Biological Sciences, School of

Dissertations and Theses in Biological Sciences

University of Nebraska - Lincoln *Year* 2010

MICROSATELLITE DEVELOPMENT, POPULATION STRUCTURE AND DEMOGRAPHIC HISTORIES FOR TWO SPECIES OF AMAZONIAN PEACOCK BASS CICHLA TEMENSIS AND CICHLA MONOCULUS (PERCIFORMES: CICHLIDAE).

Jason C. Macrander
University of Nebraska-Lincoln, Jason507@gmail.com

This paper is posted at DigitalCommons@University of Nebraska - Lincoln.
http://digitalcommons.unl.edu/bioscidiss/20

MICROSATELLITE DEVELOPMENT, POPULATION STRUCTURE AND DEMOGRAPHIC HISTORIES FOR TWO SPECIES OF AMAZONIAN PEACOCK BASS *CICHLA TEMENSIS AND CICHLA MONOCULUS* (PERCIFORMES: CICHLIDAE).

By

Jason C. Macrander

A THESIS

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Master of Science

Major: Biological Sciences

Under the Supervision of Professor Etsuko Moriyama

Lincoln, Nebraska

December, 2010

MICROSATELLITE DEVELOPMENT, POPULATION STRUCTURE AND DEMOGRAPHIC HISTORIES FOR TWO SPECIES OF AMAZONIAN PEACOCK BASS *CICHLA TEMENSIS AND CICHLA MONOCULUS* (PERCIFORMES: CICHLIDAE).

Jason Macrander, M.S.

University of Nebraska, 2010

Adviser: Etsuko Moriyama

The Neotropics of South America represent one of the most diverse assemblages of freshwater organisms in the world. The geologic and ecological changes that have occurred throughout the Amazon and Orinoco River basins have resulted in the two most diverse rivers in the world producing very heterogeneous environments potentially influencing current population structure of Neotropical species. To investigate specific ecological and geological events influencing populations in these regions, I explore population structure and demographic histories of two species from the genus *Cichla* with wide distributions among the Amazon and Orinoco. First, I describe the methods used to isolate and characterize microsatellite loci for this genus. Using these microsatellite data as well as mitochondrial DNA sequences, I study the population structures for *C. temensis*, a species found in the Orinoco and Amazon spanning the Casiquiare River (a natural hydrogeologic corridor connecting the two rivers), and *C. monoculus*, which is the most widely distributed species in this genus found throughout the Amazon and the majority of its tributaries. The most significant results from my study show that the current population distribution of *C. temensis* is likely a result of population expansion

across the Casiquare River and the contrasting water types found in the Orinoco and Amazon Rivers seem to be limiting gene flow to the immediate boundaries of the Casiquiare River. The analysis of *C. monoculus* shows strong population structure not reflective of geologic events, specifically the breaching of the Purus Arch. However, the different water types seem to be the most likely factor influencing population structure. While the analysis for both species has indicated that geologic histories of these river basins may have influenced their current population structure, ecological variation seems to have a more dramatic effect on current gene flow across their wide distributions.

Acknowledgments

I would first like to thank my advisors and committee members: Dr. Etsuko Moriyama, Dr. Guillermo Orti, Dr. Andy Benson, Dr. Ed Cahoon and Dr. Tomas Hrbek. Your guidance throughout my pursuits at the University of Nebraska has been truly beneficial, something I will be able to carry with me the rest of my life. I also learned a lot from my fellow (and former) lab members (Dr. Chenhong Li, Dr. Mike Bessert, Dr. Cal Borden, Stuart Willis, Julie Sommer, Jeremy Brozek and Shane Gibson) your continued support and enthusiasm while pursuing my undergraduate and graduate degrees has been a source of inspiration and joy over the years. Although there are too many to mention here I would also like to thank the support of those faculty and students who have helped me tremendously throughout my academic career at the University of Nebraska. Finally the most important person I met at the University of Nebraska is my amazing wife, Ashley Bowers. Although we were just peers our first year as undergraduates on the same floor of the dorms it was several years later when we really found each other. Ashley has always been there for me and I will forever be thankful for her continued support as she is the most important person in my life.

Although there have been a lot of people I should thank, one person in particular has provided the most help while pursuing the research described in this thesis, Stuart Willis. While finishing up my undergraduate pursuits at UNL (2006) I approached Stuart about starting a collaborative project investigating the population structure for various species of *Cichla*. Stuart has been working with species from the genus *Cichla* throughout his academic career. Over the years, his travels to the Neotropics have allowed him to collect a significant tissue sample collection for the genus. Unfortunately,

I was unable to travel to the Neotropics and collect specimens for this study myself. However, his hard work and the hard work of others (many cited throughout this paper) has made this manuscript possible through their established network of tissue collection by researchers from both North and South America. Our collective enthusiasm in population genetics for species from this genus led to my acquiring SBS special funds as an undergraduate and graduate student. Additionally Stuart had helped out by completing the majority of lab work necessary for the mtDNA analysis. A large portion of sequences used in this study were from his previous work. Furthermore, he was solely responsible for acquiring the mitochondrial DNA sequences newly incorporated in this study. Upon returning to the University of Nebraska to pursue my graduate degree Stuart and Shane were a big help with characterizing the microsatellite loci when doing the initial primer tests. While at George Washington University, Stuart also had access to a high throughput genotyping facility and was able to collect loci for a large portion of the samples used in the microsatellite analysis. If it was not for Stuart Willis, none of the work described in this manuscript could be done.

Table of contents

List of Figures

List of Tables

CHAPTER 1 –AN OVERVIEW OF THE AMAZON AND ORINOCO RIVER BASINS AND NEOTROPICAL POPULATIONS

1.1 – Introduction

It is estimated that one-quarter of the total freshwater icthyofauna biodiversity found in the Neotropics of South America (Schaefer 1998). The Amazon and Orinoco drainages encompass approximately 7,880,000 km^2 of northern South America. The Amazon River has the largest basin in the world and is the second longest spanning, 6,400 km, from headwaters found in the Andes to the mouth that empties into the Atlantic Ocean (Hoorn and Wesselingh 2010). The Orinoco River is one of the largest in South America (2,140 km), 1/3 the length of the Amazon and the drainage covers just over 1/10 of the total area (830,000 km^2) (Lundberg et al. 1998). Despite a large geographic separation between the Orinoco and Amazon River Basins, there is a unique biogeographic corridor, the Casiquiare River, which connects the two most diverse rivers in the world allowing species distributions to transverse the very different river basins (Winemiller et al. 2008; Willis et al. 2010).

1.2 – Geologic History

A wide range of hypotheses have been proposed regarding major geologic events in the Neotropics, which are believed to be responsible for current population distributions and the observed icthyofauna diversity (Figure 1.1)**.** These events can be traced back to deposition events that began during the Proterozoic eon (3.0×10^{12} – 1.0×10^{12} years ago), as the heterogeneous craton make up the Guyana and Amazonian

shields (Tassinari and Macambira 1999). Currently, these shields make up the majority of the Amazon and Orinoco river drainages, which would be eventually be occupied by the Amazon River as it drains into the Atlantic. In the last 90 million years (MY), however, events in landscape evolution are thought to have had the most influence on the diversity observed in the Neotropics (Lundberg et al. 1998). Major geologic events that have been attributed to influencing the emergence of such diversity include the formation of the Andes, marine incursions, weathering of ancient shields, and palaeo-arches (Hoorn 1993; Lundberg et al. 1998; Hubert and Renno 2006). While these events have been used to explain terrestrial diversity (Bush 1994; de Fatima Rossetti 2005), the influence on the aquatic environment is known to a lesser extent. In the Amazon and surrounding river basins, the icthyofauna diversity has been mostly attributed to events occurring during the Neogene (between 15 – 10 MY ago) as extensive lakes and inland seas dominated the western Amazon basin, more commonly known as the Lake Pebas or Pebas Wetland System (Lundberg et al. 1998; Wesselingh et al. 2006)

The western half of what is now the Amazon drained into the vast lake/wetland system, which emptied north into the Caribbean Sea through what is now Venezuela. The eastern Guyana and Brazilian Shield rivers drained into the Atlantic Ocean through what is now known as the lower Amazon River Channel. The western Amazon was separated from the eastern waterways by the Purus Arch. The eventual uplift of the Venezuelan Andes and the Vuapes Arch between 5 and 8 MYA blocked the drainage of Lake Pebas into the Caribbean. Ultimately, the central Andes continued to rise, leading to the Purus Arch being breached connecting what is now the upper and lower Amazon River into a single basin for the first time (Lundberg et al. 1998; Gregory-Wodzicki 2000). The

separated western and eastern Amazon basins were reconnected allowing colonization and population expansions to occur (Farias and Hrbek 2008). The Casiquiare River (often referred to as the Casiquiare Canal) was established sometime after as a direct result of erosion and deposition in the Orinoco River. This led to a 300+ kilometer portion of the upper Orinoco emptying into the Rio Negro, thus connecting the Amazon and Orinoco River basins (Sioli 1984; Lundberg 1998).

1.3 – Ecological Variation

Across these drainages, waterways can vary considerably in color, sediment load, and chemical properties, all of which are strongly influenced by local watersheds. The diverse geology, vegetation cover, tributary discharge, and seasonal precipitation have greatly influenced the heterogeneous ecosystem present in the Amazon and Orinoco River drainages (Sioli 1984; Meade et al. 1991; Lundberg 1998). During rainy seasons or extended wet periods water temperatures and conductivities tend to decrease, while water depths, velocities, and dissolved oxygen concentrations tend to increase. These extended wet periods create fluctuations in the amount of dissolved organic and inorganic nutrients, which allows for an increase in primary production (Meade 1994; Winemiller et al. 2008).

These heterogeneous landscapes contribute to the varying water types throughout the basins, which are typically described based on the appearance or coloration of the water (Figure 1.2). “White-water” rivers are actually muddy in appearance due to high concentrations of dissolved solids from sediments originating from the Andes. They are also characterized as having an alkaline to neutral pH. Lowland rivers, like the Rios

Solimõs (Upper Amazon), Madeira, and some Orinoco tributaries drain from the Andes and exhibit this white-water type. In contrast, “black-water” rivers appear tea-colored due to high concentrations of dissolved organic matter. They are described as having very little suspended sediment loads with medium transparencies and are also very dilute in dissolved ions resulting in their usually acidic pH. The combined vegetation and soil type are attributed to the formation of the black-waters with common examples being the Rio Negro and some upper Orinoco tributaries. Lastly, “clear-water” rivers have high transparencies and are clear or olive-green in coloration. Their dissolved load is typically low, but this varies across different rivers. Clear-water Rivers are known to have a wide range of pH’s, from acidic to alkaline, and low suspended sediment loads, as seen in the Tapajos, Xingu and some Orinoco tributaries (Meade 1994; Dubroeucq and Volkoff 1998; Mayorga and Aufdenkampe 2002).

Although these rivers are commonly used as examples for different water types, it is certainly not the rule. Variation in factors such as precipitation throughout the year also results in fluctuating physiochemical environments from one season to another. The dynamic landscape makes classifying rivers difficult due to seasonal fluctuations in water levels, which affect dissolved and suspended sediment loads. Additionally, the Casiquiare is a unique hydrological corridor that links two very different physiochemical rivers (the upper Orinoco, "clear-water", with the Rio Negro, "black-water") (Winemiller and Jepsen 1998). Changes brought about by seasonal variation play a key part in the life history traits for many of the Amazon species ultimately influencing the food web, fish population distributions, migrations, and floodplain lake biota throughout the year (Rodriguez and Lewis 1997; Winemiller and Jepsen 1998).

1.4 – Peacock Bass, the Genus *Cichla*

The family Cichlidae is a large species-rich, ecologically diverse family in the Neotropics (Winemiller et al. 1997). This family has been the subject of much phylogenetic research (Lopez-Fernandez et al. 2010). The charismatic peacock basses are found exclusively throughout the Neotropics of South America, and form a monophyletic lineage within this family as the genus *Cichla*. This genus contains 15 described species and is widely distributed throughout the Amazon and Orinoco River basins, as well as the small rivers draining from the Guyana Shield to the Atlantic Ocean (Kullander and Ferreira 2006). They are among the largest representatives in the family Cichlidae and are highly sought after as sport fishes due to their aggressiveness when attacking bait fishes and artificial lures (Winemiller 2001). Their popularity among anglers has directly resulted in several exotic introductions affecting ecosystems in both negative (Zaret and Paine 1973; Pelicice and Agostinho 2009) and seemingly positive ways by controlling problematic introduced species (Shafland 1995; Hill et al. 2004), for example, resulting in *C. ocellaris* being the only legally introduced exotic fish species in Florida (Shafland et al. 2008).

Despite the ecological and economic interest, the genus had remained a relatively understudied group up until the past decade, with the majority of published studies focusing on introduced species in reservoirs and not their naturally occurring counterparts (Zaret and Paine 1973; Zaret 1980; Winemiller 2001; Kullander and Ferreira 2006). Aspects regarding their phylogenetic histories have only recently been the focus amongst the scientific community (Renno et al. 2006; Willis et al. 2007). Individuals from this

genus have been described to have relatively low levels of migration or dispersal ability (Hoeinghaus et al. 2003) and are believed to be substrate spawners providing some parental care (Zaret 1980). Although they exhibit life history characteristics that seem to constrain their dispersal ability, multiple species in the genus are found throughout the Amazon and Orinoco river basins with overlapping population distributions (Kullander and Ferreira 2006). Their distribution and unique life history traits lend *Cichla* to be a potentially informative genus when investigating population structure and mechanisms that gave rise to the diverse Neotropic ichthyofauna.

Examinations of genetic diversity for several Neotropical species have shown dramatically different factors contributing to the observed population structure. Population analyses have been used to reconstruct phylogeographic distributions of several Amazonian fishes (Hrbek et al. 2005; Ready et al. 2006; Hubert et al. 2007; Farias and Hrbek 2008), widespread South American species (Sivasundar et al. 2001, Montoya-Burgos 2003, Hubert and Renno et al. 2006), and even localized populations along a single tributary (Cooke et al. 2009) to test geologic and ecological hypotheses regarding species distributions and population structure. Some species exhibit tremendous population size and high rates of migration (Sivasundar et al. 2001; Santos et al. 2007), ultimately masking any historical geologic influencing current population structure. There have also been reported signatures of population structure exhibiting isolation by distance due to their wide distribution throughout the Amazon (Hrbek et al. 2005; Hubert et al. 2007). Several assumptions have been made regarding geological processes affecting current population structure in the Neotropics (Lundberg et al. 1998); however, analyses have yet to converge on a single explanation or model for population

structure across multiple taxa. A significant factor impeding our understanding of geologic processes influencing population structure is simply due to the lack of population level studies in the Neotropics (Beheregaray 2008). Additional population studies of Neotropical species should contribute our understanding of processes leading to the observed ichthyofauna diversity.

Although there is evidence of major geologic events and ecological variation influencing ichthyofauna population across these basins, results thus far do not converge on a single event, ecological factor, or time period that can best explain processes leading to the observed diversity. Using both mitochondrial and nuclear loci to determine the population structure of two *Cichla* species I hope to shed light on factors contributing to population distributions that span the Orinoco and Amazon rivers. First, I will use these loci to test whether *Cichla* populations exhibit panmictic distributions or if there is population structure. If there is apparent structure, I hope to clarify whether geologic events are responsible for the observed population distribution, specifically the breaching of the Purus Arch and the formation of the Casiquare River. Additionally, to establish whether or not the ecological variation found throughout these river basins may play a key role I will see whether the three distinctive river types coincide with their current distribution. Although the factors that influence Neotropical fish population distributions are intertwined and complex over the following chapters I hope to aid in our understanding of processes influencing the most diverse region on the planet.

References

Beheregaray, L. B. 2008. Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology*, 17, 3754 – 3774.

Bush, M. B. 1994. Amazonian speciation: a necessarily complex model. *Journal of Biogeography*, 21: 5 – 17.

Cooke, G. M., N. L. Chao, and L. B. Beheregarary. 2009. Phylogeography of a flooded forest specialist from central Amazonia based on intron DNA: the cardinal tetra *Paracheirodon axelrodi*. *Freshwater Biology*, 54 1216 – 1232.

de Fatima Rossetti, D., P. Mann de Toledo and A. M. Goes. 2004. New geological framework for Western Amazonia (Brazil) and implications for biogeography and evolution. *Quaternary Research*, 63: 78 – 89.

Dubroeucq, D. and B. Volkoff. 1998. From oxisoils to spodosoils and histosoils: evolution of the soil mantles in the Rio Negro basin (Amazonia). *Catena*, 32: 245 – 280.

Farias, I. P. and T. Hrbek. 2008. Patterns of diversification in the discus fishes (Symphysodon spp. Cichlidae) of the Amazon basin. *Molecular Phylogenetics and Evolution*, 49: 32 – 43.

Farias, I. P. and T. Hrbek. 2008. Patterns of diversification in the discus fishes (Symphysodon spp. Cichlidae) of the Amazon basin. *Molecular Phylogenetics and Evolution*, 49: 32 – 43.

Gregory-Wodzicki, K.M. 2000. Uplift history of the central and northern Andes: A review. *Geologic Society of America Bulletin*, 112: 1091 – 1105.

Hill, J. E., L. G. Nico, C. E. Cichra and C. R. Gilbert. 2004. Prey vulnerability to peacock cichlids and largemouth bass based on predator gape and prey body depth. *Proc. Annu. Conf. SEAFWA*. 58: 47 – 56.

Hoeinghaus, D. J., C. A. Layman, D. A. Arrington, and K. O. Winemiller. 2003. Spatiotemporal variation in fish assemblage structure in tropical floodplain creeks. *Environmental Biology of Fishes*, 67: 379 – 387.

Hoorn, C. 1993. Marine incursions and the influence of Andean tectonics on the Miocene depositional history of Northwestern Amazonia: results of a palynostratigraphic study. *Paleogeography, Paleoclimatology and Paleoecology*. 105: 267 – 309.

Hoorn, C. and F. P. Wesselingh. 2010. Introduction: Amazonia, landscape and species evolution. In: *Amazonia, Landscape and Species Evolution: A Look into the Past*. (eds C. Hoorn and F. P. Wesselingh), pp. 1 – 6. Blackwell Publishing.

Hrbek T., I.P. Farias, M. Crossa, I. Sampaio, J.I.R. Porto, and A. Meyer. 2005. Population genetic analysis of the *Arapaima gigas*, one of the largest freshwater fishes of the Amazon basin: implications for its conservation. *Animal Conservation*, 8(3): 297 – 308.

Hubert N. and J-F Renno. 2006. Historical biogeography of South American freshwater fishes. *Journal of Biogeography*, 33: 1414 – 1436.

Hubert N., F. Cuponchelle, J. Nunez, C. Garcia-Davila, D. Paugy, and J-F. Renno. 2007. Phylogeography of the piranha genera *Serrasalmus* and *Pygocentrus*: implications for the diversification of neotropical ichthyofauna. *Molecular Ecology*, 16: 2115 – 2136.

Kullander, S. O. and E. J. G. Ferreira. 2006. A review of the South American cichlid genus Cichla, with descriptions of nine new species (Teleostei: Cichlidae). *Ichthyological Exploration of Freshwaters*, 17(4): 289 – 398.

Lopez-Fernandez, H., K. O. Winemiller and R. L. Honeycutt. 2010. Multilocus phylogeny and rapid radiations in neotropical cichlid fishes (Perciformes: Cichlidae: Cichlinae). *Molecular Phylogenetics and Evolution*, 55 (3): 1070 – 1086.

Lundberg, J. G., L. G. Marshall, J. Guerrero, B. Horton, M.C.S.L. Malabarba, F. Wesselingh. 1998. The stage for Neotropical fish diversification: a history of tropical South American rivers. In: *Phylogeny and Classification of Neotropical Fishes* (eds L.R. Malabarba, R. E. Reis, R. P. Vari, Z. M. Lucena, C.A.S. Lucena CAS), pp. 13–48. Edipucrs, Porto Alegre.

