Evolution and Ecology of Neotropical crocodiles

Jose Avila Cervantes

Department of Biology
McGill University, Montreal

August 2019

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

TABLE OF CONTENTS

Appendices

ABSTRACT

To understand the origin of biodiversity, it is fundamental to understand the evolutionary processes that give rise to new species. Genomic approaches have enhanced our understanding of the underlying mechanisms that contribute to speciation and divergence. This thesis presents three case-studies examining speciation, hybridization and divergence of crocodile populations using a genomic approach to understand the evolution and ecology of Neotropical crocodiles. Chapter one reassesses the phylogeny of *Crocodylus* and focuses on the global events may have driven the radiation and diversity of the genus *Crocodylus* worldwide, with emphasis on possible migrations from Africa to the Neotropics. I used only sequences of wild specimens of a highly variable mitochondrial marker and employed a Bayesian approach to estimate a rooted time-calibrated phylogeny with a relaxed molecular clock under different models of nucleotide substitution. I also used the fossil record and reconstructions of paleoclimatic, paleobathymetry and paleo oceanic currents in combination with the genetic data. I found the previously accepted hypothesis of a single dispersion event *Crocodylus* from Africa to the neotropics to be more complex, and there were multiple transoceanic dispersals from America to Africa and back. Chapter two evaluates a Caribbean insular population of *Crocodylus acutus* in Mexico previously reported to be pure. I used microsatellites, mitochondrial DNA, single digest Restriction site Associated DNA Sequencing (sRAD-seq) and demographic models to test the origin and purity of the population. I found that this insular population might be a hybrid population with two discrete genetic demes and restricted gene flow from coastal populations in the process

of speciation. Chapter three investigates the effects of the rise of the Central American Isthmus (CAI) and the closure of the Central American Seaway (CAS), on the divergence of Pacific and Caribbean crocodile populations. I used sRAD-seq and two different demographic model approaches. I found that the rise of the CAI and the climatic changes associated with it had no detectable effect on the divergence of crocodile populations. Instead, the divergence is coincident with the Last Glacial Maximum (LGM), another global climatic event. This thesis demonstrates that combining genomic-based approaches with demographic inferences, climatic events and the biology of the species contribute to the understanding of evolutionary history patterns of species divergence and speciation.

RÉSUMÉ

Pour comprendre l'origine de la biodiversité, il est fondamental de comprendre les processus évolutifs à l'origine de nouvelles espèces. Les approches génomiques ont amélioré notre compréhension des mécanismes sous-jacents qui contribuent à la spéciation et à la divergence. Cette thèse présente trois études de cas portant sur la spéciation, l'hybridation et la divergence de populations de crocodiles à l'aide d'une approche génétique. Le premier chapitre réévalue la phylogénie de Crocodylus et ses approches des événements mondiaux qui ont conduit au rayonnement et à la diversité du genre Crocodylus dans le monde entier, en mettant l'accent sur les migrations de l'Afrique vers les néotropes. J'utilise uniquement des séquences de spécimens sauvages d'un marqueur mitochondrial hautement variable et une approche bayésienne pour estimer une phylogénie enracinée dans le temps avec une horloge moléculaire détendue sous différents modèles de substitution de nucléotides. J'utilise également les reconstructions d'archives fossiles, paléoclimatiques, paléobathymétrie et courants paléo-océaniques en combinaison avec les données génétiques. J'ai trouvé l'hypothèse d'un seul événement de dispersion Crocodylus d'Afrique aux néotropes plus complexe, et il y avait de multiples dispersions transocéaniques d'Amérique-Afrique et retour. Le chapitre deux évalue une population insulaire des Caraïbes de Crocodylus acutus au Mexique déclarée pure. J'utilise des microsatellites, de l'ADN mitochondrial, du séquençage de l'ADN associé à un site de restriction (sRAD-seq) et des modèles démographiques pour tester l'origine et la pureté de la population. J'ai trouvé que cette population insulaire pourrait être une population hybride avec deux démes génétiques

distinctes et un flux de gènes restreint des populations côtières en cours de spéciation. Le chapitre trois étudie les effets de la montée en puissance de l'isthme centraméricain (IAC) sur la divergence des populations du Pacifique et des Caraïbes. J'ai utilisé sRAD-seq et deux approches de modèle démographique différentes. J'ai constaté que la montée du CAI et les changements climatiques qui y sont associés n'avaient aucun effet détectable sur la divergence des populations de crocodiles. La divergence coïncide avec le dernier maximum glaciaire (LGM), un autre événement climatique mondial. Cette thèse démontre que la combinaison d'approches basées sur la génomique avec des inférences démographiques, des événements climatiques et la biologie de l'espèce contribue à la compréhension des modèles historiques d'évolution de la divergence des espèces et de la spéciation.

AKNOWLEDGMENTS

Thanks to my parents for always supporting me in everything I do, even if it is catching crocodiles, for their knowledge and for the education I received.
Thanks to my mom for being there always to remind me that not everything is work and I should enjoy life.
Thanks to my dad for teaching me that perseverance and hard work will always take you where you want to go.
Thanks to my life companion Adriana for encouraging me to start, continue and end this PhD. For teaching me to be humble, listen to others and be patient, I will be forever grateful.
Thanks to Hans for pushing me to my limits and for teaching me to look for more than meets the eye and to aim high.
Thanks to my PhD companion these years: Christine, we had great times during the PhD.
Thanks to Owen McMillan for the support, advice and funding to the project.
Thanks to Pierre Charruau for the fun times in the field and for believing in this porject.
Thanks to Rogelio Cedeño for the samples from Yucatan and for receiving me as a family member.
Thanks to Rowan Barrett and Andrew Hendry for their input on my project.
Thanks to Carlos Arias for the support and advice to analyze next generation sequencing data.
Thanks to Martha Vargas for her support and patience while teaching me how to prepare libraries for sequencing.
Thanks to Myriam Venegas for the samples provided from Panama.
Thanks to Robyn for proof reading this manuscript, her support and for being a great friend, also for helping me to stay focused and calm.
Thanks to all the Larsson Lab members, especially Rui and Trina.
Thanks to Ananda, Jonathas, Goyo and Paulina for being there.

Special thanks to Jose Miguel “Gory” and Brandon “Carnalito” for their help on understanding how to code.
Thanks to my brave field assistants in Panama: Sebastian, Felipe and Mónica.
Thanks to Franklin Cuapio and Marcos A. Morales for their support on the field work.
Thanks to the park rangers in Coiba, Galeta and Banco Chinchorro.
Thanks to Enrique Martinez Meyer and Victor H. Reynoso Rosales for the support in Mexico City.
Thanks to Ancil G. for everything.
Thanks to Lolita for helping me to take breaks because you needed a walk to the park, and for sitting on my lap in my must desperate times.

Special thanks to funding and governmental agencies:

To the Crocodile Specialist Group for the IUCN-SSC Crocodile Specialist Group Student Research Assistance Scheme (SRAS) grant.
The Smithsonian Tropical Research Institution for the short-term fellowship.
To CONACYT for the funding through the scholarship 217923-313487, CVU-249562
To the NSERC-BESS program for providing financial support through most of my PhD.
To the Comision Nacional de Areas Naturales Protegidas (CONANP) and the Secretaria de Medio Ambiente y Recursos Naturales (SEMARNAT) for the permits and support to make this project possible.
To the Colleccion Herpetologica (ECO-CH-H) del Museo de Zoologoia de El Colegio de la Frontera Sur for facilitating me the access to their tissue bank.
To the Ministerio de Ambiente in Panama for the permits.
To the Secretaria de Educación Pública, for the complementary scholarship for the year 2014.
To the McGill-Integrated Quantitative Biology Initiative (IQBI) for the funding provided.
To the McGill Biology department for the Biology travel awards and the mobility award.
To the QCBS for the excellence awards for the field work.
To the Neo Program for the funding to develop my project in Panama.
To el Centro del Cambio Global y la Sustentabilidad (CCGS), for the funding to work in Banco Chinchorro.
To McGill T-PULSE for teaching me how to teach.

CONTRIBUTION TO ORIGINAL KNOWLEDGE

Chapter 1 uses previous studies and genetic sequences to reconstruct migrations of the order *Crocodylia* at a global scale. It focuses on the Trans-Atlantic migrations of crocodiles from Africa to America. Our study provides an original contribution since previous studies to reconstruct the phylogenetic history of the genus were based on samples from farms and zoos. Our study used only samples from wild populations expanding the genetic diversity previously found on crocodilian species. This increased our resolution and we were able to detect events that were only hypothesized in previous studies. We did not limit our inferences to the genetic data. We integrated geological and climatic events, an extensive review of the fossil record, ocean currents and the use of sequences from wild populations to reassess the phylogenetic history of the order *Crocodylia*.

Chapter 2 is the result of a collaboration of many institutions and researchers in Panama, Mexico and Canada. With the integration of samples from many regions of Mexico and Panama. This is also one of the few studies that uses multiple genetic markers, including next generation sequencing, made on crocodiles. The original contribution of this study was to demonstrate the extent of a reported hybrid zone in the Yucatan Peninsula. With the integration of different genetic markers, demographic models and ocean currents we rejected the concept of the last American crocodile pure population in the Caribbean of Mexico and the persistence of pure C. Moreletii in the Gulf of Mexico. Instead we found that it has two isolated pure hybrid populations with

limited gene flow. These populations are the perfect setting to make further research on fitness, chromosomal rearrangements and Introgression of established hybrids

Chapter 3. There are many studies about the effects of the rise of the Central American Isthmus (CAI) on the divergence of marine organisms, and the migrations of species from South America and North America. This was an event with effects at a global scale. In this study we used next generation sequencing to evaluate the effects of the rise of the Isthmus on the divergence of crocodile populations of the Caribbean Sea and the Pacific Ocean. The original contribution of this study was to demonstrate that not all species were affected by the rise of the (CAI). The species populations diverged, but not as a consequence of the rise of the CAI. With the use of demographic models, we discovered that divergence is much more recent than previously thought and is related to the Last Glacial Maximum. Crocodiles had all the adaptations to resist the effects of the rise of this land bridge, which other species did not.

CONTRIBUTION OF AUTHORS

Chapter 1 is co-authored with Hans Larsson. The hypothesis for this chapter was conceived by Hans and me. I had the idea of finding global patterns of crocodile species distribution and then relating these patterns to geological events, ocean currents and fossils. I performed all the phylogenetic analysis. Hans and I both contributed to the preparation of the text and the figures. Some of the samples and ideas used in this chapter were provided by Pierre Charruau and Rogelio Cedeño

Chapter 2 is co-authored with Hans Larsson, Owen McMillan, Pierre Charruau and Rogelio Cedeño. The motivation for this chapter was to find patterns of crocodile radiations and dispersion on a local scale. I collected some of the tissue samples and the others were provided by my co-authors. I performed all the laboratory work from, which included DNA extraction and genomic library preparation. I conducted all the bioinformatic analyses, statistical analyses, and I wrote the text and prepared the figures. Hans provided feedback on the manuscript. Pierre Charruau, Rogelio Cedeño provided the samples, logistics for the field work and feedback on the writing.

Chapter 3 is co-authored with Hans Larsson and Owen McMillan. The idea for this chapter came up as a way to investigate the potential effect of the Rise of the Isthmus of Panama on crocodilian evolutionary history. Also, to test the potential anthropogenic effect of the opening of the Panama Canal on crocodile populations. I performed all the laboratory work from DNA extraction to genomic library preparation. I conducted all the bioinformatic analyses, statistical analyses, wrote the text and prepared the figures. Hans reviewed and corrected the manuscript.

Illustration of “El Coco” by Adriana Aranda Cuevas.

INTRODUCTION

In the literature review, I discuss the current knowledge about species and speciation, the advantages of using next-generation sequencing techniques to elucidate patterns of divergence in non-model species, the origin of Neotropical crocodiles and the current knowledge about the species studied in this thesis.

Chapter 1: the goal of this chapter is to reassess the biogeography of *Crocodylus* using mitochondrial DNA. Although several previous studies have used this molecular source, we restrict our samples to only wild, geo-referenced populations and increase the number of samples from the Neotropics. Previous analyses have relied heavily on captive population samples and limited sampling within species. The pervasive hybridization between *Crocodylus* species may have confounded these prior analyses as clade support has been typically low. We also examine how geological and climatic events may have shaped the distribution of *Crocodylus*.

Chapter 2: the objectives of this chapter are: a) to evaluate whether the population of Banco Chinchorro is a pure *C. acutus* population as reported by previous studies; b) to understand the effects of ocean currents on directing the gene flow of crocodiles in the Caribbean and Mexico, since crocodiles can ride these currents to disperse; and c) to deduce the origin of Banco Chinchorro crocodiles. To address these objectives, I used microsatellites, mitochondrial DNA, and next-generation sequencing and demographic models.

Chapter 3: the objectives of this chapter are: a) to evaluate the effect of the closure of the Central American Isthmus (CAI) on the emergence of two divergent

lineages of crocodiles in the Caribbean and the Pacific; b) to estimate the time of divergence of Pacific and Caribbean crocodiles; and c) to assess the effect of the Anthropocene, with a focus on the opening of the Panama Canal as a pathway to facilitate a secondary contact of crocodiles on both sides of the Isthmus. To address these objectives, I use sRAD-seq and demographic models.

In the general discussion I place my novel findings in the context of eco-evolutionary dynamics. I discuss whether the radiation of *Crocodylus* was adaptive or non-adaptive, whether the speciation and divergence of the genus was ecological or not, and what forces have shaped the distribution of extant neotropical crocodiles?

LITERATURE REVIEW

Species and Speciation

The elusive question of what is a species has remained central to current biological research. Ernst Mayr (1963) considered 3 major species concepts and discussed the limitations of them. The Typological or Essentialist concept implies species can be recognized by their essential nature or essential characters, and it is reflected in their morphology. This is also known as the morphological species concept. But variation due to sexual dimorphism, age, polymorphisms and individual variation have made this concept all but abandoned. The Nominalistic concept states that species are defined by humans as a means to name and classify them. However, species exist as a consequence of evolution and are not simply a product of the human imagination, thus making this concept filled with subjective opinion. The Biological Species concept has withstood the greatest amount of scrutiny. It states that species are groups of interbreeding natural populations that are reproductively isolated from other such groups. For this definition, species had three fundamental properties 1) a reproductive community, 2) ecological unit, and 3) genetic unit.

The debate about species concepts and the purpose of defining species depends on the nature of the species under study. Coyne and Orr (2004), consider that the purpose of species concepts should fulfill five goals: 1) systematic classification, 2) correspondence to discrete entities in nature, 3) understanding how entities arise, 4) representative of the evolutionary history of organisms, and 5) universal, applied to a large number of organisms. The problem of species concepts relies on the fundamental

questions about the organism under study, and different species concepts may result in different species allocations. Additional species concepts include the Genotypic cluster species concept (Mallet, 1995). This concept requires that species are distinguishable groups of individuals with a few or no intermediates when in contact, based on genetics and morphology. Van Valen (1976) introduced the Ecological species concept, where a species is a lineage (or a closely related set of lineages) that occupies an adaptive zone minimally different from that of any other lineage in its range and which evolves separately from all lineages outside its range. Wiley (1978) introduced the Evolutionary species concept where a species is a lineage of ancestral descendant populations that maintains its identity from other such lineages, and which has its own evolutionary tendencies and historical fate. Finally, the Phylogenetic Species concept was introduced by De Queiroz and Donoghue (1988). This concept requires that each species is the smallest (least inclusive) monophyletic group. The phylogenetic species concept can take into account molecular and phenotypic data and is scale independent in that individual populations or groups of populations can function as species.

All these species concepts rely on lineage sorting, reproductive isolation, morphological differences and genetic differences, but what happens when these conditions can not be fulfilled, because the species are in the process of speciation? Nosil et al. (2009) presented a way to explain incomplete speciation from the perspective of divergence as a process that varies continuously. In such, some species concepts are part of stages along the speciation continuum. Hendry et al., 2009, proposed 4 states of the speciation continuum: State 1, continuous variation without

reproductive isolation; State 2, partially discontinuous variation within or between groups with partial reproductive isolation; State 3, discontinuous variation with strong but reversible reproductive isolation; and State 4, complete and irreversible reproductive isolation. Shaw and Mullen (2014) defined “speciation continuum” from a genetic perspective, as a continuous sequence of genetically based changes that occur as two lineages diverge from one another on the way to reproductive isolation. The reproductive isolation barriers are characteristics of organisms that keep individuals in one population from exchanging genes with other populations. Reproductive isolation can occur by preventing individuals of separate species from mating (premating isolation) or by selecting against hybrids (postmating isolation) (Losos, 2014). Premating isolating barriers can be on behavioral, ecological, mechanical nature or related to the mating system. On the other hand, postmating prezygotic isolation can be behavioural or gametic, whereas postmating postzygotic barriers are related to hybrid sterility and inviability (Coyne & Orr 2004).

But how has the species concept changed in the genomics era? Genomic information gathered from model and non model species has helped us understand that ancient alleles with pleiotropic effects characterize sympatric and allopatric divergence and were often acquired by interspecific hybridization (Seehausen et al., 2014). Now we can compare sister species or races that are in the early stages of divergence with the use of whole genome data, transcriptomes and/ or with thousands of single nucleotide polymorphisms (Harrison & Larson, 2014). Studies of hybrid zones using genomic data, suggest that mechanisms of reproductive isolation are polymorphic, or context dependant and that differences between hybrid fitness and reproductive isolation are

probably not controlled by single ecological factors, but rather by interacting ecological characteristics dependant on the location (Gompert et al., 2017). To study hybrid zones, it is important to consider a definition of hybridization for cases where outcrossing and gene flow occur between populations that differ at multiple heritable characters or genetic loci that affect fitness (Gompert & Buerkle, 2016). Hybridizing species have proportions of the genome resistant to introgression, suggesting that the species boundaries vary geographically and are context dependant, as a result of heterogeneous environments (Harrison & Larson, 2014). Translation of patterns of genetic variation to inferences about hybridization requires models that incorporate gene flow, selection acting on hybrid populations and recombination rates (Payseur & Rieseberg, 2016).

