-10) for any of the parameters of NLMB allele frequency, mean heterozygosity, or Q-value. This result was in-line with results found by Dumont and Lutz-Carrillo (2011) who used 7 microsatellites to evaluate if there were age differences in allele percentages in admixed populations in order to determine if a specific age needs to be targeted for estimating a population's introgression status.

Even though this result indicated that the sample population was genetically homogenized across year classes, it was still important to investigate the possibility of size at age variation in relation to genotype. A linear model analysis, as was previously conducted on the overall samples, was conducted, but for each age class separately (age 1-4). This analysis was limited to ages 1-4 because of the small sample size for age classes beyond age 4. Linear models (figures 7 through 10) showed no significant trends when slope is compared to a slope of zero, indicating that there was in fact no size advantage based on genotype. However there was a significant interaction between *Wr* and genotype at age 4 exclusively (Figure 10).

In age four fish NLMB allele frequency and Q-value were both negatively correlated to *Wr*. Likewise there was a significant positive correlation between mean heterozygosity and *Wr*. This indicated that LMB in Guntersville with a higher *Wr*, may have had more FLMB influence, an interaction that was not significant over the entire sample set, nor in the other age classes tested. This may be indicative of a body condition advantage for FLMB later in life. Without more

samples for ages beyond four years this conclusion would have to remain speculative. The lack of an apparent interaction of this nature in the analysis of year classes 1-3 as well as in the overall analysis could be an artifact of the sampling method. A scale with low sensitivity (measured to the nearest 50g) was used, causing the *Wr* calculated for smaller fish in the lower age classes to lack some sensitivity, which could have caused and possible correlation to be obscured until the older fish where the increment of measurement was a much lower percent of the average body weight.

Von Bertalanffy Growth Curves

Von Bertalanffy growth curves were calculated for two subsets of the population (high >0.787> low). The curves are plotted in Figure 11 and parameter estimates are given in table 17. The lack of significant difference in the growth curves suggests that Q-value, and, by association, genotype may not have as much influence on growth in Lake Guntersville LMB populations. It is possible that this population may not have enough genetic variation to detect differences with this small of a sample size. A larger sample from this lake (N=~2,000), may be more appropriate for this type of analysis. However, unless a larger sample size actually included a higher number of individuals in the age classes beyond age four, the analysis would be equally limited by excluding many of the oldest and potentially largest fish that may in fact be exhibiting some type of genetic advantage.

Figure 7 For Age 1 Guntersville samples: **a)** TL as a function of NLMB allele frequency **b)** Wt as a function of NLMB allele frequency **c)** *Wr* as a function of NLMB allele frequency **d)** TL as a function of mean heterozygosity **f)** Wt as a function of mean heterozygosity **g)** *Wr* as a function of mean heterozygosity **h)** TL as a function of Q-value **i)** Wt as a function of Q-value **j)** *Wr* as a function of Q-value. Reported p-values of <0.05 indicate that the slope of the regression is significantly different from zero.

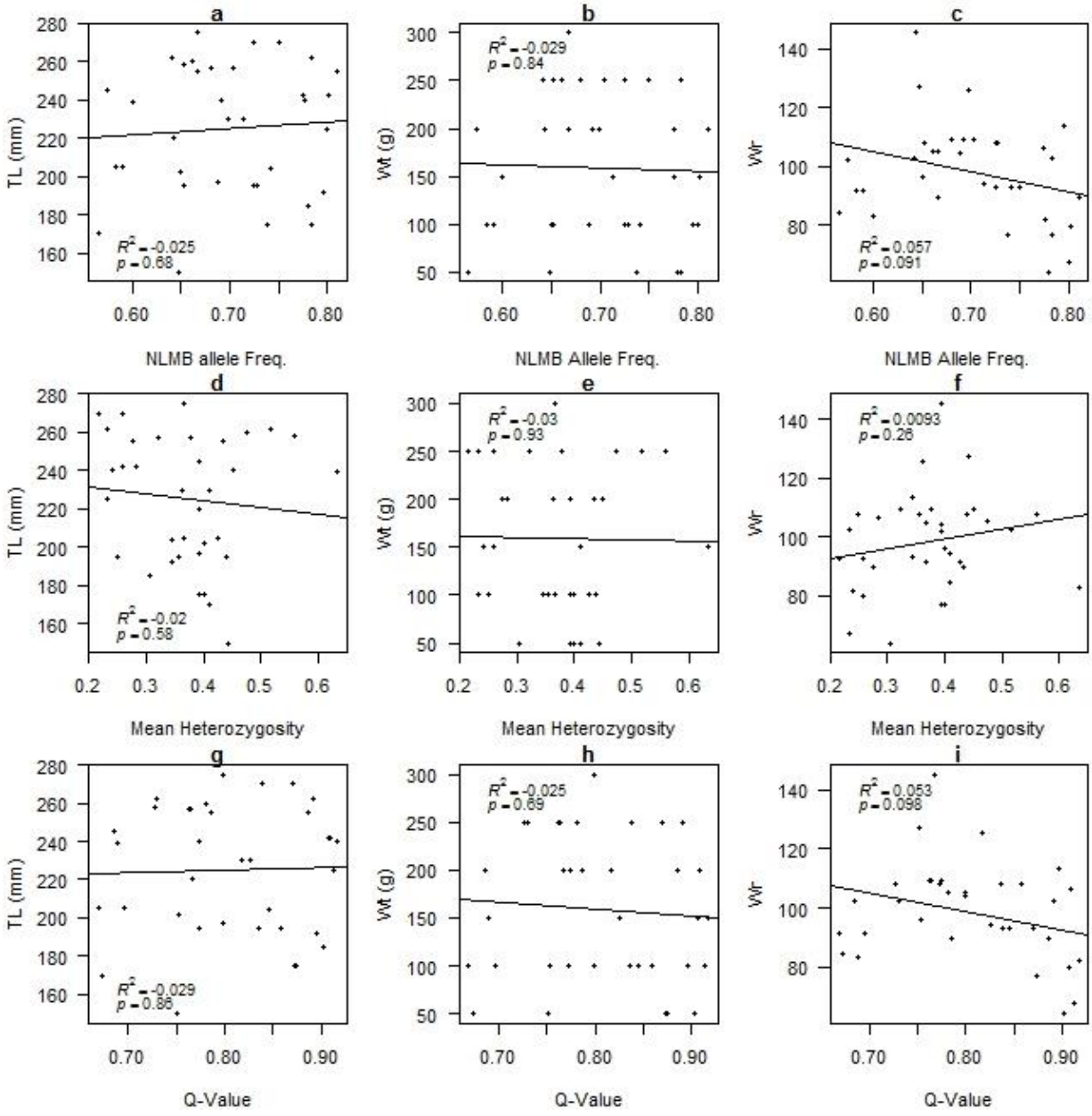


Figure 8 For Age 2 Guntersville samples: **a)** TL as a function of NLMB allele frequency **b)** Wt as a function of NLMB allele frequency **c)** *Wr* as a function of NLMB allele frequency **d)** TL as a function of mean heterozygosity **f)** Wt as a function of mean heterozygosity **g)** *Wr* as a function of mean heterozygosity **h)** TL as a function of Q-value **i)** Wt as a function of Q-value **j)** *Wr* as a function of Q-value. Reported p-values of <0.05 indicate that the slope of the regression is significantly different from zero.

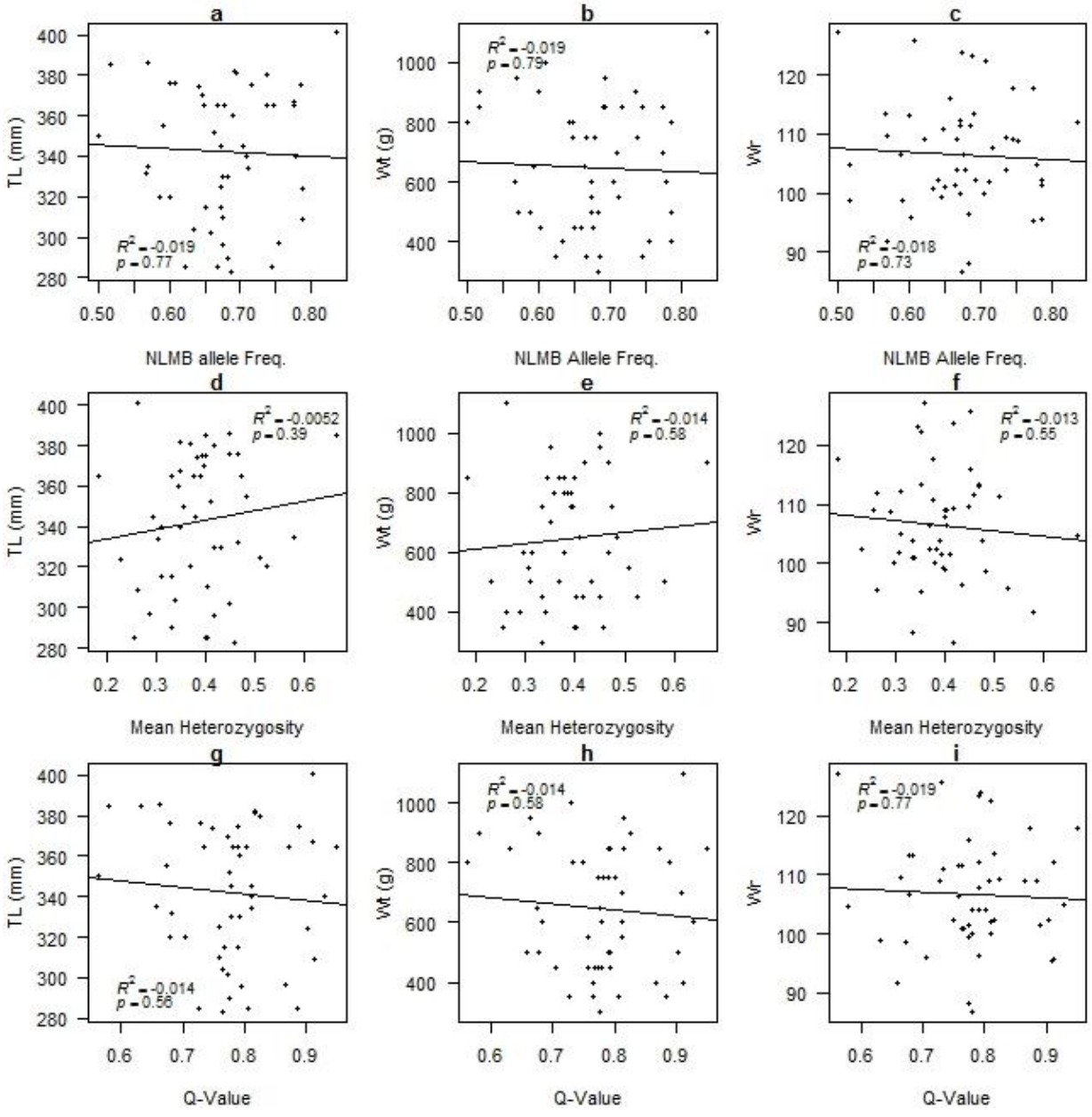


Figure 9 For Age 3 Guntersville samples: **a)** TL as a function of NLMB allele frequency **b)** Wt as a function of NLMB allele frequency **c)** *Wr* as a function of NLMB allele frequency **d)** TL as a function of mean heterozygosity **f)** Wt as a function of mean heterozygosity **g)** *Wr* as a function of mean heterozygosity **h)** TL as a function of Q-value **i)** Wt as a function of Q-value **j)** *Wr* as a function of Q-value. Reported p-values of <0.05 indicate that the slope of the regression is significantly different from zero.

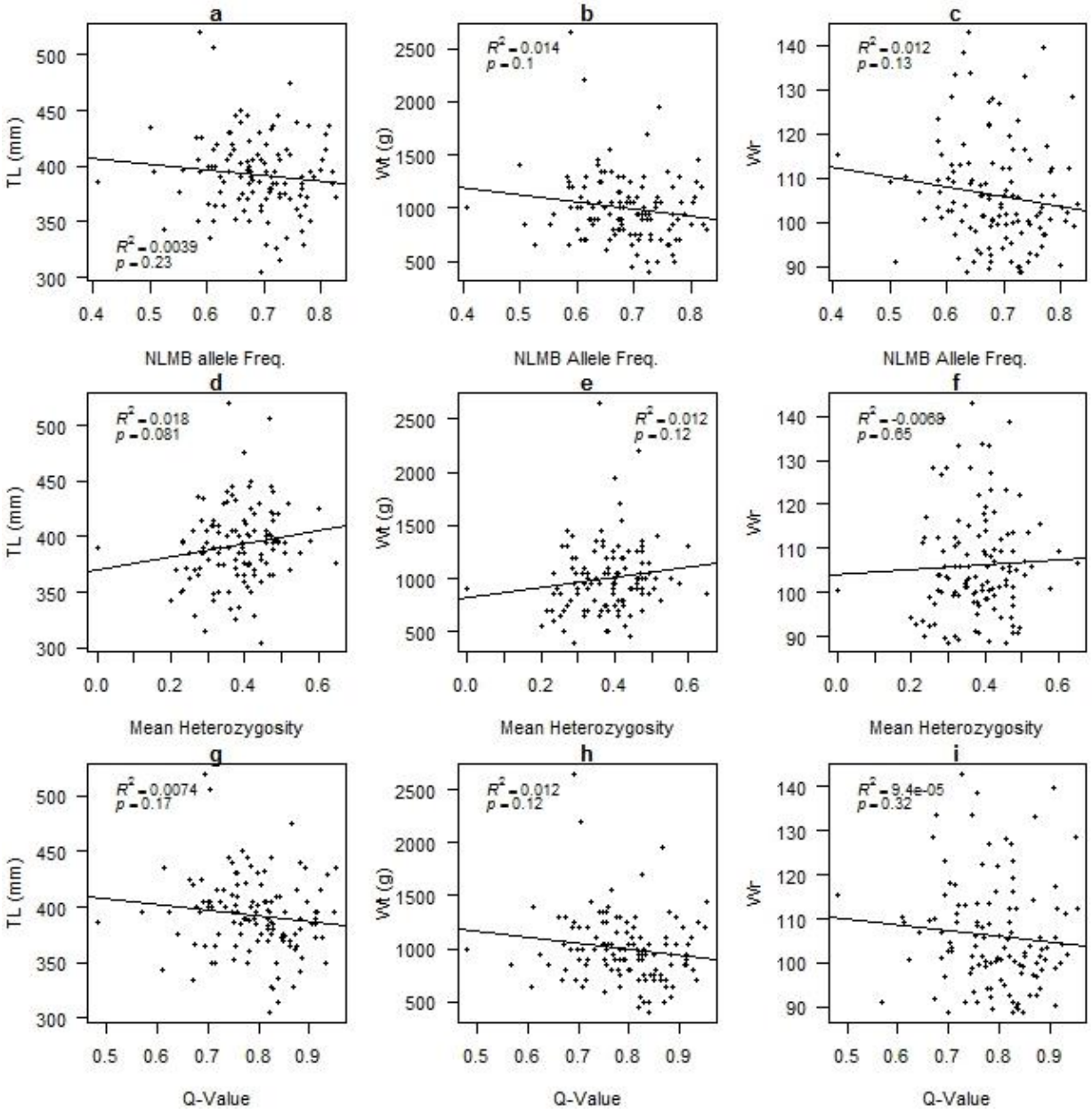


Figure 10 For Age 4 Guntersville samples: **a)** TL as a function of NLMB allele frequency **b)** Wt as a function of NLMB allele frequency **c)** *Wr* as a function of NLMB allele frequency **d)** TL as a function of mean heterozygosity **f)** Wt as a function of mean heterozygosity **g)** *Wr* as a function of mean heterozygosity **h)** TL as a function of Q-value **i)** Wt as a function of Q-value **j)** *Wr* as a function of Q-value. Reported p-values of <0.05 indicate that the slope of the regression is significantly different from zero.