Mayoraga, E. and A Aufdenkampe. 2002. Processing of bioactive elements in the Amazon river system. In: *the Ecohydrology of South American Rivers and Wetlands*. (eds H. Salz and Z. W. Kundzewicz), pp. 1 – 24. IAHS Press.

Meade, R. H. 1994. Suspended sediment of the modern Amazon and Orinoco rivers. *Quaternary international*, 21: 29 – 39.

Meade, R. H., J. M. Rayol, S. C. Da Conceicao, and J. R. G. Natividade. 1991. Backwater effects on the Amazon river basin of Brazil. *Environmental Geology*, 18 (2): 105 – 114.

Montoya-Burgos, J. I. 2003. Historical biogeography of the catfish genus *Hypostomus* (Siluriformes: Loricariidae), with implications on the diversification of neotropical ichthyofauna. *Molecular Ecology*, 12, 1855 – 1867.

Pelicice, F. M. and A. A. Agostinho. 2009. Fish fauna destruction after the introduction of a non-native predator (*Cichla kelberi*) in a neotropical reservoir. *Biological Invasions*. 11(8): 1789 – 1801.

Ready, J. S., E. J. G. Ferreira, and S. O. Kullander. 2006. Discus fishes: mitochondrial DNA evidence for a phylogeographic barrier in the Amazonian genus *Symphysodon* (Teleostei: Cichlidae). *Journal of Fish Biology*, 69: 200 – 211.

Renno J-F., Hubert N., Torrico J.P., Duponchelle F., Rodriguez J. N., Davila C. G., Willis S. C., and Desmarais E. 2006. Phylogeography of *Cichla* (Cichlidae) in the upper Madera basin (Bolivian Amazon). *Molecular Phylogenetics and Evolution*, 41: 503 – 510.

Rodriguez, M. A. and W. M Lewis. 1997. Structure of fish assemblages along environmental gradients in floodplain lakes of the Orinoco river. *Ecological Monographs*, 67 (1): 109 – 128.

Santos M.C.F., M.L. Ruffino and I.P. Farias. 2007. High levels of genetic variability and panmixia of the tambaqui *Colossoma macropomum* (Cuvier, 1816) in the main channel of the Amazon River. *Journal of Fish Biology*, 71 (Supplement A): 33 – 44.

Schaefer SA. 1998. Conflict and resolution: impact of new taxa on phylogenetic studies of neotropical cascudinos (Siluroidei: Loricariidae). In: *Phylogenetic Classification of Neotropical Fishes* (eds Malabarbra L.R., Reis R.E., Vari R.P., Lucena Z.M.S., Lucena C.A.S.) pp. 375 – 400. Edipucrs, Porto Alegre, Brazil.

Shafland, P. L. 1995. Introduction and establishment of a successful butterfly peacock fishery in southeast Florida canals. *American Fisheries Society Symposium*. 15: 443 – 451

Shafland, P. L., K. B. Gestring, and M. S. Stanford. 2008. Florida's Exotic Freshwater Fishes-2007. *Florida Scientist*. 71(3): 220 – 224.

Sioli, H. 1984. The Amazon and its main affluents: hydrography, morphology of the river courses, and river types. *The Amazon: Limnology and landscape ecology of a mighty tropical river and its basin*. (ed. by H. Sioli), pp. 127 – 165. W. Junk, Dordrect.

Sivasundar, A., E. Bermingham, and G. Orti. 2001. Population structure and biogeography of migratory freshwater fishes (Prochilodus: Characiformes) in major South American rivers. *Molecular Ecology*, 10: 407 – 417.

Tassinari, C. C. G. and M. J. B. Macambira. 1999. Chronological provinces of the Amazonian Craton. *Episodes*, 22(3): 174 – 182.

Wesselingh, F. P., J. Guerrero, M. Rasanen, L. Romero Pitmann and H. Vonhof. 2006. Landscape evolution and depositional processes in the Miocene Amazonian Pebas lake/wetland system: evidence from exploratory boreholes in northeastern Peru. *Scripta geologica*, 133: 323 – 361.

Willis S. C., M. S.Nunes, C. G.Montana, I. P.Farias, and N. R. Lovejoy 2007. Systematics, biogeography, and evolution of the Neotropical peacock basses. *Cichla* (Perciformes: Cichlidae). *Molecular Phylogenetics and Evolution*, 44: 291 – 307.

Willis, S. C., M. Nunes, C. G. Montana, I. P. Farias, G. Orti, and N. Lovejoy. 2010. The Casiquiare river acts as a corridor between the Amazonas and Orinoco river basins: biogeographic analysis of the genus Cichla. *Molecular Ecology*, 19: 1014 – 1030.

Winemiller K.O., D.C.Taphorn, and A. Barbarino-Duque. 1997. Ecology of Cichla (Cichlidae) in two blackwater rivers of Southern Venezuela. *Copeia*. 4, 690 – 696.

Winemiller, K. O. 2001. Ecology of peacock cichlids (Cichla spp.) in Venezuela. *Journal of Aquiculture and Aquatic Sciences*. Cichlid Research: State of the Art IX, 93–112.

Winemiller, K. O. and D. B. Jepsen. 1998. Effects of seasonality and fish movement on tropical river food webs. *Journal of Fish Biology*, 53: 267 – 296.

Winemiller, K. O., H. Lopez – Fernandez, D. C. Taphorn, L. G. Nico, and A. B. Duque. 2008. Fish assemblages of the Casiquiare River, a corridor and zoogeographical

filter for dispersal between the Orinoco and Amazon basins. *Journal of Biogeography*, 35: 1551 – 1563.

Zaret, T. M. 1980. Life history and growth relationships of *Cichla ocellaris*, a Predatory South American Cichlid. *Biotropica*, 12: 144 – 157.

Zaret, T. M. and R. T. Paine. 1973. Species introduction in a tropical lake. *Science*. 182 (4111): 449 – 455.

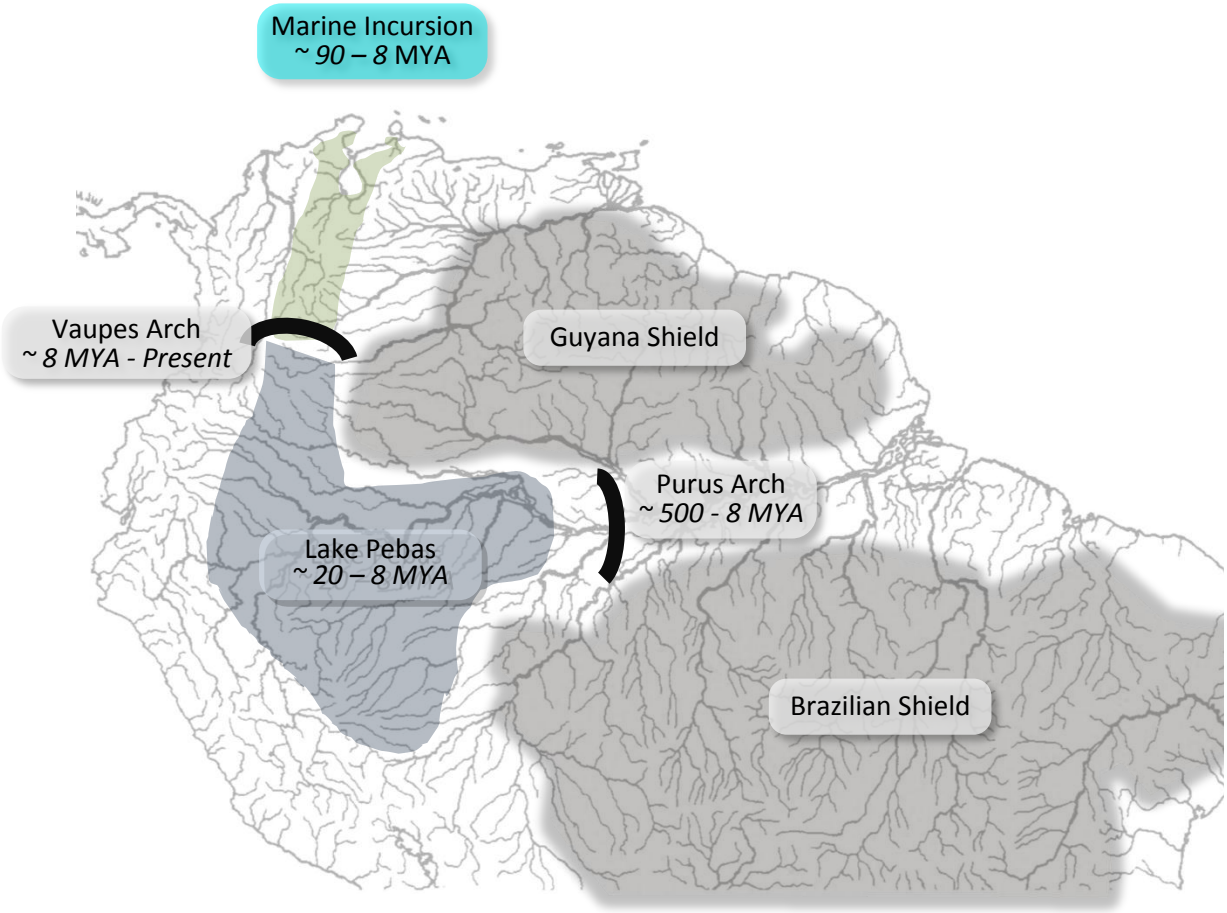

Figure 1.1 – Key geologic features influencing Neotropical landscape evolution.

Approximate time of feature existences are shown in italics.

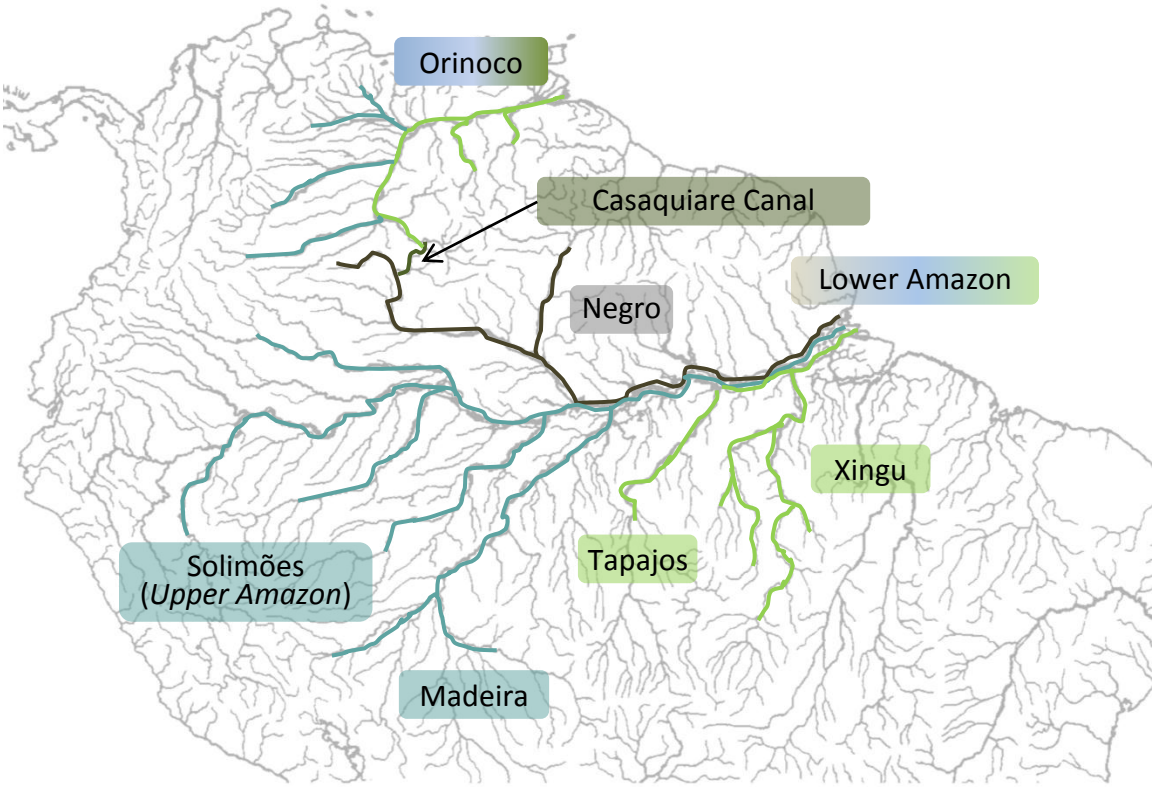

Figure 1.2 - Different water types of various Amazonian and Orinoco tributaries following examples provided by Meade (1994) and Mayorga and Aufdenkampe (2002). Blue: “white–water”; Green: “clear-water”; Black: “black-water”.

CHAPTER 2 – ISOLATION, CHARACTERIZATION, AND ANALYTICAL METHODS FOR MICROSATELLITE LOCI

2.1 – Microsatellite Isolation and Primer Design

Candidate microsatellite loci were isolated following a protocol by Tenzer et al. (1999) and subsequently modified by Farias et al. (2003). Genomic DNA from *C. orinocensis* and *C. temensis* were extracted and digested separately with the restriction enzyme *MboI* (Fermentas) for over 4 hours with a 500 – 1000 bp size range selected and gel extracted. Fragments from each species were then ligated to oligonucleotide adaptors and enriched via polymerase chain reaction (PCR) at 12 cycles over 32 PCR tubes to maintain sequence diversity. The amplified library for each species were then denatured and hybridized to biotin labeled CA_{10} probes, which were linked to DynaBeads (Invitrogen) that use a biotin – treptavidin bond. The hybridized microsatellite probes were then washed sequentially with more stringent saline sodium citrate (SSC) solutions. The enriched DNA was ligated and transformed into competent *Escherichia coli* cells using the CloneJet PCR cloning kit (Fermentas). The transformed cells were grown overnight on 1 X Luria-Beertani agar plates with 100 µg/mL of ampicillin (Sambrook and Russell 2001). White colonies were picked and regrown overnight in a 16 x 96-well culture plate. PCR amplifications were done using pJet primers (Fermentas) and colonies with PCR products between 500 – 1000 bp were ran on a gel and then sequenced.

Approximately 50% of the total sequenced library contained tandem repeating region (128/192 for *C. orinocensis* and 124/288 for *C. pinima*). The length of the microsatellite regions varied greatly across sequences (from 10 bp – 146 bp). Of the 252 candidate loci containing microsatellites, 50 were randomly selected from sequences with

flanking regions of sufficient size and sequence quality, which could be used to design primers.

Sequences containing tandem repeat regions were screened and primers were designed using the PRIMER 3 software (Rozen and Skaletsky 2000). For each microsatellite primer set annealing temperatures were determined by PCR conditions using a gradient thermal cycler (MJ Research PTC-200) with an annealing step of 48–66°C. The PCR had a total of 10 µl, which was made up of 2.0 µl of 1 mm dNTPS, 1.0 µl of 10x PCR buffer, 0.4 µl of 50 mm MgCL2, 0.4 ul of BSA (NE Bio), 0.4 µl 10 mm forward primer, 0.4 µl10 mm reverse primer, 0.1 µl (0.5 U) of Taq DNA polymerase (Invitorgen), 4.3 µl of sterile ddH2O, and 1.0 µl (50 – 100 ng) of DNA. PCR reaction steps were as followed: denaturation at 94^{o}C (30 s), annealing temperature range (30 s), 72^{o}C (40 s), followed by a final extension of 72^{o}C (2 min) over 30 cycles.

To test the ability for these microsatellite loci to amplify across the genus, total genomic DNA was extracted from muscle tissue using the DNAEasy extraction kit (Quiagen) for five species of *Cichla* (*C. monoculus*, *C. temensis*, *C.* orinocensis, *C. pinima*, and *C. intermedia*). The species selected represent diverse branches amongst the genus (Willis et al. 2010). From the candidate loci 50 microsatellite primers were screened for the ability to cross amplify across the genus using these five species. Of these primers tested, only two did not appear on an agarose gel when tested on along temperature gradients. Ultimately, the resulting 12 variable microsatellite loci primers successfully cross-amplified for all five species of *Cichla* tested.

Variations in microsatellite loci were quantified by eye on a gel across the five *Cichla* species. Those that did amplify were further characterized with 25 and 26

individuals from a single sampling locality for both *C. temensis* and *C. orinocensis*, respectively (Table 2.1). Variable loci were also screened using fluorescently labeled dUTPs (ChromaTide Rhodamine Green, Invitrogen) and genotyped using an ABI 310 with the size standard LIZ 500. Of the remaining 48 primers, only 14 primers were tested with fluorescently labeled primers (Applied Biosystems) for the initial population analysis for *C. temensis* and *C. orinocensis*. Of those remaining 14 primers, 12 amplified cleanly and were used for characterizing microsatellites in 25 *C. temensis* and 26 *C. orinocensis* individuals and the subsequent population level analysis.

Each locus was tested for Hardy-Weinberg equilibrium (HWE) using the HWE probability test, as well as linkage disequilibrium across 25 *C. temensis* and 26 *C. orinocensis* individuals, each from a single locality, using the program GENPOP (4.0) (Raymond et al. 1995) with significant p-values determined with the Bonferroni correction for multiple comparisons (Rice 1989) (Table 2.1). Of the 12 loci tested one (Ori1597-F12) showed significant departure from HWE in *C. orinocensis*, but not in *C. temensis*. It was found that Ori1597-G4 and Ori1597-B3 were in linkage disequilibrium for *C. orinocensis* ($P = 0.0017$). Excluding these loci did not significantly change any results of our further analysis. Therefore, all tests were done with these two loci to maximize statistical performance.

The apparent high rate of success of the described microsatellite isolation protocol indicates that this approach would be beneficial in characterizing loci for other studies for which microsatellites are desired. In fact, since the initial work on this genus the protocol has been used for several species with high success rates and the results have since been published (Fantin et al. 2007; Silva et al. 2008; Gravena et al. 2009; Santos et al. 2009).

There have been 10 microsatellite loci previously described for one other species, *C. piquiti* (Carvalho et al. 2009). The combination of these previously reported microsatellites in addition to the loci described in this study provides significantly high coverage across the entire genus (Willis et al. 2010). The addition of loci from this study provides 22 biparentally inherited nuclear loci. It allows us to perform more effective tests for biogeographic distributions using widely distributed species of *Cichla* as a focal taxonomic group. Microsatellite based analyses are quite powerful as they often provide inexpensive alternatives to sequence data for multiple independent unlinked nuclear loci (Beheregaray 2008).

2.2 – Population Studies

Phylogeographic and population genetic analyses aim to highlight processes responsible for contemporary population structure by using genetic variation. Although there are multiple types of molecular loci, non-coding mitochondrial or nuclear genes are often used to depict population structure as gene genealogies not directed by selective pressures are reflective of historic patterns of gene flow across their geographic distribution. Although these types of analyses are limited to the evolutionary history of the specific loci (Avise 2000), they are commonly used to infer processes that have influenced populations over time. One of the most commonly used loci are simple sequence repeats, or microsatellites, due to their rapid rates of mutation and relatively low cost for genotyping. Although multiple analytical methods exist, variation amongst microsatellite loci are used to measure F-statistics, specifically F_{ST} and R_{ST}. Both F_{ST} and R_{ST} analyses quantify genetic variation within and between populations and are often

used to detect levels of gene flow across populations. F_{ST} values are calculated to analyze the sub-population structure based on the variance in alleles (Wright 1951). Alternatively, R_{ST} accommodates for differences in allele sizes, specifically for microsatellites following a stepwise mutation model (discussed in the next section) as F_{ST} analyses may exhibit biased results as they exhibit differing mutation models (Slatkin 1995).

2.3 – Detecting Population Decline and Expansion Using Microsatellites

When the effective population size remains constant over time a population exhibits what is known as mutation-drift equilibrium (MDE). This occurs when the rate of genetic drift and mutation for specific loci are proportionately equivalent to one another. When the MDE is disrupted species may be experiencing (or recently experienced) a decline in population size or a rapid population expansion, which is often difficult to interpret when influenced by complex demographic histories.

