Transactions of the Illinois State Academy of Science
(2009), Volume 102, #3&4, pp. 205-216

received 9/15/08
accepted 3/1/09

A Nuclear Gene Investigation of the Sunfish Genus *Lepomis* (Perciformes: Centrarchidae)

[1]David K. Peyton and [2,3]Robert L. Hopkins, II
[1]Department of Biological and Environmental Science
Morehead State University, Morehead, Kentucky
[2]School of Sciences
University of Rio Grande, Rio Grande OH 45674
[3]Corresponding Author, 1-740-245-7252, rhopkins@rio.edu

ABSTRACT

In this paper, we test the phylogenetic utility of the nuclear gene beta-actin for the sunfish genus *Lepomis* using maximum parsimony and maximum likelihood analyses. Using only the beta-actin sequence all major species-clades were recovered in both analyses. However, more basal nodes were only weakly supported. Results indicate that beta-actin performs moderately well as a phylogenetic marker and should be subjected to further use. Secondly, we explore the utility of the beta-actin gene to develop a rapid, non-lethal polymerase chain reaction (PCR)-based species identification method. We developed taxon-specific primers to identify *Lepomis* species based on polymorphisms in the introns of the beta-actin nuclear gene. Products were diagnostic for all species except for the *L. cyanellus*/*L. symmetricus* and *L. minatus*/*L. punctatus* species pairs.

INTRODUCTION

The sunfish genus *Lepomis* is comprised of 12 species endemic to eastern North America: *Lepomis auritus* (redbreast sunfish); *L. gibbosus* (pumpkinseed); *L. gulosus* (warmouth); *L. microlophus* (redear sunfish); *L. cyanellus* (green sunfish); *L. macrochirus* (bluegill); *L. marginatus* (dollar sunfish); *L. megalotis* (longear sunfish); *L. miniatus* (redspotted sunfish); *L. punctatus*(spotted sunfish); *L. symmetricus*(bantam sunfish); *L. humilis* (orangespotted sunfish); (Page and Burr, 1991). The genus is about 15 million years old (Bolnick et al., 2006) and given its wide distribution and abundance has been subjected to a series of evolutionary investigations using several markers, including: allozymes, external morphology, osteology, mitochondrial DNA, and nuclear DNA (Wainwright and Lauder, 1992; Mabee, 1993; Near et al., 2004; Harris et al., 2005; Bolnick et al., 2006). Researchers have consistently recovered the more basal patterns of evolutionary divergence in *Lepomis* including the early divergence of a clade containing *L. cyanellus* and the recognition of some major species complexes such as *L. megalotis* – *L. marginatus* (Near et al., 2004; Harris et al., 2005; Bolnick et al. 2006). However, species-level relationships remain largely unresolved and certain species, such as *L. auritus* and *L. gibbosus*, repeatedly fall out in different places within phylogenies. It is unclear

why hypotheses of *Lepomis* relationships show such inconsistencies but possible reasons include rapid historical diversification, inadequate analyses of phylogenetic data, inadequate markers or taxon sampling, genetic introgression via historical and recent hybridization, or cryptic diversity (Near et al., 2004; Harris et al., 2005).

The objectives of this study are two-fold. First, we investigate the phylogenetic utility of the nuclear gene beta-actin using the sunfish genus *Lepomis* as the exemplar. We analyze the performance of the marker by estimating phylogenies using maximum parsimony and maximum likelihood tree estimation techniques. Our results are compared to published phylogenetic hypotheses of the genus.

Secondly, we explore the use of the EPIC-PCR approach to derive species-specific molecular markers for members of genus *Lepomis* using beta-actin. Given the frequency of misidentifications due to morphological similarities (e.g., juveniles often require scale counts to correctly identify species) and phenotypic plasticity (Harris et al., 2005), we evaluate the utility of our nuclear DNA primers for performing rapid PCR-based techniques that exploit variation in intron sequences for species identification. The importance of developing a technique for identifying fishes via a rapid, non-lethal genetic screening can be dramatic for some species, especially if the fish in question is endangered, threatened, or locally rare.