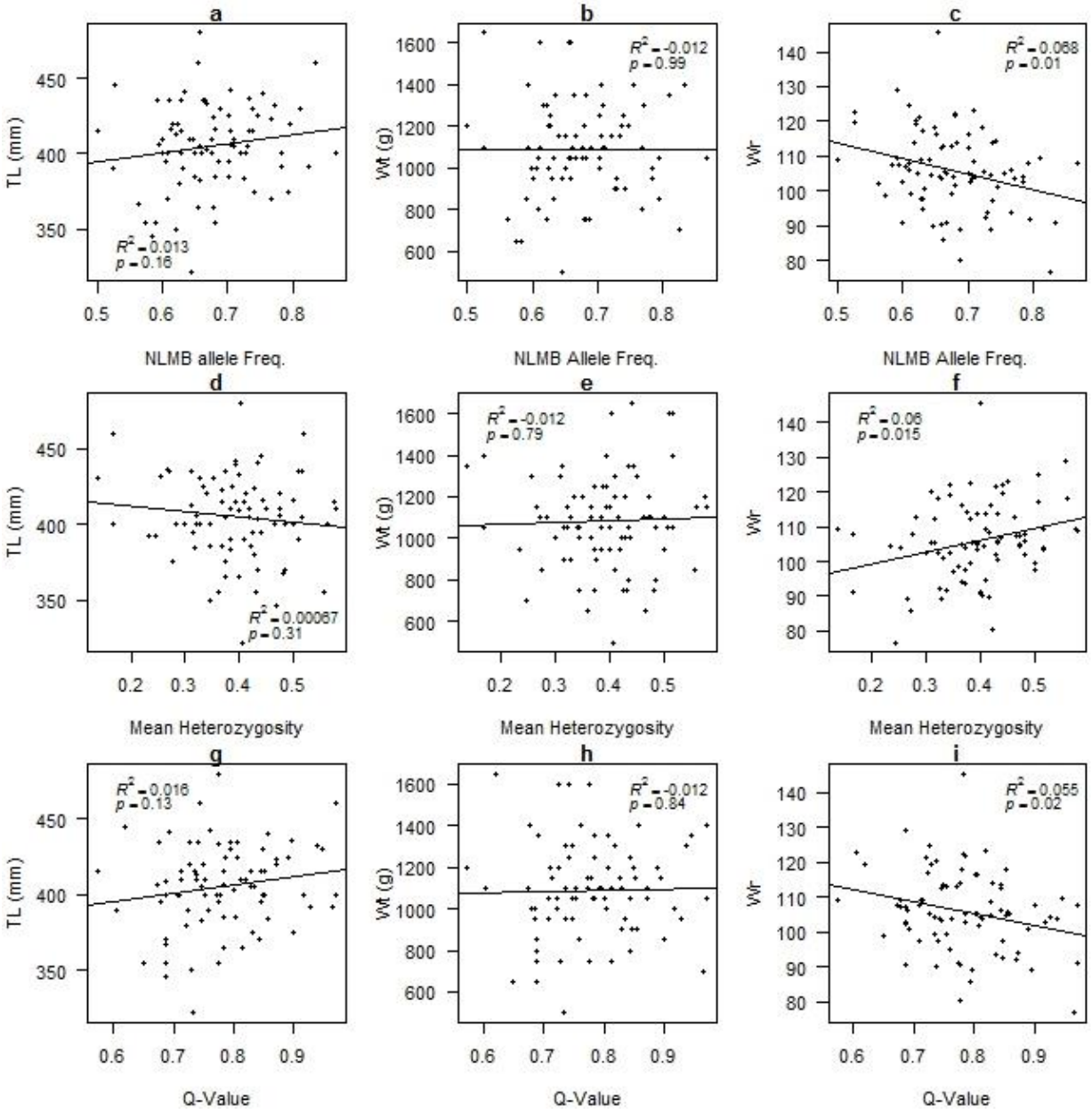


Figure 11 Von Bertalanffy Growth Curves growth curve based on length at age data for LMB collected from Lake Guntersville grouped by Q-value estimates from STRUCTURE population clustering analysis. (High >0.787> low).

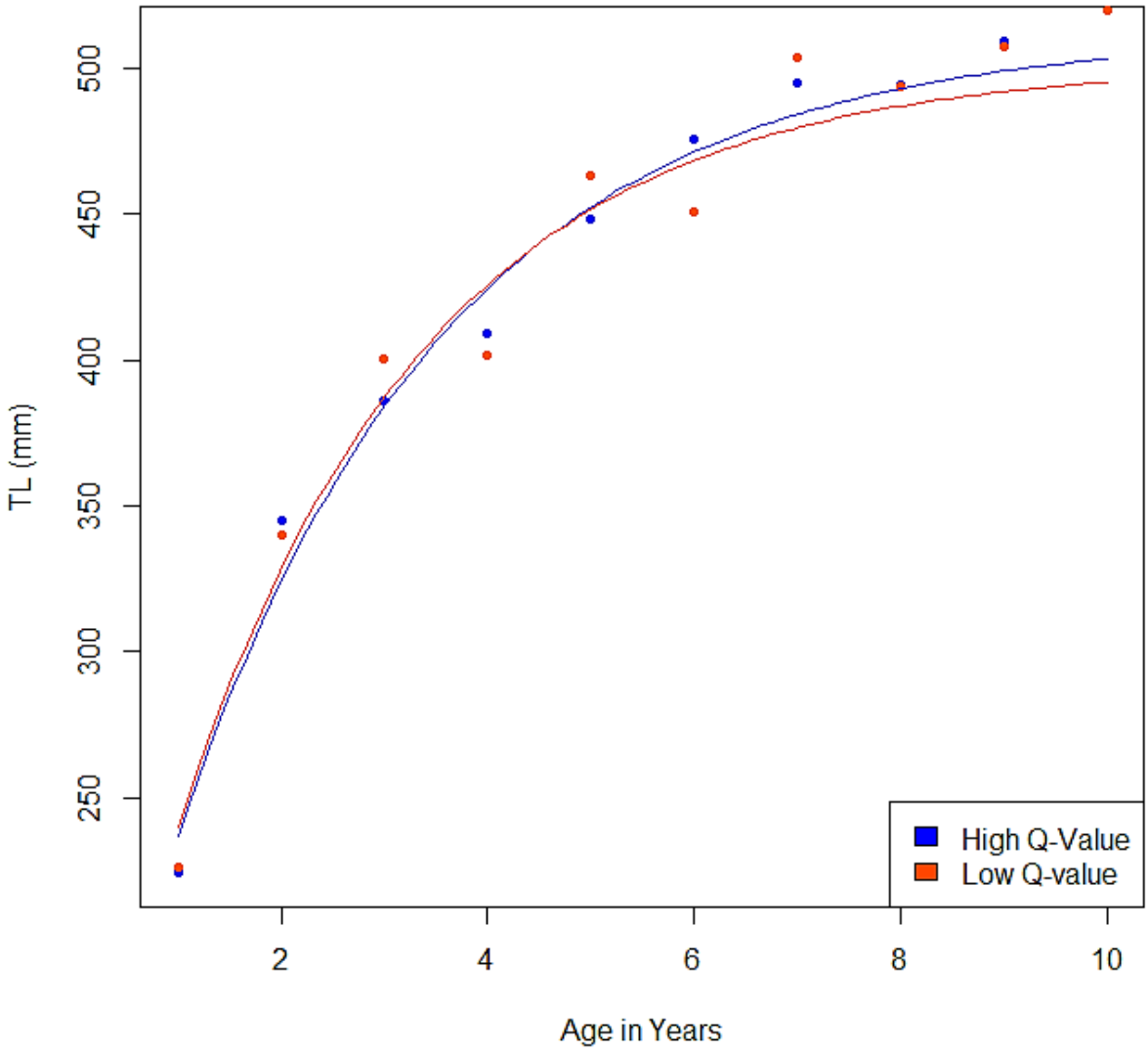


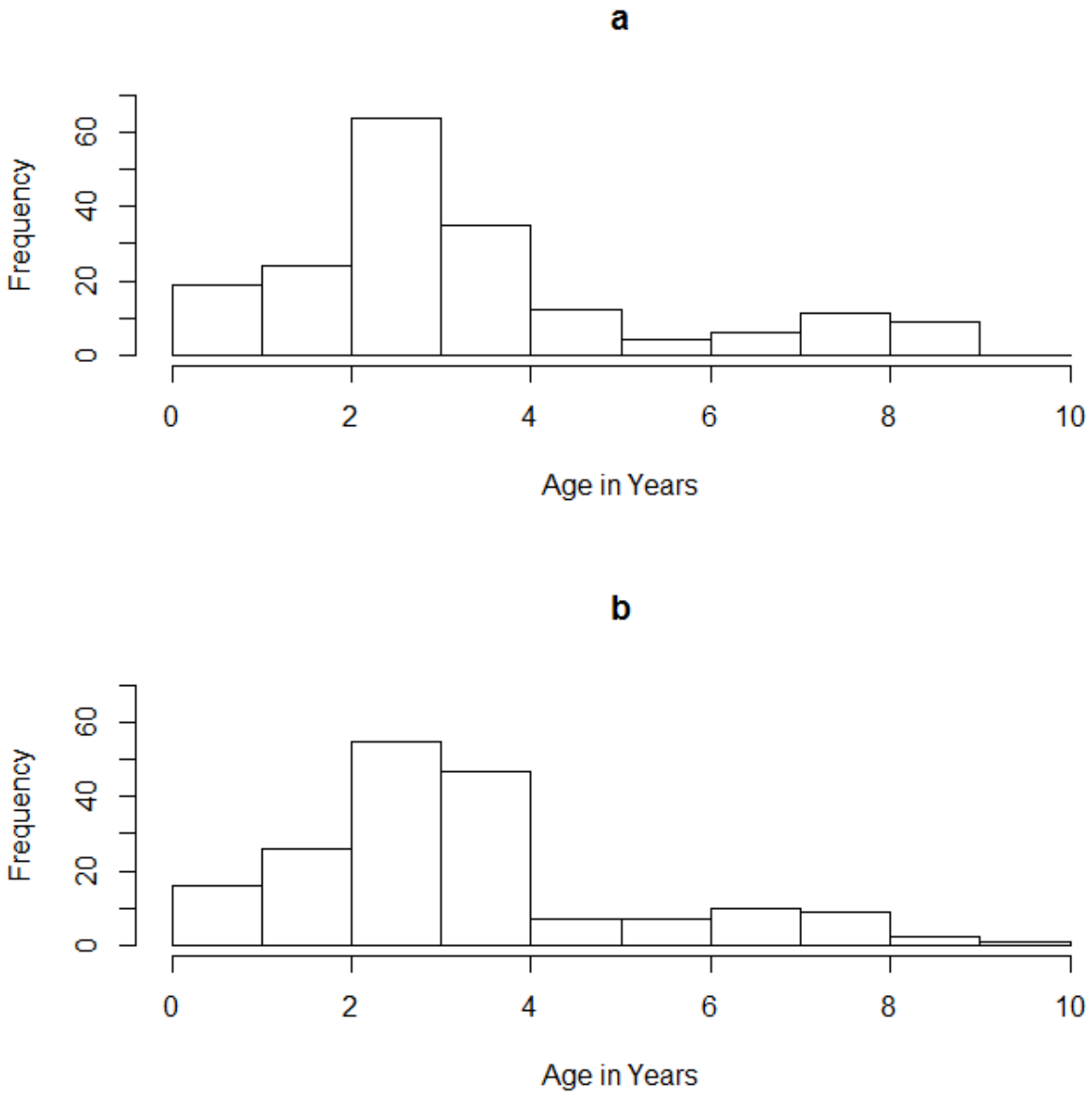
Table 17 Estimated Von Bertalanffy growth curve parameters for LMB collected from Lake Guntersville grouped by Q-value estimates from STRUCTURE population clustering analysis. (High >0.787> low). All parameters are not significantly different between populations.

Q-value	*k*	*S.E.*	*L.infinity*	*S.E.*	*t0*	*S.E.*
High	0.383	0.034	511.996	10.260	-0.619	0.116
Low	0.414	0.046	501.603	11.922	-0.570	0.191

Table 18. Frequency and mean TL and mean Wt table for all, high q-value, and low Q-value LMB sampled in spring 2015 in Lake Guntersville via electrofishing

	All Samples			High Q-value			Low Q-Value		
Age	**N**	**Mean TL (mm)**	**Mean Wt (g)**	**N**	**Mean TL (mm)**	**Mean Wt (g)**	**N**	**Mean TL (mm)**	**Mean Wt (g)**
1	35	225.11	158.57	19	224.16	150	16	226.25	168.75
2	50	342.32	646	24	670.83	345.04	26	339.81	623.08
3	119	392.54	1000.84	64	385.92	933.59	55	400.24	1079.09
4	82	405.01	1085.37	35	409.03	1084.29	47	402.02	1086.17
5	19	454.11	1636.84	12	448.67	1541.67	7	463.43	1800
6	11	460	1718.18	4	475.75	1862.5	7	451	1635.71
7	16	500.38	2081.25	6	494.83	2108.33	10	503.7	2065
8	20	494.25	2102.5	11	494.73	2127.27	9	493.67	2072.22
9	11	509.27	2250	9	509.67	2283.33	2	507.5	2100
10	1	520	1850	NA	NA	NA	1	520	1850

928 **Figure 12** Frequency by age of **a)** high Q-value LMB (more NLMB alleles) and **b)** low Q-value
929 samples (more FLMB alleles) from fish collected via electrofishing from Lake Guntersville in
930 the spring 2015.



931

Trophy bass alleles

Lastly, allele usage in LMB over 7 lbs. ("Trophy" bass) and in LMB over 5 lbs. ("memorable" bass) were estimated. Markers identified as having statistically significant association with the Trophy and Memorable groups are listed in table 19. This indicates that these alleles are more common within the larger fish in this population. Without more samples it is hard to predict if these alleles would correlate with larger fish in all populations. A wider-scale survey (covering more populations) may give more universal results that could be be a potential way to select fish suitable for hatchery projects aimed at bolstering the Trophy bass potential of a reservoir. Also, beyond the gene ontology that was done with these markers, identifying where in the coding region of the DNA they fall and if they are silent or have important biological implications would be valuable.

944

Table 19 Markers information and gene ontology for SNPs that were found to show significance (p<.05) in pairwise exact G-test between Trophy Bass (>7 lbs.), memorable bass (> 5 lbs.) and all other bass (< 5 lbs. and ≥ age 3) ~Indicates markers that were shown to have elevated pairwise Fst values between Trophy, Memorable and other bass. *Indicates markers from 38-Plex.