RAD sequencing and applications

For decades, population genetics research relied heavily on microsatellite markers. These markers have been used to detect population structure, population differentiation (F-statistics and D), migration, population diversity, kinship and effective population size (Putman & Carbone, 2014). However, microsatellites present size homoplasy and null alleles, which can affect the estimation of population parameters (Putman & Carbone, 2014). In contrast, single nucleotide polymorphisms (SNP), which have a simple character state (biallelic) and can now be sequenced in large numbers, present a novel and actual trend for genetic studies for non-model species.

The development of restriction site-associated DNA sequencing (RADseq) has been an important scientific breakthrough in the past decade. As a response to the

development of RADseq, we have seen an increase in studies of the genomics of adaptation, inbreeding and genomic diversity, effective population size, population structure, phylogeography, introgression and phylogenomics, that harnessed the power of next-generation sequencing to retrieve hundreds to thousands of polymorphic genetic markers across the genome (Andrews et al., 2016). For this study, we used the original RAD because of its low cost, the length of the loci are ± 300 bps, the PCR duplicates can be identified with paired end sequencing, it is appropriate for large genomes, and it has been proved useful for de-novo identification of loci (Andrews et al., 2016).

This technique relies on the digestion of Genomic DNA with a restriction enzyme and an adapter (P1) ligated to the fragment's overhanging ends. The P1 adapter contains forward amplification and Illumina sequencing primer sites, as well as a nucleotide barcodes 4 or 5 bp long for sample identification. The adapter-ligated fragments are subsequently pooled, randomly sheared, and size-selected. End repair and A-tailing is performed on the DNA before ligation to a second adapter (P2) that has divergent ends in a Y shape. The structure of this adapter ensures that only P1 adapter ligated RAD tags will be amplified during the final PCR amplification step (Baird et al., 2008). This process of adding adapters not only ensures high quality DNA amplification and sequencing, but also allows for multiple individuals to be sequenced concurrently. These innovations have made RADseq attractive for population studies, especially in non-model species.

Origin of Neotropical crocodiles

The study taxon used here is *Crocodylus*. This genus is a diverse clade within the extant Crocodylia. Crocodylia, in turn, lie at the end of a long lineage of Archosauria with the earliest records from the Early Triassic (Butler et al., 2011). Although the sister lineage, Avemetatarsalia, which includes pterosaurs, dinosaurs, and birds dominated Mesozoic terrestrial ecosystems and today's aerial niches, the crocodile lineage, Pseudosuchia, enjoyed similar success in terrestrial, amphibious, and aquatic environments (Brusatte et al., 2010; Nesbitt, 2011). The pseudosuchian clade Crocodyliformes had an exponential diversification through the Cretaceous and Paleocene, surviving the end Cretaceous mass extinction. Their first records of declines are during the Eocene and Plio-Pleistocene during the initiation of large-scale global cooling (Markwick, 1998; Bronzati et al., 2015). The long-term decline is strongly correlated with the decrease in temperature in the high latitudes, while lower latitudes witnessed increased aridification, sea-level changes and hydrographic rearrangements (Mannion et al., 2015). Even during this decline, Miocene crocodylian diversity in low latitudes in South America was higher than in temperate regions (Mannion, 2015). The fossil record in the Miocene of tropical and subtropical South America has more than 26 crocodyliform species presenting a variety of snout shapes adapted to different feeding behaviours (Salas-Gismondi et al., 2015; Scheyer et al., 2013). The eventual decline of this hyperdiverse crocodylian community is correlated with the Andean uplift and hydrological changes associated with this event (Salas-Gismondi et al., 2015; Salas-Gismondi et al., 2018; Scheyer et al., 2013). In spite of this high diversity, none of these fossil crocodiles were ancestral to the extant Neotropical *Crocodylus*.

The origin of the genus *Crocodylus* is dated to 13.6 - 8.3 Mya, with a rapid global radiation throughout the tropics and subtropics in a time span of ± 6 My (Oaks, 2011). The genus originated in the Indopacific, radiated to Africa 7.8 - 12.3 Mya, crossed the Atlantic Ocean and arrived in the Neotropics 4 - 6.3 Mya (Meredith et al., 2011; Oaks, 2011). Of this initial radiation, only four Neotropical species of *Crocodylus* are alive today: 1) Cuban crocodile (*C. rhombifer*), with a restricted distribution to Cuba (Targarona et al., 2018), 2) Orinoco crocodile (*C. intermedius*) restricted to the Orinoco basin (Balaguera-Reina et al., 2018), 3) Morelet's crocodile (*C. moreletii)* restricted to the Gulf of Mexico, the Yucatan Peninsula, Guatemala and Belize (Cedeño-Vázquez et al., 2012), and 4) The American crocodile (*C. acutus*) widely distributed in the Pacific and Atlantic coast and the Caribbean islands (Ponce et al., 2012). Although once widespread, all are now listed in CITES Appendices I and II, with many populations at high risk of extinction, initially due to overhunting, and currently because of habitat loss. Understanding the factors that shaped the distribution of modern crocodiles in the past can be used to predict the future of the species and design conservation strategies.

Study Species

American crocodile (Crocodylus acutus)

The American crocodile (*Crocodylus acutus*) generally inhabits coastal environments and is broadly distributed from the southeastern tip of Florida to the limits of mangrove habitat in northern Peru. Even though it is frequently found in mangrove-lined coastal lagoons or estuaries, it also inhabits a variety of environments ranging from hypersaline lakes to freshwater rivers and reservoirs (Ernst et al., 1999). American

crocodiles are also known to inhabit cays and coral atolls (Thorbjarnarson, 1989). The species was hunted for its skin from 1920 to 1970 leading to population declines. The loss of habitat related to tourism also endangered the species (Mazzotti, 1999). Thorbjarnarson and collaborators (2006) proposed eight Crocodile Conservation Units (CCU's) considering the habitat quality, nesting habitat, population size, connectivity, habitat destruction, the potential for sustainable use, killings of crocodiles and the percentage of the area under protection. These CCU's are informative because they consider many factors to rank its conservation priority.The Yucatan Peninsula ranked as highest to high priority for conservation. Since 2006, no studies have been made to assess the actual status of individual populations. The species is listed in the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) in Appendix I as most endangered and in the International Union for Conservation of Nature (IUCN) as vulnerable (VU) (IUCN, 2019).

Morelet's crocodile (Crocodylus moreletii)

The distribution of *Crocodylus moreletii* is restricted to the Gulf of Mexico, in coastal Yucatan Peninsula, Guatemala and Belize (Cedeño-Vázquez et al., 2012). This species was extensively hunted for skin trade in the 1960s until the government of Mexico protected the species and prohibited its trade (Sánchez Herrera et al., 2011). The estimated population size is of 100,000 individuals with approximately 20,000 reproductive adults. In 2010 in the conference of parts of CITES, the species was moved from Appendix I to II (Sánchez Herrera et al., 2011). Within the IUCN red list, the species is listed as least concern (LC) (IUCN, 2019). After the evaluation of *C. moreletii*

populations and the change of Appendix, a ranching program in Mexico was proposed (Barrios-Quiroz & Cremieux, 2018). The species in Belize has recovered from the over-harvesting and is not considered as threatened anymore (Platt & Thorbjarnarson, 2000). There is no updated information about the status of the populations in Guatemala.

Hybridization in Crocodylus

Hybridization between *C. siamensis* and *C. porosus,* and *C. siamensis* and *C. rhombifer* has been reported in captivity (Fitzsimmons et al., 2002). It has also been reported to happen in between *C. rhombifer* and *C. acutus* in captivity (Weaver et al., 2008), and in the wild in Cuba (Milian-Garcia et al., 2015). Hybrids of *C. acutus* and *C. moreletii* with behavioural and ecological differences have been reported in Belize (Hekkala et al., 2015).

A hybrid zone of *C. acutus* and *C. moreletii in* the Yucatan Peninsula has been under study for the last ten years (Cedeño-Vázquez et al., 2008; Machkour-M'Rabet et al., 2009; Pacheco-Sierra et al., 2016; Pacheco-Sierra et al., 2018; Rodriguez et al., 2008). The First studies suggested that hybridization was promoted by anthropogenic activities (Cedeño-Vázquez et al., 2008), yet recent studies demonstrate that hybridization is a thousand-years natural process (Pacheco-Sierra et al., 2016; Pacheco-Sierra et al., 2018). Two island populations off the Yucatan peninsula have been suggested to be the last refugees of pure *C. acutus* populations in the Caribbean of Mexico (Machkour-M'Rabet et al., 2009; Pacheco-Sierra et al., 2016; Pacheco-Sierra et al., 2018). However, significant limitations of these studies are that they did not use

pure *C. acutus* populations from the Caribbean of Central America to assess how pure these island populations really are.

Three species concepts have been used for the genus *Crocodylus*: the genotypic cluster (Mallet, 1995), the evolutionary species (Wiley, 1978) and the phylogenetic species (De Queiroz & Donoghue, 1988). The genus does not fulfill the Biological species concept (Mayr, 1969) because there does not appear to be any reproductive isolation. Studies of the genus in the Neotropics indicate that the speciation process might be in the State 2 of the speciation continuum, with partially discontinuous variation within or between groups with partial reproductive isolation (Hendry et al., 2009). With the use of next generation sequencing, the present thesis intends to give a better understanding of ecological and evolutionary processes that shape speciation in crocodiles.

CHAPTER 1

Multiple Tans-Atlantic dispersals of *Crocodylus* established African and Neotropical species.

Abstract

Previous phylogenetic relationships of extant crocodiles were analyzed with molecular methods and fossil calibrations. However, most of these studies were based on sequences from individuals that came from zoos, collections and farms. These results present a haphazard sampling error because they do not represent wild genotypes. Samples from wild populations are fundamental to reconstruct dispersions at a fine scale and relate them to geological and climatic events. In this study, we used D-loop sequences of wild crocodiles from extant species to reconstruct the phylogeny of *Crocodylus*. We also used the fossil record, geological and climatic events to reconstruct the dispersion of neotropical crocodiles. Our results are consistent with previous studies, showing the origin of the group in the Indo Pacific in the late Miocene a radiation to Africa and a westward radiation to the Neotropics. However, our results also indicate potential cryptic species in Australasia and unique lineages of *C. acutus* in the Neotropics. *C. niloticus* is strongly nested within the Neotropical radiation implying an eastward transAtlantic dispersal for this lineage. The novel paleobiogeography our study revealed a more complex history of the genus than previously suspected.

Introduction

The evolutionary history of *Crocodylus* spans much of the tropical world. Although the origin of the genus has been timed to be between 13.6 - 8.3 Ma with molecule-based estimates (Oaks, 2011). Well-supported scenarios of the evolutionary biogeography of the genus differ significantly, from a westward migration from Australasia to Africa and finally a trans-Atlantic crossing to the Americas, to an Australasian origin with a trans-Pacific crossing to the Americas with a subsequent trans-Atlantic crossing to Africa (Oaks, 2011). Molecular-based timing of these radiations place the radiation from Australasia to Africa between 8.3 - 5.3 Ma and a radiation to the New World from Africa between 5.3 - 3.0 Ma (Oaks, 2011) (Fig. 1) or 7.3 Ma (Pacheco-Sierra et al. 2018). The overlapping dates imply a relatively rapid global colonization. Confounding these scenarios is the African provenance of the sister clade to *Crocodylus*. The extant *Mecistops* and *Osteolaemus,* and all fossil representatives of Osteolaminae, are all African (Shirley et al., 2013).

The living *Crocodylus* of Africa are split into two species, *C. niloticus* and *C. suchus* (Hekkala et al., 2011). Both species used to range across Africa, but all specimens collected after 1975 have a general geographic separation. *C. suchus* is present in isolated populations in West Africa, Chad, Central African Republic, Republic of Congo, and Democratic Republic of Congo. *C. niloticus* is present throughout East Africa, Madagascar, and an isolated population in Gabon. Phylogenetic analysis of mitochondrial and nuclear DNA fragments maintains the consensus of the African *Crocodylus* species as a grade sister to the Neotropical species, with *C. niloticus* sister to the Neotropical clade (Figure 1) (Meredith et al., 2011; Oaks, 2011). Some analyses

did recover alternative phylogenetic topologies, including a sister-relationship of *C. niloticus* and *C. moreletii* (McAliley et al. 2006). The consensus grade of African *Crocodylus* sister to the Neotropical clade generally supports an African origin for the latter.

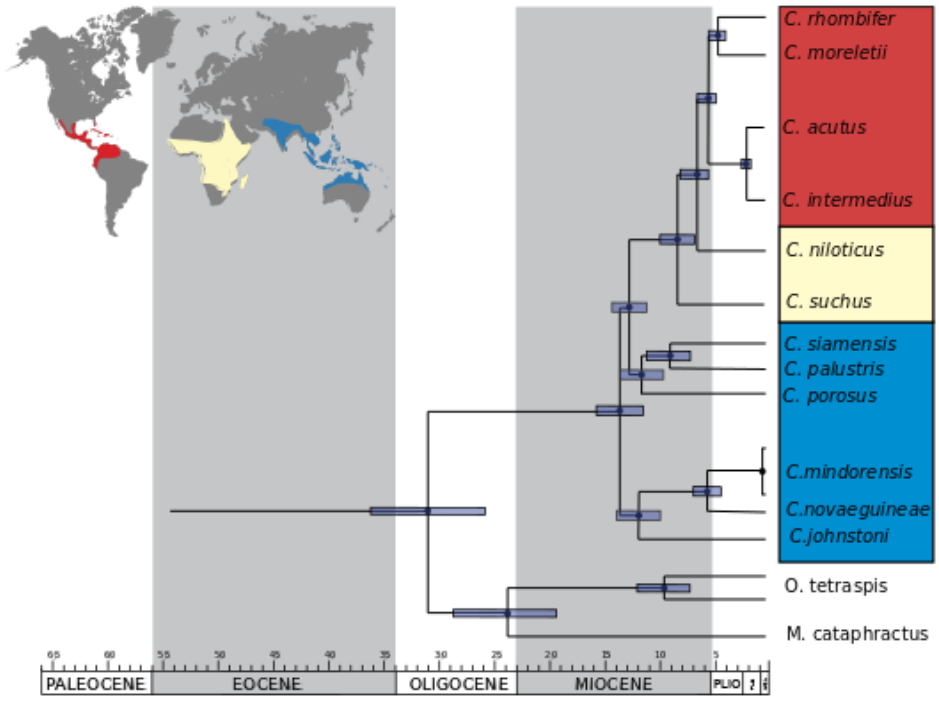

Figure 1. Majority-rule consensus tree from the BEAST posterior sample of species trees. Bars at nodes represent the 95 % highest posterior density (HPD) of the node age. Modified from Oaks (2011).

The fossil record offers some support for an Australasian origin of *Crocodylus* with a westward expansion. The oldest *Crocodylus* fossils are *C. palaeindicus* found in Late Miocene and Pliocene beds of Pakistan and India, respectively, with the oldest specimens in the Tortonian aged Chinji Formation (14.2 - 11.2 Ma) in northern Pakistan (Hussain & West, 1979; Barry et al. 2002). The oldest evidence of *Crocodylus* in Africa

is about as old and comes from the Late Miocene Nawata Formation of Kenya and the Sahabi sequence of Libya (Brochu & Storrs 2012). In both units, fossils attributable to *C. checchiai* Maccagno 1947 are well represented and both deposits are approximately 7 Ma (McDougall and Feibel, 2003; Beyer, 2008). Notable morphologies of *C. checchiai* are the presence of a rostral boss and a lacrimal lacking a posterior maxillary process on its anterolateral margin. The former is otherwise only present in extant Neotropical *Crocodylus* and the latter present in the Neotropical *C. intermedius* (Brochu & Storrs, 2012). *Crocodylus anthropophagus* is from the Pleistocene aged Middle Bed I of the Olduvai Gorge, dated to 1.84 Ma (Brochu et al. 2010). Fossils more closely allied to the extant African *Crocodylus* include *C. niloticus kaisensis* from the Pleistocene of Uganda (Swinton, 1926). Several fragmentary specimens referred to *Crocodylus* sp. have been identified from Tortonian and Messinian to Pliocene aged (11.6 - 5.3 Ma) deposits in Italy (Delfino et al. 2007, Delfino and Rossi 2013). Of particular importance is that the younger material shows evidence of a medial dorsal boss over the snout, similar to the boss present in *C. checchiai* and extant *Crocodylus* in the Americas. This similarity led Delfino et al. (2007) to suggest the Mediterranean crocodiles may have been an ancestral stock for a trans-Atlantic dispersing lineage to the Americas.

The earliest evidence of the immigration of *Crocodylus* to the New World comes from *Crocodylus falconensis* (Scheyer et al. 2013) recovered from the early Pliocene section of the San Gregorio Formation of Venezuela and a fragmentary, indeterminate *Crocodylus* from the Late Pliocene Lower Ware Formation (Moreno-Bernal et al., 2016). The former locality would make this taxon no older than approximately 5 Ma and the latter no younger than 3 Ma. *C. falconensis* exhibits the nasal boss present in many

Neotropical *Crocodylus* (Scheyer et al. 2013). This taxon has also been recovered as the sister taxon to all other Neotropical *Crocodylus* (Scheyer et al., 2013) or in an unresolved polytomy with these taxa (Salas-Gismondi et al., 2018).