Reduced genetic variation following a bottleneck event can result in the occurrence of inbreeding, fixation of deleterious alleles, and a reduction in the capacity of a population or species to adapt (Nei et al. 1975; Luikart and Cornuet 1998). Because of these reasons it is important to identify populations that have gone through (or have recently gone through) a rapid reduction in population size. The MDE disrupted as a result of a bottleneck event can be detected as the diversity of alleles are reduced faster than the heterozygosity, resulting in observed heterozygosity excess (Maruyama and Fuerst 1985) It is important to note, that in order for the a bottleneck to be detected the population must have a decreased over several generations in order for genetic drift to occur. There are also instances in which a bottleneck was known to occur with no MDE

disruption due to a rapid recovery or rapid population expansion following the bottleneck event (Rooney et al. 1999; Whitehouse and Harley 2001; Waldick et al. 2002). Additionally, the heterozygosity excess (or deficiency) will often become diluted or undetectable (Luikart et al. 1998) after many generations (25 or more) (Cornuet and Luikart 1996).

A rapid population expansion will also disrupt the MDE and result in a different type of genetic signal as a heterozygosity deficiency. Tests using sequence data to detect the occurrence of population expansion are often times ambiguous and unclear with regards to identifying population expansion events (Reich et al. 1999) and are speculative without a statistical simulation approach (Piry et al. 1999) or cladistic analysis of geographical distributions among unique haplotypes (Templeton et al 1995). By using multiple unlinked neutral loci from a nuclear genome, the possibility of selection affecting the genetic signal of population expansion is reduced and the genetic signal is stronger with multiple independent loci. Unfortunately, when a population expansion follows a population size reduction the heterozygosity deficiency may be even more significant and difficult to discern between a population bottleneck and population expansion (Cornuet and Luikart 1996).

The occurrence of population size reduction or a rapid population expansion can be inferred from allele frequency data in the form of different mutation models implemented in various statistical tests (Cornuet and Luikart 1996). The infinite allele model (IAM) (Kimura and Crow 1964; Nei et al 1976) was proposed as an explanation of mutation rates at nucleotide sequences in a population. This is the original mutation model for testing MDE and was adapted as detecting bottlenecks (Maruyama and Fuerst

1985). Unlike the IAM, where the allelic DNA segments were dependent on the rate of nucleotide substitutions, microsatellites mutate in a stepwise manner. The variation that appears in microsatellites results from the polymerase slippage when moving along the repeat region. If this slippage occurs, this event may increase or decrease the microsatellite repeat number in the next generation in a predictable fashion, otherwise known as the stepwise mutation model (SMM) (Valdes et al. 1993). Although the SMM theoretically represents microsatellite evolution, the polymerase does not always slip in a stepwise manner. On rare occasions the polymerase may transcend the stepwise restriction and move beyond a single repeat which violates the SMM. An alternative approach combines the IAM and SMM into what is commonly known as a two-phase model (TPM) (Di Rienzo et al. 1994). These models are often used in combination when inferring historical demography as they relate to distinct changes within a population and reasons behind MDE disruption.

The additional microsatellite loci described here allow for further investigation into the demographic histories of various *Cichla* species. The multiple independent nuclear loci provide a more robust analysis in comparison to a mitochondrial DNA (mtDNA) sequence. These biparentally inherited markers are able to determine the level of gene flow and when combined with mtDNA tests for genetic signal influenced by of their life history characteristics, such as exclusive male or female migration. Additionally, the rapid mutation rates of these markers also allow for the detection of various reasons behind MDE disruption in order to understand the demographic history of a species or populations. Microsatellite loci for Neotropical species would provide insight into geological and ecological factors contributing to the observed diversity.

References

Avise, J.C. 2000. Phylogeography; the history and formation of species. Harvard University Press. Cambridge, Mass.

Carvalho, D. C. D. A. A. Oliveira, I. Sampaio, L. B. Beheregaray. 2009. Microsatellite markers for the Amazon peacock bass (Cichla piquiti). *Molecular Ecology Resources*, 9: 239 – 241.

Cornuet J.M. and G. Luikart. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144:2001-2014.

Di Rienzo, A., A. C. Peterson, J. C. Garza, A. M. Valdes, M. Slatkin. 1994. Mutational processes of simple sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences USA*, 91: 3166 – 3170.

Fantin, C. C. F. Carvalho, T. Hrbek, J. W. Sites, L. A. S. Monjelo, S. Astolfi-Filho, I. P. Farias. 2007. Microsatellite DNA markers for Podcnemis unifilis, the endangered yellow-spotted Amazon River turtle. *Molecular Ecology Notes*, 7: 1235 – 1238.

Farias IP, T. Hrbek, H. Brinkmann, I. Sampaio, A. Meyer. 2003. Characterization and isolation of DNA microsatellite primers for *Arapaima gigas*, an economically important but severely over-exploited fish species of the Amazon basin. *Molecular Ecology Notes*, **3**, 128–130.

Gravena, W., T. Hrbek, V. M. S. Da Silva, S. Astolfi-Filho, I. P. Farias. 2009. Microsatellite loci for population and parentage analysis in the Amazon River dolphin (Inia geoffrensis de Blainville, 1817). *Molecular Ecology Resources*, 9(2): 600 – 603.

Kimura, M. and J. F. Crow. 1964. The number of alleles that can be maintained in a finite population. *Genetics*, 49: 725 – 738.

Luikart, G., F. W. Allendorf, J.-M. Cornuet, and W. B. Sherwin. 1998. Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity* 89(3): 238 – 247.

Lundberg, J. G., L. G. Marshall, J. Guerrero, B. Horton, M.C.S.L. Malabarba, F. Wesselingh. 1998. The stage for Neotropical fish diversification: a history of tropical South American rivers. In: *Phylogeny and Classification of Neotropical Fishes* (eds L.R. Malabarba, R. E. Reis, R. P. Vari, Z. M. Lucena, C.A.S. Lucena CAS), pp. 13–48. Edipucrs, Porto Alegre.

Maruyama T. and P. Fuerst. 1985. Population bottlenecks and nonequilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. *Genetics*, 111: 675 – 689.

Nei, M. R., Chakraborty, and P. A. Fuerst. 1976. Infinate allele model with varying mutation rate. *PNAS*, 11: 4164 – 4168.

Nei, M. R., T. Maruyama, and R. Chakraborty. 1975. The bottleneck effect and genetic variability in populations. *Evolution*, 29(1):1-10.

Piry, S., G. Luikart, and J. M. Cornuet. 1999. BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, 90(4): 502 – 503.

Raymond M. & Rousset F, 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86:248-249

Reich, D. E., M. W. Feldman, D. B. Goldstein. 1999. Statistical properties of two tests that use multilocus data sets to detect population expansions. *Molecular Biology and Evolution*. 16(4): 453 – 466.

Rice, W. R. 1989. Analyzing tables of statistical tests. *Evolution* 43(1) : 223 – 225.

Rooney, A. P., R. L. Honeycutt, S. K. Davis, J. N. Derr. 1999. Evaluating a putative bottleneck in population of bowhead whales from patterns of microsatellite diversity and genetic disequilibria. *Journal of Molecular Evolution*,49: 682 – 690.

Rozen, S. and H. J. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365-386.

Sambrook J, D. Russell. 2001. *Molecular Cloning: A Laboratory Manual*, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Santos, M. D. C. F., T. Hrbek, I. P. Farias. 2009. Microsatellite markers for the tambaqui (Colossoma macropomum, Serrasalmidae, Characiformes), an economically important keystone species of the Amazon River floodplain. *Molecular Ecology Resources*, 9(3): 847 – 876.

Silva, T. J., T. Hrbek, and I. P. Farias. 2008. Microsatellite markers for the silver arowana (*Osteoglossum bicirrhosum*, Osteoglossidae, Osteoglossiformes). *Molecular Ecology Resources*, 9(3):1019 – 1022.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, 139: 457 – 462.

Templeton, A. R, E. Routman, and C. A. Phillips. 1995. Separating population structure from population history: a cladistics analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics*, 140: 767 – 782.

Tenzer I, S. degli Ivanissevich, M. Morgante, C. Gessler. 1999. Identification of microsatellite markers and their application to population genetics of *Venturia inaequalis*. *Phytopathology*, 89, 748–753.

Valdes, A. M., M. Slatkin, and N. B. Freimer. 1993. Allele frequencies at microsatellite loci: The stepwise mutation model revisited. *Genetics*, 133: 737 – 749.

Waldick, R. C., S. Kraus, M. Brown, B. N. White. 2002. Evaluating the effects of historic bottleneck events: an assessment of microsatellite variability in the endangered, North Atlantic right whale. *Molecular Ecology*, 11(11): 2241 – 2249.

Whitehouse, A. M. and E. H. Harley. 2001. Post-bottleneck genetic diversity of elephant populations in South Africa, revealed using microsatellite analysis. *Molecular Ecology*. 10(9): 2139 – 2149.

Willis, S. C., M. Nunes, C. G. Montana, I. P. Farias, G. Orti, and N. Lovejoy. 2010. The Casiquiare river acts as a corridor between the Amazonas and Orinoco river basins: biogeographic analysis of the genus Cichla. *Molecular Ecology*, 19: 1014 – 1030.

Wright, S. 1951. The genetical structure of populations. *Annals of Eugenics*, 15: 323 – 354.

Table 2.1. Primer sequence, repeat motif and allelic diversity of 12 microsatellite markers in the genus *Cichla*, each screened in 25 and 26 individuals from a single sampling locality for the species *C. temensis* and *C. orinocensis*, respectively.

Locus	Primers (5' – 3')	Repeat motif	*C. tem*/*C. ori* Size (bp)	T_a (°C)	*C. tem*/*C. ori* k	*C. tem* H_o†/H_E	*C. ori* H_o†/H_E
Ori1597-A6	F: ACGTCCACCCTGGCTTT R: TGATTTCAAGGCCAGATTCA	$(TC)_{18}$	267 / 271 – 303	56	1/7	---/---	0.808/0.756
Ori1597-B3	F: CCCTTCTCTGCGCCAAT R: ACCCGTGGAGGGTTCC	$(GA)_{19}$	213 – 223 / 193 – 203	58	3/3	0.500/0.471	0.280‡/0.365
Ori1597-D12	F: GCTTGAGGGGCTCACTTTC R: GGGGGAATCCTTTTTCTGG	$(TC)_{19}$	154 – 162 / 154 – 168	56	2/4	0.577/0.473	0.680/0.594
Pin1869-E3	F: CCCAATCTTCCCGCTGT R: TGGGAAGAAAGAAACACCTCA	$(GT)_{27}$	276 / 286 – 312	59	1/9	---/---	0.800/0.733
Int1022	F: CTTCCTTTGGTGCAGCTTG R: GCGTGGGCGCAAGATA	$(TG)_{34}$	149 – 169 / 125 – 179	54	3/11	0.077/0.076	0.731/0.876
Pin1869-D2	F: TGGGTGCATAAGGGCAGT R: GGCAGTGGGTGGTGAGA	$(AG)_{24}$	291 – 309 / 287 – 319	59	6/12	0.885/0.802	0.769/0.769
Ori1597-G4	F: GGTGTCTGGGAAGCTGGA R: TCCGCCAGGACTGAGG	$(TC)_{26}$	310 / 290 – 296	62	1/4	---/---	0.480‡/0.482

Table 2.1. *continued.*

Locus	Primers (5' – 3')	Repeat motif	*C. tem*/*C. ori* Size (bp)	T_a (°C)	*C. tem*/*C. ori* k	*C. tem* H_o†/H_E	*C. ori* H_o†/H_E
Ori1597-F12	F: CCATCCCCATGGCAAC R: CCTGCTTCAAGCGCAGA	$(GT)_{21}$	260 – 294 / 234 – 292	58	10/20	0.885/0.827	0.850*/0.937
Ori1597-B6.2	F: CATGCCCAAATGCATCC R: CCATGGATTCAAGGGCTTT	$(TC)_{19}(CA)_{14}(CA)_7$	290 – 292 / 266 – 272	58	2/4	0.038/0.038	0.520/0.502
SM2	F: CAGGCCCAAACATGAAAATG R: TTTGTGGAAGAACAGGGTTTG	$(AG)_{20}$	248 – 254 / 231 – 261	60	2/9	0.346/0.340	0.800/0.825
Pin1869-C11	F: ACAAGGGAGCTGAAGGCATA R: GAAACACGATACACCCATTTGTT	$(GA)_{18}$	229 – 237 / 207 – 239	60	5/8	0.461/0.495	0.920/0.816
Pin1869-C1	F: GATGGTGCCAAGGCAAA R: TGACCCAAGGCCACAAA	$(AG)_{19}$	227 – 229 / 221 – 225	58	2/3	0.038/0.038	0.400/0.375

C. tem, C. temensis; C. ori, C. orinocensis; T_a, annealing temperature; *k* number of alleles; H_o, observed heterozygosity; H_E, expected heterozygosity. † Hardy-Weinberg equilibrium (HWE) for each locus was checked using HWE probability test, GENPOP (4.0) (Raymond et al. 1995) *P < 0.05. ‡ Linkage disequilibrium following Bonferroni test P < 0.016 (Rice 1989).

CHAPTER 3 - DEMOGRAPHIC HISTORY AND POPULATION STRUCTURE OF THE AMAZONIAN PEACOCK BASS *CICHLA TEMENSIS*

3.1 – Introduction

Spatial and temporal genetic variation found across widely distributed populations often correspond with major event in landscape evolution and environmental heterogeneity (Riddle et al. 2008). Analyses that use combined molecular markers exhibiting varying rates of molecular evolution are necessary in order to understand the complex processes that have guided landscape evolution in the Neotropics (Beheregaray 2008). One key hydrogeologic feature of the Neotropics, the Casiquiare River, is a unique connection between two drastically different ecosystems (Figure 1.2). The Casiquiare River (often referred to as the Casiquiare Canal) is a 300+ kilometer bifurcation of the Upper Orinoco main channel that ultimately empties into the Rio Negro, thus connecting the Amazon and Orinoco River basins. This connection appeared sometime after the Orinoco and Amazon established their current west to east drainage (Lundberg et al 1998). The Casiquiare River is truly unique in that it connects drastically different water types, the "black-water" Rio Negro and the "clear-water" Rio Orinoco. These rivers differ greatly in pH, conductivity, dissolved organic material and sediment loads (Winemiller et al. 2008). This connection not only acts as a corridor for many species (Willis et al. 2010) but it may also provide a barrier to their distribution (Winemiller et al. 2008).

Historical geologic events of the Neotropics affected not only the Casiquiare, but other regions of headwaters surrounding these basins. Toward the end of the Miocene (8

MY ago) the coastal drainage of the Orinoco shifted eastward towards its current drainage and the Amazon established its current west to east drainage towards the Atlantic (Lundberg et al 1998). The changes were eventually followed by extended periods of glacial retreat resulting in large narrow freshwater lakes formed in the lower Amazon, collectively extended up to the Rio Negro and possibly some of the upper tributaries as well (Irion et al. 1995; Irion et al. 2009). Collectively, these geologic events changed the Amazon and Orinoco River basins, potentially affecting the population distribution of various species currently found amongst these waterways.

Cichla temensis is the largest peacock bass (>500 mm) and often the target of many sport fisherman. Their current distribution spans the Rio Negro and Rio Orinoco with apparent continued gene flow between the different basins via the Casiquiare River (Willis et al. 2010) and share overlapping distributions with two other species, *C. intermedia* and *C. orinocensis* (Kullander and Ferreira 2006). Although their population structure is likely influenced by the formation of the Casiquare River (Winemiller et al. 2008; Willis et al. 2010) it is possible that their current distribution is reflective of major geologic events that have undergone less scrutiny (i.e., formation of the large lakes along the main channel of the Amazon). Although their distribution expands into the white and clear water tributaries of the upper Orinoco (Winemiller et al. 2008), they were previously characterized as being exclusively black-water (Kullander and Ferreira 2006).

To address impacts geologic events or ecological heterogeneity has on the population distribution of *C. temensis* I have used mtDNA and microsatellite loci in my analyses. These analyses allow me to determine if population structure is reflective of key geologic events (the formation of the Casiquare River), ecological heterogeneity (white

waters of the Orinoco and black waters of the Rio Negro), or a combination of thereof. Finally, I test for deviations from MDE across different population divisions and whether significant heterozygosity excess or deficiency reinforces hypotheses regarding their historic distribution.

3.2 – Materials and Methods

3.2.1 – Sampling Strategy

Individuals of *C. temensis* were collected from 20 localities throughout the Amazon and Orinoco river basins. They provide sufficient coverage across their natural distribution (Figure 3.1). The sampling strategy for individuals used in this analysis is described by Willis et al. (2007). For each fish collected a voucher photograph was taken and tissue samples were preserved in SMSO-EDTA buffer. Although I was unable to collect the specimens myself, the same individuals from Willis et al. (2007) and Willis et al. (2010) were incorporated into this analysis with additional specimens collected by collaborators doing field collections along these rivers. The specimens used in this study are a direct result from the strong collaborative collection effort established by many authors cited throughout this thesis, such as Renno et al. (2006), Willis et al. (2007) and Willis et al. (2010).

3.2.2 – mtDNA Sequencing

Sequences for mtDNA control regions from 49 new samples were generated using previously described primers and procedures (Willis et al. 2007; Willis et al. 2010). The newly acquired sequences were analyzed in combination with sequences from previous

studies (DQ841909 - DQ841929, Willis et al. 2007; GU295739 and GU295740, Willis et al. 2010). A total of 159 sequences spanning approximately 489 bp of the control region (CR) for *C. temensis* were aligned using the program MAFFT version 6 with G-INS-I alignment strategy and default parameters (Katoh and Toh 2008) resulting in 30 variable sites across their entire distribution. Although more than 159 specimens were collected, variation between mtDNA sequences amongst the different localities was relatively low. Thus a maximum of 10 per sampling locality were used in this analysis.

3.2.3 – Microsatellite Genotyping

In the population level analysis for *C. temensis*, the 12 microsatellite loci were amplified for 173 individuals. The conditions for PCR were the same as described in Chapter 2, with the only difference being that PCRs were done using MJ Tetrads and fragments were analyzed on an ABI 3730. The microsatellite size fragments were analyzed on the program GeneMapper (Applied Biosystems) using default binning parameters for each microsatellite loci across the entire population.

3.2.4 – mtDNA Analysis

The aligned mtDNA control region (CR) sequences were used to construct a median joining haplotype network (Bandelt et al. 1999) using the program NETWORK (flux-engineering.com) with each unique CR sequence representing a unique haplotype. The rdf-format file, which is required for NETWORK, was created using the program DnaSP (Librado and Rozas 2009). The haplotype network was constructed using all

default parameters and was arranged to represent *C. temensis* sampling localities associated with the different haplotype groups (Figure 3.1).

3.2.5 – Microsatellite Analyses

The genetic structure of the population was measured as a pairwise analysis of variance using sampling localities as subpopulations for *C. temensis*. To analyze the sub-population structure and to check for variance in allele size, F_{ST} (Wright 1951) values were calculated. R_{ST} accommodates for differences in allele size (Slatkin 1995), and it should be less biased for demographic patterns amongst microsatellites because it is based on the SMM (Di Rienzo et al. 1994). All of the F_{ST} and R_{ST} values were calculated using Arlequin 3.1 (Excoffier et al. 2005). To determine whether the population exhibited characteristics of isolation by distance, geographical distances were obtained as a guided pathway along waterways using Google Earth (http://earth.google.com/). F_{ST} and R_{ST} were plotted against geographical distances between sampling localities in order to determine if there was a positive correlation. The statistical significance of the correlation was tested by Mantel test (Mantel 1967) for 1000 permutations using Arlequin 3.1 (Excoffier et al. 2005).