Contig	p-value	FLMB Genotype	NLMB Genotype	F1 Genotype	Gene Name	Function
Contig16665*	0.013	A	C	C/A	Aquaporin 7	Membrane
Contig17151*~	0.000	A	T	A/T	Proto-oncogene protein c-Fos	Expression Regulation
Contig18101*	0.023	C	A	A/C	Pepatic lipase	Fatty Acid Metabolism
Contig23633*	0.028	C	T	T/C	V-ATPase subunit A	ATP Metabolism
Contig6106~	0.006	T	C	C/T	Fanconi anemia group F protein	DNA repair
Contig15421	0.014	T	A	A/T	Serine incorporator 1	Lipid biosynthesis
Contig19092	0.020	G	T	G/T	Nonspecific cytotoxic cell receptor protein-1	Immune
Contig35139	0.032	T	C	C/T	Repressor of RNA polymerase III transcription MAF1	Nutrient-dependent growth
Contig22709*	0.003	T	G	G/T	Hepatoma-derived growth factor-related protein 2	Cellular regulation
Contig28601*	0.050	A	G	A/G	Interleukin-1 receptor type 1-like	Cellular regulation
Contig5713*	0.036	A	G	A/G	Calcium-binding mitochondrial carrier protein SCaMC	Ion Transport
Contig9870*	0.023	G	A	A/G	suppressor of cytokine signaling 5	Intercellular Signaling

945

Conclusion

The growth and performance potential of natural and anthropogenic intergrade populations of NLMB and FLMB has long been argued and disputed by fisheries scientists. Convincing investigations that support all sides of the issue have often been limited by genotyping resolution (Philipp et al. 1983; Fields et al. 1987). In this study advances SNP marker technology that have been long used in other applications for non-model species were applied to help answer some important questions relating to this highly important species.

In this study SNP markers were successfully developed for distinguishing NLMB populations from FLMB populations. These markers were also capable of assessing rates of introgression in admixed populations.

Briefly, a subset of 61 SNP markers were selected from a pool of 3674 possible markers identified from LMB RNA sequencing data, and were developed into a set of diagnostic multiplex assays. These markers were validated for accuracy with microsatellite verified "pure" samples.

The utility of the markers as a high throughput technology for genotyping large numbers of samples was demonstrated by genotyping over 1500 individual hatchery and wild fish from around the state of Alabama. This widespread genotyping effectively re-evaluated the current status of populations that have not been genotyped at this scale for over 20 years (Norgren et al.1986; Maceina & Dicenzo 1995). The genotyping results had similarity to previous studies, but the increased marker number, along with some modern analysis software revealed some nuances in the populations that may have been previously looked over with other marker technology. Most notably, these results indicated that populations throughout Alabama have not had equal responses to stocking of FLMB. Genetic isolation is still apparent across geographic elements such as fall-

lines and between unique drainages, indicating that the FLMB stocking either has had little influence on these populations or have had equal influence all over, simply shifting the baseline. The former is more likely considering that stocking efforts have not been evenly distributed among populations, with some populations receiving no FLMB stocking.

The newly developed markers were also applied to an individual population of artificial intergrades in the Tennessee River impoundment of Lake Guntersville in northern Alabama. The intention was to evaluate the genetic influence on trophy bass in an environment known to produce large bass, while limiting environmental variation by analyzing samples from only one reservoir. Guntersville was chosen for its trophy LMB reputation and was shown by previous sampling to have a certain level of FLMB stocking influence.

The growth analysis, either by age specific comparison, or Von Bertalanffy growth curves, all did not indicate growth differences based on genotype in Lake Guntersville. However, a correlation of *Wr* and genotype was observed that indicated that older fish with higher levels of FLMB influence are plumper, possibly contributing to the higher proportions of FLMB alleles present in the larger of angler-caught fish. A larger sample set may also have given more conclusive results. Also this study did not evaluate daily growth of the age 1 fish. A study evaluating daily growth rings of age 1 fish with genotype may find more conclusive results relating genetic influence on early growth as well as getting estimates of temporal spawning differences as related to genotype in an admixed population.

One important observation that was made while conducting the analysis of Lake Guntersville genotypes was that bass collected by tournament anglers were significantly different with respect to genotype and weight. Tournament fish were much larger and had more FLMB alleles. In the angler captured fish (a population selected with considerable bias) a positive

correlation between size and FLMB alleles was observed as well as a correlation of size and heterozygosity, however this correlation was absent in LMB collected via electrofishing. The most outstanding observation is that the tournament fish were of a size not represented in the electrofishing samples, indicating that there is a portion of the population (the largest fish) that is being missed in electrofishing surveys. And likewise there seems to be a genotype being missed by electrofishing surveys.

In this study, the primary focus was the state of Alabama, but future studies should expand the use of these and additional SNP panels to better define the distributions of pure FLMB and pure NLMB, update the status of the intergrade hybrid zone, and evaluate the effect of continued widespread stocking of FLMB since the last range-wide study (Philipp et al. 1983). The lower cost and higher throughput nature of SNPs relative to other marker types would greatly improve the feasibility of such a study.

Developing more polymorphic (non-fixed) markers could be another future direction for this work. The fixed nature of these markers limit their use to evaluating the purity of a given population within the range of FLMB to NLMB. Developing polymorphic markers would be helpful in the elucidation of some unique genotypes in certain populations such as the Delta Bass in the Mobile-Tensaw Delta, as well as helpful in identifying gradients within near-pure populations.

This resource could also be used to conduct some more rigorous examination of allelic usage patterns in key functional genes in naturally introgressed populations when compared with populations with anthropogenic impacts on hybridization through stocking. This should be performed across the wide range of environments where bass thrive (Jiggins & Mallet 2000; Martinsen et al. 2001; Gompert et al. 2006; Payseur 2010; Carson et al. 2012; DeVries et al. 2014).

Since variables in a large wild population such as a reservoir can obscure results, another way to evaluate growth differences would be to return to the head-to-head controlled environment challenges of the past. Using these more powerful markers, future studies are planned with the Alabama State Lakes programs with controlled stocking to evaluate the trophy production and growth of pure and hybrid populations.

The observed differences between genotype in the angler-caught samples and the electrofishing samples as well as the fact that the trophy range LMB that are targeted by anglers are rare and therefore not common in electrofishing samples, is a good reason for developing an angler-driven genetic sampling program. This would be a program where anglers send a genetic sample with size metrics to be genotyped, and added to a master data-set. This way the parts of the population that the electrofishing surveys seem to be missing may have a chance to be included in ongoing studies. Work on developing a method for anglers to swab a LMB before release to capture genetic material that can be subsequently sent to the Auburn University Aquatic Genetics and Genomics lab for analysis. This will potentially open up new opportunities for identifying what genotypes are truly responsible for the fish that LMB fishery managers are striving to promote produce out of these world class fisheries in Alabama.

References

Addison, J. H., & Spencer, S. L. (1972). Preliminary evaluation of three strains of largemouth bass, Micropterus salmoides (Lacepede), stocked in ponds in south Alabama. In *Proceedings of the Annual Conference Southeastern Association of Game and Fish Commissioners* (Vol. 25, No. 1971, pp. 366-374).

Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids: setting conservation guidelines. *Trends in Ecology & Evolution* **16**, 613-622.

An L-h, Lei K, Zheng B-h (2014) Use of heat shock protein mRNA expressions as biomarkers in wild crucian carp for monitoring water quality. *Environmental toxicology and pharmacology* **37**, 248-255.

Ansari KI, Mishra BP, Mandal SS (2008) Human CpG binding protein interacts with MLL1, MLL2 and hSet1 and regulates Hox gene expression. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* **1779**, 66-73.

Austin JD, Johnson A, Matthews M, *et al.* (2012) An assessment of hatchery effects on Florida bass (Micropterus salmoides floridanus) microsatellite genetic diversity and sib-ship reconstruction. *Aquaculture Research* **43**, 628-638.

Azuma, M., and Y. Motomura. (1999) Feeding habits of largemouth bass in a non-native environment: the case of a small lake with bluegill in Japan. Environmental Biology of Fishes 52:379–389

Bai, J., Lutz-Carrillo, D. J., Quan, Y., & Liang, S. (2008). Taxonomic status and genetic diversity of cultured largemouth bass Micropterus salmoides in China.*Aquaculture*, *278*(1), 27-30

Bailey R.M., and C.L. Hubbs (1949) The black basses (*Micropterus*) of Florida, with description of a new species. *University of Michigan Occasional Papers of the Museum of Zoology* 516:1–40

Barthel BL, Lutz-Carrillo DJ, Norberg KE, *et al.* (2010) Genetic relationships among populations of Florida bass. *Transactions of the American Fisheries Society* **139**, 1615-1641.

Bender A, Desplats P, Spencer B, *et al.* (2013) TOM40 Mediates Mitochondrial Dysfunction Induced by α-Synuclein Accumulation in Parkinson's Disease. *PloS one* **8**, e62277.

Blaker-Lee A, Gupta S, McCammon JM, De Rienzo G, Sive H (2012) Zebrafish homologs of genes within 16p11. 2, a genomic region associated with brain disorders, are active during brain development, and include two deletion dosage sensor genes. *Disease models & mechanisms* **5**, 834-851.

Bolnick DI, Near TJ (2005) Tempo of hybrid inviability in centrarchid fishes (Teleostei: Centrarchidae). *Evolution* **59**, 1754-1767.

Bottroff, L. J., & Lembeck, M. E. (1978). Fishery trends in reservoirs of San-Diego county, California, following introduction of Florida largemouth bass, micropterus-salmoides-floridanus. *California Fish and Game*, *64*(1), 4-23.

Brown, M. L., & Murphy, B. R. (1994). Temporal genetic structure of an intergrade largemouth bass (Micropterus salmoides) population. *Ecology of Freshwater Fish*, *3*(1), 18-24.

Bremer, J. R. A., Zhang, L., Bulak, J. S., & Ely, B. (1998). A polymerase chain reaction–restriction fragment length polymorphism (PCR–RFLP) assay for the discrimination of mitochondrial DNA from the Florida and northern subspecies of largemouth bass. *Transactions of the American Fisheries Society*, *127*(3), 507-511.

Bugel SM, Bonventre JA, White LA, Tanguay RL, Cooper KR (2014) Chronic exposure of killifish to a highly polluted environment desensitizes estrogen-responsive reproductive and biomarker genes. *Aquatic Toxicology* **152**, 222-231.

Cai J, Wei S, Wang B, *et al.* (2013) Cloning and expression analysis of nonspecific cytotoxic cell receptor 1 (Ls-NCCRP1) from red snapper (Lutjanus sanguineus). *Marine genomics* **11**, 39-44.

Carson EW, Tobler M, Minckley W, Ainsworth RJ, Dowling TE (2012) Relationships between spatio-temporal environmental and genetic variation reveal an important influence of exogenous selection in a pupfish hybrid zone. *Molecular Ecology* **21**, 1209-1222.

Casimir M, Lasorsa FM, Rubi B, *et al.* (2009) Mitochondrial glutamate carrier GC1 as a newly identified player in the control of glucose-stimulated insulin secretion. *Journal of Biological Chemistry* **284**, 25004-25014.

Chew R (1975) The Florida largemouth bass. *Black bass biology and management. Sport Fishing Institute, Washington, DC*, 450-458.

Cichra, C. E., Neill, W. H., & Noble, R. L. (1980). Differential resistance of northern and Florida largemouth bass to cold shock [Micropterus salmoides salmoides, Micropterus salmoides floridanus]. In *Proceedings of the Annual Conference-Southeastern Association of Fish and Wildlife Agencies (USA)*.

Clugston, J. P. (1964). Growth of the Florida largemouth bass, Micropterus salmoides floridanus (LeSueur), and the northern largemouth bass, M. s. salmoides (Lacépède), in subtropical Florida. *Transactions of the American Fisheries Society*, *93*(2), 146-154.

Cooke S, Philipp D (2006) Hybridization among divergent stocks of largemouth bass (Micropterus salmoides) results in altered cardiovascular performance: The influence of genetic and geographic distance. *Physiological and Biochemical Zoology* **79**, 400-410.

Crawford, S., Coleman, W. S., & Porak, W. F. (1989). Time of annulus formation in otoliths of Florida largemouth bass. *North American Journal of Fisheries Management*, *9*(2), 231-233.

Denslow ND, Garcia-Reyero N, Barber DS (2007) Fish 'n'chips: the use of microarrays for aquatic toxicology. *Mol. BioSyst.* **3**, 172-177.

DeVries DR, *Wr*ight RA, Glover DC, *et al.* (2014) Largemouth Bass in Coastal Estuaries: A Comprehensive Study from the Mobile-Tensaw River Delta, Alabama. *American Fisheries Society Symposium.* **82**, in press

Dumont SC, Lutz-Carrillo DJ (2011) Estimating Genetic Composition in Introgressed Largemouth Bass Populations: Does Age Matter? *North American Journal of Fisheries Management* **31**, 176-181.

Dunham RA, Turner C, Reeves W (1992) Introgression of the Florida largemouth bass genome into native populations in Alabama public lakes. *North American Journal of Fisheries Management* **12**, 494-498.

Fields, R., Lowe, S. S., Kaminski, C., Whitt, G. S., & Philipp, D. P. (1987). Critical and chronic thermal maxima of northern and Florida largemouth bass and their reciprocal F1 and F2 hybrids. *Transactions of the American Fisheries Society*, *116*(6), 856-863.

Fitzpatrick BM, Johnson JR, Kump DK, *et al.* (2009) Rapid fixation of non-native alleles revealed by genome-wide SNP analysis of hybrid tiger salamanders. *BMC Evolutionary Biology* **9**, 176.

Garcia-Reyero N, Griffitt RJ, Liu L, *et al.* (2008) Construction of a robust microarray from a non-model species largemouth bass, Micropterus salmoides (Lacèpede), using pyrosequencing technology. *Journal of Fish Biology* **72**, 2354-2376.

Goh SL, Wang Q, Byrnes LJ, Sondermann H (2012) Versatile membrane deformation potential of activated pacsin. *PloS one* **7**, e51628.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., ... & Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology*, *29*(7), 644-652.

Gompert Z, Fordyce JA, Forister ML, Shapiro AM, Nice CC (2006) Homoploid hybrid speciation in an extreme habitat. *Science* **314**, 1923-1925.

Gratwicke, B., and B. E. Marshall (2001) The relationship between the exotic predators Micropterus salmoides, and Serranochromus robustus, and native stream fishes in Zimbabwe. Journal of Fish Biology 58:68–75

Hadkar V, Skidgel RA (2001) Carboxypeptidase D is up-regulated in RAW 264.7 macrophages and stimulates nitric oxide synthesis by cells in arginine-free medium. *Molecular pharmacology* **59**, 1324-1332.

Hallerman EM, Smitherman R, Reed RB, Tucker WH, Dunham RA (1986) Biochemical genetics of largemouth bass in mesosaline and freshwater areas of the Alabama River system. *Transactions of the American Fisheries Society* **115**, 15-20.

Hasselman DJ, Argo EE, McBride MC, *et al.* (2014) Human disturbance causes the formation of a hybrid swarm between two naturally sympatric fish species. *Molecular Ecology*.

Henson, J. C. (1991). *Quantitative description and development of a species-specific standard growth form for largemouth bass with application to the relative weight (Wr) index* (Doctoral dissertation, Texas A&M University).

Hohenlohe PA, Amish SJ, Catchen JM, Allendorf FW, Luikart G (2011) Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. *Molecular ecology resources* **11**, 117-122.