A westward expansion fits well with molecular and morphology based phylogenetic hypotheses. However, Oaks (2011) noted that an eastward radiation from the Indopacific to the Americas and finally to Africa was less supported scenario. The relatively young dates suggest the lineage radiated throughout the tropics via large-scale oceanic migrations. Such marine migrations are not farfetched because the genus *Crocodylus* has adaptations to hyperosmotic environments such as lingual salt-secreting glands (Taplin and Grigg 1981; Taplin et al. 1982; Taplin 1988), a heavily keratinized buccal epithelium (Taplin and Grigg, 1989), and an osmoregulatory cloaca (Pidcock, Taplin, & Grigg, 1997). The saltwater crocodile (*C.porosus*) has been recorded travelling from 200 up to 590 kilometers offshore Australia in less than a month (Campbell et al., 2010) These individuals rode surface water currents for long distance travel as a low energy cost dispersal strategy (Campbell et al., 2010).

The continental scale radiations of *Crocodylus* thus occurred during the Miocene and potentially early Pliocene. This time witnessed several global changes that included global warming, cooling, and a mass extinction event. A global warming event, called the Mid-Miocene climate optimum, occurred about 18 Ma. Mean annual temperatures (MAT) were estimated at about 18.4°C, around three degrees higher than the current MAT (Ribbe et al., 2009). At about 14 Ma, this was followed by the Middle Miocene disruption that was manifested as a terrestrial and marine extinction event. This event is also associated with the beginning of global-scale cooling that continued to recent

times. The cooling trend accelerated approximately 8 Ma and the cooler temperatures with polar ice cap expansion caused a sea level decrease of over 50 m between 5 and 2 Ma (Hansen et al. 2013).

This backdrop of climatic variation is associated with multiple faunal and floral biogeographical changes. The majority of studies assessing the effect of this climatic change were on mammals (Figueirido, Janis, Pérez-claros, Renzi, & Palmqvist, 2012; Madern & Hoek, 2015). There are reports of extinctions of ectothermic vertebrates in Europe, including gavialoids and *Dyplocynodon* (Böhme, 2003). A latitudinal worldwide range contraction of marine and terrestrial crocodylians occurred during this time (Mannion et al., 2015) as well as a diversification of Crocodyloidea (Markwick, 1998; Scheyer et al., 2013).

One possible victim of the rapid drop in global temperatures is the extinction of gavialoid crocodylians from South America. Gryposuchine gavialoids radiated throughout the continent from at least the Oligocene, based on the Greater Antilles gavialoid *Aktiogavialis puertoricensis* (Vélez-Jarbe et al. 2007) and survived to at least the Late Miocene with *Piscogavialis jugaliperforatus* Kraus 1998 from the Pisco Formation of Peru and several taxa from the Urumaco Formation of Venezuela (Salas-Gismondi et al. 2018). A recent phylogenetic analysis places *Argochampsa krepsi* (Hua and Jourve 2004) from the approximately 60 Ma Oulad Abdoun basin of Morocco as the most basal member of Gryposuchinae (Salas-Gismondi et al. 2018). This phylogeographic pattern implies a trans-Atlantic migration from Africa to South America sometime between the Paleocene and Oligocene. Although the extant *Gavialis gangeticus* is restricted to freshwater habitats, it still retains salt glands on its tongue,

suggesting an ancestral marine adaptation (Taplin, 1985; Taplin et al., 1989). However, even more basal members of Gavialoidea are present in Late Cretaceous deposits in North America, thus not rejecting their presence in the Americas but without a fossil record during most of the Cenozoic.

The goal of this present study is to reassess the biogeography of *Crocodylus* using mitochondrial DNA. Although several previous studies have used this molecular source, we restrict our samples to only wild, geo-referenced populations and increase the number of samples from the Neotropics. Previous analyses have relied heavily on captive population samples (Meredith et al., 2011: Oaks, 2011). The pervasive hybridization between *Crocodylus* species may have confounded these prior analyses (Cedeño-Vázquez et al., 2008; Fitzsimmons et al., 2002; Hekkala et al., 2015; Milian-Garcia et al., 2015; Rodriguez et al., 2011). To avoid the confounding effects of hybrid samples reducing the phylogenetic resolution or producing misleading results, all known hybrid sources of samples were excluded. We also examine how geological and climatic events may have shaped the distribution of *Crocodylu*s.

H_0: The use of samples from wild populations to reconstruct the phylogeographic history of Crocodylus around the globe follows a west ward migration from the Indopacific to Africa to the Neotropics as shown in previous studies

H_A: The use of samples from wild populations to reconstruct the phylogeographic history of Crocodylus around the globe does not follow a west ward migration and has a more complex biogeographic history.

Methods

Sampling

We used 182 samples of *C. acutus* and *C. moreletii* from 32 localities throughout its range in Mexico on the Pacific and Atlantic coasts. Caudal scales were collected in microtubes with DESS (EDTA/DMSO solution saturated with NaCl) and stored at -20°C prior to DNA isolation.

DNA extraction and Amplification

DNA was extracted using phenol:chloroform:isoamyl alcohol (Bardakci & Skibinski, 1994). The DNA quality was assessed by electrophoresis using 1% Ultra-Pure Agarose (Invitrogen) gel. The purity and quantity of the templates were measured with an Infinite 200 Pro Nanoquant (Tecan). The $tRNA^{Pro}$-$tRNA^{Phe}$-d-loop region was amplified from all samples with the primers L15459 and CR2HA (Glenn et al., 2002; Ray & Densmore, 2002). Amplification was carried out in a 50 µL reaction volume containing: 5 µL 10X PCR buffer (200 mM Tris-HCl pH 8.4, 500 mM KCl), 1.5 µL $MgCl_2$(50mM), 0.5 µL of each primer (10 mM), 1.25 U *Taq* DNA Polymerase (Invitrogen), 50 ng of template DNA. The PCR conditions were performed in a TProfessional Gradient Thermocycler (Biometra). The Thermal cycling conditions for the PCR were an initial denatutation of 3 min at 94°C followed by 33 cycles at 94°C for 45 sec, 58°C for 1 min and 72°C for 45 sec, and a final extension of 72°C for 7 min. To remove the primer excess and any other reagent from the PCR product we used the QIAquick PCR Purification Kit (Qiagen). The sequencing was performed in an Applied Biosystems 3730xl DNA Analyzer at the McGill University and Génome Québec

Innovation Centre, Montréal, Canada. The forward and reverse sequences were edited and aligned using Geneious 11.0.5 (https://www.geneious.com).

The D-loop results were used to perform a nucleotide blast (blastn) search in Genbank (Benson et al., 2018) to retrieve sequences from extant *Crocodylus* species. The sequences were then used to make a database with all the metadata available from each sample (Table A1). Then the samples were selected to keep the wild samples and discard the ones from farms. To increase the resolution, we used as many available samples and sequences from wild populations from Genbank (Benson et al., 2018) (Table A1). The sequences were trimmed, aligned and edited with Geneious 11.0.5 (https://www.geneious.com). We used *Mecistops* and *Osteolaemus* as the outgroup. Alligatoroids were excluded because they have large sequence gaps in this region of their mitochondrial genome. These gaps confounded branch length estimates and pushed the origin of Crocodylia to approximately 200 Ma.

Data preprocessing

The software jModelTest (Posada, 2008) was used to choose the models of nucleotide sequence evolution using the Akaike information criterion (AIC), and Bayesian information criterion (BIC). A total of 88 models with invariant sites (I) and Gamma categories (G) were tested (Table A2).

Dating priors

We performed a Bayesian phylogenetic analysis to reconstruct the phylogenetic history of the group using BEAST 2 (Bouckaert et al., 2014) and the 3 best substitution

models based on the AIC and BIC criterions from the jModelTest analysis. We ran 5 $X10^7$ MCMC chains, logging every 10000 chains and saving the tree logs every 5000 chains. A pre burn-in of 50000 chains and a random tree as a starting point for all the simulations. The log files were analyzed with Tracer v.1.7 (Rambaut et al, 2018) and the Maximum clade credibility (MCC) consensus trees created with a 10 percent burn-in in TreeAnnotator v.2.5.1.

Initially, a phylogenetic analysis without priors was performed to assess the simplest model. Subsequent priors were imposed for divergence times. A divergence of *Gavialis gangeticus* from *Crocodylus* was imposed as between 70 - 100 Ma. This range includes two-time calibrations used previously. Brochu and Densmore (2011) used 70 Ma for this divergence and Oaks (2011) used 90 Ma. The oldest fossil gavialoids are *Eothoracosaurus mississippiensis* (Brochu 2004) and *Thoracosaurus macrorhynchus* (Brochu, 2006). The former is from the Early Maastrichtian aged Ripley Formation of Mississippi, USA (Sohl et al., 1991; Brochu, 2004) and the latter from several Maastrichtian aged deposits in France and Sweden (Brochu, 2006). The Maastrichtian is 70.6 to 66.043 Ma, supporting a 70 Ma minimum age of divergence for this clade. The oldest Brevirostres (the clade including Crocodyloidea and Alligatoroidea) is *Leidyosuchus canadensis* from the Dinosaur Park Formation of Alberta, Canada, dated between 77.03 - 75.46 Ma (Fowler, 2017). The range of 70 - 100 Ma was used to provide a conservative 30 Ma range to account for missing fossil taxa. *Mecistops* and *Osteolaemus* divergence time was imposed between 20 - 24 Ma. based on molecular data (Oaks, 2011; Shirley et al., 2013) and fossil occurrences from the Early Miocene (McAliley et al., 2006). For the calibration of *Crocodylus* a divergence time of 10 - 14

Ma. was used based on the oldest fossil of the group, *Crocodylus palaeindicus* from the Late Miocene (Hussain & West, 1979; Barry et al., 2002), and molecular data calibrations (Oaks, 2011: Meredith et al., 2011).

Ancestral geographic range reconstruction

We used the consensus trees from the Bayesian phylogenetic analysis from BEAST2 (Bouckaert et al., 2014), to reconstruct the probability of the ancestral geographic range of each node with BioGeoBEARS (Matzke, 2014). We performed a stratified analysis using three time periods (25, 13, 9 Ma.) based on reported dispersion events from Oaks, 2011. Dispersal-extinction-cladogenesis with jumps, which emphasize dispersals, (DEC) + J models were used. We also ran an unstratified analysis with no time constraints, all areas allowed and no dispersal multipliers. A maximum likelihood ancestral geographic reconstruction was also estimated with Mesquite v.3.51 (Madison & Madison, 2018) and presented as supplementary information (Table A3).

Results

The best model of nucleotide sequence evolution model determined by AIC is TPM3 uniform + I + G and for the BIC the HKY + I + G (Table A2). The resulting consensus trees are presented in Figure 2 and the posterior probabilities and the divergence times in Table 1.

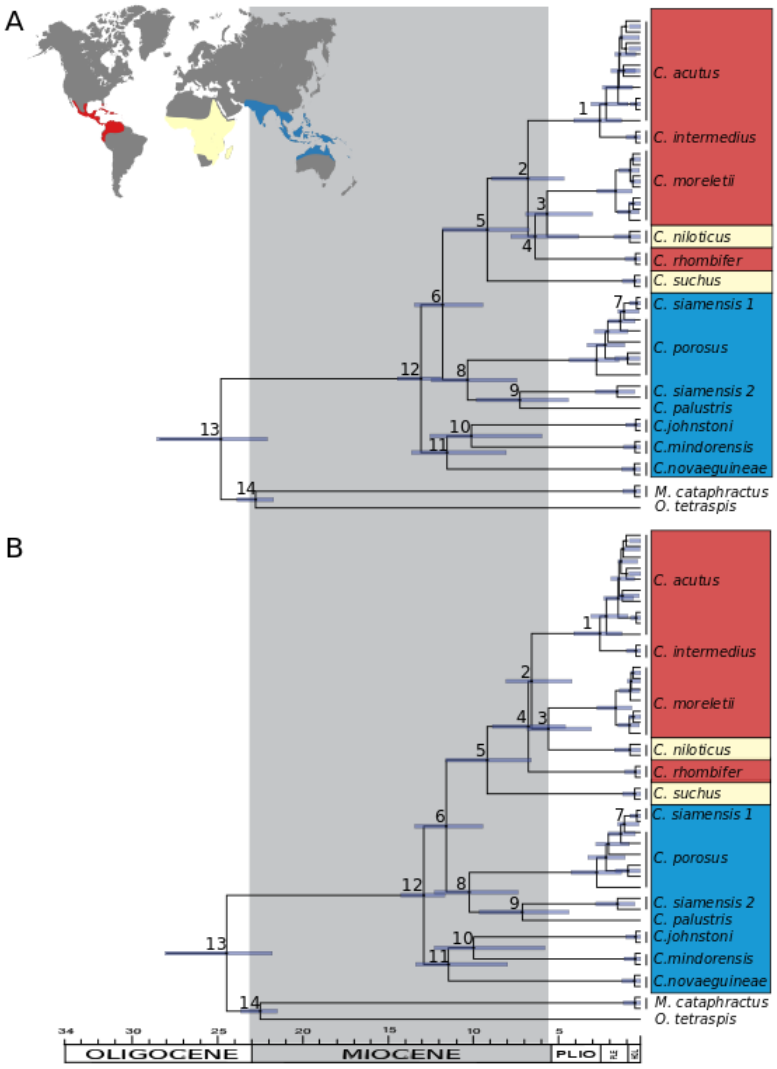

Figure 2. Maximum clade credibility tree from BEAST for the A) TPM3uf + I + G and B) HKY + I + G models. Bars at nodes represent the 95 % posterior density of the node age. The numbers and values at the nodes are shown in Table 1.

Table 1. Node statistics for the AIC and BIC models. Node age means and 95% HPD from the maximum clade credibility trees from BEAST analysis with the fossil calibrations for *Mecistops* + *Osteolaemus* and *Crocodylus* divergence. Node numbers correspond to Figure 2.

Substitution model	TPM3uf + I + G			HKY + I + G		
Node	Mean	95%HPD	Posterior probability	Mean	95%HPD	Posterior probability
1	2.36	1.1-3.88	1	2.39	1.09-3.93	1
2	6.54	4.44-8.65	0.96	6.41	4.04-7.92	0.21
3	5.43	2.79-6.66	0.46	5.42	2.9-6.67	0.48
4	6.12	3.58-7.53	0.4	6.6	4.43-8.72	0.97
5	8.9	6.47-11.5	0.99	9.02	6.43-11.45	0.99
6	11.49	9.15-13.13	0.81	11.42	9.27-13.29	0.87
7	0.96	0.08-1.32	0.41	0.96	0.1-1.36	0.42
8	10.05	7.18-12.15	0.77	10.7	7.19-12.12	0.77
9	7.01	4.18-9.53	0.94	6.96	4.21-9.49	0.95
10	9.82	5.72-12.24	0.59	9.81	5.62-12.13	0.61
11	11.24	7.8-13.3	0.72	11.3	7.82-13.22	0.72
12	12.75	11.58-14.13	1	12.75	11.51-14.1	1
13	24.37	21.66-28.1	1	24.33	21.65-27.95	1
14	22.35	21.31-23.45	1	22.36	21.35-23.52	1

There is strong support for the following phylogenetic relationships: 1) *Crocodylus* generic clade support; 2) *Mecistops* sister to *Osteaolaemus* and with a very old divergence; 3) paraphyly of Australasian *Crocodylus* species; 4) Neotropical and African *Crocodylus* monophyly; 5) independent ancestry of *C. niloticus* and *C. suchus*; and monophyly of 6) *C. mindorensis*, *C. novaeguineae* and *C. johnstoni*; 7) *C. porosus*, *C. siamensis* and *C. palustris*; 8) *C. acutus* and *C. intermedius*: 9) *C. niloticus* and *C.*

moreletii. The position of *C. rhombifer* is poorly supported. Its position, however, is consistent in all search parameters used but the HKY + I + G with and without prior dating calibrations and using Osteolaeminae as the outgroup (Figures A1 - A4). With all other models used, *C. rhombifer* is sister to a *C. niloticus* and *C. moreletti* clade. However, for the HKY + I + G model, *C. rhombifer* is sister to the remaining Neotropical crocodiles and *C. niloticus* clade, with the latter still sister to *C. moreletii*.

The topologies of the phylogenies where *Gavialis* and *Tomistoma* are used as outgroups have compatible and overlapping dates with the phylogenies. Using *Gavialis* and *Tomistoma* as outgroup taxa yielded topologies consistent with all but the HKY model. *Tomistoma* and *Gavialis* are recovered as sister species, however, the divergence time goes deep in time to the Early Cretaceous (71-73 Ma). The divergence of *Gavialis* + *Tomistoma* clade from *Mecistops* + *Osteolaemus* + Crocodylus clade is estimated to 71-85 Ma. (Figures A5 - A13)

Ancestral state reconstruction results for three biogeographic character states on the TPM3uf + I + G model are shown in figure 3. The ancestral reconstruction for *Mecistops* + *Osteolaemus* is not clear with a probability of 40% for an African origin and 22% Australasian (Node 13). The Australasian species reconstruction *C. johnstoni* + C*. mindorensis* + *C. novaeguineae* (Node 12) has a probability of 87%, supporting an origin in the Indopacific. The *C. porosus* + *C. siamensis* + *C. palustris* (Node 6) has a probability of 81% supporting an origin in Africa.

However, the African + Neotropical clade of *Crocodylus* (node 5) has a 100% probability of an African origin under the stratified DEC + J model. Yet, for the unstratified, the probability is mixed 54 % from African origin and 41% Neotropical. The Neotropical

clade, including *C. niloticus*, has an ancestral state reconstruction suggesting a Neotropical origin with a probability of 100% (node 2). The ancestral range for the *C. rhombifer* + *C. niloticus* + *C. moreletii* has the 50% probability for an African and 50% Neotropical origin under the unstratified model, but a 95% probability of a Neotropical when not constrained (node 4). The complete set of reconstructed Ancestral geographic ranges are in Figures A14 – A17 and the values for all the nodes are in Table A4.