METHODS

Tissues samples were taken from individuals collected from different locations within their respective ranges. The *L. cyanellus* × *L. microlophus* hybrid specimens were artificially produced. In a previous study, we examined the entire beta-actin gene and found that the majority of the intronic differences between *L. cyanellus* and *L. macrochirus* occurred in the first two (of four) sequenced introns (Peyton, 2004). Based on that observation, we focused this study on sequencing the first 1000 nucleotides from the remaining ten species. This genomic fragment contains all but 27 nucleotides of exon 1, all of intron 1, all of exon 2, and most of intron 2.

Phylogenetic Analyses of Beta-actin Gene Sequences

The beta-actin sequences were aligned via a combination of manual and automatic techniques using BioEdit (Hall, 1999) and the Vector NTI software suite. Phylogenies were estimated using maximum parsimony (MP) and maximum likelihood analyses (ML) using PAUP* (D.L. Swofford, Sinauer Associates, INC., Sunderland, MA, 2002). Optimal trees were found via a heuristics search of 1000 pseudoreplicates with tree bisection-reconnection. A bootstrap analysis consisting of 1000 pseudoreplicates and 100 random addition sequences was completed for each analysis. Branches with bootstrap support of less than 50% were collapsed.

Modeltest (Posada and Crandall, 1998) was used to identify the appropriate parameters for the ML analysis. The ML parameters were as follows: A = 0.2772, C = 0.2350, G = 0.1956, T = 0.2922, A-C = 0.9629, A-G = 1.8763, A-T = 0.9282, C-G = 0.4553, C-T = 1.8763, G-T = 1.000, Ti/Tv = 1.0746, proportion of invariable sites = 0 and discrete gamma distribution shape parameter = 0.8325. For the MP analysis, gaps were treated as a fifth character.

Species-specific Oligonucleotide Primers

DNA was isolated from fin clip biopsies using the Qiagen DNeasy Kit. Three microliters were used for subsequent PCR reactions. Beta-actin gene primers were designed to hybridize with the first 9 codons and a conserved region in the second intron in accordance with EPIC-PCR techniques. Primer sequences were as follows: Forward: 5' atg gat gat gaa atc gcc gca ctg gtt 3'; Reverse: 5' aga taa ggc aca cag tca aga gag acc tgt 3'. Cycle parameters were: 95° C for 2 minutes, followed by 35 cycles of (95° C for 30sec, 47° C for 30sec, 72° C for 1minute). The genomic fragment produced from each fish was approximately 1 kb and was cloned into the pGEM-T Easy vector (Promega). Candidate clones were sequenced by MWG Biotech, Inc. and compared to known beta-actin genes for verification, and the exons were determined using consensus splice sites and the virtually invariant beta-actin amino acid sequence. The sequence data for *L. cyanellus* and *L. macrochirus* were taken from a previous study (Peyton 2004).

The polymerase chain reaction was carried out with each sample using the following amounts: 3 μl genomic DNA , 1.5 μl of each primer (from a 10 ng/μl stock), 12.5 μl of 2× GoTaq PCR master mix (Promega), and water up to 25 μl. Primer sequences, annealing temperature used, and number of cycles are given in Table 1. Samples were electrophoresed on a 1.8% agarose gel and visualized using ethidium bromide staining.

RESULTS

Phylogenetic Analyses

In total there were 280 parsimony informative characters. The monophyly of *Lepomis* was well supported in each analysis and is in agreement with recent phylogenetic studies of the Centrarchidae genera (Near et al., 2004; Harris et al. 2005). The MP analysis provided a moderate level of resolution (Figure 1). Notwithstanding the placement of *L. gulosus*, *L. gibbosus*, and *L. auritus*, all traditionally recognized species complexes within *Lepomis* were resolved with the MP analysis. The ML tree (Figure 2) was slightly less resolved than the MP tree. Again, traditional species complexes were recovered but the more basal relationships were poorly supported and the major species groups were collapsed into a polytomy based on weak nodal support.

Species-Specific Primers and PCR-based Identification

The gel images in Figure 3 illustrate typical screening results for each primer set detailed in Table 1. Clear identification was possible for each species except for distinction between *L. cyanellus*/*L. symmetricus* and *L. miniatus*/*L. punctatus*, respectively. In the case of the former pair, *L. cyanellus* and *L. symmetricus*, we had clear and reproducible bands for both using multiple primer pairs, but none could robustly distinguish the subtle nucleotide differences between the two. This explains the result that the hybrid (*L. cyanellus* cross to *L. microlophus*) is identified as *L. microlophus, L. cyanellus,* and *L. symmetricus*. In the case of the latter pair, *L. miniatus* and *L. punctatus*, we found that one primer pair effectively identifies *L. miniatus*, but the corresponding pair is only marginally effective at identifying *L. punctatus*. Attempts to use alternate primer pairs, or to reduce the stringency of the reaction, resulted in products for both species.