Hoyer, M. V., Shireman, J. V., & Maceina, M. J. (1985). Use of otoliths to determine age and growth of largemouth bass in Florida. *Transactions of the American Fisheries Society*, *114*(2), 307-309.

Imamura K, Imamachi N, Akizuki G, *et al.* (2014) Long noncoding RNA NEAT1-dependent SFPQ relocation from promoter region to paraspeckle mediates IL8 expression upon immune stimuli. *Molecular cell* **53**, 393-406.

Inman, C. R., Dewey, R. C., & Durocher, P. P. (1977). Growth comparisons and catchability of three largemouth bass strains. *Fisheries*, *2*(5), 20-25.

Inuzuka M, Hayakawa M, Ingi T (2005) Serinc, an activity-regulated protein family, incorporates serine into membrane lipid synthesis. *Journal of Biological Chemistry* **280**, 35776-35783.

Isely, J. J., Noble, R. L., Koppelman, J. B., & Philipp, D. P. (1987). Spawning period and first-year growth of northern, Florida, and intergrade stocks of largemouth bass. *Transactions of the American Fisheries Society*, *116*(5), 757-762.

Jiggins CD, Mallet J (2000) Bimodal hybrid zones and speciation. *Trends in Ecology & Evolution* **15**, 250-255.

Johnson R, Fulton T (1999) Persistence of Florida largemouth bass alleles in a northern Arkansas population of largemouth bass, Micropterus salmoides Lacepede. *Ecology of Freshwater Fish* **8**, 35-42.

Kalinowski S, Novak B, Drinan D, Dem Jennings R, Vu N (2011) Diagnostic single nucleotide polymorphisms for identifying westslope cutthroat trout (Oncorhynchus clarki lewisi), Yellowstone cutthroat trout (Oncorhynchus clarkii bouvieri) and rainbow trout (Oncorhynchus mykiss). *Molecular ecology resources* **11**, 389-393.

Kassler T, Koppelman J, Near T, *et al.* (2002) Molecular and morphological analyses of the black basses: implications for taxonomy and conservation, 291-322.

Kassmeier MD, Mondal K, Palmer VL, *et al.* (2012) VprBP binds full-length RAG1 and is required for B-cell development and V (D) J recombination fidelity. *The EMBO journal* **31**, 945-958.

Lamothe, K. A., & Johnson, R. L. (2013). Microsatellite Analysis of Trophy Largemouth Bass from Arkansas Reservoirs. *Journal of the Arkansas Academy of Science*, *67*, 71.

Lamaze FC, Sauvage C, Marie A, Garant D, Bernatchez L (2012) Dynamics of introgressive hybridization assessed by SNP population genomics of coding genes in stocked brook charr (Salvelinus fontinalis). *Molecular Ecology* **21**, 2877-2895.

Lamer JT, Sass GG, Boone JQ, *et al.* (2014) Restriction site-associated DNA sequencing generates high-quality single nucleotide polymorphisms for assessing hybridization between bighead and silver carp in the United States and China. *Molecular ecology resources* **14**, 79-86.

Lapegue S, Harrang E, Heurtebise S, *et al.* (2014) Development of SNP genotyping arrays in two shellfish species. *Molecular ecology resources*.

Leitner, J., Bulak, J., & Dunham, R. (2002). A comparison of first and third year growth of two strains of largemouth bass in South Carolina. In *American Fisheries Society Symposium* (Vol. 31, pp. 365-370).

Li C, Waldbieser G, Bosworth B, *et al.* (2014) SNP Discovery in Wild and Domesticated Populations of Blue Catfish, Ictalurus furcatus, Using GBS and Subsequent SNP Validation. *Molecular ecology resources*.

Luo Y (2014) Transcriptomic profiling of differential responses to drought in two freshwater mussel species, the giant floater Pyganodon grandis and the pondhorn Uniomerus tetralasmus. *PLoS one* **9**, e89481.

Lutz-Carrillo DJ, Nice CC, Bonner TH, Forstner MR, Fries LT (2006) Admixture analysis of Florida largemouth bass and northern largemouth bass using microsatellite loci. *Transactions of the American Fisheries Society* **135**, 779-791.

Lutz-Carrillo DJ, Hagen C, Dueck LA, Glenn TC (2008) Isolation and characterization of microsatellite loci for Florida largemouth bass, Micropterus salmoides floridanus, and other micropterids. *Molecular ecology resources* **8**, 178-184.

MacCrimmon HR, Robbins WH (1975) Distribution of the black basses in North America. *Black bass biology and management. Sport Fishing Institute, Washington, DC*, 56-66.

Maceina, M. J. (1988). Simple grinding procedure to section otoliths. *North American Journal of Fisheries Management*, *8*(1), 141-143.

Maceina, M. J., & Murphy, B. R. (1988). Variation in the weight-to-length relationship among Florida and northern largemouth bass and their intraspecific F1 hybrid. *Transactions of the American Fisheries Society*, *117*(3), 232-237.

Maceina MJ, Murphy BR (1992) Stocking Florida largemouth bass outside its native range. *Transactions of the American Fisheries Society* 121(5), 686-691

Maceina MJ, Murphy BR, Isely JJ (1988) Factors regulating Florida largemouth bass stocking success and hybridization with northern largemouth bass in Aquilla Lake, Texas. *Transactions of the American Fisheries Society* **117**, 221-231.

Maceina, M. J., and V. J. DiCenzo. 1995. Long-term genetic changes and growth of largemouth bass populations in Alabama public reservoirs. Alabama Dep. Cons. Nat. Res., Montgomery, Alabama. Federal Aid to Fish and Wildlife Restoration. Fin. Rep. F-40 (Study 22). 151 pp.

Mandilaras K, Missirlis F (2012) Genes for iron metabolism influence circadian rhythms in Drosophila melanogaster. *Metallomics* **4**, 928-936

Martinsen GD, Whitham TG, Turek RJ, Keim P (2001) Hybrid populations selectively filter gene introgression between species. *Evolution* **55**, 1325-1335.

Mehinto AC, Prucha MS, Colli-Dula RC, *et al.* (2014) Gene networks and toxicity pathways induced by acute cadmium exposure in adult largemouth bass (Micropterus salmoides). *Aquatic Toxicology* **152**, 186-194.

Miller, L. W. (1965). A growth study and blood protein analysis of two subspecies of largemouth bass; the Florida bass, Micropterus salmoides floridanus (LeSueur), and the northern bass, Micropterus salmoides salmoides (Lacepede). *San Diego County, California. Calif. Dept. Fish and Game, Inland Fish. Admin. Rep*, 65-15

Mitchell, J. M., K. K. Sellers, W. D. Harvey, and L. T. Fries. 1993. Effects of stocking regime and harvest regulation on Florida largemouth bass stocking success. Proceedings of the Annual Conference Southeastern Association of Fish and Wildlife Agencies 45(1991):477–483.

Moffatt NSC, Bruinsma E, Uhl C, Obermann WM, Toft D (2008) Role of the Cochaperone Tpr2 in Hsp90 Chaperoning†. *Biochemistry* **47**, 8203-8213

Near TJ, Kassler TW, Koppelman JB, Dillman CB, Philipp DP (2003) Speciation in North American black basses, Micropterus (Actinopterygii: Centrarchidae). *Evolution* **57**, 1610-1621.

Nedbal MA, Philipp DP (1994) Differentiation of mitochondrial DNA in largemouth bass. *Transactions of the American Fisheries Society* **123**, 460-468.

Norgren K, Dunham R, Smitherman R, Reeves W (1986) Biochemical genetics of largemouth bass populations in Alabama **40**, 194-205.

Onisto M, Slongo L, Graziotto R, *et al.* (2001) Evidence for FSH-dependent upregulation of SPATA2 (spermatogenesis-associated protein 2). *Biochemical and biophysical research communications* **283**, 86-92

Page LM, Espinosa-Pérez H, Findley LT et al.(2013) Common and scientific names of fishes from the United States, Canada, and Mexico, 7th edn. American Fisheries Society Special Publication 34, Bethesda

Payseur BA (2010) Using differential introgression in hybrid zones to identify genomic regions involved in speciation. *Molecular ecology resources* **10**, 806-820.

Philipp DP, Childers WF, Whitt GS (1983) A biochemical genetic evaluation of the northern and Florida subspecies of largemouth bass. *Transactions of the American Fisheries Society* **112**, 1-20.

Philipp, D. P. (1991). Genetic implications of introducing Florida largemouth bass, Micropterus salmoides floridanus. *Canadian Journal of Fisheries and Aquatic Sciences*, *48*(S1), 58-65.

Philipp, D. P., & Whitt, G. S. (1991). Survival and growth of northern, Florida, and reciprocal Fl hybrid largemouth bass in central Illinois. *Transactions of the American Fisheries Society*, *120*(1), 58-64.

Philipp DP, Claussen JE (1995) Fitness and performance differences between two stocks of largemouth bass from different river drainages within Illinois **15**, 236-243.

Philipp, D.P., Ridgway, M.S.eds. (2002) Black Bass: Ecology Conservation and ManagementAmerican Fisheries Society, Symposium 31. Bethesda, Maryland

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

Pritchard V, Abadía-Cardoso A, Garza J (2012) Discovery and characterization of a large number of diagnostic markers to discriminate Oncorhynchus mykiss and O. clarkii. *Molecular ecology resources* **12**, 918-931.

Pritchard VL, Campbell NR, Narum SR, Peacock MM, Garza JC (2013) Discovery and characterization of novel genetic markers for use in the management of Lahontan cutthroat trout (Oncorhynchus clarkii henshawi). *Molecular ecology resources* **13**, 276-288.

Rafail, S. Z. (1973). A simple and precise method for fitting a von Bertalanffy growth curve. *Marine biology*, *19*(4), 354-358.

Redenbach Z, Taylor E (2003) Evidence for bimodal hybrid zones between two species of char (Pisces: Salvelinus) in northwestern North America. *Journal of evolutionary biology* **16**, 1135-1148.

Rice AM, Rudh A, Ellegren H, Qvarnström A (2011) A guide to the genomics of ecological speciation in natural animal populations. *Ecology Letters* **14**, 9-18.

Richter CA, Martyniuk CJ, Annis ML, *et al.* (2014) Methylmercury-induced changes in gene transcription associated with neuroendocrine disruption in largemouth bass (Micropterus salmoides). *General and comparative endocrinology*.

Rideout EJ, Marshall L, Grewal SS (2012) Drosophila RNA polymerase III repressor Maf1 controls body size and developmental timing by modulating tRNAiMet synthesis and systemic insulin signaling. *Proceedings of the National Academy of Sciences* **109**, 1139-1144.

Rodrigues SM, Farb CR, Bauer EP, LeDoux JE, Schafe GE (2004) Pavlovian fear conditioning regulates Thr286 autophosphorylation of Ca2+/calmodulin-dependent protein kinase II at lateral amygdala synapses. *The Journal of neuroscience* **24**, 3281-3288.

Rogers, M. W., Allen, M. S., & Porak, W. F. (2006). Separating genetic and environmental influences on temporal spawning distributions of largemouth bass (Micropterus salmoides). *Canadian Journal of Fisheries and Aquatic Sciences*, *63*(11), 2391-2399.

Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular ecology resources*, *8*(1), 103-106.

Sasaki, S. (1961). Introduction of Florida largemouth bass into San Diego County. *Calif. Dept. Fish and Game, Inland Fish. Admin. Rep*, *61*, 11-16.

Seyoum S, Barthel BL, Tringali MD, *et al.* (2013) Isolation and characterization of eighteen microsatellite loci for the largemouth bass, Micropterus salmoides, and cross amplification in congeneric species. *Conservation Genetics Resources* **5**, 697-701.

Shen Y, Catchen J, Garcia T, *et al.* (2012) Identification of transcriptome SNPs between Xiphophorus lines and species for assessing allele specific gene expression within F1 interspecies hybrids. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* **155**, 102-108.

Slate J, Gratten J, Beraldi D, *et al.* (2009) Gene mapping in the wild with SNPs: guidelines and future directions. *Genetica* **136**, 97-107.

Snellings, P. L. (2015). *Economic Value of the Tournament Black Bass Fishery on Lake Guntersville, Alabama* (Doctoral dissertation, Auburn University).

Stephens MR, Clipperton NW, May B (2009) Subspecies-informative SNP assays for evaluating introgression between native golden trout and introduced rainbow trout. *Molecular ecology resources* **9**, 339-343.

Taubert, B. D., & Tranquilli, J. A. (1982). Verification of the formation of annuli in otoliths of largemouth bass. *Transactions of the American Fisheries Society*,*111*(4), 531-534.

Thurston M, Field D (2005) Msatfinder: detection and characterisation of microsatellites. Distributed by the authors at *http://www.genomics.ceh.ac.uk/msatfinder/*. CEH Oxford, Mansfield Road, Oxford OX1 3SR.

Tucci S, Primassin S, Spiekerkoetter U (2010) Fasting-induced oxidative stress in very long chain acyl-CoA dehydrogenase-deficient mice. *FEBS journal* **277**, 4699-4708.

USDI (2011). U.S. Department of the Interior, Fish and Wildlife Service, and U.S. Department of Commerce, U.S. Census Bureau. 2011 National Survey of Fishing, Hunting, and Wildlife-Associated Recreation.

Wang W-A, Groenendyk J, Michalak M (2012) Calreticulin signaling in health and disease. *The international journal of biochemistry & cell biology* **44**, 842-846.

Wege, G., & Anderson, R. (1978). New Approaches to the Management of Small Impoundments, chapter Relative weight (*Wr*): a new index of condition for largemouth bass. *American Fisheries Socity Special Publication*, (5), 2.

Williams DJ, Kazianis S, Walter RB (1998) Use of random amplified polymorphic DNA (RAPD) for identification of largemouth bass subspecies and their intergrades. *Transactions of the American Fisheries Society* **127**, 825-832.

Williamson, J. H., Carmichael, G. J., Schmidt, M. E., & Morizot, D. C. (1986). New biochemical genetic markers for largemouth bass. *Transactions of the American Fisheries Society*, *115*(3), 460-465.

Williamson, J. H., & Carmichael, G. J. (1990). An aquacultural evaluation of Florida, northern, and hybrid largemouth bass, Micropterus salmoides.*Aquaculture*, *85*(1), 247-257.

Yorikawa C, Shibata H, Waguri S, *et al.* (2005) Human CHMP6, a myristoylated ESCRT-III protein, interacts directly with an ESCRT-II component EAP20 and regulates endosomal cargo sorting. *Biochem. J* **387**, 17-26.

Zhao L, Li Y, He M, *et al.* (2014) The Fanconi anemia pathway sensitizes to DNA alkylating agents by inducing JNK-p53-dependent mitochondrial apoptosis in breast cancer cells. *International journal of oncology* **45**, 129-138.

Zolczynski Jr, S. J., & Davies, W. D. (1976). Growth characteristics of the northern and Florida subspecies of largemouth bass and their hybrid, and a comparison of catchability between the subspecies. *Transactions of the American Fisheries Society*, *105*(2), 240-243.