The *C. niloticus* + *C. moreletii* ancestral state has a 98% probability of African origin for the stratified analysis and 87% probability of Neotropical origin for the unstratified DEC + J analysis (node 3). These different ancestries support a dispersal from Africa to the Neotropics, then back to Africa and again to the Neotropics. Maximum likelihood reconstructions with Mesquite, with the different character states tested are consistent with those presented in Figure 3 with the exception of the African origin of the *C. niloticus* + *C. moreletti* for the stratified analysis (Table A5, Figures A18 – A27).

Osteolaeminae (*Osteolaemus* and *Mecistops*) have a deep divergence time in the late Oligocene - early Miocene. Rapid radiation and dispersal of extant *Crocodylus* took place in the Middle to Late Miocene. The timing of the origin of *Crocodylus* is consistent in all models and it is estimated to have occurred in the Mid Miocene (Serravalian - Langhian) with the split between Australasian and African-Neotropical clades in the Serravalian - Tortonian. A vicariant event of *C. suchus* from *C. niloticus* + Neotropical species in the Tortonian. Timing of the migration of *Crocodylus* from Africa to the Neotropics is dated between 11.5 - 4.4 Ma. This large range spans from the earliest date for the origin of the Africa + Neotropical clade to the youngest date for the origin of the

Neotropical clade. Similarly, the subsequent return of *Crocodylus* from the Neotropics to Africa to establish as *C. niloticus* is dated between 7.8 - 2.8 Ma.

Discussion

The phylogenetic hypothesis using only wild-caught *Crocodylus* is largely congruent with previous hypotheses but with notable exceptions in the relationships of African and Neotropical species. Divergence times are slightly altered. For example, the divergence of *Osteolaemus* and *Mecistops* of the outgroup. Osteolaeminae, is estimated at 22.4 Ma compared with the 17.7 Ma estimate of Oaks (2011). The paraphyly and basal position of Australasian *Crocodylus* species is congruent with previous authors and supports an Australasian origin of the genus (Meredith et al., 2011; Oaks, 2011). Additionally, the phylogenetic division of samples identified as *C. siamensis*, with one (*C. siamensis* 1) sister to *C. porosus* and the other (*C. siamensis* 2) sister to *C. palustris*. Meredith et al. (2011) suggest the samples *C. siamensis*1 (Ji, Wu, Yan, & Amato, 2008), recovered as sister to *C. porosus*, might come from hybrids of the two species. This suggests that there is a *C. porosus* haplotype that has not been sampled in previous studies. Another explanation is the presence of a different morphotype of *C. porosus* similar to *C. siamensis*, in an overlapping area of distribution. This division is worth further exploration given the formerly widespread distribution of this species throughout southeast Asia with potential cryptic species.

Another peculiar phylogenetic result is the placement of *C. niloticus* within the radiation of Neotropical crocodiles. This position is well supported by all optimal

phylogenetic searches. Using both optimal search algorithms, *C. niloticus* is recovered as the sister taxon to *C. moreletii*. However, the position of *C. rhombifer* varies between the TPM3 and HKY searches. With TPM3, *C. rhombifer* is sister to a *C. niloticus* + *C. moreletii* clade whereas the HKY search yields *C. rhombifer* sister to the entire Neotropical and *C. niloticus* clade. Although clade support for many nodes within the Neotropical and *C. niloticus* clade is relatively low, support for the entire clade is high (Table 1, Figure 3).

Crocodylus karyotype does not conflict with the novel inclusion of *C. niloticus* within the Neotropical radiation. Karyotypes are conserved in the genus with 58 fundamental chromosomes for all species, except for *C. moreletii* that has 56. The conservative arrangement has been attributed to the lack of postzygotic isolation mechanisms in the genus *Crocodylus* (Srikulnath et al., 2015). The karyotype arrangements are conserved for clades: 1) *C. mindorensis* + *C. novaeguineae* + *C. johnstoni*: 2) *C. palustris* + *C. siamensis*: 3) *C. acutus* + *C. intermedius*. The karyotypes of *C. suchus* and *C. porosus* are equal even if they are not sister species, this suggests it is a karyotype present after the divergence of Australasian and African clades. Neotropical crocodiles + *C. niloticus* have a chromosome number of 2n = 32, except *C. rhombifer* which has a 2n = 30 (Srikulnath et al., 2015). The changes in the configuration of acrocentric and metacentric chromosomes do not conflict with our phylogenetic hypothesis (Figure A28). Chromosomal rearrangements (CRs) of *C. moreletii* and *C. rhombifer,* which are young species, are concordant with some predictions of the chromosomal speciation models suggested by Faria and Navarro (2010). The predictions fulfilled are 1) more karyotypic differences between sympatric

sister species (*C.moreletii* and *C.acutus*; *C.rhombifer* and *C.acutus*) than between allopatric ones; 2) these young *Crocodylus* Neotropical species have more CRs relative to molecular divergence compared with Australasian species.

Biogeography of Crocodylus

The inclusion of *C. niloticus* within the Neotropical radiation of *Crocodylus* has dramatic implications compared to earlier hypotheses where *C. niloticus* is placed sister to that radiation. This position presents intriguing possibilities of *Crocodylus* biogeography. In any scenario, multiple trans-Atlantic crossings are required. Assuming the establishment of *C. suchus* from an Indopacific ancestry, an initial trans-Atlantic crossing is estimated to have occurred between 11.5 - 4.4 Ma (Figure 4A). Using the most frequently recovered topology (7 of 8 runs) of *C. rhombifer* sister to *C. morelletii* + *C. niloticus*, maximum parsimony and ML ancestral geographic state reconstructions suggest the Neotropical lineage radiated into a southern clade of *C. acutus* and *C. intermedius* and a more northern clade of *C. rhombifer* and *C. moreletii* (figures A18 - A27). The stem lineage of *C. moreletti* evolved a migrant lineage that returned to Africa to establish *C. niloticus* after 6.7 Ma. The youngest date is poorly constrained because the youngest 95% confidence date of divergence of the two *C. niloticus* samples is Recent. This scenario requires two trans-Atlantic migrations. MP and ML reconstructions of the single run that yields *C. rhombifer* as sister to the remaining clade of Neotropical and *C. niloticus* results in a similar migration scenario.

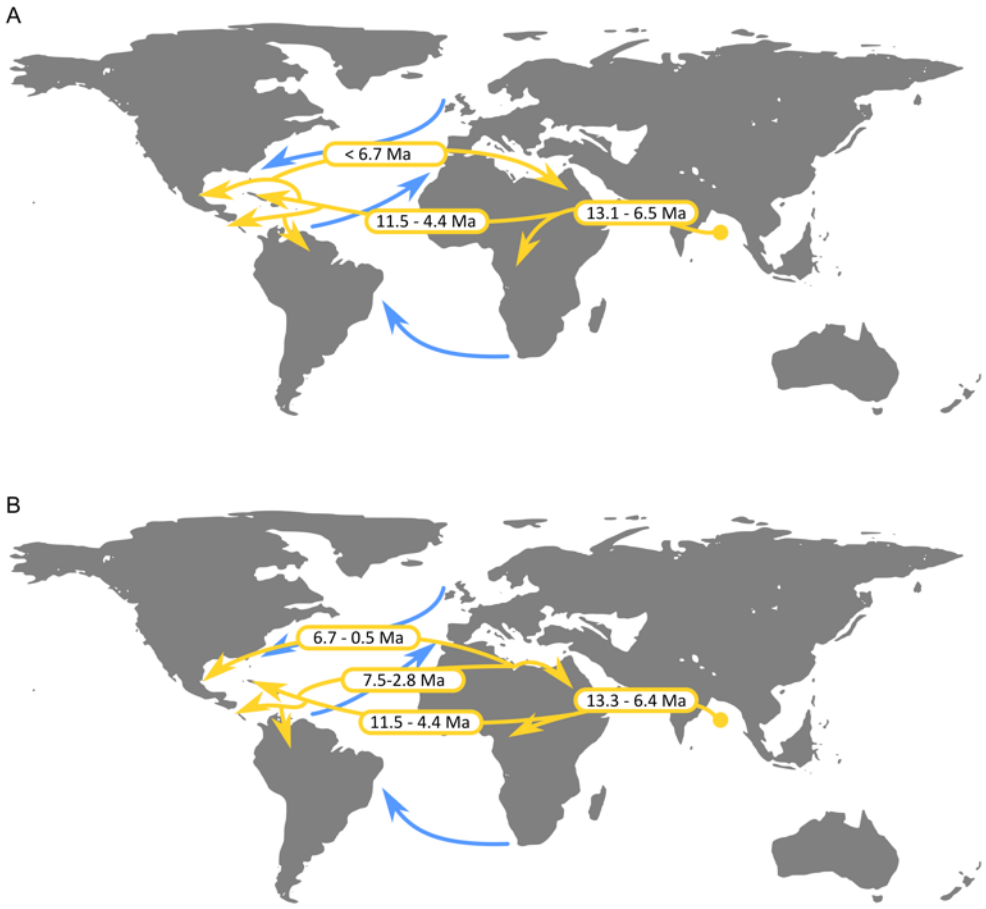

Figure 4. Estimated dispersal routes for *Crocodylus* based on (A) Maximum Likelihood and (B) a Bayesian biogeographic reconstruction. Map represents a mid-Miocene geography with blue arrows illustrating Atlantic Ocean currents at that time. Yellow arrows denote trans-Atlantic dispersals of *Crocodylus* with estimated age ranges for each dispersal.

An alternative scenario with both African species originating from a common antecedent in Africa requires a trans-Atlantic migration for, first, the *C. acutus* and *C. intermedius* clade, later for *C. rhombifer*, and a final migration of *C. moreletii*, in all eight trees. Although this requires three trans-Atlantic crossings, the difference between two

and three migrations may be minimal given the salt-water tolerance of the genus. In either scenario, multiple trans-Atlantic migrations for a saltwater tolerant genus are required.

More complex dispersal and extinction models were tested using ancestral reconstruction range with BioGeoBEARS with the most recovered chronogram; the one maintaining a *C. rhombifer* sister to *C. niloticus* + *C.moreletii* clade. The temporally stratified and unstratified stratified DEC + J model both yield an African origin for the non-Australasian *Crocodylus*, although stratified without jumps suggests a Neotropical origin is nearly as likely (Figure 3, Figures A14 - A17). Both also support a Neotropical origin of the Neotropical clade that includes *C. niloticus*. The *C. acutus* and *C. intermedius* radiation is recovered as a Neotropical origin in all models. The unstratified model maintains the return to Africa migration for the *C. niloticus* lineage. However, the stratified model with and without jumps strongly supports a return to Africa for the *C. niloticus* and *C. moreletii* clade with a subsequent return to the Neotropics by the *C.moreletii* lineage. Such a complex scenario is probably facilitated by the short branch lengths about these nodes. Although seemingly less likely, an African origin and return to the Neotropics by the *C. moreletii* lineage may reflect its unique karyotype. This more complex dispersal scenario is illustrated in Figure 4B. Although the timing for the initial trans-Atlantic dispersal is the same (11.5 - 4.4 Ma), the return to Africa of the *C. niloticus* + *C. moreletii* clade is estimated between 7.5 - 2.8 Ma. The final return of the *C. moreletii* lineage to the Neotropics is estimated between 6.7 - 0.5 Ma.

Regardless of the origin of the *C. moreletii* lineage, a Neotropical origin for the immediate ancestry of *C. niloticus* is well supported. This ‘there and back again’ model

suggests a secondary African radiation of a Neotropical *Crocodylus* lineage. It is tempting to conclude that the Tortonian *Crocodylus* fossils in the Mediterranean and *C. checchiai* from Africa, that include the nasal boss characteristic of many Neotropical crocodiles, may actually be descendants of this earlier radiation in the Americas.

A series of another well-established Cenozoic trans-Atlantic terrestrial vertebrate have occurred in the past. Most are westward and occurred during the Paleogene. A westward migration of the burrowing blind snakes has been postulated from the Late Palaeocene to Early Oligocene (Vidal et al. 2010). Platyrrhine monkeys and caviomorph rodents are endemic to South America yet these clades are sister to African catarrhine monkeys and phiomorph rodents, respectively (Poux et al. 2006). An ancient trans-Atlantic dispersal from Africa for these groups was first postulated by Lavocat (1969). The oldest fossil platyrrhine, *Perupithecus* is 37 Ma (Bond et al. 2015). The oldest caviomorph, *Cachiyacuy contamanensis* is 41 Ma (Antoine et al. 2011). Molecular and morphological phylogenies and the fossil record indicate westward trans-Atlantic migrations of these mammalian clades from African ancestries during the Middle to Late Eocene. Caviomorph dispersal, based on fossil evidence, is nearly coincident with the Mid-Eocene Climate Optimum (Antoine et al. 2011) and molecular-estimated origins of platyrrhines overlap this age (Poux et al. 2006). A concurrent migration of these lineages from Africa via rafting and/or island hopping cannot be ruled out.

An additional migration of a terrestrial bird species has also been hypothesized. *Lavocatavis* is a late Early or early Middle Eocene phorusrhacoid bird from Algeria (Mourer-Chauviré et al. 2011) and *Eleutherornis* a Middle Eocene phorusrhacoid from France (Angst et al. 2013). The latter is well dated to 43.5 - 41.2 Ma (Angst et al. 2013).

All other members of this clade are South American and flightless and range from Eocene to the present day, with only the extant lineage capable of some flight. The Eocene origin of this clade implies either a westward migration to South America or an eastward migration to Africa around this time. In all cases, these trans-Atlantic dispersals were probably facilitated by island hopping. A string of emerged islands has been postulated to have been present along the Walvis Ridge and or Rio Grande Rise from Cretaceous to Late Miocene times (Sclater et al. 1977, Perez-Diaz & Eagles 2017). Prevailing ocean currents have been modelled as circulating westward from southern Africa to present day Brazil along the Walvis Ridge to the Rio Grande Rise, westward from Iberia to the Caribbean, and eastward along the Equator during the Miocene (Herold et al. 2012). Such topographies and ocean currents would have facilitated both terrestrial dispersals and eastward and westward amphibious dispersals.

These terrestrial trans-Atlantic dispersals precede the predicted *Crocodylus* dispersals by about 30 million years. However, the frequent terrestrial vertebrate dispersals highlight the propensity of trans-Atlantic migrations. Moreover, the presence of Atlantic islands and recirculating tropical currents during the Miocene would have facilitated dispersal of the amphibious and marine adapted *Crocodylus*. The short branch lengths recovered between nodes inferred to have made these westward and eastward migrations suggests there were potentially frequent dispersals along these routes. Alternatively, the radiation of *Crocodylus* throughout the tropics may be older than the fossil record and molecular clock estimates reveal.

Climatic changes and geological events have shaped the distribution, diversification and extinction of crocodylians through space and time (Bronzati, Montefeltro, & Langer, 2015; Mannion et al., 2015; Markwick, 1998). However, some extinction events have also been associated with the radiation of *Crocodylus*. One is the Late Miocene extinction of South American gavialoids, which is coincident with both the arrival of *Crocodylus* and declining climatic temperatures (Gasparini 1968; Buffetaut 1982; Scheyer et al. 2013; Dıaz de Gamero 1996; Quiroz & Jaramillo 2010; Salas-Gismondi et al. 2018). Although neither can be directly implicated, sympatry of extant *Gavialis* with *Crocodylus* (Choudhary et al. 2017) indicates they may have persisted together in South America, at least until temperatures declined. Another possible extinction associated with the arrival of *Crocodylus* is the loss of Osteolaeminae on Madagascar. *Voay robustus* is the last surviving Osteolaeminae on that island during the early-middle Pleistocene, possibly related to changes in habitat and food availability (Bickelmann & Klein, 2009). This crocodile was a top predator in Madagascar and a close relative to *Osteolaemus* (Brochu, 2007). Some have suggested the coexistence and later ecological displacement of *V. robustus* by *C. niloticus* (Bickelmann & Klein, 2009), or the establishment of *C. niloticus* after *V. robustus* extinction (Brochu, 2007). The fossil record offers little resolution on this interaction because the oldest fossil of *C. niloticus* from Madagascar is only between 460 - 310 years old (Mathews and Samonds, 2016). However, this extinction event is roughly coincident with our estimated time of divergence of *C. niloticus* to Madagascar (0.6 Ma), supporting a potential cause for the extinction of Osteolaeminae on the island.

During the Middle Miocene and Pliocene, an event changed the dynamics of worldwide ocean circulations and terrestrial migrations between North and South America. The rise of the Central American Isthmus (CAI) is associated with the initiation of the thermohaline circulation (THC), the Northern Hemisphere Glaciation (NHG), the formation of the Caribbean Sea, and the Great American Biotic Interchange (GABI). The Central American seaway that connected the Pacific and Atlantic Oceans closed around 10 Ma, but shallow waters and islands remained until approximately 3.5 Ma (Jaramillo, 2018). These dates coincide with dates estimated for the multiple trans-Atlantic *Crocodylus* dispersals. The dates of the complete formation of the CAI overlap with the divergence of the Pacific and Caribbean *C. acutus* populations (node 1, Fig. 3). This species is widely distributed in the Neotropics with populations in the Caribbean and Pacific coasts (Thorbjarnarson et al., 2006), and the rise of CAI may have contributed to the divergence and differentiation of Pacific and Atlantic *C. acutus* populations. Our data shows four main *C. acutus* clades (Fig. 3), from recent to old: 1) Caribbean islands in Yucatan, 2) Northern Yucatan Peninsula, 3) Pacific, and 4) the Caribbean.