DISCUSSION

Phylogenetic Analyses

The MP tree was similar to nearly all published hypotheses of species relationships and recovered essentially all major species groups despite being constructed using only a single nuclear gene (Avise et al., 1977; Wainright and Lauder, 1992; Mabee, 1993; Near et al., 2004; Harris et al., 2005). The position of *L. gibbosus* was most enigmatic and was placed basal to all other *Lepomis* species. Other hypotheses (Near et al., 2004; Harris et al., 2005) place *L. gibbosus* in a more terminal position and sister to a clade consisting of *L. microlophus*, *L. minatus*, and *L. punctatus*. The position of *L. auritus* was also different from other hypotheses, however, Harris et al. (2005) noted that the recovery of a consistent placement of *L. auritus* has not been achieved and attributed much of this due to historical genetic introgression with other *Lepomis*. The MP analysis also resolved *L. gulosus* as sister to a clade containing *L. cyanellus*, *L. symmetricus*, *L. macrochirus*, and *L. humilis*. Other hypotheses (Near et al., 2004; Harris et al., 2005) consistently place *L. gulosus* as embedded within this clade and sister to the complex including *L. cyanellus* and *L. symmetricus*. In contrast, the position of *L. gulosus* was different in the ML analysis and that species was shown sister to the *L. cyanellus – L. symmetricus* clade which is in general agreement with other hypotheses. The position of *L. gibbosus* remained uncertain and these specimens comprised one of the four collapsed branches on the tree. *Lepomis auritus* was still shown as belonging to the *L. punctatus – L. minatus* clade.

In summary, the recovery of major species complexes and moderate resolution of more basal relationships within the genus *Lepomis* suggests that beta-actin is a potentially useful phylogenetic marker. Because we were primarily focused on generally testing the phylogenetic utility of beta-actin, our sample size is inadequate for developing a robust hypothesis of relationships with *Lepomis*. Additional sampling is needed to test for intraspecific variation not accounted for in the current study. The results, however, are encouraging and this marker should be subjected to further scrutiny using more comprehensive sampling and perhaps combining the data with that of previously published hypotheses.

Species-specific Oligonucleotide Primers

In the *Lepomis* specimens we examined we found ample polymorphic sites to distinguish the majority of the twelve species. Even among the species that were not reliably distinguished, *L. cyanellus*/*L. symmetricus* and *L. miniatus*/*L. punctatus*, there were sufficient SNPs to distinguish the species by sequence. It was our design to not rely on SNPs because of the probability that these substitutions could occur by chance in another species, eliminating the position as a species-specific character. All of the primers utilized in this study were designed to take advantage of multiple mismatches occurring in succession, or when possible to anneal to an insertion present in only a single case. Polymorphisms involving multiple nucleotide mutations are most likely the result of several mutational events, occurring over the course of more than one generation. An important consideration about the efficacy of this molecular test is whether the differences observed in the *Lepomis* sequences are stable across populations. In this study we used a limited number of specimens and further research is needed with a geographically expanded sample size to rigorously test for intraspecific stability. However, in a previous study our

primers recognized *L. macrochirus* taken from disparate geographical locations in the eastern US (Peyton, 2004).

The *L. miniatus*/*L. punctatus* complex provides an interesting example of how similar two sequences can be for separate species. *Lepomis miniatus*, was only recently elevated to species status, having previously been a subspecies of the spotted sunfish, *L. punctatus* (Warren, 1992). Their phylogenetic closeness is evident in each of our phylogenetic analyses: a close analysis of the alignment between the two sequences shows that the divergence consists of SNPs spaced out over the length of the two fragments. This is what we would expect as the result of random genetic drift occurring over an evolutionarily short period of time, as compared to the gaps and insertions seen among the other species, which would take longer as mentioned. A 1.7% difference in sequence composition is sufficient, hypothetically, to design allele-specific primers but we found the results of our allele-specific primers for these species to be less than satisfactory.