The model based clustering method STRUCTURE (Pritchard et al. 2000) was used to estimate the number of genetically distinct populations (*k*). The first 50,000 generations of the Markov Chain Monte Carlo (MCMC) were discarded as burn-in and the remaining 50,000 were used to estimate the proportion of membership. The true number of populations (*k*) is often identified by selecting the highest likelihood value based on the average of multiple *k* values over several runs (Pritchard et al. 2000). However,

sometimes the likelihood of the highest *k* value does not deviate from other likelihood values significantly and the correct *k* may be difficult to determine without any *apriori* knowledge regarding population structure, which was the case for this study. Therefore the optimal *k* was determined by following a protocol described by Evanno et al. (2005), which is based on an *ad hoc* measure Δk, determined by the second order rate of change between log probability for various *k* values over several runs, $\ln[\Pr(X|k)]$. Based on the Bayesian estimations of conformity every individual from this analysis was assigned to a designated number of populations across several possible *k* values (1-11 and 24-25). The resulting likelihood values were recorded and put into the following equation, ($\Delta k = m(|L''k|)/s[L(k)]$), where $m(|L''k|)$ is the mean of the variance of *k* across runs and $L''k$ is calculated as $\Pr(X|k+1) - 2\Pr(X|k) + \Pr(X|k-1)$, and $s[L(k)]$ is the standard deviation of the likelihood value for the focal *k*.

To test for potential deviations from MDE in *C. temensis* we used the program BOTTLENECK 1.2.02 (Piry et al. 1999). It implements two tests to determine whether the population had experienced a bottleneck event or a population expansion. The program tests for changes in population that reflect MDE disruptions based on the number of unique alleles. It simulates a population that would have the same number of alleles following the sign test (Cornuet and Luikart 1996) and a Wilcoxon signed-rank test (Luikart et al. 1998). The simulation results are then compared to the observed heterozygosity in the population and significant heterozygote deficiencies or excesses occurring within the population are identified based on a p-value below 0.05. The program implements these statistical analyses following the IAM (Maruyama and Fuerst 1985), as well as the SMM (Cornuet and Luikart 1996) and the TPM (Di Rienzo et al.

1994). Additionally, the program monitors the allele frequency distribution, which can distinguish between many bottlenecked and stable populations. The program does this by determining if the distribution of allele frequencies observed graphically is L-shaped or a deviation from L-shaped, a "mode-shift" distortion (Luikart et al. 1997). Individuals from the population were divided according to both microsatellite STRUCTURE results, the mtDNA haplotype network, and by sampling locality. The program tests for a significant heterozygosity excess and deficiency comparing the observed allele frequency and the population simulations across the described tests for different mutation models.

3.3 – Results

3.3.1 – mtDNA analysis

Every locality was represented with the mtDNA analysis. The median-joining haplotype network resulted in the 159 sequences being placed in 21 haplotypes. The haplotypes were resolved into five groups by a single mutation and separated from other haplotype groupings by three or more mutations (Figure 3.1). For the most part the resolved groups are separated by geographically adjacent sampling localities. When a single locality had individuals found in two distinctly different haplotype groups, those individuals in the haplotype network and locality in the map was highlighted with an alternative color. This was seen with individuals from Casiquiare (groups 1, 4, and 5), Caura (1 and 2), and Barcelos (4 and 3). Individuals from the Casaquiare were the only locality found amongst more than two haplotype groupings. Although the haplotype groups seemed to have little overlapping between groups, their arrangement across the Rio Orinoco and Rio Negro did not follow their current distribution patterns. For

example, some of the individuals from Caura were grouped more closely with individuals from Ventuari and Atabapo than with individuals from adjacent localities within the Rio Negro, placing these individuals furthest away. Similar non-conforming patterns were found between localities along Rio Negro, such as specimens from the Igapo Acu being found in the same haplotype grouping as specimens from the Uaupes.

3.3.2 – Microsatellite analyses

Every locality was represented with the microsatellite data. Individuals that had more than four failed PCRs out of 12 were removed from the analysis, leaving a total of 174 genotypes for the *C. temensis* population. Overall, the 20 sampling localities had a moderately high level of genetic differentiation (F_{ST} = 0.238; R_{ST} = 0.228), showing strong population structure by sampling locality (Table 2.2). The most geographically distant localities were Caura and Urubu, which are 3625 km away from each other. They exhibited high genetic difference (F_{ST} = 0.378 and R_{ST} = 0.421), although not the highest. The sampling locality with the highest average F_{ST} values was Pirara (F_{ST} = 0.355; R_{ST} = 0.397). The highest F_{ST} and R_{ST} between sampling sites were between Pirara and Parguaza (2,130 km) for F_{ST} (0.507) and Pirara and Caura (2,590 km) for R_{ST} (0.699). Only one specimen was collected from Cunavichito, which resulted in almost all of the F_{ST} and R_{ST} values from that locality not to be significant. The population structure of *C. temensis* was also analyzed to determine if there is a correlation among allele frequencies and geographic distance (Figure 3.2). There was a strong positive correlation of isolation by distance for both F_{ST} (R = 0.6340) and R_{ST} (R = 0.5888). Both correlations are significant ($P < 0.001$ for F_{ST} and R_{ST}) according to the Mantel test (Mantel 1967).

Likelihood values were recorded for values of *k* ranging from 1 – 11 and 24 – 26 (Figure 3.3) as an output from the program STRUCTURE (Pritchard et al. 2000). These values were used to calculate the Δ*k* statistic (Δ *k* = m(|L"K|)/s[L(K)]) (Figure 3.3) as proposed by Evanno et al. (2005). While the highest average likelihood values were at *k* = 6, Δ*k* statistic (Δ*k* = 249) recognized only a single division (*k* = 2) corresponding to the Rio Orinoco and Rio Negro populations with the specimens from the Casiquiare being placed into both. The STRUCTURE output for *k* = 2 exhibited definite separation between the Orinoco and Rio Negro sampling localities with the Casiquiare samples in between. Additionally, the STRUCTURE output for *k* = 6 repeated this pattern. However, there was significant variation throughout localities within the Rio Negro and apparently very little variation between localities in the Orinoco River. For both *k* =2 and *k* = 6 the STRUCTURE results did not coincide with patterns of genetic similarity across their geographic distribution when compared to the haplotype network.

There were no significant results that would indicate excess of heterozygosity in any of the population divisions. Additionally, all of the allele frequencies were L-shaped, as frequency of intermediate alleles did not exceed low and high allele frequency classes (Figure 3.4), which is an indication of long term stable population size (Luikart et al. 1997). There was, however, a significant heterozygosity deficiency (Table 3.1) for various population divisions. The IAM was not significant for either the Sign test or the Wilcoxon two-tailed test, although, for both the SMM and varying parameters of the TPM there was a significant heterozygosity deficiency. The only population divisions that did not show significant heterozygosity deficiency were the Casiquiare (when dividing the population according to the microsatellite analysis) and when haplotype

groups 1 and 2 combined (from the haplotype network analysis) for the TPM under the sign test (TABLE 2.2). When testing sampling localities only Capanaparo, Cinaruco, Tapera, Igapo Acu, Tapera, and Urubu were determined as having gone through experienced a population expansion for both tests (data not shown).

3.4 – Discussion

Inferences made from both maternally and biparentally inherited markers with differing rates of mutation allow one to apply comparative data to diverse analytical methods. We used this approach to investigate the population structure for *Cichla temensis* to test whether the formation of the Casiquiare River and ecological heterogeneity may be factors influencing gene flow across their distribution.

3.4.1- Population Structure and the Casiquiare River

One major goal of phylogeographic analyses is to determine what the genetic signal is in a population and how this reflects historical events influencing the population's distribution. Events that have influenced an extant species distribution will often be reflective in population genetic analyses using multiple unlinked neutral molecular markers that exhibit varying rates of molecular evolution. The mtDNA haplotype network alone seems to exhibit inconclusive results and do not allow one to accept or reject the potential influences the Casiquare had on their population distribution. Individuals collected from the Caura, Atabapo, Ventuari localities and some Casiquiare are grouped more closely to individuals that were collected from throughout the Amazon (3000+ km) than nearby collection sites in the Orinoco River (300 – 500

km). Extensive work has been done with how the Amazon and Orinoco Rivers have evolved; however, to my knowledge no research thus far has indicated a historical alternative connection or corridor in this area between the two water basins. The previously mentioned periods of glacial retreat may have established connections between ancestral populations (Irion et al. 1995; Irion et al. 2009), however, this is very unlikely as Pirara, Parguaza and Caura exhibit low levels of gene flow, with the highest F_{ST} (PI and PZ: 0.531) and R_{ST} (PI and CA: 0.699). With only a single mtDNA sequence we cannot adequately differentiate between the haplotype network illustrating historical population distributions or simply a gene genealogy that does not coincide with their current distribution, due to migration or high rates of gene flow for maternally inherited DNA.

It is only when we are able to compliment this data with the rapidly evolving microsatellite loci does the demographic history of *C. temensis* population become apparent. The microsatellite analysis highlights that there is strong genetic signal for reduced gene flow between the basins when the population is separated into 2 groups, as well as the analysis separating the population into 6 groups. This was observed with the strong delineation between basins when the sampling localities for the Orinoco and the Rio Negro (Figure 3.4 and 3.5). Additionally, the average F_{ST} and R_{ST} values for the Orinoco (F_{ST} = 0.206; R_{ST} = 0.212) and Rio Negro (F_{ST} = 0.135; R_{ST} = 0.116) were much lower when measured across sampling localities within each basin when compared to average F_{ST} (0.324) and R_{ST} (0.336) of localities between the basins.

When the population is divided into 2 and 6 groups there is an apparent lack of mixing between the Orinoco and the Rio Negro, with individuals found within the

Casiquiare River exhibiting high levels of gene flow between the two rivers. Contrary to these results, when the population is separated into 6 groups there is an apparent lack of genetic variation in the Orinoco River in contrast to the geographically similar distributions in the Rio Negro that exhibit some population structuring and proportionately higher levels of genetic diversity. Although individuals from the Rio Negro exhibit some differentiation throughout their distribution (Figure 3.5) according to the F_{ST} and R_{ST} values they have less genetic diversity (F_{ST} = 0.135; R_{ST} = 0.116) when compared to the Orinoco (F_{ST} = 0.206; R_{ST} = 0.212).

This overall signature of separation between the Orinoco and Rio Negro is most likely due to an established population in the Rio Negro going through a colonization event after the Casiquiare River was established. If this colonization were to occur, individuals that had populated to the Orinoco basin could have rapidly lost diversity from the Rio Negro due to genetic drift. Also through genetic drift more common alleles could have been lost, resulting in the higher average F_{ST} and R_{ST} values for individuals found in the Orinoco. Additionally, as the mutation rate of mitochondrial DNA evolves much slower the unresolved pattern found in the haplotype network is most likely a gene genealogy reflective of their migration and colonization events. Investigations for species with overlapping distributions should be done to determine whether similar population structure is replicated in other population distributions.

3.4.2- Population Expansion

Overall the statistical tests for disruptions of MDE show significant p-values for population expansion across almost all population divisions. When the population is

divided based on the haplotype network there is significant heterozygote deficiency in haplotype groups 3, 4 and 5 as well as in the Orinoco and Negro samples when the population is divided according to the microsatellite analysis. The signal contradicts potential population expansion mentioned earlier, as there were signals for population expansion in both the Orinoco and Rio Negro (Table 3.2). The only group of haplotypes that did not show significant heterozygote deficiency were combined haplotype groups 1 and 2 for the TPM under the sign test and the Casiquiare River samples when looking at the results from the microsatellite analysis (Table 3.2). Although gene flow (most likely to occur in the Casiquiare) or small sample size (combined haplotype groups 1 and 2) may influence these analyses (Piry et al.1999), the ability to appropriately test for variation in MDE following population expansion heavily relies on correct assumptions for mutation model and population structure. The population divisions tested in for deviations from MDE were inferred from mtDNA and microsatellite analyses, as well as sampling locality. Although these are our best indicators of population division, the defined subpopulations may be incorrect and therefore exhibit a genetic signal for population expansion. When the population was divided based on sampling localities (data not shown) only some of the localities had significant p-values indicating areas going through a population expansion. These apparent signals of population expansion were found in both rivers, with the majority of the sampling localities found along the Rio Negro. These results did not correlate with geologic or environmental patterns that might indicate a population expansion. It is important to interpret the results with great scrutiny as significant p-values may simply be an artifact an incorrectly divided population. Although a portion of *C. temensis'* population exhibited some signal of

population expansion these results are unable to determine whether this may be linked directly to the described population structure.

3.4.3- Black and White Water

The results from these analyses thus far support that the Casiquiare River has influenced their current geographic distribution; however, we have not addressed to what extent the ecological gradient across these rivers might be influencing ongoing gene flow in this population. With populations established in both the Amazon and Orinoco River basins there is a dramatic difference between the contrasting water types of these rivers. There seems to be no correlation between the haplotype network and water type, which is most likely due to the slower rate of evolution in comparison to the microsatellite loci, however, the microsatellite analyses suggest there is restricted gene flow between the river types. A strong separation between the Orinoco and the Amazon River, with the Casiquiare River split between both groups for $k = 2$ (Figure 3.4). Additionally, when the population is divided according to the highest average likelihood value ($k = 6$) there is a strong split between the two different river types. If the geologic events were to play a key role in their population expansion, but the ecological gradient did not act as a biological filter we would most likely see continued gene flow between these rivers showing a gradient in the average likelihood values resembling isolation by distance across these rivers, however, there is a strong separation between these two supporting that the ecological variation between these watersheds. This further supports the hypothesis that the Casiquiare River acts as a biological filter between the two river types (Winemiller et al. 2008).

3.5. Conclusion

These results indicate that the current population structure of *C. temensis* has been influenced by the formation of the Casiquiare River and the contrasting water types found between the two rivers. These analyses indicate that the formation of the Casiquiare River was likely followed by a population expansion of *C. temensis* into the Orinoco River. Additionally, these results support that the contrasting water types exhibit low levels of gene flow between them. The differing water types may exhibit alternative selective pressures on species whose distribution spans these contrasting environments and most likely play a role influencing rates of speciation throughout the Neotropics. Similar analyses with populations that share overlapping distributions for *C. temensis* (i.e., *C. intermedia* and *C. orinocensis*) would allow for further testing of the influence the formation of the Casiquiare River and the contrasting water types have on population distributions.

References

Bandelt H-J, P. Forster, and A. Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16:37-48

Beheregaray, L. B. 2008. Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology*, 17, 3754 – 3774.

Cornuet J.M. and G. Luikart. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144:2001-2014.

Di Rienzo, A., A. C. Peterson, J. C. Garza, A. M. Valdes, M. Slatkin. 1994. Mutational processes of simple sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences USA*, 91: 3166 – 3170.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14:2611–2620.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1: 47 – 50.

Irion, G., J. Muller, J. Nunes de Mello, W. J. Junk. 1995. Quaternary geology of the Amazonian lowland. *Geo-Marine Letters*, 15 (3-4): 172 – 178.

Irion, G., J. Muller, J. O. Morais, G. Keim, J. Nunes de Mello, W. J. Junk. 2009. The impact of quaternary sea level changes on the evolution of the Amazonian lowland. *Hydrological Processes*, 23(22): 3168 – 3172.

Katoh, K. and H. Toh. 2008. Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics*, 2(4) 286 – 298.

Kullander, S. O. and E. J. G. Ferreira. 2006. A review of the South American cichlid genus Cichla, with descriptions of nine new species (Teleostei: Cichlidae). *Ichthyological Exploration of Freshwaters*, 17(4): 289 – 398.

Librado, P. and J. Rozas. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451 – 1452.

Luikart, G, F. W. Allendorf, J. M. Cormuet, and W. B. Sherwin. 1997. Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity*, 89 (3): 238 – 247.

Luikart, G., F. W. Allendorf, J.-M. Cornuet, and W. B. Sherwin. 1998. Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity* 89(3): 238 – 247.

Lundberg, J. G., L. G. Marshall, J. Guerrero, B. Horton, M.C.S.L. Malabarba, F. Wesselingh. 1998. The stage for Neotropical fish diversification: a history of tropical South American rivers. In: *Phylogeny and Classification of Neotropical*

Fishes (eds L.R. Malabarba, R. E. Reis, R. P. Vari, Z. M. Lucena, C.A.S. Lucena CAS), pp. 13–48. Edipucrs, Porto Alegre.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* 27: 209 – 220.

Maruyama T. and P. Fuerst. 1985. Population bottlenecks and nonequilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. *Genetics*, 111: 675 – 689.

Piry, S., G. Luikart, and J. M. Cornuet. 1999. BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, 90(4): 502 – 503.

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

Renno J-F., Hubert N., Torrico J.P., Duponchelle F., Rodriguez J. N., Davila C. G., Willis S. C., and Desmarais E. 2006. Phylogeography of *Cichla* (Cichlidae) in the upper Madera basin (Bolivian Amazon). *Molecular Phylogenetics and Evolution*, 41: 503 – 510.

Riddle, B. R., M. N. Dawson, E. A. Daly, D. J. Hafner, M. J. Hickerson, S. J. Mantooth, and A. D. Yoder. 2008. The role of molecular genetics in sculpting the future of integrative biogeography. *Progress in Physical Geography*, 32(2): 173 – 202.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, 139: 457 – 462.

Willis S. C., M. S.Nunes, C. G.Montana, I. P.Farias, and N. R. Lovejoy 2007. Systematics, biogeography, and evolution of the Neotropical peacock basses. *Cichla* (Perciformes: Cichlidae). *Molecular Phylogenetics and Evolution*, 44: 291 – 307.

Willis, S.C., M.,Nunes, C.G.Montaña, I.P.Farias, , & N.R. Lovejoy. 2010. Testing historical hypotheses using an integrated genetic approach: The Casiquiare River acts as a dispersal corridor between the Amazonas and Orinoco Rivers. *Molecular Ecology*, 19: 1014 – 1030.

Winemiller, K. O., H. Lopez – Fernandez, D. C. Taphorn, L. G. Nico, and A. B. Duque. 2008. Fish assemblages of the Casiquiare River, a corridor and zoogeographical filter for dispersal between the Orinoco and Amazon basins. *Journal of Biogeography*, 35: 1551 – 1563.

Wright, S. 1951. The genetical structure of populations. *Annals of Eugenics*, 15: 323 – 354.