1347

Appendices

Appendix 1 List of 63 multiplex primers for Sequenom MassARRAY both 25-Plex and *38-Plex SNP panels

Contig ID	Forward Primer Sequence	Reverse Primer Sequence	Extended Primer Sequence
Contig8751	ACGTTGGATGTTCTCTGGTGATCTGTGTGG	ACGTTGGATGAGTGCTGTTCTTAGCCGTTC	AACACCAGCCCTGCG
Contig33105	ACGTTGGATGGTTTTGCACAGCGACCACAG	ACGTTGGATGTAAAGTGAGCCAGTCCTCCG	GCGCTGTGCGGCGTT
Contig2930	ACGTTGGATGTCATTGTCCCTGACATGTGG	ACGTTGGATGGGCTGTGAAAAAGCTAGACC	TGTGGCTGAAGGACG
Contig4716	ACGTTGGATGACAACTGGGAAAAGCTCAGG	ACGTTGGATGCAAAGAGCTCGTCCAACTTG	TACCACACCCTCACTCG
Contig35112	ACGTTGGATGGGAGCCTCCTATGATCACTA	ACGTTGGATGCTCCCTTGTAGTCAGGGTTG	GGGTTGTCTATCTGCTT
Contig5903	ACGTTGGATGCGTTCTAGTTCTTATGTCAC	ACGTTGGATGTCCCTCCTAATCTGAACCTG	AGGGAAAGAGGATTGTC
Contig2993	ACGTTGGATGAAGACATAACGTGGTGCAAG	ACGTTGGATGCTCCCACAGTATGTCAACTC	tctACATGCTGTCCCAGG
Contig31857	ACGTTGGATGACTAAAACCACTGGCCAGTC	ACGTTGGATGGGGAAAGTGGCCTCTAAAAG	cAAACCCTCCCTGCGTGCT
Contig20911	ACGTTGGATGGTGCCTGTCCAGTTTTGAAT	ACGTTGGATGGCAAATGACAACAAATCCA	CTGTAAAGTCTGTCTTGGC
Contig6106	ACGTTGGATGGGAGACTTTGATGTCAGAGG	ACGTTGGATGTCCAAGACTGACTGAGTCAC	cctTGAGTGTTTGGGGAGA
Contig33087	ACGTTGGATGATGGTGACAGCCACAGCATC	ACGTTGGATGTGTGAAGTTCGGAGCTGATG	aacacGCTCGTGCCCTCATG
Contig11367	ACGTTGGATGCAATGTTGGGCAGTTCAGAG	ACGTTGGATGTCATCCTTAGCTTTGACCAG	ccctCCCGGGCTGTCCGTGA
Contig35139	ACGTTGGATGTCTCACGTCTGTATGTGCTG	ACGTTGGATGTTTGTTGACCCCGTGGTAAC	ACCATACAGGAGAATGAAGG
Contig31992	ACGTTGGATGTTTAGTTCCACTGCCACCAC	ACGTTGGATGCTGGTGGATCTGGATCCAGT	ccACAAATGATCCAGGAAAAC
Contig10770	ACGTTGGATGTTCTGTGGTGCATGCATAAA	ACGTTGGATGCCTTACCTAACTACTACAAG	TGGTGCATGCATAAAAATGGC
Contig26936	ACGTTGGATGTGCTCATGGTAAGCGTCCTC	ACGTTGGATGAGAGGCAACAGGTAGAGAAG	gggaGTGAGGCTCGTGAGAAG
Contig5885	ACGTTGGATGCCAACATTTGTACCATTATC	ACGTTGGATGTTTAAGAGGCCCAAGAGCAC	AGCACAAACTGTAATCACAACTT
Contig25196	ACGTTGGATGAAGGTCCTTCACTCCCTCAG	ACGTTGGATGGTCAGAAAGGTGACACAGAG	aaccaCTCAGGAAGTCCTATTGC
Contig32455	ACGTTGGATGGAGATCATTCGTCACTCCTG	ACGTTGGATGAATGTTGACCACTGGTAGAG	TGGTAGAGTTAATTAACACAAAT
Contig17385	ACGTTGGATGCACTGCTAATGGAGTAGCTG	ACGTTGGATGCAGCAGTGGCTTGGAAATAC	ctGCTTGGAAATACTAGACTACAA
Contig25677	ACGTTGGATGGCCTTTTCTTATCGAAAGAG	ACGTTGGATGGCATGTTTAAGTGTTACCTC	aggggAAGTGTTACCTCAGAAAGA
Contig15421	ACGTTGGATGCATCTACTACACCCACACTG	ACGTTGGATGTGCAGAGGAGCATGTTGATG	ctcccACACCCACACTGATGGTTGC
Contig19092	ACGTTGGATGGCAAAAATGTGTTAAACGGG	ACGTTGGATGGCTACTAGGCTGAAACCTAC	ctgaAGACATAGTTGCTGACCTACT
Contig34438	ACGTTGGATGAACATGGAGATTCAAGTGAG	ACGTTGGATGGGAGATTGTGATACTATGAG	gAGTGAGATAAACTTTTGTGAAAAAT
Contig11367-1	ACGTTGGATGTCCCAGGATTTGCCTATCAG	ACGTTGGATGGGAGGAGATTGTTGTGTGAG	ggggaGGGGATAAGGATGTTCTTGTC
FLContig4919*	ACGTTGGATGGTAGGCCAAGCGTTAAAAGG	ACGTTGGATGTACTTTTGTTAAGTGCCACC	TGCCACCACTTCAAT
FLContig21621*	ACGTTGGATGATAAGCGTGGATGACTCAGC	ACGTTGGATGTGTAGGCCAATAGGATGGTG	ACCCACGACCAACAG

Appendix 1 continued List of 63 multiplex primers for Sequenom MassARRAY both 25-Plex and *38-Plex SNP panels

Contig ID	Forward Primer Sequence	Reverse Primer Sequence	Extended Primer Sequence
2NBContig8717*	ACGTTGGATGTTGTAGGCTCCGTCGTACTG	ACGTTGGATGATGGATGAGGAGCTGGAGAG	GGACGAGCGGGTCTT
FLContig15950*	ACGTTGGATGAGAGAAGCTGCAGAGGAATC	ACGTTGGATGTCTCTGTACAAAGTTGCGGG	GTTGCGGGTGATGCG
2FLContig5713*	ACGTTGGATGAGGTGATGAAAACCCGTCTG	ACGTTGGATGTTCACTCCCTCCTTCTTCAG	AGTACTGTCCGGTCTT
2FLContig13020*	ACGTTGGATGAAGGAGTGATGTTTGATGGG	ACGTTGGATGGGCTCAAAGTCAGAATCATC	TCTTCGTCATCCTCGCC
2FLContig22803*	ACGTTGGATGCAACACTGTTCCTCTTTCTC	ACGTTGGATGATCTGTCTGGCAGCCTCTTC	AGCCTCTTCCTCAACAA
2FLContig19961*	ACGTTGGATGTGATGAACAACCCGTCTAGG	ACGTTGGATGAGAAAGAGTCCAGCAGCCAG	TAAGCAACATTCAGGGA
2FLContig1240*	ACGTTGGATGAACGGTTTGGACTCAACAGG	ACGTTGGATGTTTCAAATTGGAGTTGAGGG	tAGTTGAGGGGATCCAT
2FLContig2242*	ACGTTGGATGGGATAGACATTGAGCACAGG	ACGTTGGATGCCATCAGGCACGAGTACATC	ttCAGACAACCACCGAAA
FLContig11272*	ACGTTGGATGAAGGAAGGCGATGTCTTCTG	ACGTTGGATGGTTTGAGCAGCAGATCAAGG	gggGAAAACCGGCAGTGA
2FLContig18667*	ACGTTGGATGGATCACCCTGTGGAAAACTG	ACGTTGGATGACTGCAGAGAACTTCAGAGC	ggaGCTGTGCACAGGGGA
2FLContig12388*	ACGTTGGATGTTCTTCTTTCCAGCGACGTG	ACGTTGGATGGCGGAGGACATGTTGTAAAC	TTAACTCTCAAAATGTCCG
2FLContig4936*	ACGTTGGATGTACTTCCCCGCCATTACCAC	ACGTTGGATGACCGACCTGTGCCAAAAAGC	TGCCAAAAAGCAACATGAC
NBContig12358*	ACGTTGGATGCCAATTCCATTTCCACTGGG	ACGTTGGATGTTCCAGTGGACCAAGTGGAG	ctgacCAGCCTCTTGTCCAC
FLContig21676*	ACGTTGGATGGGGATTTGAAAATGAGCACAC	ACGTTGGATGTAAGAGGTCCAGAGACTTTG	aCCAGAGACTTTGTCTGAAC
2FLContig36172*	ACGTTGGATGTCAACACTTGTGGCCCAATG	ACGTTGGATGAATACCTTCTCTGCAGGAGC	ggTTTGATCTTGATCAGGGC
FLContig3296*	ACGTTGGATGGTGAGTAAACAGGAAGACGG	ACGTTGGATGTGTGTTGTGGTCTGCATAGC	cctaTGTCTGTCTGCATGTCC
FLContig23008*	ACGTTGGATGTTTCTCCACAGCCTCATAAG	ACGTTGGATGTGGCAGTTCTAAATGGCATC	AATGGCATCAAAATCACAAGC
FLContig2635*	ACGTTGGATGTACTGGACAAGACATGGTGG	ACGTTGGATGTGGATTGACTGTCAGTGCTC	AAGCTAAATGTTGTACAAATTG
FLContig20908*	ACGTTGGATGAGTATCACCCAAATCTGCCG	ACGTTGGATGTGAGTGCGCTGAGACCAAAC	GAATGTGCCATAAAAACAATGA
FLContig18101*	ACGTTGGATGAAAAGATGTCCTCTCCTGCG	ACGTTGGATGCGGCCAGAAGAGAAAACCAC	ctccGAGAAAACCACTCCTAACT
FLContig1826*	ACGTTGGATGTCAAAGAGAGCGAGAGGTAG	ACGTTGGATGGTCATGTTCACACTCTGCTC	ttttaTGCTCCAACTGGATAATT
FLContig4773*	ACGTTGGATGTGTTTCCCTGTGCTTCTGTG	ACGTTGGATGTTCTTTCCCACTGAACCACC	gaagATAGCAATCTGAGGATGAT
2FLContig3379*	ACGTTGGATGCCAGCATTCTGTTGTTCACC	ACGTTGGATGTGAGGCCTTCAAACACAGAG	agagACACAGAGTGAAAGAGTAC
2FLContig9758*	ACGTTGGATGTGTCTGCACAGCTGCTTCTC	ACGTTGGATGGCCCCGCAGGAATTTAAAAG	AGTTTCACATGATAACAGTAACAG
FLContig16665*	ACGTTGGATGGAACCACTGCCTTAGATTAC	ACGTTGGATGGAGTGCATGCAAATCTGATG	ggcgGGCATTAAACCAACAAGCTA
FLContig23633*	ACGTTGGATGATCATGCTCACATTGTAGCC	ACGTTGGATGATCTAACATGCCTGTGGCTG	cctgCACACTGTCCGAATACTTCCG
2FLContig22709*	ACGTTGGATGAAAGGAGCCATCACCAGAAG	ACGTTGGATGTGTCTGGGTTGTCCACTTTG	cccCTTTGAGTGCAAACTTTATATC
2FLContig28601*	ACGTTGGATGTTCTACTGTAAAGTTAGGC	ACGTTGGATGTCCAATGGCCACTTAAGGAG	aatcAGGAGGGCTCCAAAAGAAATT
2FLContig2880*	ACGTTGGATGTTGCTCCATACTAGGTGTAG	ACGTTGGATGTGTGGGATCTGACGACAAAG	gagtTGAGCTATGAGTCTGACTAAG

Appendix 1 continued List of 63 multiplex primers for Sequenom MassARRAY both 25-Plex and *38-Plex SNP panels

Contig ID	Forward Primer Sequence	Reverse Primer Sequence	Extended Primer Sequence
FLContig3616*	ACGTTGGATGTCCAGTCTTTGAACCCTGTC	ACGTTGGATGCTGCTTTTAGCTGGTGTCTG	AGCTCATTTATTAAATGACTGACATC
FLContig21917*	ACGTTGGATGTCTTACAGATACAGAGCGCC	ACGTTGGATGTTGTTGGGTGGATGGATTGG	AAAACAAATAAAAACAATAAAGCAAA
FLContig17151*	ACGTTGGATGAGGAGTCTCTGGATCTGCTG	ACGTTGGATGTTGTGTAGAGGGAGTTGGAC	cccCGGGCACTGACCGCGCCGTCTCCA
FLContig6127*	ACGTTGGATGATGTGAAAAGCATGTACTGG	ACGTTGGATGCCAGGAATTGCCTTTGACTG	cctgCAAAGTTTTAAAAAGGGTTTACA
2FLContig6920*	ACGTTGGATGTCCCACCTTTCTTGAACTCC	ACGTTGGATGGCATCAAGGTAGCAAAGGAG	gggtCAAAGGAGAAGTTTGATTCAGCC
2FLContig9870*	ACGTTGGATGCTTAGAATCTGTCCCTCCTG	ACGTTGGATGTGGCGAATGATTTTGTGAGC	ccgTTGTGAGCAAGCTCAACTTTATTCC
2FLContig31979*	ACGTTGGATGAATGAATGCACAGGCTTGTC	ACGTTGGATGGGGCTAAGATGTATTCACAG	AAACCTAAATAAACAGAGAAAAAACAAA

1348

1349

Appendix 2 P-values from Mann-Whitney pairwise comparison of percent NLMB allele frequency by River System (significant at $p<0.05$). Genotypes include Florida largemouth (FLMB), Northern largemouth (NLMB). River systems include the Chattahoochee (CHA), Tallapoosa (TAL), Coosa (COO), Tennessee (TEN), Black Warrior (BLW), Alabama (ALA), Tombigbee (TOM), Mobile-Tensaw (MOT), and four smaller rivers that drain directly into the Gulf of Mexico (Other GOM).

	ALA	**BLW**	**CHA**	**COO**	**F1**	**FLMB**	**MOT**	**NLMB**	**Other GOM**	**TAL**	**TEN**
BLW	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CHA	<0.001	<0.001	NA	NA	NA	NA	NA	NA	NA	NA	NA
COO	<0.001	<0.001	<0.001	NA	NA	NA	NA	NA	NA	NA	NA
F1	<0.001	<0.001	0.003	0.043	NA	NA	NA	NA	NA	NA	NA
FLMB	<0.001	<0.001	<0.001	<0.001	<0.001	NA	NA	NA	NA	NA	NA
MOT	1	1	<0.001	<0.001	<0.001	<0.001	NA	NA	NA	NA	NA
NLMB	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	NA	NA	NA	NA
Other GOM	<0.001	<0.001	<0.001	1	0.004	<0.001	<0.001	<0.001	NA	NA	NA
TAL	<0.001	<0.001	<0.001	<0.001	1	<0.001	<0.001	<0.001	<0.001	NA	NA
TEN	1	1	<0.001	<0.001	<0.001	<0.001	1	<0.001	<0.001	<0.001	NA
TOM	1	1	<0.001	<0.001	<0.001	<0.001	1	<0.001	<0.001	<0.001	1

Appendix 3 Mann-Whitney p-values of pairwise comparison of mean heterozygosity by river system (significant at $p<0.05$). Genotypes include Florida largemouth (FLMB), Northern largemouth (NLMB). River systems include the Chattahoochee (CHA), Tallapoosa (TAL), Coosa (COO), Tennessee (TEN), Black Warrior (BLW), Alabama (ALA), Tombigbee (TOM), Mobile-Tensaw (MOT), and four smaller rivers that drain directly into the Gulf of Mexico (Other GOM).