The use of molecular tests for species identification is becoming more widespread (Hebert at al., 2004; Noël et al., 2008). In the case of DNA barcoding (Herbert et al., 2004), a standard genomic position is sequenced and compared to a database. Each individual specimen must be sequenced for this region and potential problems arise when examining specimens of hybrid origin or taxa with extensive historical genetic introgression. Noël et al. (2008) used a technique similar to the one used here to design specific primers for identifying morphologically ambiguous individuals from unisexual populations of rare *Ambyostoma* salamanders. In their study, the primers they designed did not discriminate between species and only produced a positive/negative result for unisexual population origin.

The development of species or strain-specific tests for fishes has lagged due to continued use of allozyme tests and microsatellites. The former requires lethal sampling of tissues and organs, the latter is primarily suited for intraspecific population studies (DeWoody et al., 2000). Because microsatellites exhibit a degree of hypervariability, that likely makes them unsuitable for species identification. A solution to this problem is to find a gene that can be easily amplified and sequenced from any species, and to use polymorphisms within that gene which are likely to be conserved across populations.

The beta-actin gene is a good candidate for species identification because it is highly conserved at the amino acid level, suggesting that universal primers will be functional across multiple taxa (Baldauf et al., 2000). Our experience has indicated that the same beta-actin primers used in this study also can amplify beta-actin genes from species within Esocidae and Cyprinidae. However, potential complications arise when dealing with closely related species (e.g., *L. minatus*/*L. puntatus* and *L. cyanellus*/*L. symmetricus*).

Moreover, fishes present problems with contemporary hybridization and historical genetic introgression common in many groups (Hopkins and Eisenhour, 2008). However, in most natural *Lepomis* populations, there are perhaps relatively few hybrids between recognized species that are not identifiable as the F_1 generation (Avise and Saunders, 1984). As demonstrated, the primers described in this paper can be used to identify F_1 hybrids. More extensive hybridization cannot be distinguished with our primers and would certainly confound the screening results. In fact, hybridization and introgression

that extends beyond the first generation would confound even multi-gene screening tests. Determination of parentage for deeply introgressed specimens would approach speculation considering the randomizing powers of meiotic cross-over events, independent assortment of chromosomes, and the potential for instraspecific variation from individual to individual. A larger survey of multiple populations of sunfish would be necessary to determine if our primers can rigorously work on sunfish from distinct geographical regions.

In summary, we have utilized the intronic polymorphisms of the highly conserved beta-actin gene to develop oligonucleotide primers that can distinguish among most species of *Lepomis* sunfish and can be used to verify the parental species of a putative F_1 hybrid via a rapid PCR-based screening. The conserved nature of the beta-actin sequence makes this test amenable to development for other genera as well, and may have applications for identifying fish species from degraded samples (e.g., stomach contents), species with uninformative morphological features, matching larval fishes to adult species, or non-lethal identification of rare species.

ACKNOWLEDGEMENTS

The authors would like to thank David Eisenhour and Ben Brammell for help in collecting specimens, and Kristopher Fultz, Rebecca Yates, and Brent Kidd for help in preparing DNA and performing PCR reactions.

MATERIALS EXAMINED

Specimens were collected in the following locations: *Lepomis auritus* (MOSU 2340), Bell County, KY; *L. gibbosus* (MOSU 2333), Horry County, SC; *L. gulosus* (MOSU 2347) and *L. microlophus* (MOSU 2343), Clear Creek Lake in Bath County, KY; *L. cyanellus* (MOSU 2344) and *L. macrochirus* (MOSU 2345), South Elkhorn Creek in Woodford County, KY; *L. macrochirus* (SIUC 37963), Illinois; *L. marginatus* (MOSU2346), Graves County, KY; *L. megalotis* (MOSU 2348), Rowan County, KY and (UF146998), Monroe County, Georgia; *L. miniatus* (MOSU 2342), Mud River in Logan County, KY; *L. punctatus* (MOSU 2339), Horry County, SC; *L. symmetricus* (MOSU 2341), Graves County, KY; *L. humilis* (University of Tennessee Tissue Collection INHS 42594, SIUC 37962(2), IL; *M. dolomieu* (MOSU 2349), Rowan County, KY; *M. punctulatus* (MOSU 2350), Rowan County, KY; *L. cyanellus* × *L. microlophus* hybrid specimens were artificially produced as described.

REFERENCES

Avise JC, Straney DO, Smith MH. 1977. Biochemical genetics of sunfish IV. Relationships of centrarchid genera. Copeia 2:250-258.