This timed biogeography is consistent with the proposed westward migration of *Crocodylus* (Meredith, 2011: Oaks, 2011), where the ancestral *C. acutus* lineage arrived or was established in the Caribbean and represented in our data set with a population at Xcalak, on the Yucatan coast. This origination timing for *C. acutus* before the rise of the CAI suggests the closure may have separated the Pacific and Caribbean populations. Our dataset estimates the divergence of these populations between 0.7 to 2.8 Ma (mean = 2 Ma), which overlaps with other estimations of the closure of the CAI (O'Dea et al., 2016). Although the complete formation of the CAI is estimated to have occurred

about 3.5 Ma, the region would have retained widespread mangroves and shallow waters that the amphibious *Crocodylus* would likely have exploited.

We found four *C. acutus* haplotypes for the Yucatan - Caribbean islands region. This is one more haplotype than previously reported by Cedeño-Vazquez and collaborators (2008) for this region. The presence of a unique haplotype in Pacific Mexico suggests it originated from one maternal lineage from the Atlantic that crossed before the rise of the CAI to the Pacific. Similar patterns of the Atlantic as a source for the Pacific, and a lower genetic diversity in the Pacific as an effect of closure of the CAI, has been reported for mangroves (Cerón-Souza et al., 2015), fish (Galván-Quesada et al., 2016), and bivalves (Marko & Moran, 2009). Our estimates of a 2.8 Ma divergence are consistent with catfish estimates that used genomic data, a more robust approach for divergence times (Stange et al., 2018). Although we cannot conclude that the rise of the CAI caused the divergence of Atlantic and Pacific populations of *C. acutus*, the timings are nearly coincident.

DNA sequences have been used to identify cryptic species/lineages within the order Crocodylia. An example is the recognition of two African *Crocodylus* species, *C. niloticus* from Eastern Africa and *C. suchus* from Western Africa (Hekkala et al., 2011) and the discovery of three subspecies of the African dwarf Crocodile (Eaton et al., 2009; Franke et al., 2013). Studies in the South American dwarf caimans, *P. trigonatus* (Bittencourt et al., 2019) and *P. palpebrosus* (Muniz et al., 2018) confirmed the existence of multiple widespread lineages. The broad distribution of *C. acutus* coupled with the presence of the different mitochondrial haplotypes found in our study suggests a more complex evolutionary history of the species. To further evaluate the presence of

cryptic species or lineages within *C. acutus*, a more intensive sampling in its geographic range and a higher genomic resolution are necessary.

Conclusions

This study emphasizes the careful use of wild, georeferenced genetic data. Although the phylogeny of *Crocodylus* is largely congruent with previous results of a general Australasian, to Africa, to the Neotropics dispersal trend during the Miocene, the higher sampling in the Neotropics and exclusion of captive individuals reveals some differences. These include potential cryptic species in Australasia with *C. siamensis* and unique lineages of *C. acutus* in the Neotropics. *C. niloticus* is firmly nested within the Neotropical radiation implying, minimally, an eastward transAtlantic dispersal for this lineage. These multiple transAtlantic dispersals suggest a more complex biogeographic history for the genus and place high emphasis on conserving unique, isolated populations of the genus.

Bridging text

In Chapter 1 we reassessed the phylogenetic history of the Genus Crocodylus. With the integration of extensive fossil record review, paleo bathymetry, paleo-ocean currents and wild caught samples we were able to confirm events previously hypothesized. In this chapter we focused on the genus and the effects of ocean currents on its radiation at a global scale.

For Chapter 2, we moved the focus from a global scale of chapter 1, to a regional scale. We concentrated on two species of crocodiles: *Crocodylus acutus* and *Crocodylus moreletii*. The chapter evaluates effect of ocean currents on crocodile migrations. This chapter continues the use of genetics but increasing the number and type of markers as well as demographic models to predict the direction of the gene flow from one island in the Caribbean to the coast. The chapter used novel next generation sequencing and modelling for the analysis.

CHAPTER 2

Origin and evolutionary history of Banco Chinchorro saltwater crocodiles

Abstract

Banco Chinchorro Biosphere Reserve is an atoll located 30 km off the coast of Yucatan. It has many species of reptiles protected by CITES, IUCN and the Mexican government. One of them is *Crocodylus acutus*, which is widely distributed in Pacific, Atlantic and Caribbean coasts and islands. The ecology, morphology, and behaviour of the Banco Chinchorro *Crocodylus* population have been studied for the past two decades and results suggested it is an isolated population with potential molecular and morphological differentiation. Recent genetic population studies also suggested that it is a pure *C. acutus* population. Our study goals were to answer the following questions: 1) Are Banco Chinchorro crocodiles a pure population? 2) How do ocean currents direct the gene flow of crocodiles in the Caribbean and Mexico? 3) What is the origin of Banco Chinchorro crocodiles? We collected 218 scale samples from 15 localities and two species of crocodiles (*C. acutus* and *C. moreletii*) in Mexico and Panama. We used Microsatellites, Mitochondrial DNA sequences and Single Digest Restriction Site Associated DNA Sequencing (sRAD-Seq) to genotype all the samples. The population's structure results revealed the presence of two genetic clusters in Banco Chinchorro. Demographic models showed that the Banco Chinchorro populations diverged from Caribbean *C. acutus* populations. We propose the Yucatan current may be the cause of secondary contact of Banco Chinchorro populations with other crocodile populations. Our results indicate that Banco Chinchorro crocodiles are an isolated hybrid population,

with population subdivision, and are a source of genetic variation to populations in the Yucatan Peninsula Coast and the Gulf of Mexico.

Introduction

Islands are natural test tubes because they have a small size, clear boundaries and geographic isolation which makes it easier to observe and to infer ecological and evolutionary patterns of their inhabitants, and their evolutionary history (Losos & Ricklefs, 2009). Island populations also have lower genetic variation caused by isolation but can contribute to mainland genetic variation by the amount and uniqueness of their divergent genetic variation (Wilson et al, 2009). Divergence occurs in allopatry and barriers to interbreeding arise as a result of selection, with or without gene flow from a parent or daughter population (Losos et al., 2010). There are four general models to explain speciation in birds in islands, the closest extant group to crocodiles : I) Speciation in allopatry II) Divergence as a cause of Secondary contact with selective reinforcement of traits evolved in allopatry III) Exchange of genes in sympatry but with introgressive hybridization, the results is new genes that enhance response to selection IV) Speciation in sympatry due to assortative mating (Losos et al., 2010).

Banco Chinchorro is a false atoll located in the Yucatan Channel, which is located between the Yucatan Basin and the Yucatan Borderland. The metamorphic rocks of the Yucatan Channel originated in the Late Triassic (59.3 - 92.5 mya), and the actual circulation of the Yucatan Channel, Straits of Florida, and the Gulf Stream have existed since the Cretaceous (López-Ramos, 1975). The currents that affect the region originate

from the Caribbean current, which becomes the Cayman current and finally transforms into the Yucatan Current. This current is divided in a southern region with southward weak coastal currents moving to Belize, a transitional region where the Cayman Current intrudes upon the Yucatan Peninsula, and the northern coastal currents which flow strongly northward along the Yucatan coast into the Gulf of Mexico and develops into the Loop Current which moves all the way up to Florida. (Carrillo et al. 2015) (Figure 1).

The Banco Chinchorro Biosphere Reserve is part of the Mesoamerican Barrier Reef System (MBRS), which is the most extensive reef system of the Atlantic Ocean (Carrillo et al., 2015). The reserve is located 30.8 km offshore from Mahahual, and it is separated from the coast by a thousand-meter deep trench (UNEP/IUCN, 1988). It has an area of approximately 145 thousand hectares with three main cays (Lobos, Center, and North). The cays represent 0.4% of the total area of the reserve (INE, 2000). There are 778 species reported for the reserve, 58% are marine fauna, 14% terrestrial fauna, 18% marine flora and 10% land flora (INE, 2000). Thirteen species of reptiles are reported inhabit the island (Charruau et al., 2016). One of them is the American crocodile (*Crocodylus acutus*), a widely distributed species in the Neotropics that is present in the Pacific, Atlantic and Caribbean. The American crocodile inhabits a variety of environments from hypersaline to freshwater (Ernst et al., 1999). An assessment of the status of the American crocodile for conservation described 69 areas in eight distinct crocodile bioregions as Crocodile Conservation Units (CCU), and Banco Chinchorro was not considered under risk (Thorbjarnarson et al., 2006). This study was based on the habitat quality, nesting habitat, CCU size, crocodile population size, habitat connectivity, habitat sustainable use, and crocodile killings more than ten years ago. But the

development of Banco Chinchorro for tourism has increased, especially for diving in the reef and many shipwrecks present in the reserve (INE, 2000; Ardisson et al., 2011). Also, the biology of the population on the island has been extensively studied during the last decade. We have knowledge of the nesting ecology (Charruau, 2012), population status (Charruau et al., 2005), ethology (Henaut & Charruau, 2012) and growth (Charruau et al., 2010). There are two relevant aspects of the population on the island that make it unique, one of them being its unique morphology. Labarre et al. (2017) suggest the existence of a broader-snouted morphotype of *C. acutus* that is typical of island populations of the Yucatan Peninsula. The other one is the genetic signature, it has been proposed that the last remnants of pure *C. acutus* in Caribbean Mexico are located in Cozumel and Banco Chinchorro (Machkour-M'Rabet et al., 2009; Cedeño-Vázquez et al., 2008; Pacheco-Sierra et al., 2016). These two Caribbean island populations are considered to be pure and isolated from coastal populations based on mitochondrial DNA (Cedeño-Vázquez et al., 2008; Pacheco-Sierra et al., 2016), microsatellites and GBS (Pacheco-Sierra et al., 2018).

We used samples from localities that are not in the reported hybrid zone to polarize and differentiate hybrids from no hybrids.

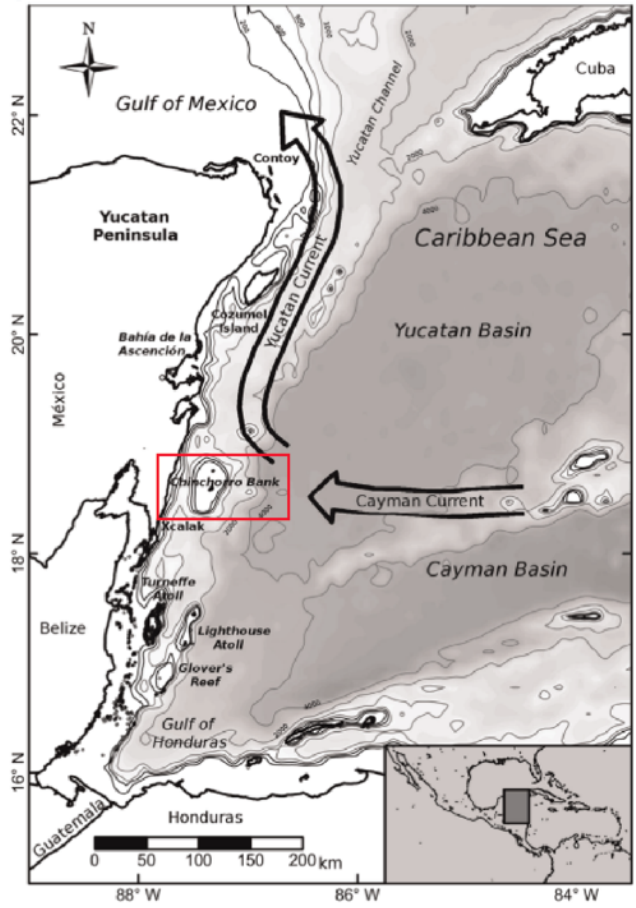

Figure 1. Ocean Currents of Yucatan. Banco Chinchorro is highlighted in red (modified from Carrillo et al. 2015).

Our study questions for this study: 1) Are Banco Chinchorro crocodiles a pure population? 2) How do ocean currents direct the gene flow of crocodiles in the Caribbean and Mexico? 3) What is the origin of Banco Chinchorro crocodiles?

H_0: The use of more genetic markers will show that Banco Chinchorro population is a pure C. acutus in the Caribbean of Mexico as suggested by previous studies

H_A: The use of more genetic markers will show that Banco Chinchorro population is not a pure C. acutus in the Caribbean of Mexico as suggested by previous studies

Methods

Sampling

Tissue samples were acquired from 15 localities in Mexico and Panama, some samples were donated from El Colegio de la Frontera Sur and the National Collection of Amphibians and Reptiles (UNAM) collections (Figure 2) Our sampling included 251 individuals from *C. acutus* (206) and *C. moreletii* (45). Species assignment was based on the morphological characters defined by Platt and Rainwater (2005) and Ross and Ross (1974).

DNA extraction

We extracted the DNA using phenol:chloroform:isoamyl alcohol (Bardakci & Skibinski, 1994). The DNA quality was assessed by electrophoresis using 1% Ultra-Pure Agarose gel (Invitrogen). The purity and quantity of the templates were measured with a Qubit Fluorometer (Thermofisher).

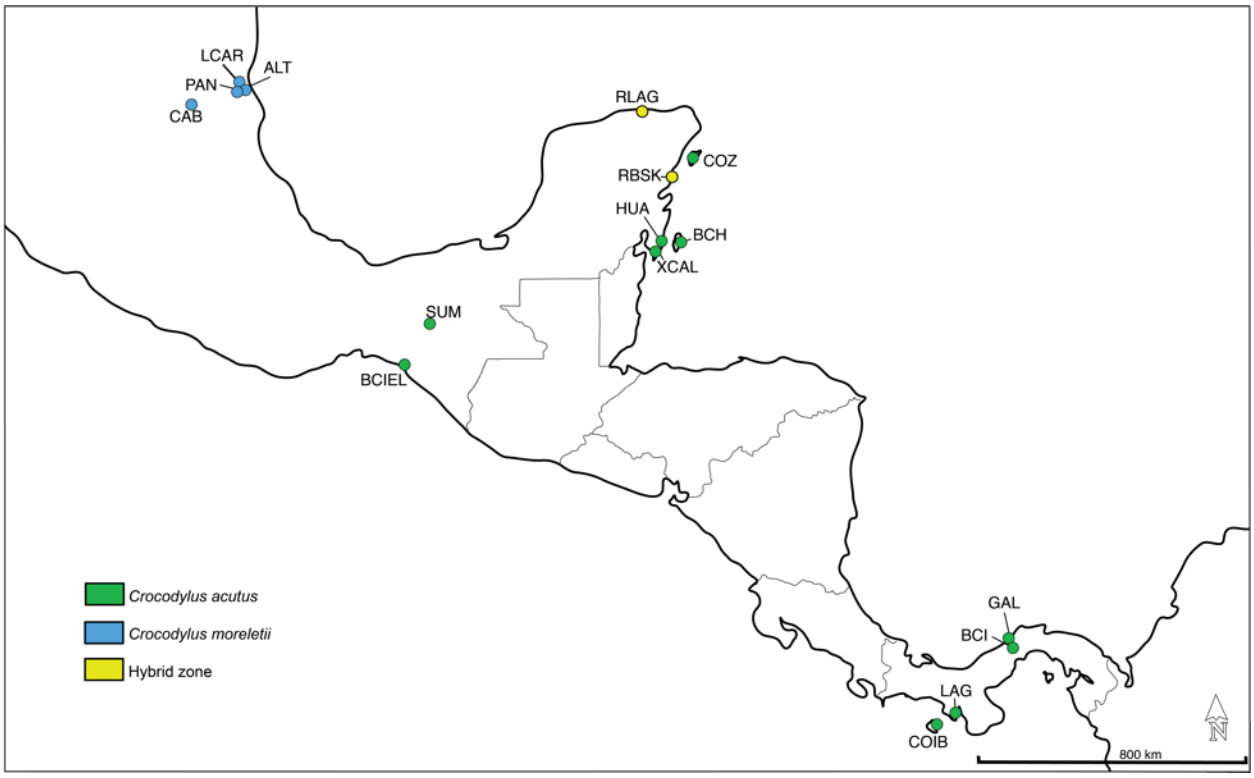

Figure 2. Map of localities sampled for the study. Abbreviations refer to: ALT, Altamira; BCH, Banco Chinchorro; BCI, Barro Colorado island; BCIEL, Boca del Cielo; CAB, Cienega de Cabezas; COIB, Coiba island; COZ, Cozumel; GAL, Galeta; HUA, Huach; LAG, Lagartero; LCAR, Laguna del Carpintero; PAN, Panuco; RBSK, Reserva de la biosfera de Sian Ka'an; RLAG, Ria Lagartos; SUM; Canon del Sumidero; and XCAL, Xcalak.

Mitochondrial DNA

The tRNAPro-tRNAPhe-D-loop region was amplified in 122 samples from 12 localities (20 *C. moreletii,* 102 *C. acutus)* with the primers L15459 and CR2HA (Glenn et al., 2002; Ray & Densmore, 2002). Amplification was carried out in a 50 µL reaction volume containing: 5 µL 10X PCR buffer (200 mM Tris-HCl pH 8.4, 500 mM KCl), 1.5 µL

MgCl2(50mM), 0.5 µL of each primer (10 mM), 1.25 U Taq DNA Polymerase (Invitrogen), 50 ng of template DNA. The PCR conditions were performed in a TProfessional Gradient Thermocycler (Biometra). The Thermal cycling conditions for the PCR were an initial denaturalization of 3 min at 94°C followed by 33 cycles at 94°C for 45 sec, 58°C for 1 min and 72°C for 45 sec. A final extension of 72°C for 7 min. Bands were visualized on a 1% agarose gel. To remove the primer excess and any other reagent from the PCR product we used the QIAquick PCR Purification Kit (Qiagen). The sequencing was performed on an Applied Biosystems 3730xl DNA Analyzer at the McGill University and Génome Québec Innovation Centre, Montréal, Canada. The forward and reverse sequences were edited and aligned using GENEIOUS 11.0.5 (https://www.geneious.com).