	ALA	**BLW**	**CHA**	**COO**	**F1**	**FLMB**	**MOT**	**NLMB**	**Other GOM**	**TAL**	**TEN**
BLW	<0.001	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CHA	<0.001	<0.001	NA	NA	NA	NA	NA	NA	NA	NA	NA
COO	<0.001	<0.001	<0.001	NA	NA	NA	NA	NA	NA	NA	NA
F1	<0.001	<0.001	<0.001	<0.001	NA	NA	NA	NA	NA	NA	NA
FLMB	<0.001	<0.001	<0.001	<0.001	<0.001	NA	NA	NA	NA	NA	NA
MOT	0.019	<0.001	<0.001	<0.001	<0.001	<0.001	NA	NA	NA	NA	NA
NLMB	<0.001	<0.001	<0.001	<0.001	<0.001	1	<0.001	NA	NA	NA	NA
Other GOM	<0.001	1	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	NA	NA	NA
TAL	<0.001	<0.001	1	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	NA	NA
TEN	<0.001	<0.001	0.033	1	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	NA
TOM	0.144	0.128	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.213	<0.001	<0.001

Appendix 4 P-values from Mann-Whitney pairwise comparison of Q-value by River System (significant at $p<0.05$) Genotypes include Florida largemouth (FLMB), Northern largemouth (NLMB). River systems include the Chattahoochee (CHA), Tallapoosa (TAL), Coosa (COO), Tennessee (TEN), Black Warrior (BLW), Alabama (ALA), Tombigbee (TOM), Mobile-Tensaw (MOT), and four smaller rivers that drain directly into the Gulf of Mexico (Other GOM).

	ALA	**BLW**	**CHA**	**COO**	**F1**	**FLMB**	**MOT**	**NLMB**	**Other GOM**	**TAL**	**TEN**
BLW	0.407	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CHA	<0.001	<0.001	NA	NA	NA	NA	NA	NA	NA	NA	NA
COO	<0.001	<0.001	<0.001	NA	NA	NA	NA	NA	NA	NA	NA
F1	<0.001	<0.001	0.0147	0.011	NA	NA	NA	NA	NA	NA	NA
FLMB	<0.001	<0.001	<0.001	<0.001	<0.001	NA	NA	NA	NA	NA	NA
MOT	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	NA	NA	NA	NA	NA
NLMB	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	NA	NA	NA	NA
Other GOM	<0.001	0.009	<0.001	0.001	<0.001	<0.001	<0.001	<0.001	NA	NA	NA
TAL	<0.001	<0.001	<0.001	<0.001	0.155	<0.001	<0.001	<0.001	<0.001	NA	NA
TEN	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1	<0.001	NA
TOM	1	0.128	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Appendix 5 Pairwise Fst estimates for each river system and genotype sampled. Higher values indicate genetic distance and lower values indicate genetic similarity. Genotypes include Florida largemouth bass (FLMB), northern largemouth bass (NLMB), and F1 hybrids (F1). River systems include the Chattahoochee (CHA), Tallapoosa (TAL), Coosa (COO), Tennessee (TEN), Black Warrior (BLW), Alabama (ALA), Tombigbee (TOM), Mobile-Tensaw (MOT), and four smaller rivers that drain directly into the Gulf of Mexico (Other GOM).

	FLMB	**F1**	**NLMB**	**CHA**	**TAL**	**COO**	**TEN**	**BLW**	**ALA**	**TOM**	**MOT**
F1	0.7606	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NLMB	0.9869	0.7345	NA	NA	NA	NA	NA	NA	NA	NA	NA
CHA	0.3808	0.0626	0.6018	NA	NA	NA	NA	NA	NA	NA	NA
TAL	0.5201	0.0653	0.4768	0.0633	NA	NA	NA	NA	NA	NA	NA
COO	0.5923	0.1414	0.4107	0.1587	0.0349	NA	NA	NA	NA	NA	NA
TEN	0.6106	0.1354	0.2023	0.2574	0.121	0.0798	NA	NA	NA	NA	NA
BLW	0.832	0.363	0.5665	0.3333	0.1623	0.0626	0.1016	NA	NA	NA	NA
ALA	0.8461	0.4518	0.6151	0.3825	0.2053	0.0841	0.1384	0.0252	NA	NA	NA
TOM	0.8007	0.4284	0.5326	0.406	0.221	0.0895	0.1334	0.0176	0.0084	NA	NA
MOT	0.866	0.4925	0.6572	0.4058	0.2325	0.107	0.1533	0.0385	0.015	0.0164	NA
Other GOM	0.7635	0.2873	0.593	0.255	0.1177	0.0454	0.136	0.0598	0.0607	0.058	0.0627

Appendix 6 P-values from Mann-Whitney Pairwise comparisons of NLMB allele frequencies by location. P< 0.05 is considered significant. Genotype populations include Northern largemouth bass from Illinois (ILL), Northern largemouth bass from Minnesota (MNN), Florida largemouth bass from the Florida Bass Conservation Center (FBCC) and F1 hybrids (F1). Population samples include Lake Harding (HRD), Lake Eufaula (EUF), Harris Reservoir (HRS), Lake Martin (LAM), Yates Reservoir (YTS), Weiss Reservoir (WEI), Logan Martin Reservoir (LOM), Neely Henry Reservoir (NEE), Lay Lake (LAY), Lake Guntersville (GUN), Wheeler Reservoir (WHL), Wilson reservoir (WIL), Pickwick Reservoir (PIC) Bear Creek Reservoir (BCR), Jones Bluff (JON), Miller's Ferry (MLF), Claiborne (CLA), Lewis Smith Reservoir (LWS), Lake Tuscaloosa (TUS), Sipsey River (SIP), Demopolis (DEM), Big Bayou Canot (BBC), Crab Creek (CCR), D'Olive Bay (DOB), Tensaw Lake (TNSW), Fish River (FIS), Dog River (DOG), Fowl River (FWL), and Styx River (STX).

	F1	ILL	BCR	BBC	CLA	CCR	DOB	DEM	DOG	EUF	FIS	FBCC	FWL	GUN
ILL	<0.001	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
BCR	0.015	<0.001	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
BBC	0.006	<0.001	1.000	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CLA	0.010	<0.001	1.000	1.000	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CCR	0.005	<0.001	1.000	1.000	1.000	NA	NA	NA	NA	NA	NA	NA	NA	NA
DOB	0.029	<0.001	1.000	1.000	1.000	1.000	NA	NA	NA	NA	NA	NA	NA	NA
DEM	0.001	<0.001	1.000	1.000	1.000	1.000	1.000	NA	NA	NA	NA	NA	NA	NA
DOG	0.038	<0.001	1.000	1.000	1.000	1.000	1.000	1.000	NA	NA	NA	NA	NA	NA
EUF	0.002	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	NA	NA	NA	NA	NA
FIS	1.000	0.039	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.478	NA	NA	NA	NA
FBCC	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.008	NA	NA	NA
FWL	0.164	<0.001	0.003	<0.001	<0.001	<0.001	0.004	<0.001	0.033	<0.001	1.000	<0.001	NA	NA
GUN	<0.001	<0.001	0.186	0.213	0.715	0.529	1.000	<0.001	1.000	<0.001	1.000	<0.001	0.157	NA
HRD	0.415	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	<0.001	<0.001	<0.001
HRS	1.000	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	<0.001	0.016	<0.001
JOB	0.023	<0.001	1.000	1.000	1.000	1.000	1.000	1.000	1.000	<0.001	1.000	<0.001	0.002	1.000
LAM	1.000	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	<0.001	0.007	<0.001
LAY	1.000	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.002	<0.001	1.000	<0.001	1.000	<0.001
LWS	0.003	<0.001	1.000	1.000	1.000	1.000	1.000	1.000	0.939	<0.001	1.000	<0.001	<0.001	<0.001
LOM	0.020	<0.001	<0.001	<0.001	<0.001	<0.001	0.004	<0.001	0.021	<0.001	1.000	<0.001	1.000	0.070
MLF	0.001	<0.001	1.000	1.000	1.000	1.000	1.000	1.000	1.000	<0.001	1.000	<0.001	<0.001	0.003
NEE	0.003	<0.001	0.029	<0.001	0.003	<0.001	0.028	<0.001	0.150	<0.001	1.000	<0.001	1.000	1.000
PIC	0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.757	<0.001	<0.001	<0.001
SIP	<0.001	<0.001	1.000	1.000	1.000	1.000	1.000	1.000	0.557	<0.001	1.000	<0.001	<0.001	<0.001
STX	0.061	<0.001	0.027	<0.001	0.004	<0.001	0.032	<0.001	0.076	<0.001	1.000	<0.001	1.000	0.744
MNN	0.008	0.042	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.342	<0.001	<0.001	<0.001
TNSW	0.004	<0.001	1.000	1.000	1.000	1.000	1.000	1.000	1.000	<0.001	1.000	<0.001	<0.001	0.514
TUS	0.009	<0.001	1.000	1.000	1.000	1.000	1.000	1.000	1.000	<0.001	1.000	<0.001	0.017	1.000
WEI	0.001	<0.001	0.071	0.002	0.004	0.004	0.035	<0.001	1.000	<0.001	1.000	<0.001	0.662	1.000
WHL	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.427	<0.001	<0.001	<0.001
WIL	0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.303	<0.001	<0.001	<0.001
YTS	1.000	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	<0.001	0.002	<0.001

1379

Appendix 6 Continued P-values from Mann-Whitney Pairwise comparisons of NLMB allele frequencies by Location. P< 0.05 is considered significant. Genotype populations include Northern largemouth bass from Illinois (ILL), Northern largemouth bass from Minnesota (MNN), Florida largemouth bass from the Florida Bass Conservation Center (FBCC) and F1 hybrids (F1). Population samples include Lake Harding (HRD), Lake Eufaula (EUF), Harris Reservoir (HRS), Lake Martin (LAM), Yates Reservoir (YTS), Weiss Reservoir (WEI), Logan Martin Reservoir (LOM), Neely Henry Reservoir (NEE), Lay Lake (LAY), Lake Guntersville (GUN), Wheeler Reservoir (WHL), Wilson reservoir (WIL), Pickwick Reservoir (PIC) Bear Creek Reservoir (BCR), Jones Bluff (JON), Miller's Ferry (MLF), Claiborne (CLA), Lewis Smith Reservoir (LWS), Lake Tuscaloosa (TUS), Sipsey River (SIP), Demopolis (DEM), Big Bayou Canot (BBC), Crab Creek (CCR), D'Olive Bay (DOB), Tensaw Lake (TNSW), Fish River (FIS), Dog River (DOG), Fowl River (FWL), and Styx River (STX).

	HRD	**HRS**	**JOB**	**LAM**	**LAY**	**LWS**	**LOM**	**MLF**	**NEE**	**PIC**
HRS	<0.001	NA	NA	NA	NA	NA	NA	NA	NA	NA
JOB	<0.001	<0.001	NA	NA	NA	NA	NA	NA	NA	NA
LAM	<0.001	1.000	<0.001	NA	NA	NA	NA	NA	NA	NA
LAY	0.315	1.000	<0.001	1.000	NA	NA	NA	NA	NA	NA
LWS	<0.001	<0.001	1.000	<0.001	<0.001	NA	NA	NA	NA	NA
LOM	<0.001	<0.001	0.002	<0.001	0.302	<0.001	NA	NA	NA	NA
MLF	<0.001	<0.001	1.000	<0.001	<0.001	1.000	<0.001	NA	NA	NA
NEE	<0.001	<0.001	0.024	<0.001	<0.001	<0.001	1.000	<0.001	NA	NA
PIC	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	NA
SIP	<0.001	<0.001	1.000	<0.001	<0.001	1.000	<0.001	1.000	<0.001	<0.001
STX	<0.001	0.023	0.008	0.005	1.000	<0.001	1.000	<0.001	0.676	<0.001
MNN	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
TNSW	<0.001	<0.001	1.000	<0.001	<0.001	1.000	<0.001	1.000	0.004	<0.001
TUS	<0.001	<0.001	1.000	<0.001	<0.001	1.000	0.026	1.000	1.000	<0.001
WEI	<0.001	<0.001	0.212	<0.001	<0.001	<0.001	0.851	<0.001	1.000	<0.001
WHL	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000
WIL	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000
YTS	0.171	1.000	<0.001	1.000	1.000	<0.001	<0.001	<0.001	<0.001	<0.001

1380

1381

Appendix 6 Continued P-values from Mann-Whitney Pairwise comparisons of NLMB allele frequencies by Location. P< 0.05 is considered significant. Genotype populations include Northern largemouth bass from Illinois (ILL), Northern largemouth bass from Minnesota (MNN), Florida largemouth bass from the Florida Bass Conservation Center (FBCC) and F1 hybrids (F1). Population samples include Lake Harding (HRD), Lake Eufaula (EUF), Harris Reservoir (HRS), Lake Martin (LAM), Yates Reservoir (YTS), Weiss Reservoir (WEI), Logan Martin Reservoir (LOM), Neely Henry Reservoir (NEE), Lay Lake (LAY), Lake Guntersville (GUN), Wheeler Reservoir (WHL), Wilson reservoir (WIL), Pickwick Reservoir (PIC) Bear Creek Reservoir (BCR), Jones Bluff (JON), Miller's Ferry (MLF), Claiborne (CLA), Lewis Smith Reservoir (LWS), Lake Tuscaloosa (TUS), Sipsey River (SIP), Demopolis (DEM), Big Bayou Canot (BBC), Crab Creek (CCR), D'Olive Bay (DOB), Tensaw Lake (TNSW), Fish River (FIS), Dog River (DOG), Fowl River (FWL), and Styx River (STX).

	SIP	**STX**	**MNN**	**TNSW**	**TUS**	**WEI**	**WHL**	**WIL**
STX	<0.001	NA	NA	NA	NA	NA	NA	NA
MNN	<0.001	<0.001	NA	NA	NA	NA	NA	NA
TNSW	1.000	0.001	<0.001	NA	NA	NA	NA	NA
TUS	1.000	0.032	<0.001	1.000	NA	NA	NA	NA
WEI	<0.001	0.567	<0.001	0.005	1.000	NA	NA	NA
WHL	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	NA	NA
WIL	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.317	NA
YTS	<0.001	0.002	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

1382

Appendix 7 P-values from Mann-Whitney Pairwise comparisons of mean heterozygosity by Location. P< 0.05 is considered significant. Genotype populations include Northern largemouth bass from Illinois (ILL), Northern largemouth bass from Minnesota (MNN), Florida largemouth bass from the Florida Bass Conservation Center (FBCC) and F1 hybrids (F1). Population samples include Lake Harding (HRD), Lake Eufaula (EUF), Harris Reservoir (HRS), Lake Martin (LAM), Yates Reservoir (YTS), Weiss Reservoir (WEI), Logan Martin Reservoir (LOM), Neely Henry Reservoir (NEE), Lay Lake (LAY), Lake Guntersville (GUN), Wheeler Reservoir (WHL), Wilson reservoir (WIL), Pickwick Reservoir (PIC) Bear Creek Reservoir (BCR), Jones Bluff (JON), Miller's Ferry (MLF), Claiborne (CLA), Lewis Smith Reservoir (LWS), Lake Tuscaloosa (TUS), Sipsey River (SIP), Demopolis (DEM), Big Bayou Canot (BBC), Crab Creek (CCR), D'Olive Bay (DOB), Tensaw Lake (TNSW), Fish River (FIS), Dog River (DOG), Fowl River (FWL), and Styx River (STX).