Avise JC, Saunders NC. 1984. Hybridization and Introgression Among Species of Sunfish (*Lepomis*): Analysis by Mitochondrial DNA and Allozyme Markers. Genetics 108:237- 255.

Baldauf SL, Roger AJ, Wenk-Siefert I, Doolittle WF. 2000. A Kingdom-Level Phylogeny of Eukaryotes Based on Combined Protein Data. Science 290:972-977.

Bolnick DI, Near TK, Wainwright PC. 2006. Body size divergence promotes post-zygotic reproductive isolation in centrarchids. Evolutionary Ecology Research 8:903-913.

DeWoody JA, Fletcher DE, Wilkins SD, and Avise JC. 2000. Parentage and nest guarding in the Tessellated Darter (*Etheostoma olmstedi*) assayed by microsatellite markers. Copeia 3:740-747.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41:95-98.

Harris PM, Roe KJ, Mayde RL. 2005. A mitochondrial DNA perspective on the molecular systematics of the sunfish genus *Lepomis* (Actinopterygii: Centrachidae). Copeia 2:340- 346.

Hebert, PDN, Penton EH, Burns JM, Jansen, DH, and Hallwachs, W. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proceeding of the National Academy of Sciences 101(4):14812-14817.

Hopkins RL, Eisenhour DJ. 2008. Hybridization of *Lythrurus fasciolaris* (Cope) and *Lythrurus umbratilis* (Girard) (Cypriniformes: Cyprinidae) in the Ohio River basin. Copeia 1:162- 171.

Mabee PM. 1993. Phylogenetic interpretation of ontogenetic change: sorting out the actual and artifactual in an empirical case study of centrarchid fishes. Zool. J. Linnearn Soc., 107:175-291.

Near TJ, Bolnick DI, Wainright PC. 2004. Investigating phylogenetic relationships of sunfishes and black basses (Actinopterygii: Centrarchidae) using DNA sequences from mitochondrial and nuclear genes. Molecular Phylogenetics and Evolution 32:344-357.

Noël S, Dumoulin J, Ouellet, M, Galois P, Lapointe FJ. 2008. Rapid identification of salamanders from the Jefferson complex with taxon-specific primers. Copeia 1:158-161.

Page, LM and Burr, BM. 1991. A Field Guide to Freshwater Fishes. Houghton Mifflin Company. New York, NY

Peyton DK. 2004. A PCR-Based F1 Hybrid Screen Using the Beta-actin Genes from the Sunfishes *Lepomis cyanellus* and *L. macrochirus* (Centrarchidae). The Journal of the Kentucky Academy of Sciences 65:104-107.

Posada D, Crandall KA. 1998. Model Test: testing the model of DNA substitution. Bioinformatics 14:817-818.

Swofford DL, Olson GJ, Waddell, Hillis DM. 1996. Phylogenetic inference, p. 407-517. *In*: Molecular Systematics. DM Hillis, C Moritz, and BK Mable (eds). Sinauer Associates, Sunderland, Massachusetts.

Wainright PC, Lauder GV. 1992. The evolution of feeding biology in sunfishes (Centrarchidae). In: Mayden RL (Ed), Systematics, Historical Ecology, and North American Freshwater Fishes. Standford University Press. Stanford, CA. pp. 472-491.

Warren, ML. 1992. Variation of the Spotted Sunfish, *Lepomis punctatus* Complex (Centrarchidae): Meristics, Morphometrics, Pigmentation and Species Limits. Alabama Museum of Natural History Bulletin 12.

Table 1. Sequences, annealing temperature, and cycle number for species-specific oligonucleotide primers.