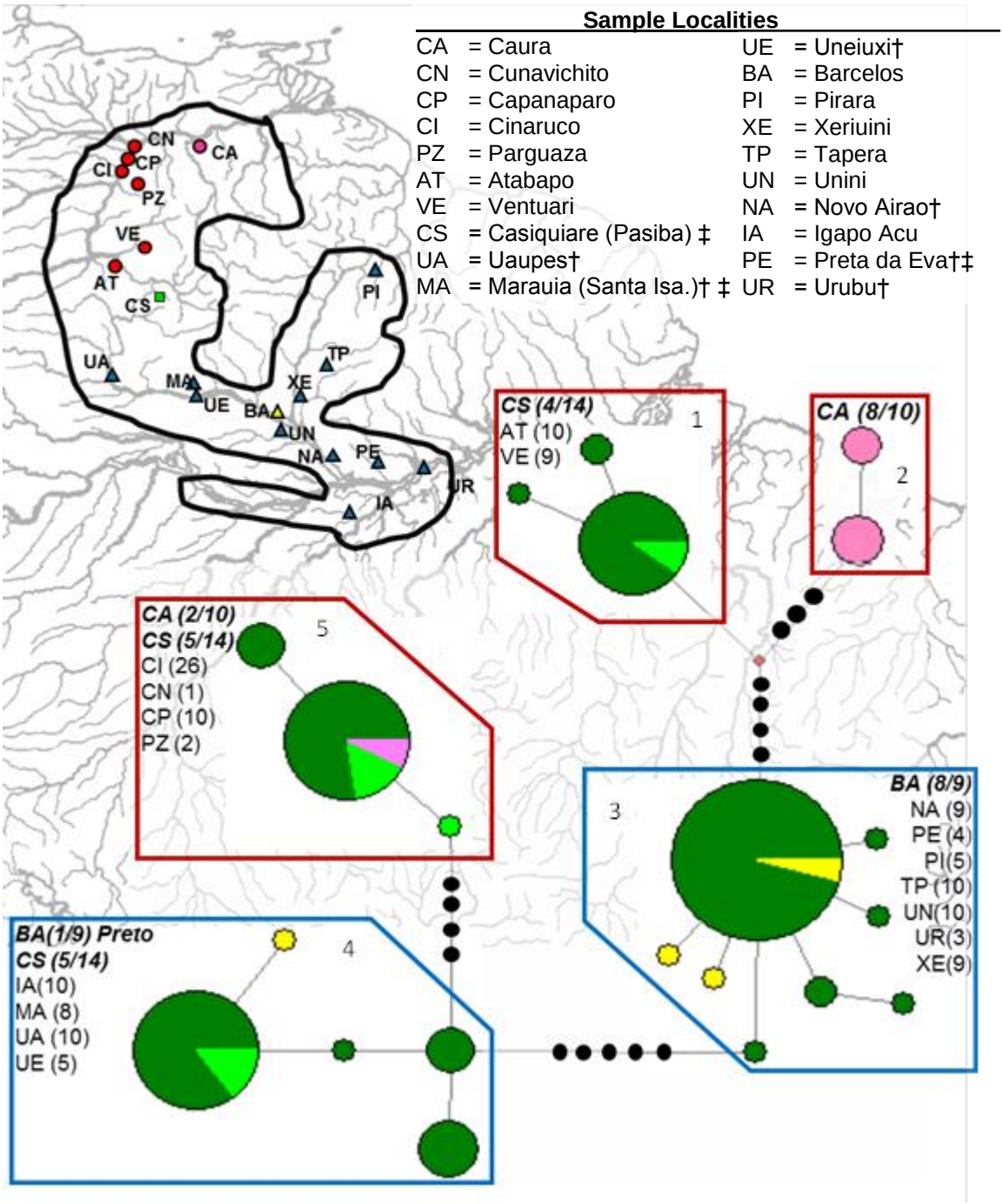

Figure 3.1 - A map of the outlining the natural distribution of C. temensis in the Orinoco and Rio Negro basins with localities collected highlighted. Sample sites are color-coded based on sampling river basin. Sampling localities found in different haplotype groups are highlighted in different colors. (†) Sampling localities new to this study, (‡) Sampling

localities that have unique haplotypes to this study. A median joining haplotype network constructed using the program NETWORK (flux-engineering.com) using haplotypes from the control region (dloop) for samples of *C. temensis* is shown at the bottom. Haplotype groups are identified with numbers 1-5 and are enclosed with all haplotypes corresponding to the appropriate sampling locality. The color for each haplotype group corresponds to the river basin from which the individuals were collected. Numbers represent individuals found within the designated haplotype group with unlabeled black circles representing transitional (missing) haplotypes. The numbers in parenthesis represent the individuals (n) from that sampling locality in the specified haplotype group. Italicized localities are split between more than one haplotype group and identified with an alternative color than the default green.

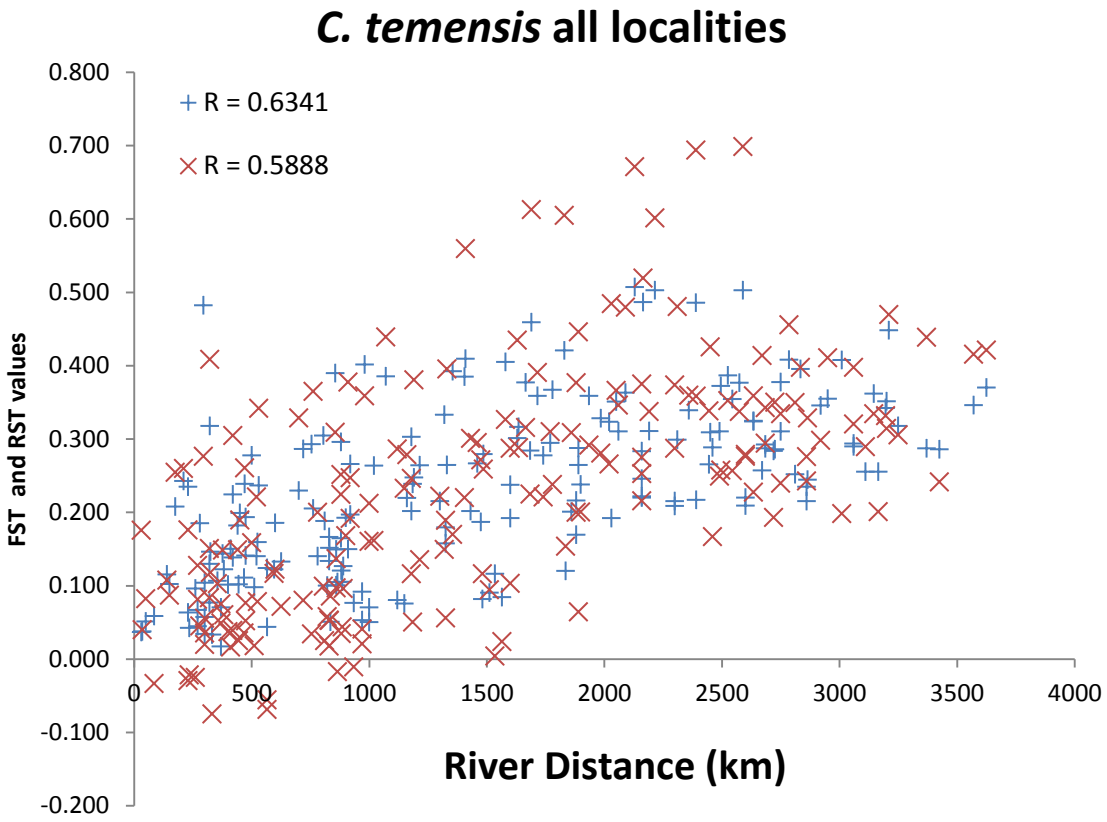

Figure 3.2 – A scatter plot showing the correlation between geographic distance and genetic distance for F_{ST} (+) and R_{ST}(×) values.

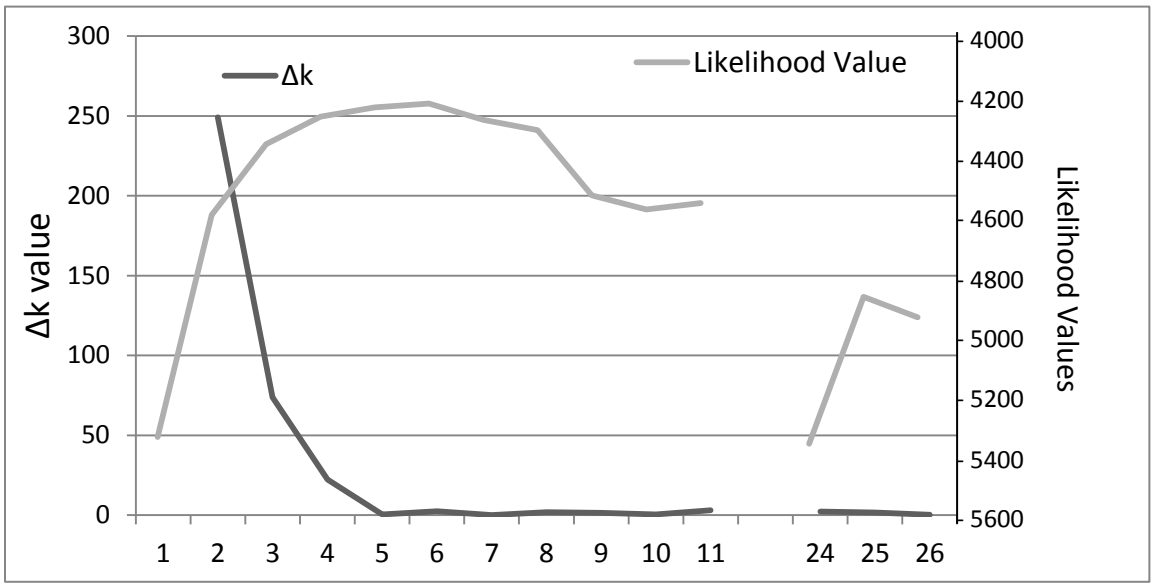

Figure 3.3 - Average likelihood and Δk according to Evanno et al. 2005. The arrow indicates the number of populations selected based on both the average likelihood and Δk values.

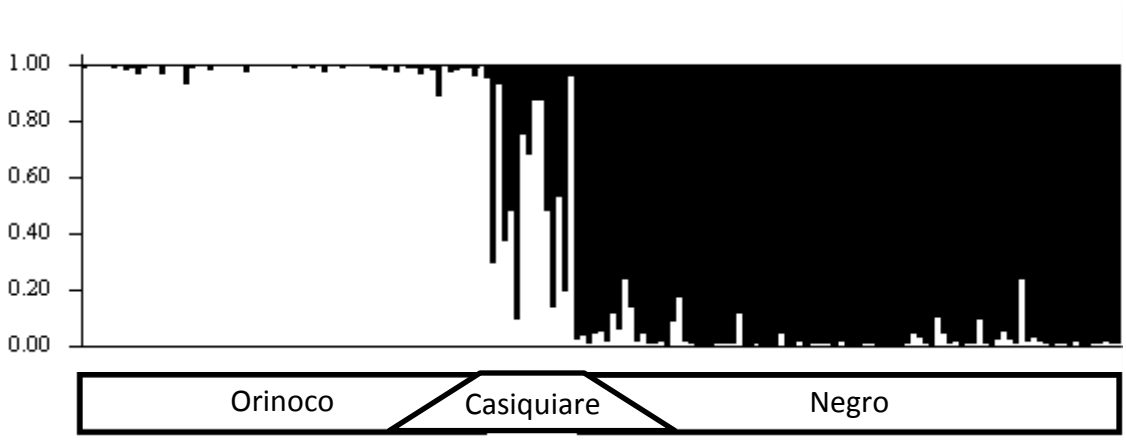

Figure 3.4 - STRUCTURE results from microsatellite loci with the likelihood of individuals being attributed to one population or the other. Each line represents a single individual in the analysis with sampling sites represented by different rivers in which they are established. This shows the population structure based on the highest Δk value (2).

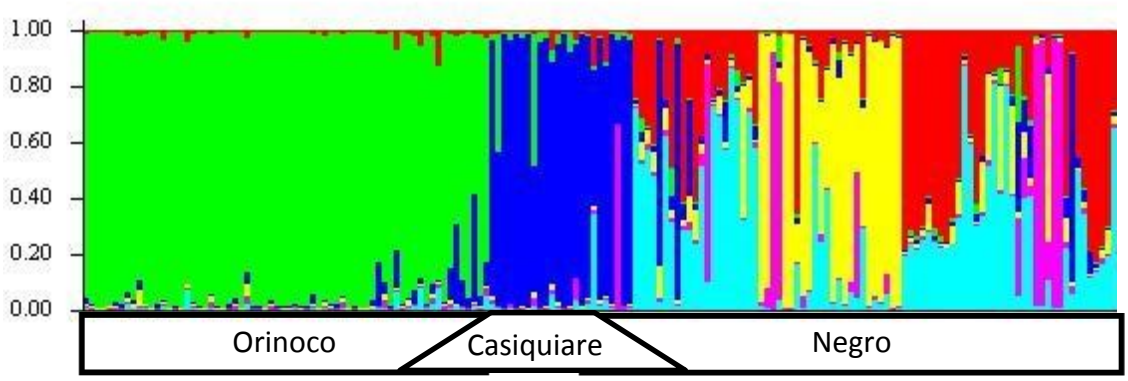

Figure 3.5 - STRUCTURE results from microsatellite loci with the likelihood of individuals being attributed to one population or the other. This figure shows the population structure based on the highest average likelihood (6).

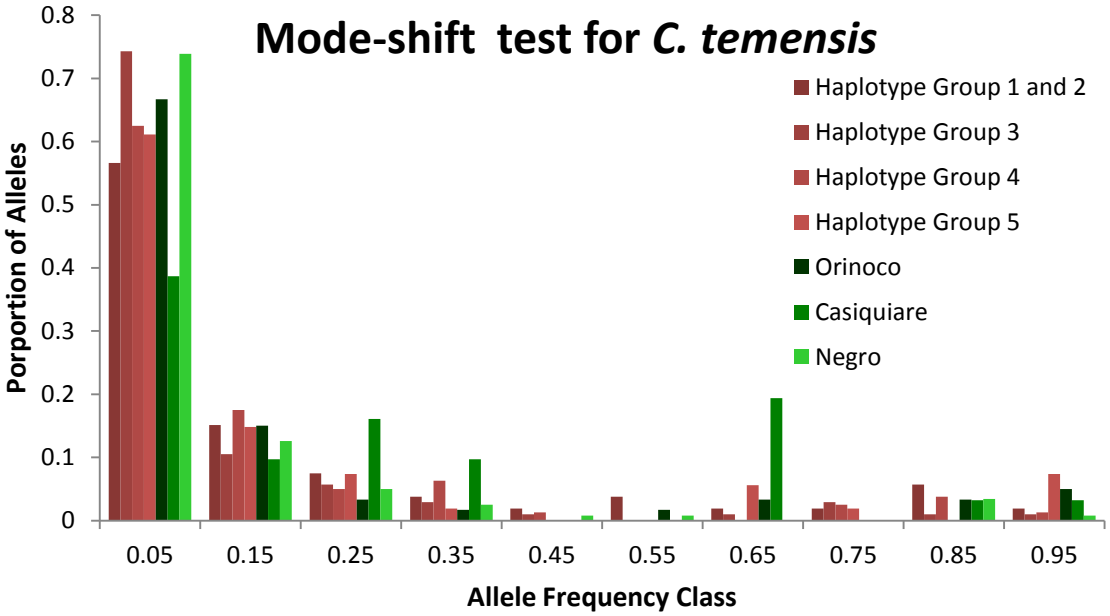

Figure 3.6 – A distribution of allele frequencies for microsatellite loci following the mode-shift test in the program BOTTLENECK.

Table 3.1 – Estimates of FST (below diagonal) and RST (above diagonal) for microsatellites for all sample sites of C. temensis.

	CA	CN*	CP	CI	PZ	AT	VE	CS	UA	MA	UE	BA	PI	XE	TP	UN	NA	IA	PE	UR
CA	---	**0.41**	**0.04**	**0.03**	**0.08**	**0.06**	0.17	0.12	0.24	0.35	0.48	0.36	0.70	0.30	0.41	0.32	0.40	0.42	0.47	0.42
CN*	**0.33**	---	**0.25**	**0.26**	**0.28**	**0.36**	**0.33**	0.36	**0.33**	**0.31**	**0.45**	**0.28**	**0.69**	**0.19**	**0.33**	**0.28**	**0.24**	**0.44**	**0.20**	**0.24**
CP	0.11	**0.24**	---	0.04	**-0.03**	**0.04**	0.08	0.10	0.22	0.22	0.39	0.28	0.60	0.26	0.34	0.29	0.34	0.33	0.40	0.31
CI	0.10	**0.23**	0.04	---	**0.08**	0.06	0.05	0.03	0.17	0.29	0.32	0.29	0.52	0.26	0.35	0.23	0.36	0.33	0.46	0.31
PZ	**0.15**	**0.38**	**0.08**	**0.04**	---	**-0.07**	**0.15**	**0.08**	**0.15**	**0.10**	0.43	**0.20**	0.67	**0.17**	0.25	0.28	0.28	0.29	**0.24**	**0.20**
AT	0.10	**0.22**	0.12	0.12	**0.04**	---	0.11	**0.04**	0.16	0.22	0.40	0.29	0.60	0.25	0.34	0.29	0.37	0.35	0.43	0.33
VE	0.19	**0.23**	0.16	0.19	0.18	0.12	---	**0.04**	0.22	0.28	0.38	0.30	0.61	0.27	0.37	0.22	0.37	0.41	0.48	0.35
CS	0.31	**0.41**	0.31	0.29	0.30	0.23	0.19	---	0.12	0.25	0.38	0.25	0.56	0.22	0.31	0.20	0.38	0.36	0.48	0.34
UA	0.37	**0.42**	0.39	0.39	0.34	0.27	0.30	0.19	---	0.12	**0.10**	0.07	0.31	**0.05**	0.14	**0.06**	0.19	0.15	0.27	**0.06**
MA	0.32	**0.24**	0.30	0.32	0.21	0.23	0.24	0.17	0.13	---	0.18	**0.06**	0.34	0.10	0.10	0.21	0.16	0.09	0.23	**0.02**
UE	0.37	**0.36**	0.35	0.35	0.33	0.25	0.23	0.14	0.09	**0.01**	---	**0.08**	0.16	**0.02**	0.14	**0.02**	**0.04**	**0.12**	0.29	**0.00**
BA	0.32	**0.21**	0.30	0.33	0.21	0.22	0.24	0.19	0.15	**0.02**	**0.02**	---	0.18	**-0.03**	**-0.03**	0.08	**0.05**	**0.03**	0.22	**-0.01**
PI	0.51	**0.53**	0.51	0.48	0.53	0.43	0.46	0.42	0.40	0.27	0.29	0.22	---	**0.05**	0.12	0.25	0.19	0.30	0.44	0.26
XE	0.35	**0.29**	0.36	0.37	0.29	0.29	0.32	0.28	0.24	0.09	0.12	0.08	0.17	---	**-0.02**	**0.04**	**0.02**	**0.02**	0.19	**-0.02**
TP	0.36	**0.32**	0.38	0.38	0.31	0.31	0.35	0.30	0.27	0.11	0.13	0.11	0.13	0.04	---	0.15	0.08	**0.05**	0.26	**0.04**
UN	0.29	**0.24**	0.29	0.32	0.22	0.21	0.22	0.22	0.18	0.06	0.05	0.07	0.27	0.10	0.13	---	0.13	0.20	0.30	**0.09**
NA	0.30	**0.23**	0.29	0.32	0.22	0.22	0.22	0.17	0.16	0.04	**0.02**	**0.02**	0.21	0.06	0.08	0.04	---	**0.02**	**0.09**	**-0.07**
IA	0.35	**0.32**	0.35	0.36	0.28	0.26	0.26	0.22	0.13	0.10	**0.06**	0.11	0.22	0.19	0.16	0.14	0.10	---	0.15	**-0.06**
PE	0.45	**0.42**	0.41	0.41	**0.38**	0.31	0.30	0.20	0.19	0.08	0.09	0.11	0.40	0.20	0.24	0.14	0.10	0.15	---	**0.02**
UR	0.38	**0.30**	0.33	0.35	**0.25**	0.25	0.29	0.27	0.26	0.08	0.10	0.07	0.30	0.10	0.10	**0.04**	**0.04**	0.13	0.15	---

Note: Nonsignificant pairwise F_{ST} and R_{ST} (Excoffier et al. 2005) values are in bold. *There was only one specimen of *C. temensis* from Cunavichito (CN) which would result in those FST values being nonsignificant.

Table 3.2 – Sign test and Wilcoxon two tailed test with 1000 replicates for *C. temensis* with the population divided by both the mtDNA and microsatellite data. M'sat = microsatellite. (Note: Significant p-values are in bold. Parameters for the two-phase model were set according to the authors recommendations [variance (v) = 12 and proportion of SMM (p) = 95% (Piry et al 1999)]. When alternative v and p values changed from significant to non-significant they are noted by (*) for v=12 and p =70; and by (**) for v = 30 and p = 70 when only the Orinoco population under the sign test remained significant. When v = 30 and p = 95 there were no changes observed between significant and non-significant values).