	F1	**ILL**	**BCR**	**BBC**	**CLA**	**CCR**	**DOB**	**DEM**	**DOG**	**EUF**	**FIS**	**FBCC**	**FWL**	**GUN**
ILL	<0.001	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
BCR	0.001	<0.001	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
BBC	0.006	<0.001	<0.001	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CLA	0.009	<0.001	<0.001	1.000	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CCR	0.005	<0.001	<0.001	1.000	1.000	NA	NA	NA	NA	NA	NA	NA	NA	NA
DOB	0.006	<0.001	<0.001	1.000	1.000	1.000	NA	NA	NA	NA	NA	NA	NA	NA
DEM	<0.001	<0.001	<0.001	0.019	0.020	0.003	0.035	NA	NA	NA	NA	NA	NA	NA
DOG	0.007	<0.001	<0.001	1.000	1.000	1.000	1.000	1.000	NA	NA	NA	NA	NA	NA
EUF	0.002	<0.001	1.000	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	NA	NA	NA	NA	NA
FIS	0.332	0.040	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	NA	NA	NA	NA
FBCC	<0.001	1.000	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.008	NA	NA	NA
FWL	0.003	<0.001	1.000	<0.001	0.005	<0.001	0.004	0.168	0.032	0.002	1.000	<0.001	NA	NA
GUN	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	1.000	<0.001	<0.001	NA
HRD	0.001	<0.001	0.002	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.377	1.000	<0.001	<0.001	1.000
HRS	0.002	<0.001	0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.010	1.000	<0.001	<0.001	1.000
JOB	0.002	<0.001	<0.001	1.000	1.000	0.606	1.000	1.000	1.000	<0.001	1.000	<0.001	0.269	<0.001
LAM	0.002	<0.001	0.016	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	1.000	<0.001	<0.001	1.000
LAY	<0.001	<0.001	1.000	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	1.000	<0.001	1.000	1.000
LWS	0.001	<0.001	<0.001	0.002	0.015	0.003	0.018	1.000	1.000	<0.001	1.000	<0.001	0.449	<0.001
LOM	0.001	<0.001	1.000	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	1.000	<0.001	1.000	0.034
MLF	0.001	<0.001	<0.001	1.000	1.000	1.000	1.000	0.883	1.000	<0.001	1.000	<0.001	<0.001	<0.001
NEE	0.001	<0.001	1.000	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	1.000	<0.001	0.573	<0.001
PIC	0.001	<0.001	0.066	0.105	0.172	0.080	0.203	1.000	1.000	<0.001	1.000	<0.001	1.000	<0.001
SIP	<0.001	<0.001	<0.001	1.000	0.977	0.364	1.000	1.000	1.000	<0.001	1.000	<0.001	<0.001	<0.001
STX	0.010	<0.001	1.000	0.095	0.453	0.131	0.342	1.000	1.000	0.306	1.000	<0.001	1.000	0.001
MNN	0.007	1.000	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.297	1.000	<0.001	<0.001
TNSW	0.004	<0.001	<0.001	1.000	1.000	1.000	1.000	0.002	1.000	<0.001	1.000	<0.001	<0.001	<0.001
TUS	0.009	<0.001	1.000	0.005	0.019	0.003	0.021	1.000	0.240	0.002	1.000	<0.001	1.000	<0.001
WEI	0.001	<0.001	1.000	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	1.000	<0.001	0.019	<0.001
WHL	<0.001	<0.001	<0.001	0.363	0.334	0.067	0.424	1.000	1.000	<0.001	1.000	<0.001	1.000	<0.001
WIL	0.001	<0.001	1.000	<0.001	<0.001	<0.001	<0.001	0.019	0.008	<0.001	1.000	<0.001	1.000	<0.001
HRD	0.003	<0.001	1.000	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	1.000	<0.001	0.030	1.000

1383

Appendix 7 Continued P-values from Mann-Whitney Pairwise comparisons of mean heterozygosity by Location. P< 0.05 is considered significant. Genotype populations include Northern largemouth bass from Illinois (ILL), Northern largemouth bass from Minnesota (MNN), Florida largemouth bass from the Florida Bass Conservation Center (FBCC) and F1 hybrids (F1). Population samples include Lake Harding (HRD), Lake Eufaula (EUF), Harris Reservoir (HRS), Lake Martin (LAM), Yates Reservoir (YTS), Weiss Reservoir (WEI), Logan Martin Reservoir (LOM), Neely Henry Reservoir (NEE), Lay Lake (LAY), Lake Guntersville (GUN), Wheeler Reservoir (WHL), Wilson reservoir (WIL), Pickwick Reservoir (PIC) Bear Creek Reservoir (BCR), Jones Bluff (JON), Miller's Ferry (MLF), Claiborne (CLA), Lewis Smith Reservoir (LWS), Lake Tuscaloosa (TUS), Sipsey River (SIP), Demopolis (DEM), Big Bayou Canot (BBC), Crab Creek (CCR), D'Olive Bay (DOB), Tensaw Lake (TNSW), Fish River (FIS), Dog River (DOG), Fowl River (FWL), and Styx River (STX).

	HRD	**HRS**	**JOB**	**LAM**	**LAY**	**LWS**	**LOM**	**MLF**	**NEE**	**PIC**
HRS	1.000	NA	NA	NA	NA	NA	NA	NA	NA	NA
JOB	<0.001	<0.001	NA	NA	NA	NA	NA	NA	NA	NA
LAM	1.000	1.000	<0.001	NA	NA	NA	NA	NA	NA	NA
LAY	1.000	1.000	<0.001	1.000	NA	NA	NA	NA	NA	NA
LWS	<0.001	<0.001	1.000	<0.001	<0.001	NA	NA	NA	NA	NA
LOM	0.022	<0.001	<0.001	0.300	1.000	<0.001	NA	NA	NA	NA
MLF	<0.001	<0.001	1.000	<0.001	<0.001	0.105	<0.001	NA	NA	NA
NEE	<0.001	<0.001	<0.001	<0.001	1.000	<0.001	1.000	<0.001	NA	NA
PIC	<0.001	<0.001	1.000	<0.001	0.001	1.000	0.001	1.000	<0.001	NA
SIP	<0.001	<0.001	1.000	<0.001	<0.001	1.000	<0.001	1.000	<0.001	1.000
STX	0.008	0.003	1.000	0.018	1.000	1.000	1.000	0.370	1.000	1.000
MNN	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
TNSW	<0.001	<0.001	0.610	<0.001	<0.001	0.001	<0.001	1.000	<0.001	0.027
TUS	<0.001	<0.001	1.000	<0.001	1.000	1.000	0.502	0.004	0.029	1.000
WEI	<0.001	<0.001	<0.001	0.001	1.000	<0.001	1.000	<0.001	1.000	<0.001
WHL	<0.001	<0.001	1.000	<0.001	<0.001	1.000	<0.001	1.000	<0.001	1.000
WIL	<0.001	<0.001	0.186	<0.001	0.031	0.324	0.006	<0.001	<0.001	1.000
YTS	1.000	1.000	<0.001	1.000	1.000	<0.001	1.000	<0.001	0.768	0.003

1384

1385

Appendix 7 Continued P-values from Mann-Whitney Pairwise comparisons of mean heterozygosity by Location. P< 0.05 is considered significant. Genotype populations include Northern largemouth bass from Illinois (ILL), Northern largemouth bass from Minnesota (MNN), Florida largemouth bass from the Florida Bass Conservation Center (FBCC) and F1 hybrids (F1). Population samples include Lake Harding (HRD), Lake Eufaula (EUF), Harris Reservoir (HRS), Lake Martin (LAM), Yates Reservoir (YTS), Weiss Reservoir (WEI), Logan Martin Reservoir (LOM), Neely Henry Reservoir (NEE), Lay Lake (LAY), Lake Guntersville (GUN), Wheeler Reservoir (WHL), Wilson reservoir (WIL), Pickwick Reservoir (PIC) Bear Creek Reservoir (BCR), Jones Bluff (JON), Miller's Ferry (MLF), Claiborne (CLA), Lewis Smith Reservoir (LWS), Lake Tuscaloosa (TUS), Sipsey River (SIP), Demopolis (DEM), Big Bayou Canot (BBC), Crab Creek (CCR), D'Olive Bay (DOB), Tensaw Lake (TNSW), Fish River (FIS), Dog River (DOG), Fowl River (FWL), and Styx River (STX).

	SIP	**STX**	**MNN**	**TNSW**	**TUS**	**WEI**	**WHL**	**WIL**
STX	1.000	NA	NA	NA	NA	NA	NA	NA
MNN	<0.001	0.002	NA	NA	NA	NA	NA	NA
TNSW	0.451	0.101	<0.001	NA	NA	NA	NA	NA
TUS	0.014	1.000	<0.001	0.002	NA	NA	NA	NA
WEI	<0.001	0.552	<0.001	<0.001	0.009	NA	NA	NA
WHL	1.000	1.000	<0.001	0.052	1.000	<0.001	NA	NA
WIL	<0.001	1.000	<0.001	<0.001	1.000	<0.001	1.000	NA
YTS	<0.001	0.920	<0.001	<0.001	0.014	1.000	<0.001	<0.001

1386

1387

Appendix 8 P-values from Mann-Whitney Pairwise comparisons of Q-value by Location. P< 0.05 is considered significant. Genotype populations include Northern largemouth bass from Illinois (ILL), Northern largemouth bass from Minnesota (MNN), Florida largemouth bass from the Florida Bass Conservation Center (FBCC) and F1 hybrids (F1). Population samples include Lake Harding (HRD), Lake Eufaula (EUF), Harris Reservoir (HRS), Lake Martin (LAM), Yates Reservoir (YTS), Weiss Reservoir (WEI), Logan Martin Reservoir (LOM), Neely Henry Reservoir (NEE), Lay Lake (LAY), Lake Guntersville (GUN), Wheeler Reservoir (WHL), Wilson reservoir (WIL), Pickwick Reservoir (PIC) Bear Creek Reservoir (BCR), Jones Bluff (JON), Miller's Ferry (MLF), Claiborne (CLA), Lewis Smith Reservoir (LWS), Lake Tuscaloosa (TUS), Sipsey River (SIP), Demopolis (DEM), Big Bayou Canot (BBC), Crab Creek (CCR), D'Olive Bay (DOB), Tensaw Lake (TNSW), Fish River (FIS), Dog River (DOG), Fowl River (FWL), and Styx River (STX).

	F1	ILL	BCR	BBC	CLA	CCR	DOB	DEM	DOG	EUF	FIS	FBCC	FWL	GUN
ILL	<0.001	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
BCR	0.003	<0.001	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
BBC	0.007	<0.001	<0.001	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CLA	0.012	<0.001	<0.001	1.000	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CCR	0.006	<0.001	<0.001	1.000	1.000	NA	NA	NA	NA	NA	NA	NA	NA	NA
DOB	0.033	<0.001	<0.001	1.000	1.000	1.000	NA	NA	NA	NA	NA	NA	NA	NA
DEM	<0.001	<0.001	<0.001	0.019	1.000	0.021	1.000	NA	NA	NA	NA	NA	NA	NA
DOG	0.048	<0.001	0.007	1.000	1.000	1.000	1.000	1.000	NA	NA	NA	NA	NA	NA
EUF	0.002	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	NA	NA	NA	NA	NA
FIS	1.000	0.005	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.278	NA	NA	NA	NA
FBCC	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	NA	NA	NA
FWL	0.049	<0.001	1.000	<0.001	<0.001	<0.001	0.002	<0.001	0.004	<0.001	1.000	<0.001	NA	NA
GUN	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	<0.001	1.000	NA
HRD	1.000	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	<0.001	<0.001	<0.001
HRS	1.000	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	<0.001	<0.001	<0.001
JOB	0.002	<0.001	0.694	<0.001	0.031	<0.001	0.040	1.000	1.000	<0.001	1.000	<0.001	0.008	<0.001
LAM	0.166	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	<0.001	<0.001	<0.001
LAY	1.000	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	<0.001	0.868	0.015
LWS	0.003	<0.001	<0.001	0.003	0.469	0.003	0.501	1.000	1.000	<0.001	1.000	<0.001	<0.001	<0.001
LOM	0.001	<0.001	0.077	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	<0.001	1.000	1.000
MLF	0.001	<0.001	<0.001	0.276	1.000	0.465	1.000	1.000	1.000	<0.001	1.000	<0.001	<0.001	<0.001
NEE	0.003	<0.001	0.250	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	<0.001	1.000	0.010
PIC	0.001	<0.001	0.001	1.000	1.000	1.000	1.000	1.000	1.000	<0.001	1.000	<0.001	<0.001	<0.001
SIP	<0.001	<0.001	<0.001	1.000	1.000	1.000	1.000	0.773	1.000	<0.001	1.000	<0.001	<0.001	<0.001
STX	0.015	<0.001	1.000	<0.001	<0.001	<0.001	0.002	<0.001	0.003	<0.001	1.000	<0.001	1.000	1.000
MNN	<0.001	1.000	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
TNSW	0.004	<0.001	<0.001	1.000	1.000	1.000	1.000	0.151	1.000	<0.001	1.000	<0.001	<0.001	<0.001
TUS	0.012	<0.001	1.000	<0.001	0.002	<0.001	0.013	0.164	0.058	<0.001	1.000	<0.001	1.000	<0.001
WEI	0.001	<0.001	0.089	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	<0.001	1.000	0.152
WHL	<0.001	<0.001	<0.001	1.000	1.000	1.000	1.000	1.000	1.000	<0.001	1.000	<0.001	<0.001	<0.001
WIL	0.001	<0.001	0.225	<0.001	0.023	<0.001	0.092	1.000	1.000	<0.001	1.000	<0.001	<0.001	<0.001
YTS	1.000	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.000	<0.001	<0.001	<0.001

1388

Appendix 8 Continued P-values from Mann-Whitney Pairwise comparisons of Q-value by Location. P< 0.05 is considered significant. Genotype populations include Northern largemouth bass from Illinois (ILL), Northern largemouth bass from Minnesota (MNN), Florida largemouth bass from the Florida Bass Conservation Center (FBCC) and F1 hybrids (F1). Population samples include Lake Harding (HRD), Lake Eufaula (EUF), Harris Reservoir (HRS), Lake Martin (LAM), Yates Reservoir (YTS), Weiss Reservoir (WEI), Logan Martin Reservoir (LOM), Neely Henry Reservoir (NEE), Lay Lake (LAY), Lake Guntersville (GUN), Wheeler Reservoir (WHL), Wilson reservoir (WIL), Pickwick Reservoir (PIC) Bear Creek Reservoir (BCR), Jones Bluff (JON), Miller's Ferry (MLF), Claiborne (CLA), Lewis Smith Reservoir (LWS), Lake Tuscaloosa (TUS), Sipsey River (SIP), Demopolis (DEM), Big Bayou Canot (BBC), Crab Creek (CCR), D'Olive Bay (DOB), Tensaw Lake (TNSW), Fish River (FIS), Dog River (DOG), Fowl River (FWL), and Styx River (STX).