Species	Primers	Annealing Temp	Cycles
L. auritus	*Laur 113F cca ggc atc agg tga gca caa* *Laur 270R aga cct cat tag caa agc aat ttt cag tta tc*	55°C	30
L. cyanellus	*Start atg gat gat gaa aty gcc gca ctg gtt* *CyanR2 tgg tta gac ctc att aga tgt cag cat atg*	55°C	30
L. gibbosus	*Lgib 318F aaa taa gta ctg tat tat agg aaa tat tac cta gac a* *Lgib 458R ccc acc atc act ccc tga aca aga cat aat*	50°C	35
L. gulosus	*Lgul 168F ata aga act tgc tga tta tgg att tta ata ctt* *Lgul 440R tca ctc cct gaa gaa gac atg aca aac ttg tta*	50°C	35
L. humilis	*Start atg gat gat gaa aty gcc gca ctg gtt* *Lhum 421R ata tgt agc agt tcc taa tta aaa aag gta ctt*	47°C	40
L. macrochirus	*Lmac 146F ata aag cca cac cgt ttt tta tgg at* *SFI2R1 aga taa ggc aca cag tca aga gag acc tgt*	50°C	35
L. marginatus	*Lmar 113F cca ggc atc agg tga gtg agt gat* *Lmar 423R gat atg tag cag ttc cta att aaa ggg aaa gag g*	53°C	35
L. megalotis	*Lmeg 396F taa tta gga act gct aca tat cat ggt gg* *SFI2R1 aga taa ggc aca cag tca aga gag acc tgt*	53°C	35
L. microlophus	*Lmic 263F agc tct aac tgc taa gca aca ttt aca acc tg* *Lmic 399R aac agt tct tca tta aag gta aag agg taa at*	53°C	35
L. miniatus	*Lmin 360F caa gtt aca tat agt cag gat ctt tat g* *Lminpun 725R gtg caa ctc tgc atg tgc aga aag ggt ac*	48°C	35
L. punctatus	*Lpun 157F gaa ctt gct gat tat gta tta ata cat* *Lpun 416R cat gga tat gta aca gtt ctt aat taa aca*	47°C	35
L. symmetricus	*Lcyasym 264F tag ctg aaa att gct ttg ctc ata tgc tga cat c* *Lsym 401R gca tgc att tgt gtc tag gca ata caa tga*	53°C	35

Figure 1. Maximum parsimony tree of *Lepomis* spp. Based on nDNA beta-actin gene sequence. Bootstrap support is shown for each branch. State localities for each sample are abbreviated.