		Sign test			Wilcoxon Test (two tails)		
		IAM	TPM	SMM	IAM	TPM	SMM
mtDNA population divisions	Hap-1 & 2	0.50243	0.06384	**0.01831**	0.69531	**0.03223**	**0.00977**
	Hap-3	0.31403	**0.00497**	**0.00532**	0.67725	**0.00610**	**0.00342**
	Hap-4	0.56283	**0.00369**	**0.00053**	0.96973	**0.00342**	**0.00122**
	Hap-5	0.11513	**0.02147**	**0.01975**	0.62500	**0.01367**	**0.01367**
M'sat population divisions	Orinoco	0.32752	**0.02142****	**0.00452**	0.79102	**0.01050**	**0.00244**
	Casiquiare	0.58247	0.31714	0.32984	0.82031	0.30078	0.30078
	Negro	0.28748	**0.01677***	**0.00267**	0.69531	**0.00488***	**0.00293**

CHAPTER 4 –POPULATION STRUCTURE OF THE WIDESPREAD PEACOCK BASS *CICHLA MONOCULUS*

4.1 – Introduction

Major geologic events that have occurred throughout the Neotropics are expected to have influenced phylogeographic patterns and population histories of many Amazonian organisms (Lundberg et al. 1998). There have been several attempts to explain population distributions and biogeographic structure using geologic events as a basis for testing distributions in South America (Sivasundar et al. 2001; Montoya-Burgos 2003; Hrbek et al. 2005; Hubert and Renno 2006; Farias and Hrbek 2008). Major geological events that have occurred throughout the Cenozoic (starting ~ 65 MYA) are believed to have had the greatest influence on this diversity (Lundberg et al. 1998), specifically major geologic events leading to the breaching of the Purus Arch and establishment of the current Amazon basin (Ready et al. 2006). Alternative hypotheses also propose that species distributions have been affected by ecological heterogeneity in the Amazon River basin (Cooke et al. 2009) or simply exhibit a lack of gene flow due to their distribution spanning across large geographic distances (Hubert et al. 2007b). These results, however, have yet to converge on a unifying theme, event, or even time period, which can be used to best explain the observed diversity. Additionally, some populations restricted to the main Amazon channel may exhibit panmictic population structure or high levels of gene flow despite the population spanning thousands of kilometers (Santos et al. 2007).

The species *C. monoculus* is naturally distributed throughout the Amazon River Basin and many of its tributaries and has overlapping distribution with five other *Cichla*

species (*C. pinima*, *C. temensis*, *C. intermedia*, *C. orinocensis*, and *C. vazzoleri*; Kullander and Ferreira 2006). It exhibits a monophyletic mtDNA group and has relatively well established species boundaries although there have been described rare hybridization events to occur between *C. monoculus* and both *C. temensis* and *C. orinocensis* (Willis et al. 2010). Previous studies characterizing the genetic structure of Amazon fish species with distributions similar to *C. monoculus* have shown contradicting patterns of genetic diversity in this region (Ready et al. 2006; Santos et al. 2007, Hrbek 2005). Here I describe combined mtDNA and microsatellite analyses and their possible sources of genetic diversity regarding the population structure in the most widely distributed peacock bass, *C. monoculus*. The combined analysis will test whether the population exhibits a panmictic distribution across their wide geographic range or if there is population structure. Additionally, if there is population structure I will contrast these results to what might be expected following the breaching of the Purus Arch and the variation of water types throughout the basin to determine whether these events are reflected in the genetic variation across their population distribution.

4.2 Materials and Methods

4.2.1 – Sampling Strategy

Individuals of *C. monoculus* were collected from 27 localities throughout the Amazon River and its tributaries to provide broad coverage across their natural distribution spanning thousands of kilometers (Figure 4.1). The sampling strategy for individuals used in this analysis is described by Renno et al. (2006) and Willis et al. (2007). For each fish collected, a voucher photograph was taken and tissue samples were

preserved in SMSO-EDTA buffer. Although the same individuals from Renno et al. (2006), Willis et al. (2007), and Willis et al. (2010) were incorporated into these analyses additional samples have been collected since then, thanks to the strong collaborative collection effort established by many of the authors cited throughout this paper (Renno et al. 2006; Willis et al. 2007; Willis et al. 2010). Although hybrids of *C. monoculus* have been identified in previous studies (Willis et al. 2007) only individuals exhibiting the *C. monoculus* morphotype and "*C. monoculus*" mtDNA sequences were used in this study.

4.2.2 – Sequencing and Genotyping

Sequences for mtDNA control regions from 88 new samples collected by collaborators using previously described primers and procedures (Willis et al. 2007; Willis et al. 2010). These sequences were used in combination with control region (CR) sequences from previous analyses (DQ841872 - DQ841899, Willis et al. 2007; GU295709 - GU295732, Willis et al. 2010). A total of 249 sequences spanning approximately 550 bases of the CR were aligned for *C. monoculus* using the online version of the program MAFFT version 6, following the G-INS-I alignment strategy with all other parameters set to the default (Katoh & Toh 2008). The conditions for PCR were the same as described in Chapter 2 using MJ Tetrads and alleles sizes measured on an ABI 3730 and then genotyped using the program GeneMapper (Applied Biosystems).

4.2.3 – mtDNA Analysis

The aligned mtDNA CR for *C. monoculus* was used to construct a median joining haplotype network (Bandelt et al. 1999) implemented in the program NETWORK (flux-

engineering.com). The input for NETWORK was produced using the program DnaSP (Librado and Rozas 2009). The haplotype network was constructed using all default parameters and the haplotype network was arranged within the program to represent *C. monoculus* sampling localities (Figure 4.2).

4.2.4 – Microsatellite Analyses

In addition to analyses done with mitochondrial sequence, population structure was also described using microsatellite based analyses. Genetic variation amongst sampling localities was estimated based on microsatellite genotypes for both F_{ST} and R_{ST} values using the program Arlequin v. 3.1 (Excoffier et al. 2005). The Mantel Test (Mantel 1967) was also done for the population as a whole and just the Rio Negro samples. The methods used to investigate genetic differentiation across sampling sites and conduct the Mantel Test are described in Chapter 3, section 3.2.5. Additionally, the number of genetically distinct populations was determined using the likelihood of individuals belonging to a specific number of populations in STRUCTURE (Pritchard et al. 2000). For this analysis, the first 50,000 generations of the Marcov Chain Monte Carlo (MCMC) were discarded as burn-in and the remaining 50,000 were used to estimate the proportion of membership to a certain population. The number of populations (k) was estimated using the protocol described by Evanno et al. (2005) as described in Chapter 3, section 3.2.5. The highest average likelihood values and Δk statistics were then compared for subsets of *C. monoculus* population (Figure 4.3 A - D).

To test for deviations of MDE across the *C. monoculus* population we used the program BOTTLENECK 1.2.02 (Piry et al. 1999) to implement two tests to determine

whether the population had experienced a bottleneck event or a population expansion. As the effective population size decreases during a bottleneck the allele numbers and heterozygosities at polymorphic loci are greatly reduced (Cornuet and Luikart 1996). The program simulated these changes over 1000 replicates based on the number of unique alleles and compares the simulated heterozygosity deficiency with the observed. The program simulates these changes in a population with statistical analyses following the IAM (Maruyama and Fuerst 1985), as well as heterozygosity excess or deficiencies using SMM (Cornuet and Luikart 1996) and the TPM (Di Rienzo et al. 1994). These are then tested for significant deviations from expected MDP with significant p-values less than 0.05. Additionally, the genetic signal indicates an apparent mode-shift distortion of allele frequencies when MDE is disrupted (Luikart et al. 1997) and can be visualized graphically to interpret whether or not the population is stable. The population was separated according to STRUCTURE results from the microsatellite analysis, the mtDNA haplotype network as well as sampling localities.

4.3 - Results

4.3.1 – mtDNA analysis

No specimens used in this study were identified as potential hybrids based on morphology of *C. monoculus* and mtDNA sequences. Every locality on the map was represented except for Mavaca and PretaDEva in the mtDNA analysis. Sampling localities with more than 10 individuals did not have all mtDNA CR's sequenced as there was little variation amongst these localities beyond the first 10 sequences. From the 88 new mtDNA sequences there were 18 unique haplotypes that have not been previously

reported. The median-joining haplotype network reconstructed by the program NETWORK (flux-engineering.com) placed the 249 sequences amongst 70 haplotypes (Figure 4.2). The haplotypes were sorted amongst the three different groups, corresponding with the Upper Amazon, Lower Amazon, and Rio Negro. For the most part the species haplotype network corresponds with their geographic distribution, with some exceptions mentioned below. Haplotypes shared between these rivers were found in close proximity to where the Upper Amazon, Lower Amazon and Rio Negro (Middle Amazon) meet (Figure 4.2). This occurred with individuals collected from Canacari, Borba, Novo Airao, and Manacapuru. Additionally, some haplotypes observed in the Xingu, Itatiuba, Jari, Araguari sampling localities did not seem to follow this trend. They were from the Lower Amazon, but have six missing haplotypes between the Upper Amazon and twelve missing haplotypes from the nearest Lower Amazon haplotype. Additionally, a single individual from Tabatinga was placed amongst the Rio Negro haplotypes, which are 1,500 km (Novo Airao) and 2,100 km (Barcelos) away.

4.3.2 – Microsatellite analyses

Every locality on the map was represented except for Jurua with the microsatellite loci due to poor quality DNA from that sampling locality. Additionally, some individuals were omitted from the microsatellite analysis due to an excess of failed PCR amplifications. Ultimately, specimens that had more than four out of 12 missing genotypes were removed from the analysis, leaving a total of 224 genotypes for the *C. monoculus* population. Collectively, the 27 sampling localities had a moderate level of global genetic variation (F_{ST} = 0.2405 and R_{ST} = 0.2844), both showing strong population

structure by sampling locality (Table 4.1). The sampling locality with the highest average F_{ST} values was Marauia (F_{ST} = 0.4054) and the highest R_{ST} value was Mavaca (R_{ST} = 0.6556), even though Eirunepe (F_{ST} = 0.2447/ R_{ST} = 0.2932) and Cruzeiro (F_{ST} = 0.2809/ R_{ST} = 0.3377) are the furthest away from all other localities by an average of 3,610 and 2,812 km respectively. The highest F_{ST} was between Marauia and Iquitos (F_{ST} = 0.6284) and R_{ST} was between Mavaca and Iquitos (R_{ST} = 0.8832). Finally average F_{ST} and R_{ST} values were calculated for the Upper Amazon (F_{ST} = 0.2016/ R_{ST} = 0.2154), Lower Amazon (F_{ST} = 0.2134/ R_{ST} = 0.2574), and Rio Negro/Middle Amazon (F_{ST} = 0.2810/ R_{ST} = 0.3681). Overall there is a strong association of genetic variation between geographic distance and localities with more R_{ST} values indicated as being not significant than F_{ST}. All correlations are significant ($P < 0.001$ for F_{ST} and R_{ST}) according to the Mantel test (Mantel 1967) with a significant correlation coefficient for the population as a whole for both F_{ST} (R = 0.6021) and R_{ST} (R = 0.6053), as well as when just the individuals from the Rio Negro F_{ST} (R = 0.6793) and R_{ST} (R = 0.6873).

For the structure analysis likelihood values were averaged across five runs for values of *k* ranging from 1 – 11 and 26 – 28 from STRUCTURE following a protocol described by Evanno et al. (2005). First analyses were done based on the *k* range of 1-11 following the graphical output produced in structure and overall trends observed with Δ*k* and the average likelihood values. Next, a *k* range of 26 – 28 was analyzed as there were 27 sample sites, to test whether the highest Δ*k* was associated with sampling locality. The likelihood from the different *k* values were used to calculate the Δ*k* statistic (Δ*k* = m(|L"K|)/s[L(K)]) (Evanno et al. 2005). The highest Δ*k* statistic recognized a subdivision of *k* = 2 populations corresponding to the Amazon River as a whole and the Rio Negro

sample sites (Figure 4.3A). As the Amazon population spans thousands of kilometers, further analyses were done to see whether population subdivisions could be detected within the Amazon. Sampling localities found along the Rio Negro were analyzed separately from the main channel (Figure 4.3B), the main channel without the Middle Amazon and Rion Negro sampling localities (Figure 4.3C), and the main channel Amazon without the furthest west sampling localities and the southern tributaries (Figure 4.3D) as they were separated in the Amazon main channel analysis (Figure 4.3C).

There was no indication of significant excess of heterozygosities for all tests and population subsets in the program BOTTLENECK. There were significant p-values for heterozygosity deficiency when the population was divided for both mtDNA and microsatellite analyses (Table 4.2), which indicates possible population expansion. The IAM analysis was significant for the Xingu River according to the sign test. For the most part, the SMM and TPM were significant for the subdivisions according to both mtDNA haplotype groups and the microsatellite analysis. The exceptions were samples from the Xingu and Cruzeiro/Eirunepe, which was not significant for any test and individuals from Itaituba were not significant for the Wilcoxon-signed rank test. When the parameters were changed for the two phase model, changes in significant values are also noted (Table 4.2). Finally, all of the allele frequencies were L-shaped, as intermediate allele frequency did not exceed low and high allele frequency classes (Figure 4.6), which indicates a long term stable population size (Luikart et al. 1997).

4.4 – Discussion

We combined both mtDNA sequence data and nuclear microsatellite loci to investigate population structure of the peacock bass *C. monoculus* throughout its native range in the Amazon River basin. We found a high degree of population structure with the mitochondrial sequences illustrating population structure throughout their basin and identify localities that exhibit gene flow between the different basins from the middle of the Amazon. Additionally the microsatellite analysis both complements and contradicts the mitochondrial data, highlighting areas exhibiting significant population structure as well as other areas resembling high rates of gene flow across this widely dispersed population.

4.4.1- Evidence for Panmixia?

The haplotype network shows a strong indication of population structure correlating with geographical distributions when divided into three sections: the Upper Amazon, Middle Amazon/Rio Negro and Lower Amazon (Figure 4.1 and Figure 4.2). Despite several localities in the Middle Amazon geographically closer to other localities in the Upper or Lower Amazon they were grouped with other samples from the Rio Negro. Many individuals from the main channel localities share haplotypes showing gene flown throughout this portion of the population. The STRUCTURE results reinforced the evidence of gene flow with apparent panmixia along several hundred kilometers of the main Amazon basin (Figure 4.3 D). Despite these results the haplotype network and remaining STRUCTURE analyses depict structure that would not be observed in a panmictic species with the Xingu and Itaituba separate from the remaining main channel sampling localities (Figure 4.3 C).

4.4.2 - Life History Characteristics

The rate of dispersal for *Cichla* has been described for both introduced (Zaret 1980) and native distributions (Hoeinghaus et al. 2003), illustrating that most fishes remained close to their natal or capture locations. *Cichla* have been observed to be substrate spawners and even mouth brooders providing some parental care for up to several months (Zaret 1980). Such life history characteristics of species are often overlooked when attempting to understand processes that may influence genetic variation at the population level in large rivers like the Amazon. One major caveat to phylogeographic investigations with mtDNA is the potential for artifactual results due to its strict maternal inheritance. If, for instance, female fishes were site-fidelous and males would show a tendency to migrate between localities, mtDNA could portray significant population structure even if nuclear genes remained homogeneous showing panmixia (Avise 2000).

The results for *C. monoculus* show a high number of unique mtDNA haplotypes exhibiting strong population structure (Figure 4.2), however, the microsatellite analysis shows high levels of gene flow along the main channel of the Amazon (Figure 4.3C & D). These contrasting patterns may be the result of biases due to maternal inheritance as the distribution of mtDNA haplotypes constrained by sex-biased migration in *C. monoculus*, with bi-parentally inherited loci indicating less structured genetic diversity. The contradicting population structure between molecular marker types (maternally vs. bipaternally inherited) is not sufficient evidence to reject the possibility of male-biased migration within the main channel Amazon. These results may provide some insight into

population genetics of many widely dispersed Amazon fishes that exhibit similar life history characteristics.

4.4.3- Population Expansion

Overall the statistical tests using microsatellites show a strong signal for population expansion in several population divisions (Table 4.2). Although the results indicate higher p-values overall for the Rio Negro and Lower Amazon (except for the Xingu and Itatuba) the signal is not much different than the Upper Amazon. The ability to appropriately test for variation in MDE following population expansion heavily relies on correct assumptions for mutation model and population structure. Similarly to our results from Chapter 3 dividing the population according to both mtDNA and microsatellite structure provided overwhelmingly significant results. The inferred subpopulations could be incorrect producing misleading results. When the population was divided based on sampling localities (data not shown) only some of the localities had significant p-values for population expansion. Despite these analyses, the results could not strongly support any hypotheses relating to geologic events followed with colonization or individuals of this species potentially moving to new environments followed by population expansions. As *C. monoculus* shares its distribution with many other Amazon species comparative tests across species may shed light on possible historical changes in population size affecting the Neotropical icthyofauna.

4.4.4- Geologic and Ecological Influences on Population Structure

The haplotype network and STRUCTURE results do not depict any significant representations that overall corroborate with the breaching of the Purus Arch or any other major events in landscape evolution thought to influence population structure (Lundberg et al. 1998). The identified genetic differentiation highlighted in the Tapajos and Xingu Rivers (Figure 4.3C) could possibly be remnant genetic signals of large lakes due to periods of glacial retreat in the Amazon when large narrow freshwater lakes developed and covered the entire Lower Amazon (Irion et al. 1995; Irion et al. 2009). Although some haplotype groups consist of several localities separated by long geographic distances they are found alongside those which exhibit more population structure, the results reported here do not significantly coincide with what one may be expected based on the geologic history of the region.

On the other hand, evidence of genetic variation from mtDNA and microsatellite data due to ecological heterogeneity throughout the Amazon River basin (i.e., water types) is more difficult to dismiss. Both the haplotype network and STRUCTURE results show some correlation with river types with samples exclusively from the Rio Negro dividing the population (Figure 4.3B). In the haplotype network, the headwaters or Upper Amazon localities (indicated with yellow haplotype in Figure 4.2) are mostly "white-water" in their description, the Rio Negro localities (indicated in with green haplotypes) are mostly "black-water", and the Tapajos and Xingu samples are from "clear-water" rivers. In both the mtDNA and microsatellite analyses, there is a strong genetic signal for differentiation between the "black-water" and all other water types (Figures 4.2 & 4.3A). As well as separation between clear water and other water types (Figures 4.2 & 4.3 A). Additionally, the average pairwise F_{ST} and R_{ST} values in the Negro were much higher in

comparison to the Upper and Lower Amazon even though the average distance between all other sampling localities on the Rio Negro was 1,497 km compared to 1,962 km and 1,401km for the Upper and Lower Amazon sample localities respectively. Although the observed genetic variation is higher over shorter distance the F_{ST} and R_{ST} spatial autocorrelation via Mantel test (Mantel 1967) (Figure 4.5 and 4.6) support the hypothesis of isolation by distance indicating gradual gene flow throughout the Rio Negro into the Middle Amazon channel. These results suggest that the genetic variation between the Rio Negro and the main Amazon channel are influenced by the ecology of the different water types, but still exhibit population structure due to geographic distance.

4.5 Conclusion

The combined analysis from mtDNA and microsatellite loci provides insight into genetic variation at the population level. Collectively the results indicated that *C. monoculus* does not exhibit high levels of gene flow throughout their entire distribution, with a strong separation between the Rio Negro and the remaining individuals from the Amazon, therefore not exhibiting panmictic gene flow throughout their entire distribution. The wide distribution and life history characteristics make *C. monoculus* an ideal candidate for these types of population genetic studies and additional nuclear data may allow for direct testing of potential historical demography from geologic data as well as ecological heterogeneity. Which are two main factors thought to have shaped the diversity of icthyofauna throughout the Amazon.

References

Avise, J.C. 2000. Phylogeography; the history and formation of species. Harvard University Press. Cambridge, Mass.