Microsatellites

We used 11 fluorescently labelled dinucleotide microsatellites in 141 samples from 10 populations in Mexico (30 *C.moreletii,* 111 *C.acutus)* developed for *Crocodylus* spp. (Fitzsimmons et al., 2000) (Table B1). We performed 3 multiplex reactions (Table B1) with the Type-it Microsatellite PCR kit (Qiagen) in a 12.5 µl reaction volume containing 6.25 µl of 2X Type-it multiplex PCR master mix, 1.25 µl of 10X primer mix with 2 µM of each primer, 3µl of DNA template and 2 µl of RNase-free water. The PCR thermocycler conditions for all the reactions were an initial HotStar Taq Plus DNA polymerase activation of 5 min at 95 °C, 28 cycles of 30s at 95 °C for denaturation, 90s at 59 °C annealing and 30s at 72 °C extension. With a final extension of 30 min at 60 °C. The presence of the products was visualized by electrophoresis in a 1.5% agarose gel.

We used 2 µl of PCR product and added 8.35µl of Applied Biosystems™ Hi-Di™ Formamide and 0.15 µl of GeneScan™ 500 LIZ™ Size Standard for a final genotyping volume of 10.5 µl. The genotyping was performed on an Applied Biosystems 3730xl DNA Analyzer at the McGill University and Génome Québec Innovation Centre, Montréal, Canada. Fragments were analyzed using GENEIOUS 11.0.5 (https://www.geneious.com). We used MICRO-CHECKER (Van Oosterhout et al., 2004) to identify null alleles, stutter peak scoring and typographic errors.

Restriction site Associated DNA (RAD) library preparation and sequencing

We used Single digest Restriction site Associated DNA Sequencing (sRAD-Seq) (Baird et al., 2008) to create genomic scans of populations from Mexico and Panama and generate Single Nucleotide Polymorphisms (SNP's). We used the restriction enzyme SbfI-HF (New England Biolabs) to digest 500 ng of DNA template of each sample. We used custom dual index adapters with unique combinations for each sample. We used a Covaris S2 for DNA fragmentation and size selected for 300 - 600 bp. Sixteen samples were pooled pre-enrichment at equimolar concentrations for each library. For each sequencing line, we pooled three post enriched libraries at equimolar concentrations. The libraries were sequenced with a HiSeq Illumina 2500 V4 paired-end 125bp at the McGill University and Génome Québec Innovation Centre, Montréal, Canada. The raw Illumina reads were demultiplexed and cleaned using *process_radtags.py* from the STACKS 1.46 package (Catchen et al., 2013). The RAD tags were aligned to the *C. porosus* reference genome (Green et al., 2014) using the BOWTIE2 aligner (Langmead & Salzberg, 2012). We used SAMTOOLS V.1.9 (Li et al.,

2009) to sort and filter sequences with a mapping quality score under 20. The SNP's were called using the *ref_map.pl* pipeline from stacks 1.46 requiring at least 4 reads to form a putative allele. To analyze the genotypes and generate populations statistics we used the populations module from STACKS 1.46.

Mitochondrial

We used the R software (R Core Team, 2019) to create the haplotype networks and calculate the frequency haplotypes with the packages PEGAS (Paradis, 2010) and APE (Paradis & Schliep, 2019). We used Arlequin 3.5 (Excoffier & Lischer, 2010) to calculate the nucleotide diversity (π) and haplotype diversity (H) and the number of mutation steps between haplotypes.

Microsatellites

We used the R software (R Core Team, 2019) to calculate population differentiation, inbreeding coefficient, heterozygosity, alleles per population and per loci with the packages PEGAS (Paradis, 2010) and HIERFSTAT (Goudet, 2005). We tested for deviations from Hardy–Weinberg equilibrium (HWE) with GENEPOP (R version 1.1.2; Rousset, 2008). To correct for the effect of sample size on allelic richness we produced unbiased estimates using the rarefaction method with the program HP-Rare (Kalinowski, 2004). To infer the number of genetic clusters and assign individuals to these clusters we used the Bayesian algorithm approach implemented in the package STRUCTURE 2.3.4 (Pritchard et al., 2000). We used the Admixture model with allele frequencies correlated clusters and without sample location priors. We used 6 replicates

for each genetic cluster (K) ranging from 1 to 20, with a 1 000 000 Markov chain Monte Carlo (MCMC) and an initial burn-in of 100 000 chains. The STRUCTURE result files were processed in STRUCTUREHARVESTER (Earl & vonHoldt, 2012) and executed the “Evanno” method (Evanno et al. 2005). To summarize and align the replicates for each K we used CLUMPP 1.1.2 (Jakobsson & Rosenberg 2007). We also ran a discriminant analysis of principal components (DAPC) to infer the number of clusters with the R software (R Core Team, 2019) and the package ADEGENET (Jombart, 2008). This method relies on data transformation using Principal Component Analysis (PCA) as a prior step to discriminant analysis (DA). The DA partitions genetic variation into a between-group and a within-group component and attempts to find groups that minimize the within-group variation (Jombart et al., 2010).

Restriction site Associated DNA (RAD)

We used the R software (R Core Team, 2019) to calculate population differentiation, number of alleles per population, and heterozygosity with the packages PEGAS (Paradis, 2010) and HIERFSTAT (Goudet, 2005). We ran a discriminant analysis of principal components (DAPC) to infer the number of clusters with the R software (R Core Team, 2019) and the package ADEGENET (Jombart, 2008). The mapped SNP’s were analyzed with the program FASTSTUCTURE to infer the population structure (Raj et al., 2014). We simulated from 1 to 20 genetic clusters, with a convergence criteria of 10 e-6, under the logistic prior which means that at a given locus, the population-specific allele frequency is generated by a logistic normal distribution, with the normal distribution having a locus-specific mean and a population-

specific variance (Raj et al., 2014). We used the *chooseK.py* option to choose the number of model components that explain the structure of our data.

Demographic Inference

For demographic inference, we used the Diffusion Approximation Demographic Inference (δaδi) (Gutenkust et al, 2009), which is based on the Allele Frequency Spectrum. To create the dataset for this analysis we used the *populations* module from STACKS, to a call a SNP, the SNP had to have a coverage of at least 8X, be present in all the populations and in 80% of the individuals. We retrieved 12,112 SNPs and removed individuals with more than 20% missing data. The joint SFS was projected to n-1 individuals in each population to avoid missing genotypes with the script easySFS (https://github.com/isaacovercast/easySFS). We used the folded SFS because we do not have an appropriate outgroup to determine the ancestral state of segregating sites. The demographic analysis with δaδi was performed pairwise on Banco Chinchorro with the other populations. We considered eight demographic models of historical divergence: Strict Isolation (SI); Isolation-with-Migration (IM); Ancient Migration (AM); Secondary Contact (SC) and their heterogeneous migration rates versions: IM2m; AM2m and SC2m, and one demographic model with no divergence: the standard neutral model (SNM) assuming the samples were collected in an unstructured, constant-sized population. To calculate the times of split and secondary contact, we used a mutation rate of 7.90E-09 and a generation time of 20 to 25 years (Green et al., 2014).

Results

Mitochondrial

We sequenced a total of 692 base pairs for the tRNAPro-tRNAPhe-D-loop region of the mitochondria. We found 34 polymorphic sites for all the populations and identified 11 haplotypes, six for *C.acutus* and five for *C. moreletii* (Figure 3). Two of the haplotypes were only observed once (singletons). There were 26 mutation steps between *C. moreletii* and *C. acutus*. Haplotype diversity (H) was non-existent in Cozumel (COZ), Sumidero (SUM) and Banco Chinchorro (BCH). On the other hand, H was large in the Populations of Altamira (ALT), Ria Lagartos (RLAG) and Sian Ka'an Biosphere reserve (RBSK), with the last two localized in the Hybrid zone (Table 1). Nucleotide diversity (π) was low for all populations but for the populations in the hybrid zone RLAG and RBSK (Table 1) RLAG presented haplotypes V, VIII and X of C. *acutus* and haplotype X of C. moreletii. The populations of RBSK, Huach (HUA) and Xcalak (XCAL) presented the haplotype VIII of *C. acutus* and haplotype II of *C. moreletii* (Figure 3).

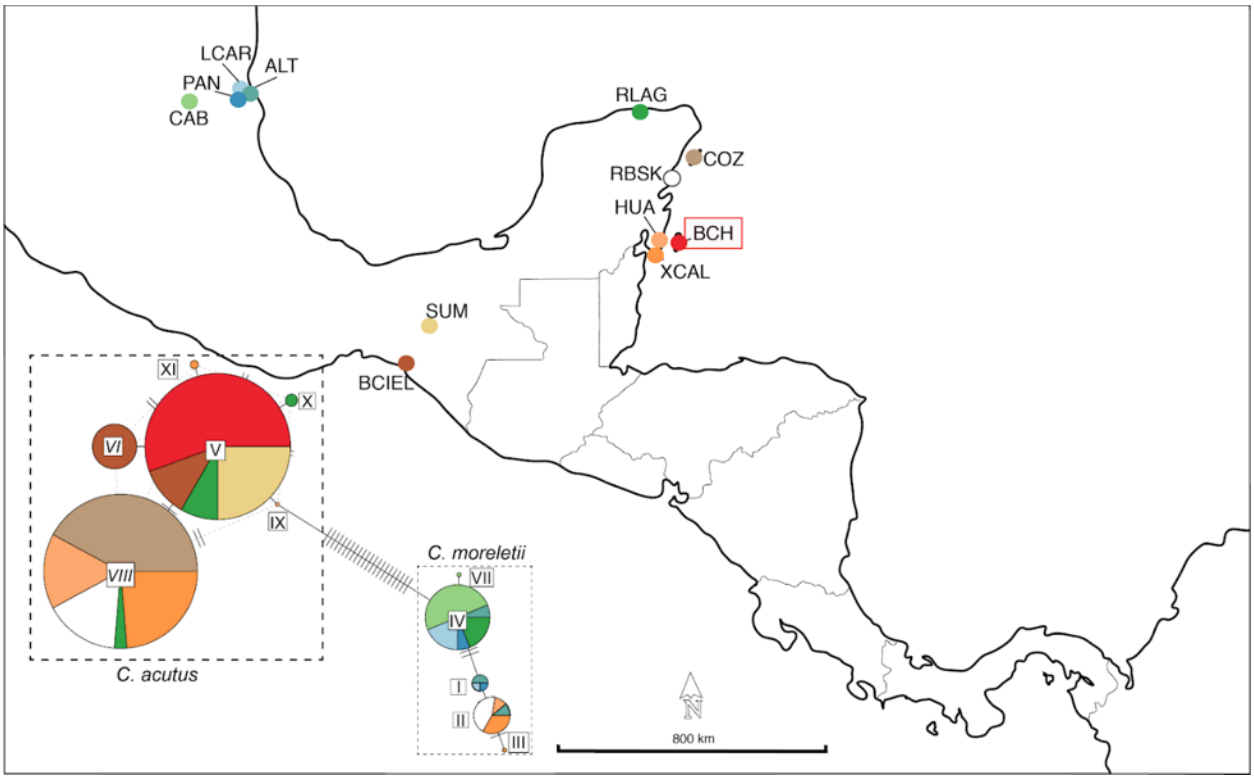

Figure 3. D-Loop haplotype network. The colors of the populations correspond to the colors of the contribution in the haplotype network. Abbreviations for population localities are shown in figure 2.

Table 1. D-Loop, sample size (n), number of haplotypes (k), number of polymorphic sites (PS), haplotype diversity (H) ± SD and nucleotide diversity (π) ± SD per population.

Species	Population	n	k	PS	H	π
C. moreletii	Altamira	5	4	7	0.900 +/- 0.161	0.0044 +/- 0.0032
	Panuco	2	2	3	1.00 +/- 0.50	0.0044 +/- 0.0051
	Carpintero	4	2	3	0.500 +/- 0.265	0.0022 +/- 0.0019
	Cabezas	9	2	1	0.222 +/- 0.166	0.0003 +/- 0.0005
C. acutus	Ria Lagartos	10	4	28	0.800 +/- 0.076	0.0192 +/- 0.0107
	Cozumel	16	1	0	0.0 +/- 0.0	0.0 +/- 0.0
	RBSK	11	3	29	0.618 +/- 0.104	0.0217 +/- 0.0119
	Banco Chinchorro	20	1	0	0.0 +/- 0.0	0.0 +/- 0.0
	Huach	7	2	29	0.286 +/- 0.196	0.0121 +/- 0.0073
	Xcalak	14	3	29	0.560 +/- 0.124	0.0159 +/- 0.0086
	Sumidero	9	1	0	0.0 +/- 0.0	0.0 +/- 0.0
	Boca del Cielo	15	2	1	0.419 +/- 0.113	0.0006 +/- 0.0006

Microsatellites

MICRO-CHECKER (Van Oosterhout et al., 2004) indicated potential null alleles and/or genotyping errors for the locus CUJ-131 and thus I removed this locus from the analysis. The number of alleles per population ranged from 24 to 69, mean number of alleles was 46.2. The number of alleles per loci ranged from 9 to 21, the mean number of alleles was 14.5. Observed heterozygosity per population ranged from 0.29 to 0.48, the inland populations Cienega de Cabezas (CAB) and Cañon del Sumidero (SUM) had the lowest heterozygosity. We detected deviations from Hardy Weinberg Equilibrium (HWE), showing an excess of homozygotes. We searched for a correlation between F_{IS}

and F_{ST} and we found a positive relationship between them (r^2=0.5118 p=0.01205), suggesting a Wahlund effect (Waples & Allendorf, 2015). Allelic richness after correction with rarefaction (Kalinowski, 2004) was of 3.39 ± 1.08, for the private alleles 0.35 ± 0.44. The lowest allelic richness was in the inland populations of CAB, SUM and Banco Chinchorro 2 (BCH2) (Table B2). The coastal populations have the highest allelic richness, and the island populations have lower allelic richness than coastal populations, but higher than inland populations (Table B2). Private allelic richness was high in COZ, Altamira (ALT) and RLAG (Table B2).

Structure

The Δk method identified three genetic clusters (mean Δk = 987.072) as the highest hierarchical structure, followed by four genetic clusters (mean Δk = 555.261). The two assignation plots revealed a unique *C. moreletii* cluster which includes the reported hybrid zone of RLAG (Cedeño-Vázquez et al, 2008; Rodriguez et al., 2008). *C. acutus* has a complex genetic structure composed of many genetic clusters. The analysis revealed that each island BCH1 and COZ has two distinct genetic clusters with limited gene flow (Figure 4 A and B).

Discriminant analysis of principal components (DAPC)

The DAPC identified nine genetic clusters and assigned most individuals to their putative population of origin, suggesting each population is differentiated. Some

individuals were assigned to geographically adjacent populations. *Crocodylus moreletii* is composed of two genetic clusters, one of these clusters is present in the Hybrid zone of Ria Lagartos. *Crocodylus acutus* has a complex structure with many different genetic clusters Cozumel island has a unique genetic cluster. Banco Chinchorro exhibits two discrete genetic clusters, which is consistent with the results from STRUCTURE (Figure 4C).

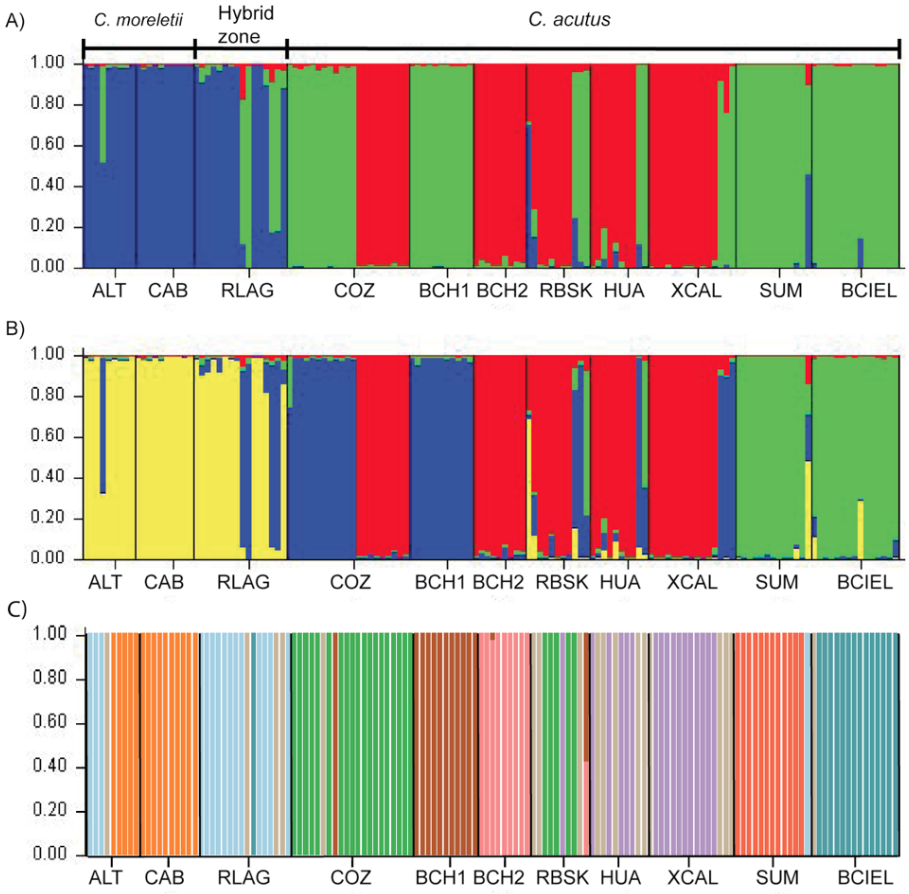

Figure 4. Individual assignation to genetic cluster based on the microsatellite results A) three genetic clusters; B) four genetic clusters with the Structure analysis C) nine genetic clusters with the DAPC analysis. The localities are group by species and the previously reported hybrid zone of Ria Lagartos (Cedeño-Vázquez et al., 2008; Pacheco-Sierra et al., 2016; Rodriguez et al, 2008).