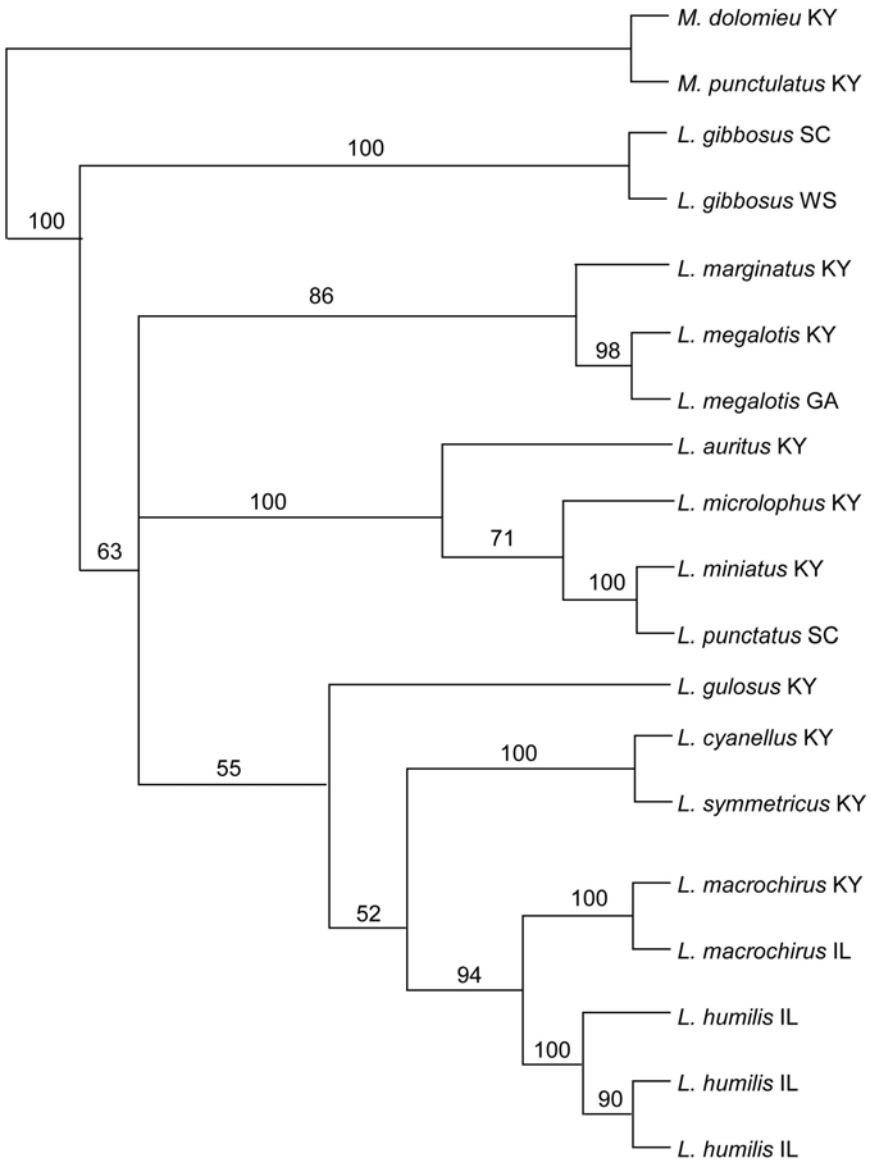

Figure 2. Maximum likelihood tree of *Lepomis* spp. Based on nDNA beta-actin gene sequence. Bootstrap support is shown for each branch. State localities for each sample are abbreviated.

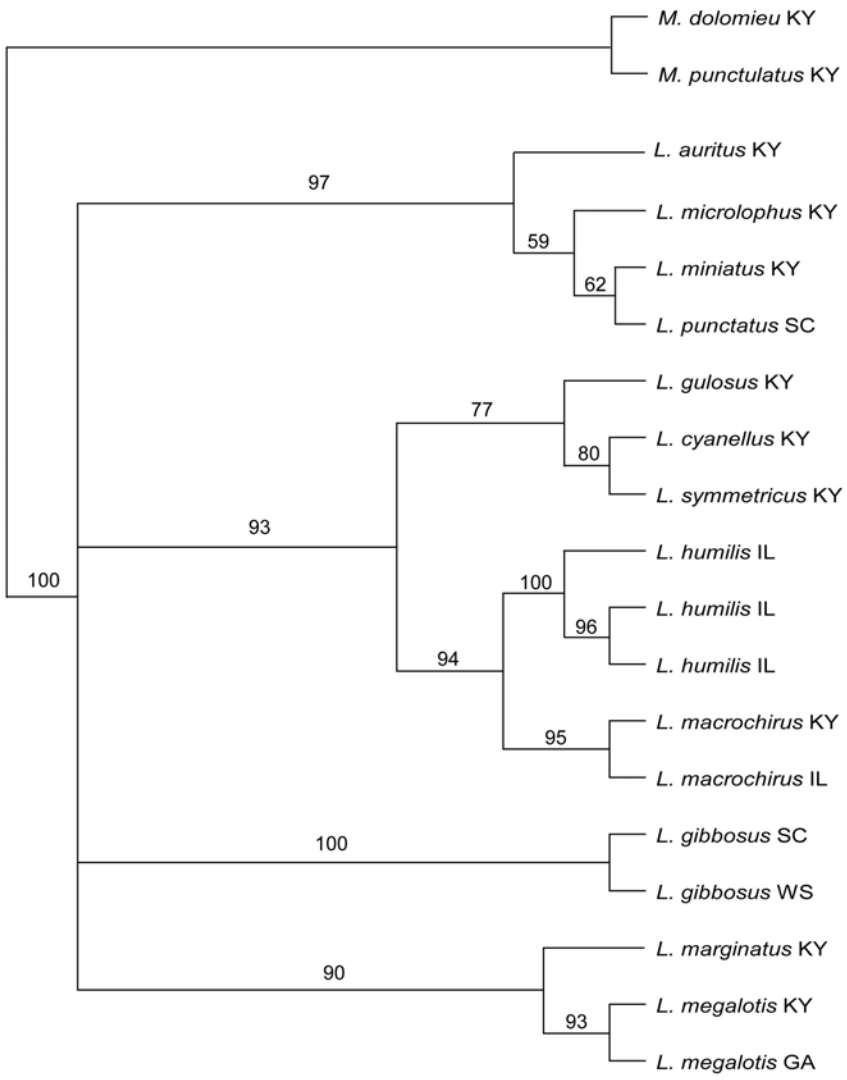

Figure 3. Products of species-specific PCR reactions to identify each of the twelve *Lepomis* species and one hybrid (*L. cyanellus* × *L. microlophus*). The first lane in each gel is a 100bp DNA ladder. The primer pair used in each case is indicated to the left of the photograph.