Bandelt H-J, P. Forster, and A. Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16:37-48

Cooke, G. M., N. L. Chao, and L. B. Beheregarary. 2009. Phylogeography of a flooded forest specialist from central Amazonia based on intron DNA: the cardinal tetra *Paracheirodon axelrodi*. *Freshwater Biology*, 54 1216 – 1232

Cornuet J.M. and G. Luikart. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144:2001-2014.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14:2611–2620.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1: 47 – 50.

Farias, I.P. and T Hrbek. 2008. Patterns of diversification in the discus fishes (*Symphysodon* spp. Cichlidae) of the Amazon basin. *Molecular Phylogenetics and Evolution*, 49: 32 – 43.

Hoeinghaus, D. J., C. A. Layman, D. A. Arrington, and K. O. Winemiller. 2003. Spatiotemporal variation in fish assemblage structure in tropical floodplain creeks. *Environmental Biology of Fishes*, 67: 379 – 387.

Hrbek T., I.P. Farias, M. Crossa, I. Sampaio, J.I.R. Porto, and A. Meyer. 2005. Population genetic analysis of the *Arapaima gigas*, one of the largest freshwater fishes of the Amazon basin: implications for its conservation. *Animal Conservation*, 8(3): 297 – 308.

Hubert N. and J-F Renno. 2006. Historical biogeography of South American freshwater fishes. *Journal of Biogeography*, 33: 1414 – 1436.

Hubert, N., F. Cuponchelle, J. Nunez, C. Garcia-Davila, D. Paugy, and J-F. Renno. 2007a. Phylogeography of the piranha genera *Serrasalmus* and *Pygocentrus*: implications for the diversification of neotropical ichthyofauna. *Molecular Ecology*, 16: 2115 – 2136.

Hubert, N., F. Duponchelle, J. Nunez, R. Rivera, F. Bonhomme and J Renno. 2007b. Isolation by distance and Pleistocene expansion of the lowland populations of the white piranha *Serrasalmus rhombeus*. *Molecular Ecology*, 16: 2488 – 2503.

Irion, G., J. Muller, J. Nunes de Mello, W. J. Junk. 1995. Quaternary geology of the Amazonian lowland. *Geo-Marine Letters*, 15 (3-4): 172 – 178.

Irion, G., J. Muller, J. O. Morais, G. Keim, J. Nunes de Mello, W. J. Junk. 2009. The impact of quaternary sea level changes on the evolution of the Amazonian lowland. *Hydrological Processes*, 23(22): 3168 – 3172.

Katoh, K. and H. Toh. 2008. Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics*, 2(4) 286 – 298.

Kullander, S. O. and E. J. G. Ferreira. 2006. A review of the South American cichlid genus Cichla, with descriptions of nine new species (Teleostei: Cichlidae). *Ichthyological Exploration of Freshwaters*, 17(4): 289 – 398.

Librado, P. and J. Rozas. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451 – 1452.

Luikart, G., F. W. Allendorf, J.-M. Cornuet, and W. B. Sherwin. 1998. Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity* 89(3): 238 – 247.

Lundberg J.G., L. G. Marshall, J. Guerrero, B. Horton, M.C.S.L. Malabarba and F. Wesselingh. 1998. The stage for Neotropical fish diversification: a history of tropical South American rivers. *Phylogeny and classification of Neotropical fishes.* (eds Malabarbra L.R., Reis R.E., Vari R.P., Lucena Z.M.S., Lucena C.A.S.) pp. 375 – 400. Edipucrs, Porto Alegre, Brazil.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27: 209 – 220.

Maruyama T. and P. Fuerst. 1985. Population bottlenecks and nonequilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. *Genetics*, 111: 675 – 689.

Montoya-Burgos, J. I. 2003. Historical biogeography of the catfish genus Hypostomus (Siluriformes: Loricariidae), with implications on the diversification of Neotropical ichthyofauna. *Molecular Ecology*, 12, 1855 – 1867.

Piry, S., G. Luikart, and J. M. Cornuet. 1999. BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, 90(4): 502 – 503.

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

Ready, J. S., E. J. G. Ferreira, and S. O. Kullander. 2006. Discus fishes: mitochondrial DNA evidence for a phylogeographic barrier in the Amazonian genus *Symphysodon* (Teleostei: Cichlidae). *Journal of Fish Biology*, 69: 200 – 211.

Renno J-F., Hubert N., Torrico J.P., Duponchelle F., Rodriguez J. N., Davila C. G., Willis S. C., and Desmarais E. 2006. Phylogeography of *Cichla* (Cichlidae) in the upper Madera basin (Bolivian Amazon). *Molecular Phylogenetics and Evolution*, 41: 503 – 510.

Santos M.C.F., M.L. Ruffino and I.P. Farias. 2007. High levels of genetic variability and panmixia of the tambaqui *Colossoma macropomum* (Cuvier, 1816) in the main channel of the Amazon River. *Journal of Fish Biology*, 71 (Supplement A): 33 – 44.

Sivasundar, A., E. Bermingham, and G. Orti. 2001. Population structure and biogeography of migratory freshwater fishes (Prochilodus: Characiformes) in major South American rivers. *Molecular Ecology*, 10: 407 – 417.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, 139: 457 – 462.

Willis S. C., M. S.Nunes, C. G.Montana, I. P.Farias, and N. R. Lovejoy 2007. Systematics, biogeography, and evolution of the Neotropical peacock basses *Cichla* (Perciformes: Cichlidae). *Molecular Phylogenetics and Evolution*, 44: 291 – 307.

Willis, S. C., M. Nunes, C. G. Montana, I. P. Farias, G. Orti, and N. Lovejoy. 2010. The Casiquiare river acts as a corridor between the Amazonas and Orinoco river basins: biogeographic analysis of the genus Cichla. *Molecular Ecology*, 19: 1014 – 1030.

Wright, S. 1951. The genetical structure of populations. *Annals of Eugenics*, 15: 323 – 354.

Zaret, T. M. 1980. Life history and growth relationships of *Cichla ocellaris*, a Predatory South American Cichlid. *Biotropica*, 12: 144 – 157.

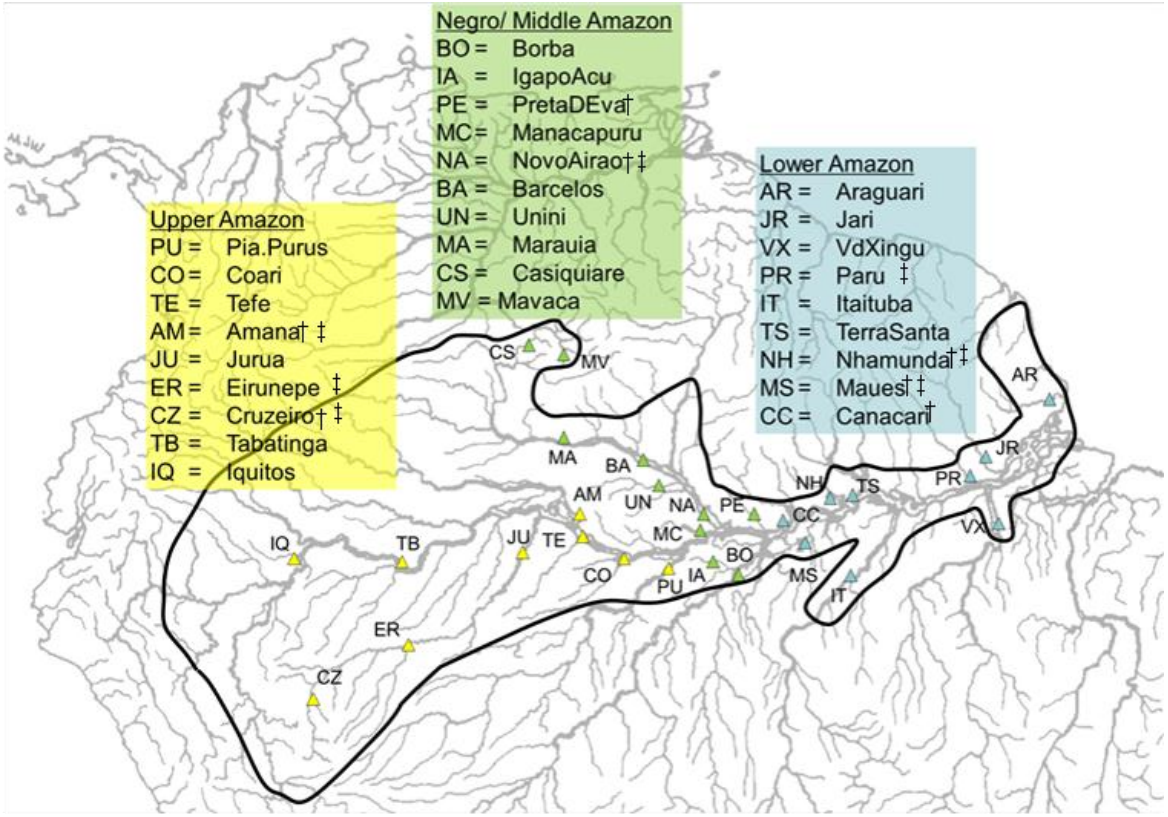

Figure 4.1 –Map of the outlining the natural distribution of *C. monoculus* in the Amazon and surrounding River basins with localities collected highlighted. Sample sites are color-coded based on sampling region. Note: (†) Sampling localities new to this study, (‡) Sampling localities that have unique haplotypes to this study. Italicized localities represent sites from which only microsatellite data was obtained and underlined localities indicate those from which only mtDNA data was obtained.

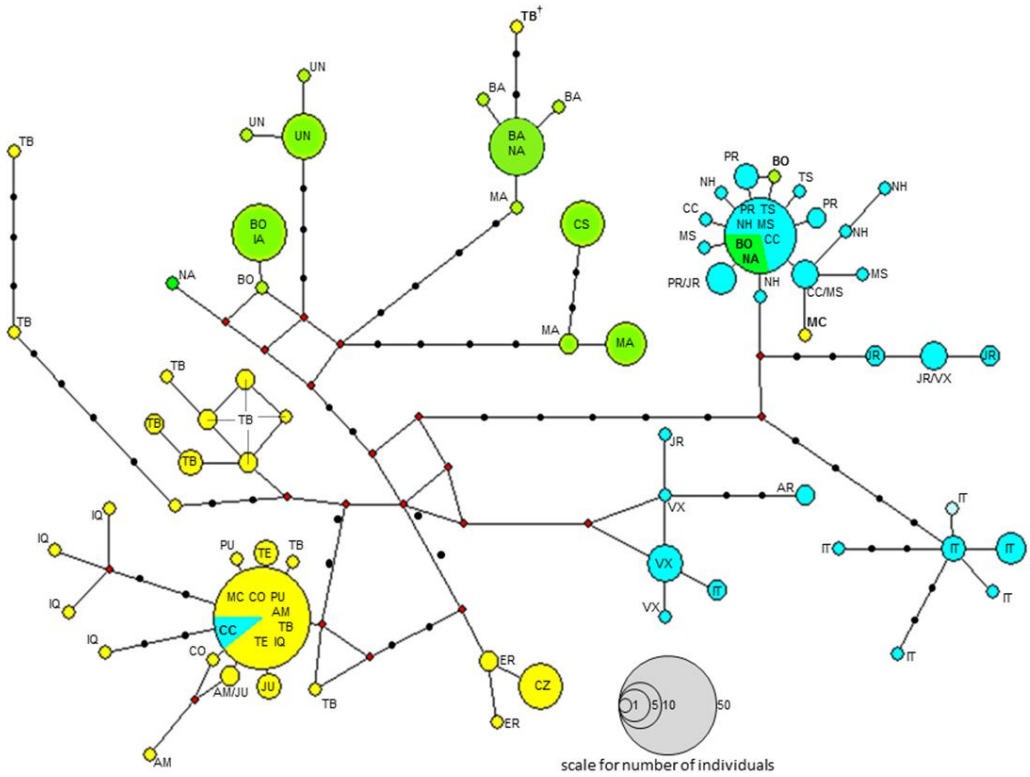

Figure 4.2 - Median joining haplotype network using haplotypes from the control region (dloop) for samples of *C. monoculus*. Circle size is proportional to the number of individuals with the allele. Unlabeled circles represent transitional (missing) haplotypes.

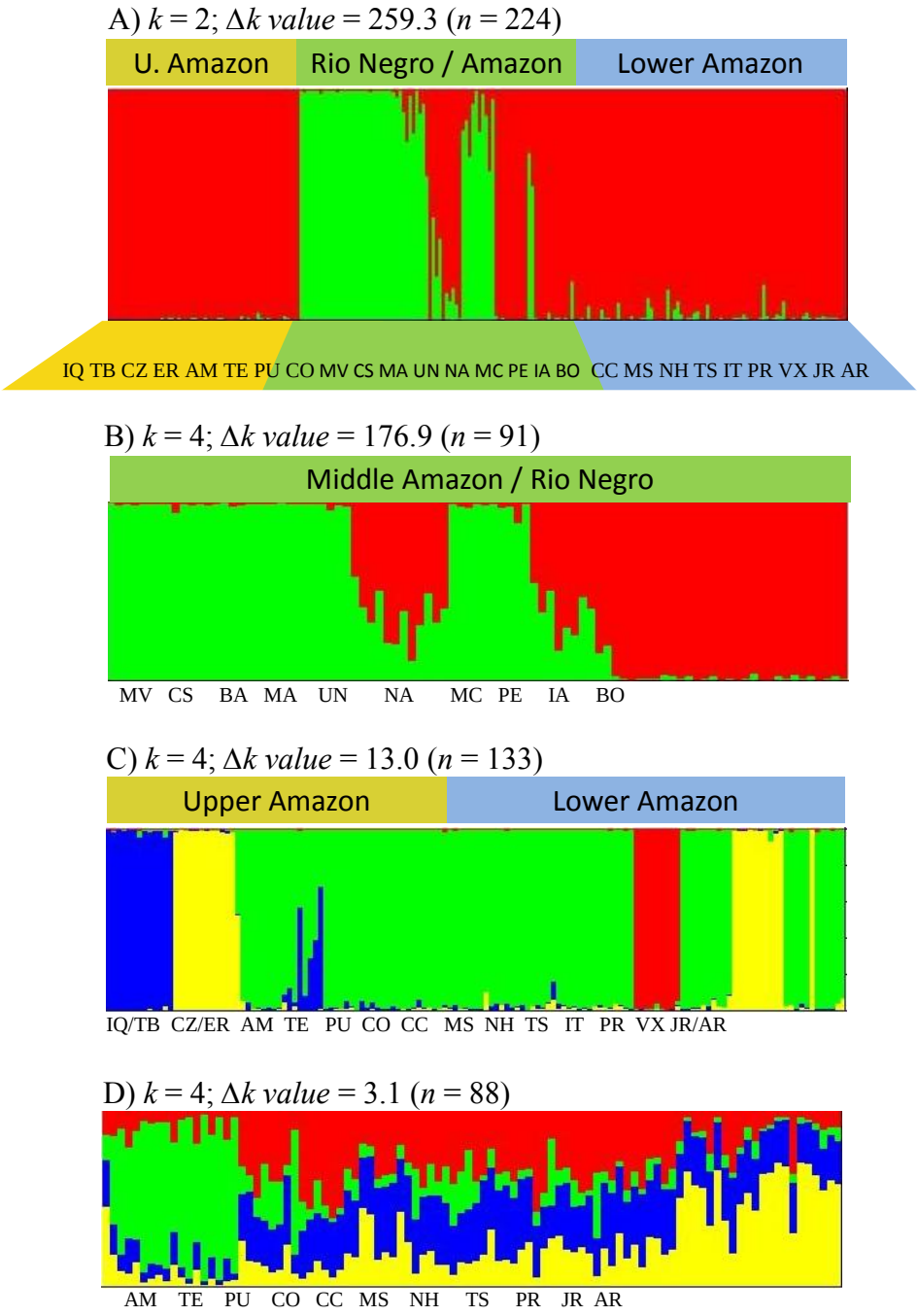

Figure 4.3 – The appointment of each individual genotype to different groups according to STRUCTURE (Pritchard et al. 2000) results for the highest value of *Δk* (Evanno et al. 2005) at various levels of population analysis. A) Strong population structure for k=2

between the Rio Negro and Amazon basin. B) A significantly lower *Δk,* but still relevant population structure based on sampling localities. C) Strong population structure of Rio Negro individual. D) A significantly lower *Δk,* but apparent panmixia for the main channel Amazon localities.

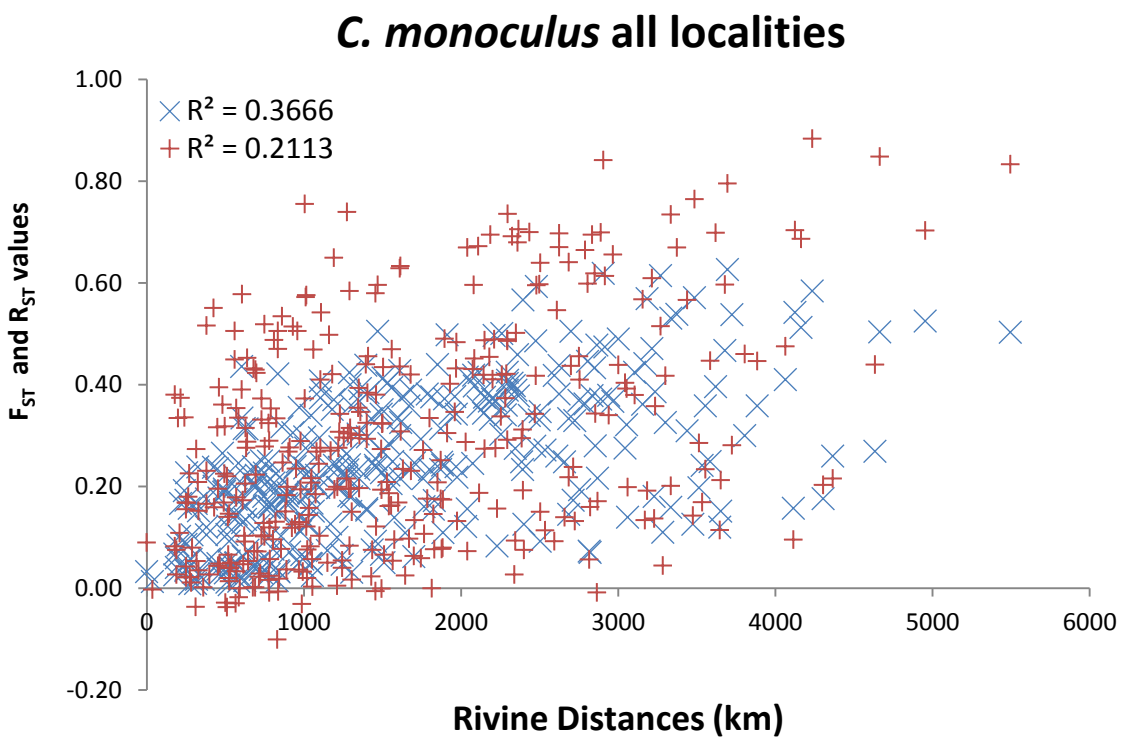

Figure 4.4 – A scatter plot showing the correlation between geographic distance and genetic distance for F_{ST} (+) and R_{ST}(×) values for all *C. monoculus* sampling localities.