	HRD	**HRS**	**JOB**	**LAM**	**LAY**	**LWS**	**LOM**	**MLF**	**NEE**	**PIC**
HRS	<0.001	NA	NA	NA	NA	NA	NA	NA	NA	NA
JOB	<0.001	<0.001	NA	NA	NA	NA	NA	NA	NA	NA
LAM	<0.001	1.000	<0.001	NA	NA	NA	NA	NA	NA	NA
LAY	0.043	1.000	<0.001	1.000	NA	NA	NA	NA	NA	NA
LWS	<0.001	<0.001	1.000	<0.001	<0.001	NA	NA	NA	NA	NA
LOM	<0.001	<0.001	<0.001	<0.001	1.000	<0.001	NA	NA	NA	NA
MLF	<0.001	<0.001	1.000	<0.001	<0.001	1.000	<0.001	NA	NA	NA
NEE	<0.001	<0.001	<0.001	<0.001	0.061	<0.001	1.000	<0.001	NA	NA
PIC	<0.001	<0.001	1.000	<0.001	<0.001	1.000	<0.001	1.000	<0.001	NA
SIP	<0.001	<0.001	0.041	<0.001	<0.001	1.000	<0.001	1.000	<0.001	1.000
STX	<0.001	<0.001	0.003	<0.001	1.000	<0.001	1.000	<0.001	1.000	<0.001
MNN	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
TNSW	<0.001	<0.001	0.005	<0.001	<0.001	0.051	<0.001	1.000	<0.001	1.000
TUS	<0.001	<0.001	1.000	<0.001	0.003	0.010	0.007	0.002	0.002	0.268
WEI	<0.001	<0.001	<0.001	<0.001	0.052	<0.001	1.000	<0.001	1.000	<0.001
WHL	<0.001	<0.001	1.000	<0.001	<0.001	1.000	<0.001	1.000	<0.001	1.000
WIL	<0.001	<0.001	1.000	<0.001	<0.001	1.000	<0.001	0.422	<0.001	1.000
YTS	0.024	1.000	<0.001	1.000	1.000	<0.001	<0.001	<0.001	<0.001	<0.001

1389

1390

Appendix 8 Continued P-values from Mann-Whitney Pairwise comparisons of Q-value by Location. P< 0.05 is considered significant. Genotype populations include Northern largemouth bass from Illinois (ILL), Northern largemouth bass from Minnesota (MNN), Florida largemouth bass from the Florida Bass Conservation Center (FBCC) and F1 hybrids (F1). Population samples include Lake Harding (HRD), Lake Eufaula (EUF), Harris Reservoir (HRS), Lake Martin (LAM), Yates Reservoir (YTS), Weiss Reservoir (WEI), Logan Martin Reservoir (LOM), Neely Henry Reservoir (NEE), Lay Lake (LAY), Lake Guntersville (GUN), Wheeler Reservoir (WHL), Wilson reservoir (WIL), Pickwick Reservoir (PIC) Bear Creek Reservoir (BCR), Jones Bluff (JON), Miller's Ferry (MLF), Claiborne (CLA), Lewis Smith Reservoir (LWS), Lake Tuscaloosa (TUS), Sipsey River (SIP), Demopolis (DEM), Big Bayou Canot (BBC), Crab Creek (CCR), D'Olive Bay (DOB), Tensaw Lake (TNSW), Fish River (FIS), Dog River (DOG), Fowl River (FWL), and Styx River (STX).

	SIP	**STX**	**MNN**	**TNSW**	**TUS**	**WEI**	**WHL**	**WIL**
STX	<0.001	NA	NA	NA	NA	NA	NA	NA
MNN	<0.001	<0.001	NA	NA	NA	NA	NA	NA
TNSW	1.000	<0.001	<0.001	NA	NA	NA	NA	NA
TUS	<0.001	0.362	<0.001	0.002	NA	NA	NA	NA
WEI	<0.001	1.000	<0.001	<0.001	<0.001	NA	NA	NA
WHL	1.000	<0.001	<0.001	1.000	0.118	<0.001	NA	NA
WIL	0.010	<0.001	<0.001	0.011	1.000	<0.001	0.603	NA
YTS	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

1391

Appendix 8 Pairwise Fst estimates for each location sampled. Higher values indicate genetic distance and lower values indicate genetic similarity. Genotype populations include Northern largemouth bass (NLMB), Florida largemouth bass (FLMB), and F1 hybrids (F1). Population samples include Lake Harding (HRD), Lake Eufaula (EUF), Harris Reservoir (HRS), Lake Martin (LAM), Yates Reservoir (YTS), Weiss Reservoir (WEI), Logan Martin Reservoir (LOM), Neely Henry Reservoir (NEE), Lay Lake (LAY), Lake Guntersville (GUN), Wheeler Reservoir (WHL), Wilson reservoir (WIL), Pickwick Reservoir (PIC) Bear Creek Reservoir (BCR), Jones Bluff (JON), Miller's Ferry (MLF), Claiborne (CLA), Lewis Smith Reservoir (LWS), Lake Tuscaloosa (TUS), Sipsey River (SIP), Demopolis (DEM), Big Bayou Canot (BBC), Crab Creek (CCR), D'Olive Bay (DOB), Tensaw Lake (TNSW), Fish River (FIS), Dog River (DOG), Fowl River (FWL), and Styx River (STX).

	FLMB	**F1**	**NLMB**	**HRD**	**EUF**	**HRS**	**LAM**	**YTS**	**WEI**	**NEE**	**LOM**	**LAY**	**GUN**
F1	0.7606	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NLMB	0.9869	0.7345	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
HRD	0.4859	0.0522	0.6229	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
EUF	0.4271	0.0999	0.7363	0.0327	NA	NA	NA	NA	NA	NA	NA	NA	NA
HRS	0.6196	0.0553	0.5551	0.044	0.119	NA	NA	NA	NA	NA	NA	NA	NA
LAM	0.6602	0.0861	0.5976	0.0622	0.1329	0.0327	NA	NA	NA	NA	NA	NA	NA
YTS	0.6671	0.0941	0.6671	0.034	0.1007	0.0195	0.0217	NA	NA	NA	NA	NA	NA
WEI	0.7578	0.1802	0.4829	0.1719	0.2874	0.0629	0.0964	0.1044	NA	NA	NA	NA	NA
NEE	0.7619	0.196	0.509	0.179	0.2906	0.0679	0.0908	0.1008	0.0019	NA	NA	NA	NA
LOM	0.7341	0.1638	0.5414	0.136	0.2425	0.0337	0.0609	0.0599	0.0128	0.0097	NA	NA	NA
LAY	0.5805	0.1039	0.5629	0.0612	0.1302	0.0155	0.0455	0.025	0.0795	0.0713	0.0452	NA	NA
GUN	0.5598	0.0813	0.2352	0.1555	0.2476	0.0673	0.0862	0.1116	0.0436	0.0549	0.0608	0.1085	NA
WHL	0.8274	0.3466	0.1978	0.404	0.5177	0.2797	0.3103	0.3598	0.1635	0.1851	0.2262	0.3208	0.0756
WIL	0.8523	0.303	0.2639	0.3496	0.4679	0.2273	0.2631	0.309	0.124	0.1449	0.1806	0.2672	0.0619
PIC	0.8772	0.3463	0.2912	0.3806	0.4993	0.2595	0.2928	0.3437	0.1567	0.1771	0.2146	0.2987	0.0754
BCR	0.7732	0.1903	0.4301	0.2142	0.3297	0.0898	0.1202	0.1395	0.0161	0.0185	0.0358	0.1066	0.0381
LWS	0.8616	0.3732	0.6151	0.3082	0.4326	0.1705	0.2105	0.2274	0.0587	0.044	0.0708	0.1589	0.1107
TUS	0.8955	0.2866	0.6861	0.2527	0.3687	0.1316	0.1686	0.1809	0.0465	0.0329	0.0635	0.1192	0.0954
JOB	0.8753	0.3537	0.6696	0.2791	0.4024	0.1487	0.196	0.2	0.0575	0.0435	0.0656	0.1279	0.1183
MLF	0.8934	0.4503	0.6968	0.3456	0.4727	0.2126	0.2659	0.2747	0.0898	0.0745	0.1108	0.1788	0.1455
CLA	0.938	0.4094	0.7877	0.3121	0.4331	0.1867	0.2329	0.2429	0.0823	0.0686	0.1016	0.1598	0.146
SIP	0.8597	0.4546	0.6197	0.3767	0.5004	0.2329	0.2843	0.2946	0.1006	0.0799	0.1261	0.1983	0.1464
DEM	0.8032	0.3921	0.5393	0.3467	0.463	0.1983	0.2519	0.257	0.0753	0.0589	0.0969	0.1681	0.1292
BBC	0.9364	0.4383	0.7862	0.3334	0.454	0.2058	0.2583	0.2761	0.1048	0.0916	0.1243	0.1758	0.1528
CCR	0.9344	0.4412	0.7879	0.3306	0.4529	0.2074	0.2587	0.2717	0.1087	0.0943	0.1243	0.1754	0.1537
DOB	0.92	0.3912	0.7659	0.3016	0.4194	0.1768	0.2324	0.238	0.0925	0.0797	0.1032	0.1486	0.144
TNSW	0.9147	0.4103	0.7401	0.3185	0.439	0.1915	0.238	0.2482	0.0922	0.0747	0.1073	0.1628	0.1399
DOG	0.9059	0.3453	0.7583	0.2747	0.3818	0.1511	0.2131	0.2109	0.0891	0.0783	0.0889	0.1304	0.1419
FIS	0.9389	0.1974	0.8624	0.1141	0.2095	0.0349	0.0591	0.0455	0.0384	0.0167	0.0063	0.0117	0.0883
FWL	0.8262	0.256	0.6982	0.201	0.298	0.0851	0.1478	0.1362	0.0723	0.0617	0.0521	0.0668	0.1266
STX	0.8557	0.2301	0.7283	0.1768	0.2683	0.0737	0.1224	0.1128	0.061	0.0535	0.0443	0.0647	0.1058

1392

Appendix 8 Continued Pairwise Fst estimates for each location sampled. Higher values indicate genetic distance and lower values indicate genetic similarity. Genotype populations include Northern largemouth bass (NLMB), Florida largemouth bass (FLMB), and F1 hybrids (F1). Population samples include Lake Harding (HRD), Lake Eufaula (EUF), Harris Reservoir (HRS), Lake Martin (LAM), Yates Reservoir (YTS), Weiss Reservoir (WEI), Logan Martin Reservoir (LOM), Neely Henry Reservoir (NEE), Lay Lake (LAY), Lake Guntersville (GUN), Wheeler Reservoir (WHL), Wilson reservoir (WIL), Pickwick Reservoir (PIC) Bear Creek Reservoir (BCR), Jones Bluff (JON), Miller's Ferry (MLF), Claiborne (CLA), Lewis Smith Reservoir (LWS), Lake Tuscaloosa (TUS), Sipsey River (SIP), Demopolis (DEM), Big Bayou Canot (BBC), Crab Creek (CCR), D'Olive Bay (DOB), Tensaw Lake (TNSW), Fish River (FIS), Dog River (DOG), Fowl River (FWL), and Styx River (STX).

	WHL	**WIL**	**PIC**	**BCR**	**LWS**	**TUS**	**JOB**	**MLF**	**CLA**	**SIP**
WIL	0.0074	NA	NA	NA	NA	NA	NA	NA	NA	NA
PIC	-7E-04	0.0073	NA	NA	NA	NA	NA	NA	NA	NA
BCR	0.1207	0.086	0.1125	NA	NA	NA	NA	NA	NA	NA
LWS	0.2211	0.1871	0.2204	0.0485	NA	NA	NA	NA	NA	NA
TUS	0.2117	0.1651	0.2045	0.0312	0.0217	NA	NA	NA	NA	NA
JOB	0.2474	0.2024	0.2481	0.0504	0.028	0.0215	NA	NA	NA	NA
MLF	0.2742	0.2381	0.2842	0.0839	0.036	0.0313	0.0052	NA	NA	NA
CLA	0.2712	0.2315	0.2782	0.0775	0.0371	0.0229	0.0139	-0.001	NA	NA
SIP	0.2682	0.2342	0.2723	0.0801	0.0361	0.0238	0.016	0.0165	0.0127	NA
DEM	0.2457	0.2055	0.2394	0.0631	0.0246	0.0074	0.0096	0.0101	0.0088	0.0085
BBC	0.2801	0.2392	0.2847	0.0983	0.0567	0.0383	0.0349	0.0224	0.018	0.0327
CCR	0.2877	0.2488	0.2966	0.0987	0.0507	0.0554	0.0243	0.0196	0.0362	0.0381
DOB	0.2838	0.2385	0.2866	0.086	0.043	0.0279	0.0216	0.0167	0.0116	0.0276
TNSW	0.2673	0.2241	0.27	0.0809	0.035	0.0222	0.0164	0.007	0.0025	0.0169
DOG	0.2926	0.2377	0.2904	0.0827	0.0684	0.0397	0.033	0.0446	0.0399	0.0519
FIS	0.3144	0.2586	0.317	0.0577	0.0987	0.0602	0.0628	0.1347	0.1338	0.1372
FWL	0.3135	0.2563	0.306	0.088	0.0926	0.0692	0.0607	0.1009	0.0944	0.1129
STX	0.2944	0.2389	0.2835	0.0757	0.0894	0.0672	0.0681	0.1096	0.1016	0.1269

1393

1394

Appendix 8 Continued Pairwise Fst estimates for each location sampled. Higher values indicate genetic distance and lower values indicate genetic similarity. Genotype populations include Northern largemouth bass (NLMB), Florida largemouth bass (FLMB), and F1 hybrids (F1). Population samples include Lake Harding (HRD), Lake Eufaula (EUF), Harris Reservoir (HRS), Lake Martin (LAM), Yates Reservoir (YTS), Weiss Reservoir (WEI), Logan Martin Reservoir (LOM), Neely Henry Reservoir (NEE), Lay Lake (LAY), Lake Guntersville (GUN), Wheeler Reservoir (WHL), Wilson reservoir (WIL), Pickwick Reservoir (PIC) Bear Creek Reservoir (BCR), Jones Bluff (JON), Miller's Ferry (MLF), Claiborne (CLA), Lewis Smith Reservoir (LWS), Lake Tuscaloosa (TUS), Sipsey River (SIP), Demopolis (DEM), Big Bayou Canot (BBC), Crab Creek (CCR), D'Olive Bay (DOB), Tensaw Lake (TNSW), Fish River (FIS), Dog River (DOG), Fowl River (FWL), and Styx River (STX).

	DEM	**BBC**	**CCR**	**DOB**	**TNSW**	**DOG**	**FIS**	**FWL**
BBC	0.0169	NA	NA	NA	NA	NA	NA	NA
CCR	0.0247	0.007	NA	NA	NA	NA	NA	NA
DOB	0.0119	0.0036	0.0063	NA	NA	NA	NA	NA
TNSW	0.0067	0.0109	0.0182	-0.001	NA	NA	NA	NA
DOG	0.0239	0.0298	0.0457	0.01	0.0256	NA	NA	NA
FIS	0.0869	0.1523	0.155	0.0919	0.1089	0.0456	NA	NA
FWL	0.0709	0.0938	0.0989	0.0635	0.0876	0.0295	0.0084	NA
STX	0.0754	0.1103	0.1158	0.075	0.0809	0.0432	-0.012	0.0266