Restriction site Associated DNA (RAD)

After filtering and data cleaning we used the genotypes of 218 crocodiles, the populations of Sian Ka'an was removed from the analysis. Because all the sequenced individuals failed the quality filters, which was likely caused by degraded DNA used for library preparation. The average coverage per sample was 13.82X. We retrieved more 12,112 SNP's for the 16 populations. We assumed two genetic clusters for Banco Chinchorro based on the microsatellite results for the RADseq analysis. The highest percentage of polymorphism was present in the inland populations of ALT, CAB, RLAG, HUA and SUM (table B3). The nucleotide diversity ($\boldsymbol{\pi}$) per population ranged from 0.048 to 0.309. Central America *C. acutus* populations have low nucleotide diversity (0.71-0.99). BCH2 and Boca del Cielo (BCIEL) had the lowest nucleotide diversity (Table B3). The inbreeding coefficient (F_{IS}) for the inland populations of CAB and SUM, and coastal populations of HUA and ALT was high, this suggests an excess of homozygotes in the populations under conditions of Hardy-Weinberg proportions. All the other populations had values close to zero suggesting no deviations from Hardy-Weinberg equilibrium (Table B3).

Pairwise F_{ST} values were high for the inland population of CAB compared to the other *C. moreletii* populations (F_{ST} = 0.35 - 0.38). F_{ST} values were lower compared to populations on the coast of Yucatan, COZ and BCH1, but was the highest when compared to BCH2 (F_{ST} = 0.43). The *C. moreletii* populations of PAN, LCAR and PAN were well differentiated from all populations (F_{ST} = 0.37 - 069), except from the RLAG hybrid zone (Table 2). Central America and Pacific *C. acutus* populations were differentiated from *C. moreletii*, but not from CAB, COZ, HUA and XCAL. BCH1 highly

differentiated from the Yucatan Coast and COZ (F_{ST} = 0.174 - 0.23) and it was even less differentiated from the *C. moreletii* population of Cabezas (F_{ST} = 0.069). BCH2 was opposite to BCH1, with high differentiation from *C. acutus* and very low differentiation from *C. moreletii* (F_{ST} = 0.097- 0.109) (Table 2). The BCH2 population was well-differentiated from Cabezas which is *C. moreletii* too. The differentiation of Banco Chinchorro populations (F_{ST} = 0.69) is as high as the differentiation between *C. moreletii* and Central America *C. acutus* (Table 2).

Table 2. Restriction site Associated DNA pairwise population F_{ST}.

Population		C. moreletii				Hybrid Zone	Yucatan C. acutus					Central America C. acutus				Pacific C. acutus	
		ALT	PAN	LCAR	CAB	RLAG	COZ	BCH1	BCH2	HUA	XCAL	GAL	BCI	LAG	COI	SUM	BCIEL
C. moreletii	ALT		0.032	0.036	0.305	0.110	0.465	0.548	0.097	0.376	0.583	0.640	0.640	0.680	0.678	0.593	0.698
	PAN	0.032		0.033	0.362	0.149	0.494	0.616	0.109	0.436	0.646	0.733	0.738	0.770	0.756	0.681	0.789
	LCAR	0.036	0.033		0.387	0.157	0.531	0.634	0.097	0.457	0.664	0.732	0.735	0.769	0.760	0.686	0.786
	CAB	0.305	0.362	0.387		0.191	0.143	0.069	0.432	0.100	0.199	0.241	0.236	0.299	0.305	0.220	0.289
Hybrid Zone	RLAG	0.110	0.149	0.157	0.191		0.331	0.388	0.220	0.219	0.406	0.440	0.433	0.487	0.495	0.397	0.498
Yucatan C. acutus	COZ	0.465	0.494	0.531	0.143	0.331		0.174	0.576	0.111	0.151	0.169	0.163	0.244	0.255	0.178	0.220
	BCH1	0.548	0.616	0.634	0.069	0.388	0.174		0.690	0.169	0.230	0.292	0.291	0.378	0.381	0.290	0.354
	BCH2	0.097	0.109	0.097	0.432	0.220	0.576	0.690		0.513	0.718	0.795	0.800	0.829	0.817	0.750	0.846
	HUA	0.376	0.436	0.457	0.100	0.219	0.111	0.169	0.513		0.065	0.164	0.158	0.233	0.236	0.150	0.205
	XCAL	0.583	0.646	0.664	0.199	0.406	0.151	0.230	0.718	0.065		0.176	0.174	0.285	0.286	0.192	0.226
Central America C. acutus	GAL	0.640	0.733	0.732	0.241	0.440	0.169	0.292	0.795	0.164	0.176		0.035	0.167	0.157	0.190	0.214
	BCI	0.640	0.738	0.735	0.236	0.433	0.163	0.291	0.800	0.158	0.174	0.035		0.180	0.168	0.194	0.222
	LAG	0.680	0.770	0.769	0.299	0.487	0.244	0.378	0.829	0.233	0.285	0.167	0.180		0.087	0.297	0.362
	COI	0.678	0.756	0.760	0.305	0.495	0.255	0.381	0.817	0.236	0.286	0.157	0.168	0.087		0.290	0.346
Pacific C. acutus	SUM	0.593	0.681	0.686	0.220	0.397	0.178	0.290	0.750	0.150	0.192	0.190	0.194	0.297	0.290		0.167
	BCIEL	0.698	0.789	0.786	0.289	0.498	0.220	0.354	0.846	0.205	0.226	0.214	0.222	0.362	0.346	0.167	

FASTSTRUCTURE

The model complexity that maximized the likelihood is to explain the population structure was ten genetic clusters (K=10) (Figure 5A). The FASTSTRUCTURE result shows a cluster that includes *C. moreletii* populations, except Cienega de Cabezas which has signs of admixture., because some individuals were assigned to multiple genetic clusters (Figure 5A) The RLAG population has a *C. moreletii* assignation with some admixture. Pacific and Central America *C. acutus* have similar cluster compositions. The Yucatan Peninsula coastal populations are admixed included the island of COZ. BCH1 is grouped in another well-defined cluster with a composition like

the one in Central America and Pacific *C acutus,* however, BCH2 is grouped as a pure cluster with a *C. moreletii* genetic signature (Figure 5A).

DAPC

The DAPC recognized 9 genetic clusters (Figure 5B) based on the Bayesian information criterion (Kstat =1522.35), the proportion of conserved variance is 0.956 and is explained by 3 discriminant functions. One clear cluster was *C. moreletii* excluding Cienega de Cabezas population. Ria Lagartos has its own genetic signature, different from all the populations, this may be a consequence of the hybridization present in the area. Cozumel Island is also a distinct genetic cluster, probably caused by the isolation of the island from coastal populations. Yucatan Coastal populations of Huach and Xcalak were accounted as a single genetic cluster. Central America *C. acutus* is divided into the Caribbean and the Panama Canal (Barro Colorado Island) and the Pacific (Lagartero and Coiba). Pacific *C. acutus* is a single genetic cluster. The results revealed two genetic clusters with limited gene flow in Banco Chinchorro, which is consistent with the FastStructure results. The difference is that BCH1 is grouped with CAB population cluster (Figure 5B).

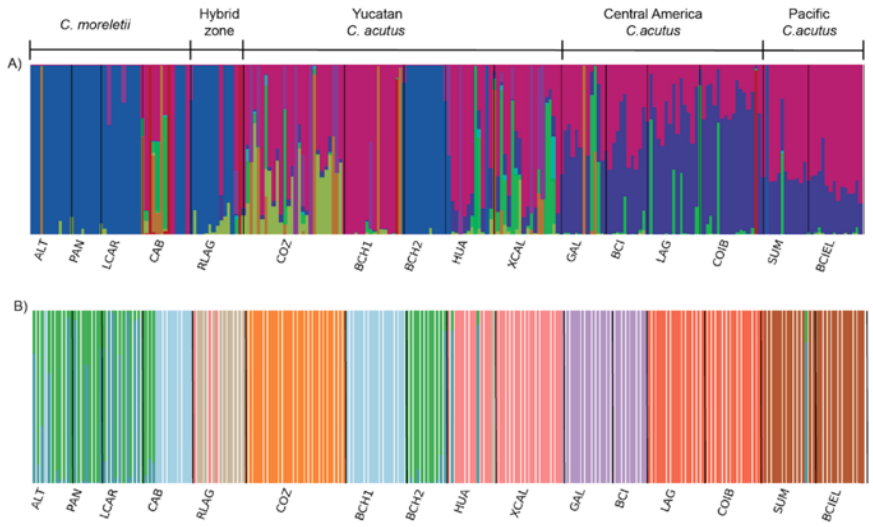

Figure 5. Restriction site Associated DNA individual assignation of A) Ten genetic clusters (K=10) of FASTSTRUCTURE; B) Nine genetic clusters (K=9) of the DAPC.

Demographic inference δaδi

The result for pairwise demographic analysis with δaδi is presented in Table 3. The pairwise analysis showed different demographic reconstructions for Banco Chinchorro populations. The calculated divergence time for population BCH1 from Central America *C. acutus* is 88.75 - 110.93 Kyr (Figure 6.A) and 153.32 - 29.65 Kyr for the BCH2 (Figure 6. B). For both populations (BCH1 and BCH2) the best model when compared to *C. moreletii* was a secondary contact with heterogeneous migration rates (SC2m). For the BCH1 the time of separation from *C. moreletii* was 8 times longer than the period of secondary contact. The migration rates and gene flow low, corresponding to less than one effective migrant per generation in the two directions (N1me12=0.018, N2me21=0.498). In comparison, the BCH2 time of separation from *C. moreletii* is only two times longer than the secondary contact. The migration rates and gene flow are high

(N1me12=6.617, N2me21=0.04), with six effective migrants per generation from C. moreletii to BCH2 (Table 3).

The direction of the migration for the other pairwise comparisons was from BCH1 to COZ (4.92), and from CAB to BCH1 (1.91) and Yucatan coast to BCH1 (8.0). The migration from BCH2 was in the direction out of the island (4.41 - 9.5) (Table 3). The times and direction of secondary contact are shown in Figure 6.

Table 3. Best model per population for the Demographic pairwise inference with δaδi. Abbreviations are: Akaike information criteria (AIC), the maximum-likelihood estimate, and the theta parameter for the ancestral population before the split. Following are the inferred values for the model parameters: the effective population size population 1 (N1) and population 2 (N2), the migration from population 2 to 1 (m12) and from 1 to 2 (m21), the effective migration in genomic islands from population 2 to 1 (me12) and from 1 to 2 (me21), the time of the split (Ts), the time of the secondary contact (Tsc) and the proportion of the genome evolving neutrally (P).

Population pairs BCH 1	Best Model	AIC	Log-likelihood	theta	N1	N2	m12	m21	me12	me21	Ts	Tsc	P
BCH1 - Moreletii	SC2M	2819.98	-1399.99	4494.31	0.019	0.010	2.214	0.005	0.957	48.826	0.025	0.003	0.000
BCH1 - Cienega de Cabezas	IM	4638.2	-2313.1	1954.23	0.010	0.080	1.908	0.207	-	-	0.005	-	-
BCH 1 - Cozumel	IM	1206.85	-597.42	1828.01	0.019	0.033	0.292	4.920	-	-	0.009	-	-
BCH1 - Yucatan Coast	IM	2385.05	-1186.53	2017.39	0.010	0.057	8.008	0.037	-	-	0.006	-	-
BCH1 - C.A. Caribbean C. acutus	SC2M	891.54	-435.77	3997.35	0.037	0.021	28.370	0.001	1.241	48.643	0.045	0.032	0.473
BCH2													
BCH2 - Moreletii	SC2M	1859.86	-919.93	509.47	0.206	3.271	3.470	0.602	32.200	0.012	1.023	0.440	0.982
BCH2 - Cienega de Cabezas	SC	2567.56	-1226.78	1347.25	0.117	0.733	0.581	4.409	-	-	1.961	0.067	-
BCH 2 - Cozumel	SC	1396.76	-691.38	1693.13	0.112	0.043	0.284	4.811	-	-	2.876	0.086	-
BCH2 - Yucatan Coast	SC	2970.01	-1478.01	4157.86	0.033	0.057	0.447	9.543	-	-	0.102	0.016	-
BCH2 - C.A. Caribbean C. acutus	SC2M	647.82	-313.91	4311.82	0.029	0.060	48.076	4.404	5.631	0.031	0.121	0.001	0.307

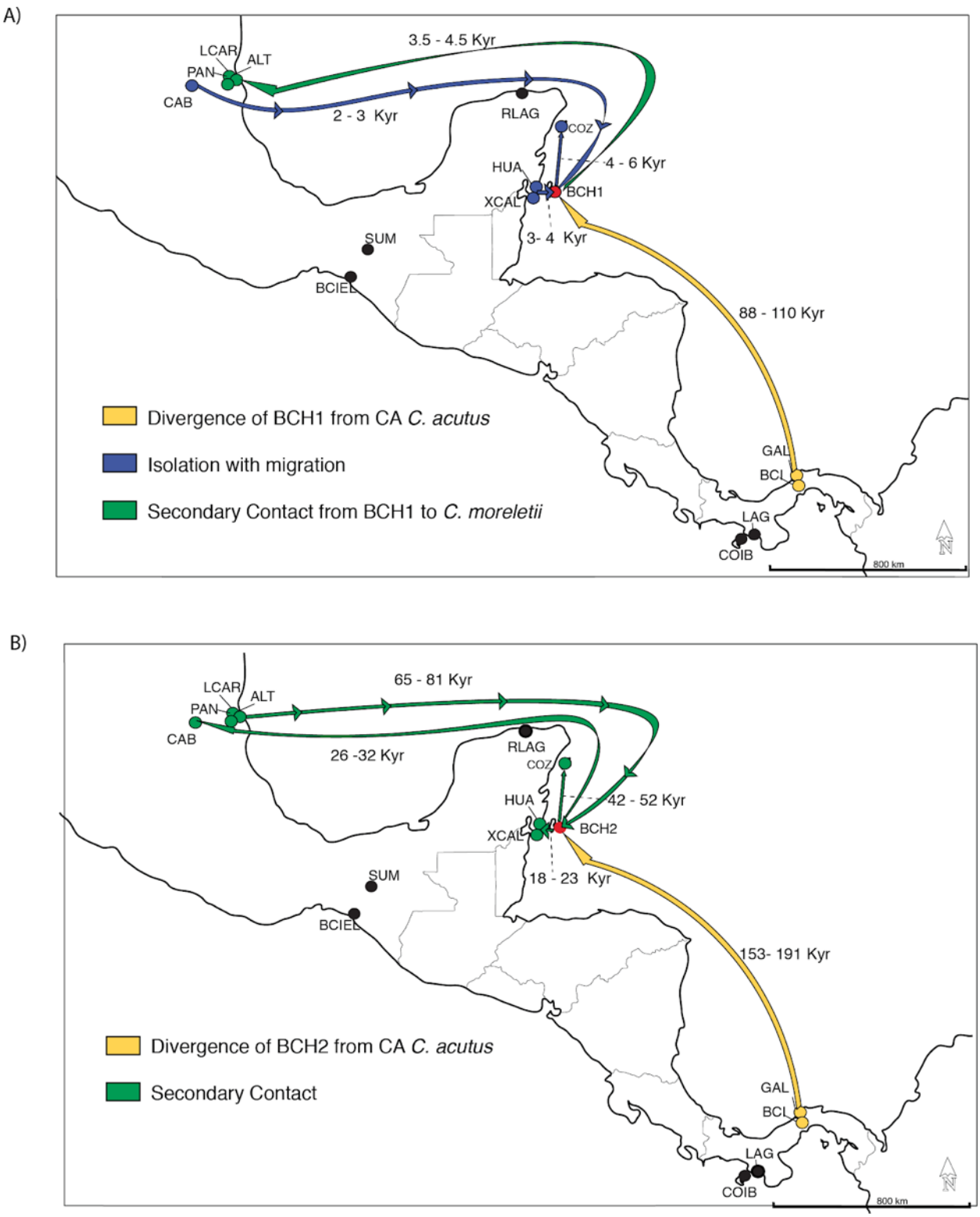

Figure 6. Demographic reconstructions with δaδi for Banco Chinchorro pairwise comparison, population A) Banco Chinchorro 1 and B) Banco Chinchorro 2.

Discussion

Our results show low levels of geographic differentiation for *C. acutus* and, considering it is a widely distributed species, the haplotype and nucleotide diversity for the D-loop are low, except for the populations on the Coast of the Yucatan Peninsula (Table 1), where the hybrid zone of *C. acutus* and *C. moreletii* is reported. The presence of haplotypes for the two species present in all the populations of the coast of Yucatan, confirms the extent of the hybrid zone reported in previous studies (Cedeño-Vázquez et al., 2008; Pacheco et al., 2018). The differences between haplotypes of the same species are a single nucleotide difference or two (Figure 3). Haplotype diversity is higher in other species of the order Crocodylia in the Neotropics but is consistent with what is found in widely distributed species of the genus *Crocodylus* in Africa and the Indopacific (Table 4).