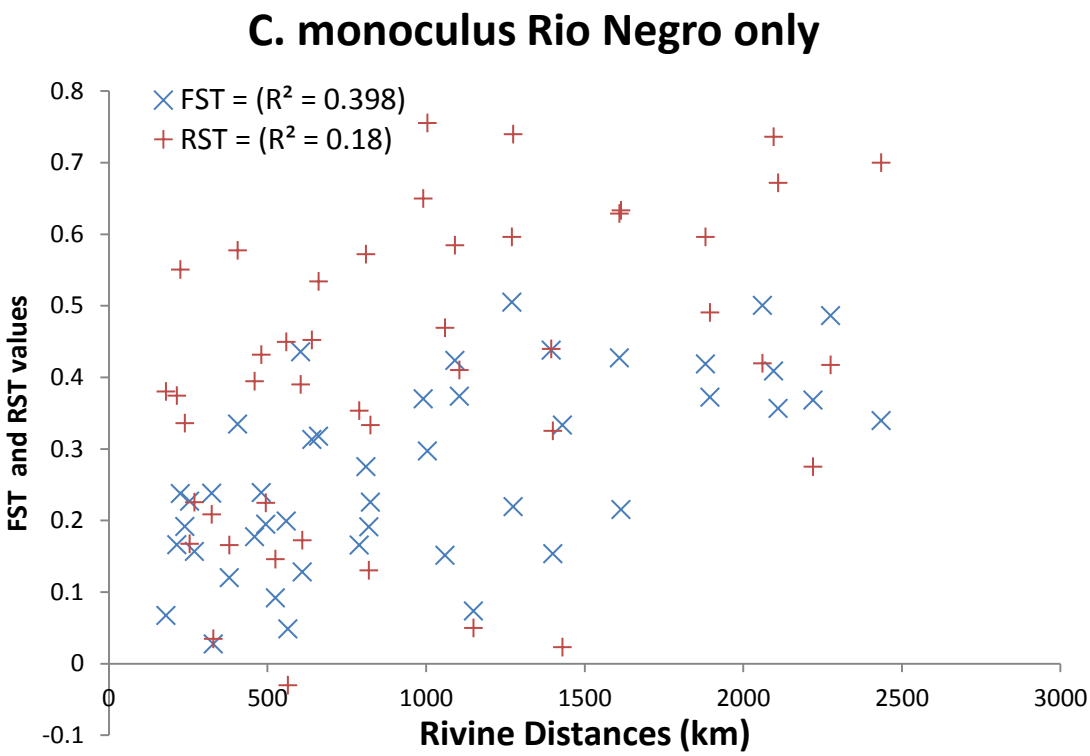

Figure 4.5 – A scatter plot showing the correlation between geographic distance and genetic distance for F_{ST} (+) and R_{ST}(×) values for *C. monoculus* found along the Rio Negro only.

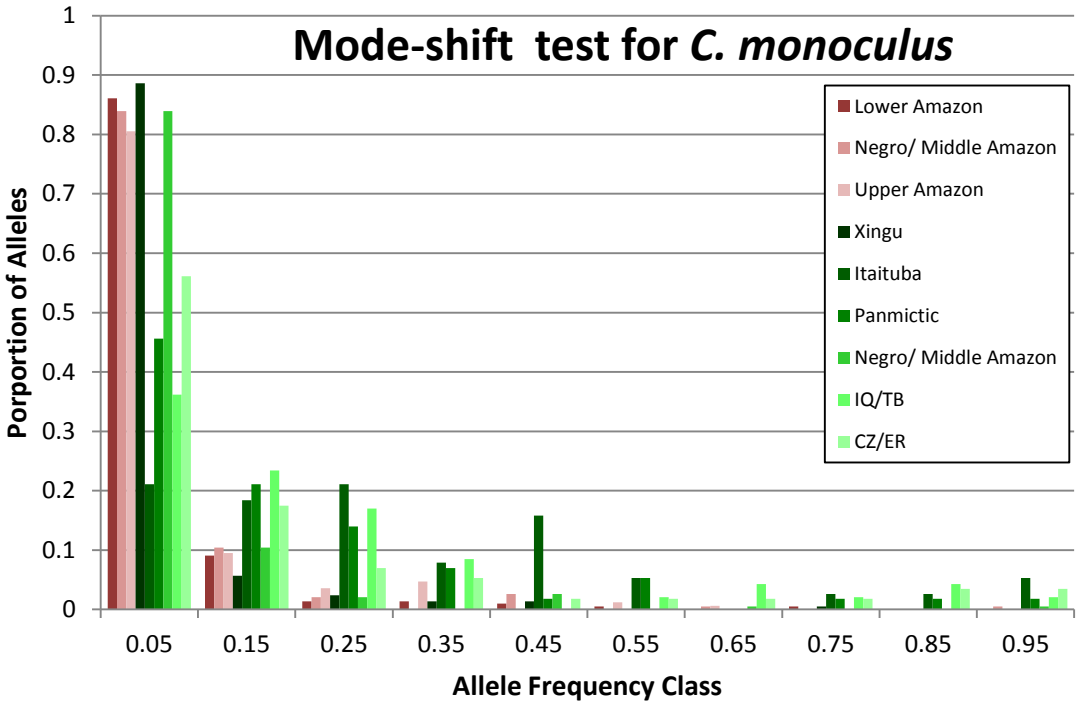

Figure 4.6 – A distribution of allele frequencies for microsatellite loci following the mode-shift test in the program BOTTLENECK.

Table 4.1. Estimates of F_{ST} (below diagonal) and R_{ST} (above diagonal) for microsatellites for all sample sites of *C. monoculus*.

	AR	JR	VX	PR	IT	TS	NH	MS	CC	BO	IA	PE	MC	NA	UN	BA	MA	CS	MV	CO	PU	TE	AM	ER	CZ	TB
AR	-----	**0.05**	**0.10**	**0.07**	0.28	**0.02**	**0.06**	**0.16**	**0.20**	0.32	0.42	**0.16**	**0.10**	0.31	0.18	0.41	0.60	0.51	0.76	**0.11**	**0.15**	0.13	0.29	**0.46**	0.44	**-0.01**
JR	0.14	-----	0.15	**0.03**	0.32	**0.00**	0.03	0.10	**0.04**	0.27	0.39	**0.00**	0.05	0.19	0.19	0.25	0.27	0.34	0.57	0.07	**0.00**	0.02	0.06	0.14	0.20	0.11
VX	0.31	0.15	-----	0.17	0.47	0.10	0.12	0.25	0.14	0.34	0.46	**0.22**	0.15	0.35	0.17	0.40	0.49	0.44	0.61	0.19	0.16	0.13	0.27	**0.17**	0.22	0.09
PR	0.15	0.04	0.17	-----	0.33	**-0.03**	0.04	**0.05**	**-0.01**	0.29	0.50	**0.08**	**0.00**	0.24	0.20	0.23	0.35	0.46	0.66	**0.04**	**0.02**	**-0.01**	**0.07**	**0.04**	0.09	0.09
IT	0.41	0.25	0.42	0.25	-----	0.32	0.36	0.33	0.37	0.51	0.54	0.13	0.37	0.21	0.29	0.23	0.30	0.44	0.61	0.19	0.30	0.38	0.38	0.36	0.48	0.29
TS	0.13	0.04	0.16	0.03	0.23	-----	**0.00**	**0.02**	**-0.04**	0.22	0.43	**0.01**	**-0.02**	0.21	0.15	0.20	0.27	0.37	0.60	**0.02**	**-0.01**	**-0.03**	0.08	**0.06**	0.11	**0.08**
NH	0.17	0.04	0.16	0.04	0.27	**0.01**	-----	**0.01**	**0.05**	0.14	0.42	**0.04**	**0.00**	0.28	0.18	0.28	0.32	0.42	0.64	**0.06**	**0.02**	**0.03**	0.22	0.16	0.21	**0.08**
MS	0.19	0.08	0.21	0.06	0.25	0.03	0.06	-----	0.08	0.23	0.51	**0.08**	**0.04**	0.32	0.28	0.27	0.35	0.49	0.71	**0.05**	**0.07**	0.08	0.28	**0.22**	0.28	**0.06**
CC	0.15	0.05	0.19	0.03	0.22	**0.01**	**0.02**	0.03	-----	0.33	0.52	**0.15**	**0.02**	0.27	0.18	0.27	0.42	0.48	0.69	**0.05**	**0.02**	**0.00**	0.10	0.15	0.20	**0.05**
BO	0.27	0.13	0.28	0.14	0.31	0.10	0.12	0.10	0.08	-----	0.38	0.39	0.17	0.55	0.43	0.57	0.58	0.60	0.74	0.35	0.31	0.29	0.50	0.55	0.57	0.10
IA	0.39	0.23	0.38	0.24	0.39	0.20	0.22	0.22	0.17	0.07	-----	**0.45**	0.45	0.58	0.53	0.65	0.60	0.42	0.42	0.52	0.49	0.50	0.58	0.66	0.70	0.21
PE	**0.28**	0.10	0.26	0.13	0.35	**0.05**	**0.04**	0.10	**0.09**	**0.18**	0.31	-----	0.15	**-0.03**	**0.13**	**0.05**	**0.02**	**0.28**	0.70	**0.02**	**0.02**	0.12	**0.19**	**0.41**	0.45	**0.00**
MC	0.14	0.05	0.18	0.02	0.24	0.01	**0.02**	0.05	**0.01**	0.12	0.20	0.09	-----	0.34	0.22	0.33	0.41	0.49	0.67	0.07	**0.04**	**0.03**	0.20	0.16	0.20	0.08
NA	0.33	0.18	0.30	0.20	0.34	0.16	0.16	0.19	0.17	0.24	0.33	**0.05**	0.19	-----	0.17	**0.03**	0.17	0.32	0.63	0.16	0.17	0.29	0.22	0.34	0.42	0.21
UN	0.31	0.22	0.33	0.22	0.36	0.18	0.18	0.20	0.17	0.24	0.32	0.19	0.19	0.23	-----	0.21	0.39	0.44	0.63	**0.07**	0.13	0.20	0.23	0.13	0.23	0.18
BA	0.37	0.23	0.34	0.23	0.38	0.20	0.20	0.22	0.19	0.27	0.37	0.07	0.23	0.03	0.24	-----	0.23	0.47	0.74	0.12	0.18	0.31	0.30	0.39	0.45	0.19
MA	0.59	0.38	0.50	0.38	0.50	0.34	0.35	0.39	0.35	0.42	0.50	0.33	0.37	0.13	0.44	0.16	-----	0.35	0.76	0.31	0.30	0.43	0.47	0.73	0.69	0.19
CS	0.61	0.38	0.49	0.37	0.50	0.34	0.36	0.39	0.34	0.42	0.50	0.37	0.37	0.15	0.44	0.15	0.17	-----	0.37	0.45	0.42	0.49	0.50	0.70	0.70	0.19
MV	0.57	0.36	0.47	0.37	0.47	0.33	0.35	0.38	0.33	0.41	0.49	0.34	0.36	0.22	0.43	0.22	0.30	0.17	-----	0.67	0.64	0.69	0.70	0.85	0.83	0.28
CO	0.16	0.08	0.23	0.06	0.24	0.05	0.04	0.07	**0.03**	0.15	0.23	**0.12**	0.06	0.20	0.20	0.22	0.40	0.40	0.38	-----	**0.00**	0.11	0.18	**0.07**	0.17	**0.10**
PU	0.16	0.07	0.21	0.03	0.22	0.03	0.03	0.05	**0.01**	0.13	0.21	0.10	0.04	0.19	0.18	0.21	0.37	0.34	0.33	**0.02**	-----	**0.03**	0.09	**0.07**	0.14	0.12
TE	0.22	0.11	0.25	0.09	0.30	0.07	0.06	0.10	0.04	0.16	0.25	0.13	0.06	0.22	0.18	0.24	0.41	0.40	0.38	0.10	0.05	-----	0.09	**0.08**	0.14	**0.04**
AM	0.12	0.06	0.22	0.05	0.25	0.05	0.05	0.07	0.03	0.15	0.22	0.11	0.04	0.20	0.18	0.23	0.39	0.39	0.37	0.05	**0.01**	0.03	-----	0.17	0.24	0.13
ER	**0.30**	0.13	0.24	0.11	0.39	0.07	0.07	0.14	0.09	0.26	0.36	**0.19**	0.08	0.28	0.33	0.32	0.53	0.54	0.50	0.16	0.13	0.17	0.11	-----	**-0.10**	**0.03**
CZ	0.27	0.18	0.26	0.16	0.41	0.12	0.15	0.20	0.16	0.31	0.40	0.25	0.14	0.33	0.36	0.36	0.51	0.53	0.50	0.22	0.18	0.25	0.17	**0.01**	-----	0.13
TB	0.48	0.27	0.42	0.27	0.42	0.25	0.24	0.29	0.25	0.36	0.44	0.38	0.24	0.40	0.38	0.42	0.57	0.57	0.54	0.29	0.24	0.17	0.24	0.40	0.44	-----
IQ	0.54	0.27	0.42	0.28	0.45	0.25	0.23	0.30	0.24	0.37	0.46	**0.40**	0.23	0.41	0.39	0.44	0.62	0.63	0.58	0.30	0.24	0.16	0.24	0.43	0.47	**0.01**

Note: Nonsignificant pairwise F_{ST} and R_{ST} (Excoffier et al. 2005) values are in bold.

Table 4.2 – Sign test and Wilcoxon two tailed test with 1000 replicates for *C. monoculus* with the population divided by both the mtDNA and microsatellite data. M'sat = microsatellite. (Note: Significant p-values are in bold. Parameters for the two-phase model were set according to the authors recommendations [variance (v) = 12 and proportion of SMM (p) = 95% (Piry et al 1999)]. When alternative v and p values changed from significant to non-significant they are noted by (*) for v=12 and p =70; and by (**) for v = 30 and p = 70 when only the Orinoco population under the sign test remained significant. When v = 30 and p = 95 there were no changes observed between significant and non-significant values).

					Sign test			Wilcoxon Test (two tails)		
				n	IAM	TPM	SMM	IAM	TPM	SMM
mtDNA	population	divisions	L. Amazon	75	0.43908	**0.00002**	**0.00001**	0.26611	**0.00024**	**0.00024**
			Negro	91	0.57093	**0.00003**	**0.00002**	0.85010	**0.00024**	**0.00024**
			U. Amazon	58	0.43874	**0.00041**	**0.00037**	0.09229	**0.00073**	**0.00073**
	Microsatellite	population divisions	Xingu	10	**0.04270**	0.46008	0.46247	0.12305	0.51953	0.63770
			Itaituba	9	0.31535	**0.03483***	**0.03915**	0.96582	0.08301	0.05371
			Panmictic	88	0.23736	**0.01769****	**0.00251**	0.09229	**0.00342**	**0.00244**
			Negro	91	0.42601	**0.00002**	**0.00002**	0.85010	**0.00024**	**0.00024**
			IQ/TB	13	0.34242	**0.00825**	**0.00910**	0.46484	**0.00488**	**0.00342**
			CZ/ER	13	0.18225	0.54418	0.55880	0.32031	1.00000	1.00000

CHAPTER 5 – GEOLOGICAL AND ECOLOGICAL INFLUENCES ON POPULATION STRUCTURE OF *CICHLA TEMENSIS* AND *C. MONOCULUS*: CONCLUSIONS

There have been multiple hypotheses proposed to explain the observed population structure for diverse Neotropical species. The combined mtDNA and microsatellite analyses for both *C. temensis* and *C. monoculus* highlight that their life history characteristics may be key factors influencing the level of gene flow throughout their distribution. High levels of genetic variation and panmixia were reported in tambaqui (*Colossoma macropomum*), suggesting high rates of gene flow due to dispersal for feeding and reproductive migration (Santos et al. 2007). High rates of gene flow were also observed by Sivasundar et al. (2001) in *Prochilodus*, which exhibit high rates of migration. Although these studies were based on the mtDNA control region, our combined mtDNA and microsatellite analysis depict strong population structure most likely reflective of their low dispersal rates and parental care for offspring (Zaret 1980; Hoeinghaus et al. 2003). Furthermore, the contradictory results between the mtDNA and microsatellite analyses in *C. monoculus* highlight potential incidence of male biased gene flow (Avise 2000) along their main channel population.

The results from these analyses strongly support that the contrasting water types are primarily responsible for the low levels of gene flow across both species distributions. These barriers to gene flow are apparent between the Rio Negro and the Orinoco River in *C. temensis* and between the Rio Negro and Amazon River in *C. monoculus*. Although there is a major ecological gradient between black water and white water, they often

occur adjacent to one another (Sioli, 1984; Winemiller et al. 2008) resulting in ecological heterogeneity within and between river basins. This has been proposed as an explanation for population structure in the genus *Symphysodon*, which exhibits population structure reflective of the ecological heterogeneity (Ready et al. 2006). The lack of gene flow most likely played a key role in speciation events as alternative selective pressures may act upon species spanning these contrasting environments. Ecological heterogeneity has been attributed to variation in coloration ultimately leading to reproductive isolation due to selective pressures (Seehausen et al. 1997). Future investigations should encompass species with populations overlapping with *C. temensis* and *C. monoculus*. Similar investigations with multiple taxa would allow for further testing of the influence contrasting water types have on population distributions.

In both species ecological heterogeneity seems to be the greatest influence to gene flow, however, in *C. temensis* the formation of the Casiquiare River played a key role in their current population distribution. The results reported here support previous hypotheses regarding the role the Casiquiare River plays on species as there is a strong contrast between ecological heterogeneity influencing species distributions (Winemiller et al. 2008) and the roll that the Casiquiare River plays in dispersal and range expansion (Willis et al. 2010). Despite evidence indicating geologic events influencing the population distribution of *C. temensis* there is no correlation between *C. monoculus* and major geologic events potentially influencing their distribution, specifically the breaching of the Purus Arch.

Comparative inferences between species in this genus would provide insight into potential areas that have experienced limited gene flow as well as panmixia throughout

the Amazon and Orinoco River basins. Although *C. monoculus and C. temensis* are not a focus of conservation, remote areas of the Neotropics are becoming more accessible and could impact these heavily-harvested food fish. It is necessary that we understand the level of genetic variation observed in this species before potential problems due to overfishing or anthropogenic barriers, such as dams, influence gene flow significantly decrease genetic diversity.

References

Avise, J.C. 2000. Phylogeography; the history and formation of species. Harvard University Press. Cambridge, Mass.

Hoeinghaus, D. J., C. A. Layman, D. A. Arrington, and K. O. Winemiller. 2003. Spatiotemporal variation in fish assemblage structure in tropical floodplain creeks. *Environmental Biology of Fishes*, 67: 379 – 387.

Ready, J. S., E. J. G. Ferreira, and S. O. Kullander. 2006. Discus fishes: mitochondrial DNA evidence for a phylogeographic barrier in the Amazonian genus *Symphysodon* (Teleostei: Cichlidae). *Journal of Fish Biology*, 69: 200 – 211.

Santos M.C.F., M.L. Ruffino and I.P. Farias. 2007. High levels of genetic variability and panmixia of the tambaqui *Colossoma macropomum* (Cuvier, 1816) in the main channel of the Amazon River. *Journal of Fish Biology*, 71 (Supplement A): 33 – 44.

Seehausen, O., J. J. M. van Alphen, and F. White. 1997. Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science*, 227: 1808 – 1811.

Sioli, H. 1984. The Amazon and its main affluents: hydrography, morphology of the river courses, and river types. *The Amazon: Limnology and landscape ecology of a mighty tropical river and its basin*. (ed. by H. Sioli), pp. 127 – 165. W. Junk, Dordrect.

Sivasundar, A.E., Bermingham, and G. Orti. 2001. Population structure and biogeography of migratory freshwater fishes (*Prochilodus: Characiformes*) in major South American rivers. *Molecular Ecology*, 10: 407 – 417.

Willis, S. C., M. Nunes, C. G. Montana, I. P. Farias, G. Orti, and N. Lovejoy. 2010. The Casiquiare river acts as a corridor between the Amazonas and Orinoco river basins: biogeographic analysis of the genus Cichla. *Molecular Ecology*, 19: 1014 – 1030.

Winemiller, K. O., H. Lopez – Fernandez, D. C. Taphorn, L. G. Nico, and A. B. Duque. 2008. Fish assemblages of the Casiquiare River, a corridor and zoogeographical filter for dispersal between the Orinoco and Amazon basins. *Journal of Biogeography*, 35: 1551 – 1563.

Zaret, T. M. 1980. Life history and growth relationships of *Cichla ocellaris*, a Predatory South American Cichlid. *Biotropica*, 12: 144 – 157.