Table 4. Haplotypes for species of the order Crocodylia

Species	Number of Haplotypes	Study
Caiman crocodylus	27	Venegas-Anaya et al., 2008
Paleosuchus trigonatus	36	Bittencourt et al,. 2019
Paleosuchus palpebrosus	22	Muniz et al., 2018
Crocodylus suchus	11	Hekkala et al., 2007
Crocodylus niloticus	5	Hekkala et al., 2007
Crocodylus porosus	10	Russello et al., 2007

Microsatellites show one clear *C. moreletii* cluster and a complex *C. acutus* that belongs to more than one genetic deme. The Ria Lagartos, Sian Ka'an, Huach and Xcalak populations show signs of admixture, individuals were assigned to more than one genetic cluster. This is consistent with results from previous studies (Pacheco et al., 2016; Rodriguez et al., 2008), but populations between Mexico and Belize were not

studied in the past, but now we confirm that the hybrid zone runs all over the coast of the Yucatan Peninsula and extends all the way down to South Belize where it was reported by Hekkala et al. (2015).

Banco Chinchorro Population

With microsatellites we detected the presence of two non-admixed genetic demes with limited gene flow, consistent with the results of Machkour-M'Rabet et al (2009). They used 77 inter-simple sequence repeat (ISSR) markers to assess the genetic status of Banco Chinchorro and found two genetic clusters. The biparental and mitochondrial markers support the idea of the population of Banco Chinchorro as isolated and pure populations. With the use of RAD sequencing, a technique that has been used successfully in ecological and evolutionary studies of non-model organisms (Andrews et al., 2016). we increased from a few markers to more than 12 thousand. We corroborated the presence of the two previously reported lineages with limited gene flow in Banco Chinchorro. All the individuals share the same *C. acutus* D-loop haplotype (V), which provides evidences that there is only one maternal lineage from one species. But the SNP's suggest a different story, where one of the lineages, BCH1, has a higher nucleotide diversity, higher number of polymorphic loci than BCH2 and a genetic signature like *C. acutus*. The second lineage, BCH2 has lower nucleotide diversity and less polymorphic loci than BCH1 and a genetic signature like *C. moreletii*. Additionally, the two populations are highly differentiated (F_{ST} = 0.69) for a small island.

The direction of hybridization between *C. acutus* and *C. moreletii* goes both ways in the Yucatan Peninsula (Rodriguez et al., 2008). For Banco Chinchorro island, there is

only one mitochondrial haplotype present, and is from *C. acutus*. The population BCH2 has a genetic signature of *C. moreletii* and a maternal lineage of *C. acutus*. Morphology of the skull is also atypical; crocodile populations present a broader-snouted cranial morphotype reported as unique for Banco Chinchorro. Hybridization might be contributing to this morphotype differences (Labarre et al., 2017). This could be evidence of unidirectional hybridization, which could be caused by prezygotic factors, like single hybridization event, size differences, sneak fertilizations, forced copulations, ecological and behavioural bias and the difference in discrimination intensity and preference for males of the other species (Wirtz, 1999). There are no postzygotic barriers reported for the genus *Crocodylus*, hybrids are viable, fertile and not maladapted (Hekkala et al., 2015; Milián-García et al., 2011; Pacheco-Sierra et al., 2016; Rodriguez et al., 2008; Weaver et al., 2008).

The size differences in *Crocodylus* seem to play a major role in mating: dominant males are territorial and would engage in confrontation for territory, and the largest males tend to be the dominant ones (Garrick, Lang, Zoologist, & Winter, 2008). Male *C. acutus* can reach lengths of 6–7m, in contrast, male *C. moreletii* can reach lengths of 3.6–4.0m (Perez-Higareda et al., 1991). But males in Banco Chinchorro only reach 3.5m (Personal communication with Pierre Charruau), suggesting that *C. moreletii* adults from the coast can outcompete *C. acutus* males of Banco Chinchorro.

Ocean Currents

The Genus *Crocodylus* has many adaptations to hyperosmotic environments. such as lingual salt-secreting glands (Taplin and Grigg 1981; Taplin et al. 1982; Taplin

1988), a heavily keratinized buccal epithelium (Taplin and Grigg, 1989), and an osmoregulatory cloaca (Pidcock, Taplin, & Grigg, 1997). The saltwater crocodile (*C.porosus*) has been recorded travelling from 200 up to 590 kilometres offshore Australia in less than a month (Campbell et al., 2010). These individuals rode surface water currents for long-distance travel as a low energy cost dispersal strategy (Campbell et al., 2010). The origin of banco Chinchorro crocodiles has been suggested to be Central America *C. acutus* (Machkour-M'Rabet et al., 2009), a possible explanation is that they drifted on the Caribbean current and settled on the atoll.

Banco Chinchorro Crocodiles might still be using the currents of the Mesoamerican Barrier Reef System (MBRS) for dispersal to other populations. The currents in the Mesoamerican Barrier Reef System are divided into a southern region with southward coastal currents moving to Belize and the Yucatan current which flows strongly from south to north into the Gulf of Mexico (Carrillo et al. 2015).

Demographic models

The demographic models support the origin of the populations in Banco Chinchorro from the Caribbean, facilitated by the Caribbean currents that move north to south as suggested by Machkour-M'Rabet et al. (2009). We estimated a split of BCH1 from Central America Caribbean *C. acutus* 88 - 110 kyrs and from BCH2 153 - 191 kyr. This suggests two colonization events on the island. The Yucatan current also facilitates gene flow from Banco Chinchorro to the coast, but the demographic stories for the two populations are different.

BCH1 is an isolated population with gene flow to Cozumel and from the coast of Yucatan. Hybrids from the coast have a higher tolerance to salinity (Cedeño-Vázquez et al., 2008), increasing their chances to migrate to Banco Chinchorro. There was gene flow from an ancestral *C. moreletii* with no hybridization, however the lack of samples from its distribution on the Gulf of Mexico limits our inferences. All these events happened in the last 2 to 6 kyrs (Figure 6A).

BCH 2 also had a secondary contact with the coast of Yucatan, Cozumel, and Cienega de Cabezas but from the *C. moreletii* coastal populations. These events have a larger time span than what happened to the BCH1 population (18 = 81 kyrs) (Figure 6B). The direction of the gene flow follows the Yucatan current direction from South to North. Carrillo et al. (2015) reported a loop that is formed in the southern part of the Yucatan Peninsula, which can also be moving individuals from Banco Chinchorro to Belize and Guatemala. The combination of maternal, nuclear polymorphic markers and SNP's show strong evidence of two discrete lineages, with a common origin but different demographic stories.

The estimated times of our calibrated populations split coincide with start of the Last Glacial Maximum (LGM), fluctuations on the sea level started 150 Kya and continued until 20 Kya (Rohling et al., 2017). The divergence of crocodile populations in the Neotropics might be associated to these changes in the sea level, that have been associated to divergence of crocodyilians (Mannion et al., 2015).

Taxonomy and Conservation

Crocodylus moreletii can tolerate salinities up to 22 ppt but is generally distributed on freshwater bodies and mainland coastal habitats (Escobedo Galvan et al., 2008: Hekkala et al., 2015); in comparison, *C. acutus* inhabits marine habitats with salinity as high as 34 ppt (Platt et al., 2013) but phenotypic hybrids of *C.acutus* can tolerate up to 41ppt (Cedeño-Vázquez et al., 2008). The mean salinity in Banco Chinchorro is 52.9 ppt (40-65ppt); which means crocodiles in this atoll can tolerate the highest salinity reported for any *Crocodylus* species (Charruau et al., 2005). Hybrid crocodiles also deposit larger clutches with larger eggs, and neonates have higher fitness than non-hybrid crocodiles in Belize (Hekkala et al., 2015). However, the clutch size (9-27), egg size (length= 61–81mm, width=40–48mm), and egg mass (60–102g) for Banco Chinchorro are lower compared to *C. acutus*. But the crocodiles on the island have a high nesting success (73%) due to the absence of nest predators (Charruau et al., 2010). The hybrids also have faster growth rates, larger adult sizes and enhanced survivorship (Cedeño-Vázquez et al., 2008). The appearance of novel phenotypes with extreme adaptations in hybrids is called transgressive segregation, where the hybrids have higher phenotypic values than parental populations (Rieseberg et al., 1999). These also fits on the model of Island speciation where there is exchange of genes at a sympatric stage (secondary contact), with introgressive hybridization (*C. moreletii* to *C. acutus*). The exchange of genes generates new genes that increase responsiveness to selection (Losos et al., 2010) in this case higher hyperosmotic tolerance than parental species.

Banco Chinchorro crocodiles could be considered a transgressive crocodile hybrid, due to their high tolerance to hyper-osmotic environments, higher nesting success, enhanced survivorship and no selection against the hybrids. A possible

explanation for these transgressive hybrids might be the chromosome number variation and the complementary action of additive alleles that are dispersed between the parental lines (Rieseberg et al., 1999). The fundamental number of chromosomes for *C. acutus* and *C. moreletii* are 56 and 58, respectively (Srikulnath, et al 2015). The chromosomal rearrangements (CRs) of *C. moreletii* are concordant with some predictions of the chromosomal speciation models suggested by Faria and Navarro (2010). The predictions fulfilled are 1) more karyotypic differences between sympatric sister species (*C. moreletii* and *C. acutus*) than between allopatric ones; and 2) these young *Crocodylus* Neotropical species have more CRs relative to molecular divergence compared with Australasian species. The rearrangement and chromosome fundamental numbers for the hybrids are unknown but studying them could reveal how additive alleles are complimented to produce these novel phenotypes. This high tolerance to salinity may increase the capacity of these hybrids to use surface currents to colonize new environments. This transgressive phenotype might be the source of all the hybridization reported in the Gulf of Mexico and Yucatan facilitated by the Ocean currents.

Banco Chinchorro is a protected area, with many protected reptile species by the International Union for Conservation of Nature (IUCN), Convention on International Trade in Endangered Species of wild fauna and flora (CITES) and the Mexican government (Charruau et al., 2015). *Crocodylus acutus* is listed as Vulnerable on the IUCN red list and Appendix 1 in CITES. The results of this study suggest that Banco Chinchorro has a population that should be considered as unique, and special protection efforts are needed to conserve these populations. Banco Chinchorro provides an

appropriate study system to study a naturally isolated pure hybrid population with high introgression from *C. moreletii*, and it is of great value to study how fitness, adaptability, and hybrid success can give rise to a new species?

Conclusions

We found two discrete lineages with limited gene flow in Banco Chinchorro with signs of a maternal *C. acutus* origin and Introgression from male *C. moreletii*. Our analysis reveals that it is not a pure *C. acutus* as reported in previous studies. The demographic history reconstructions reveal two colonization events from Central America *C. acutus* and secondary contact from coastal populations in Yucatan and facilitated by ocean currents in the Caribbean. Banco Chinchorro crocodiles are isolated, with population subdivision, and are a source of genetic variation to coastal populations. The characteristics of the hybrids suggest a transgressive hybrid population with higher fitness and enhanced dispersion capabilities generated by its high tolerance to hyperosmotic environments.

Bridging text

In Chapter 2 we evaluated the status of a previously reported pure population of *C. acutus* in Banco Chinchorro island in the Caribbean of Mexico. We used multiple genetic markers to reconstruct the history and demography of the population in the island. The use of multiple markers was needed considering the island is located near a hybrid zone. We rejected the idea of the island population as a hybrid and propose it as a pure hybrid population. We analyzed the effects of hybridization of two species in an isolated island, but with information for the species at a regional level. We found two populations with restricted gene flow, possibly in the process of speciation.

For Chapter 3, we moved from a regional scale to a local scale and focused our study on other species of Crocodile. The chapter focus is to evaluate the divergence of a single crocodile species in relation to the rise of the Panama isthmus. We use next generation sequencing, demographic modelling, the species biology and integrated geologic and climatic events to evaluate when and why populations of *C. acutus* diverged in Panama.

CHAPTER 3

Divergence of *Crocodylus acutus* in the Central American Isthmus

Abstract

Climatic and Geological events have shaped life on Earth throughout its history, the rise of the Central American Isthmus (CAI) is one of these events. It changed global circulation patterns, set the start of a glaciation, connected the biotas of South and North America and formed the Caribbean Sea. The nature of this event makes it the perfect natural scenario to test vicariance, divergence and speciation by allopatry. Many studies have shown the effect of the formation of this land bridge on marine and terrestrial species, but no studies have been made on semi-aquatic ones. The American crocodile is an amphibious species that arrived in the Neotropics before the complete closure of the CAI, and a candidate to test if the rise of the Isthmus had a divergent effect on the populations of the Pacific and the Caribbean. We used Single RAD sequencing on populations in Panama to: A) Detect the structure in the populations of the Caribbean and the Pacific B) Estimate divergence times and migration rates, and C) Evaluate the effect of the opening of the Panama Canal on *C. acutus*. We sampled individuals from 4 populations: the Caribbean coast, the Panama Canal, the Pacific coast and the Pacific Island of Coiba. We retrieved more than 17,000SNPs per population. We found 3 genetic demes: 1) Caribbean and the Panama Canal, 2) Pacific Coast, and 3) Pacific Island. The divergence times are not related to the rise of the CAI. We postulate the biology of the species played an important role on the resilience of the

species to this event. Rather, the divergence of the genetic demes coincides with the Last Glacial Maximum (LGM), an event that caused a sea level drop of 121 meters. Hydrological changes have shaped Crocodylian distribution and diversity since the appearance of the group and *C. acutus* is no exception. The LGM potentially affected the nesting and nursery sites restricting and isolating crocodile populations in Panama. We did not find alterations in the population structure caused by the reconnection of the Pacific and Caribbean, but mutation rates and long generation times of crocodiles may be masking this process.

Introduction

The rise of the Central American Isthmus (CAI) was a geological event that had a global scale impact. The main events were I) the beginning of thermohaline circulation (THC), II) the onset of northern hemisphere glaciation (NHG), III) The formation of the Caribbean Sea and IV) the Great American Biotic Interchange (GABI) (Jaramillo, 2018).

Thermohaline circulation changed as a consequence of the gradual closure of the Central America Seaway (CAS). In the late Miocene (11-5 Ma) when the uplift of the Isthmus was above 200 m there was water flow from the Pacific to the Caribbean via restricted lagoons or a shallow seaway (Sepulchre et al., 2014). The final closure of the CAS (3.6 - 2.7 Mya) increased the surface water salinity in the Caribbean, which overturned the thermohaline circulation and increased the salt transport to the North Atlantic (Mikolajewicz & Crowley, 1997). The changes in the thermohaline circulation favoured an early Pliocene warming of the Northern Hemisphere introducing moisture to

the Northern hemisphere, a condition that facilitated the ice-sheet growth (Haug & Tiedemann, 1998).

Isolation of the Caribbean by the uplift of the CAI (4.2 - 3.4 Ma), is reflected in today's high salinity in the Caribbean, small interannual and seasonal variability, lack of upwelling, and low planktonic productivity. In contrast, the Pacific of Panama has low salinity, small interannual and seasonal variability, upwelling and high planktonic productivity (O'Dea et al., 2007). These and the reorganization of South American drainages at the beginning of the NHG have been associated with changes in composition and extensive reef development in the Caribbean (Jaramillo, 2018).

The rise of the CAI connected South America and North America and divided the Atlantic and Pacific Oceans. This triggered many biogeographical events like the Great American Biotic Interchange (GABI) (Jaramillo, 2018). The effects were a split of marine species into lineages in the Pacific and Atlantic Oceans, and increased dispersal of terrestrial species between North and South America. The dispersal events across the isthmus are related to environmental factors rather than biological, like land availability, sea and freshwater corridors and suitable climates and environment for migrant species (Bacon et al., 2015). But the Isthmus per se does not seem to be the sole factor for the divergence of species and other aspects like physical environment (salinity, upwelling, tides evapotranspiration), biotic environment (shallow waters) and the evolutionary history and the diversification rate specific to each group also account for divergence (Lessios, 2008). But what happens with semi-aquatic organisms, as in this case with the American crocodile (*Crocodylus acutus*)?

Crocodylus acutus is distributed in the Tropics and subtropics. Populations are present in the Pacific Coast from Northern Mexico to Peru and in Florida, the Caribbean islands and the Caribbean coast all the way down to Colombia. It inhabits coastal lagoons, estuaries, hypersaline lakes, freshwater bodies and rivers, cays and islands (Thorbjarnarson et al., 2006). The conditions of these habitats vary in salinity, land availability, climates and environments.

Crocodilian diversity in Northern South America was at its peak during the Miocene, with representatives of giant caimanines, crocodylids, gavialoids and sebecids. These occupied a variety of niches and their diets were differentiated by their ecomorphology. Of this diversity only four species remain in the area (Riff et al., 2009). The diversity peak came to an end with the massive extinction, probably by changes in the hydrography caused by the Andean uplift (Scheyer et al., 2013). Gryposuchines, a subfamily of extinct gavialids, inhabited Panama and Venezuela, and may have had adaptations to osmoregulate and prosper in marine habitats. A high diversity of this group may have been present in the Caribbean during the Miocene, where *Crocodylus* is the dominant group at present (Salas-Gismondi et al., 2018). Extant *Gavialis* can only inhabit freshwater environments (Taplin, 1998). The closure of the CAS and the associated increase of salinity may have played an important role in the extinction of gryposuchines, and the subsequent establishment of *Crocodylus*, but this is merely speculative.

During this period *Crocodylus* dispersed to the Neotropics from Africa and presumably occupied the niches left by other crocodilian species (Oaks, 2011). The arrival of *Crocodylus* in the Netropics (2.8 -8.3 Ma) is estimated before the total rise of

the CAI 2.8 Ma and closure of the Central American Seaway (CAS)(O'Dea, 2016). The rise of this land bridge that separates the Pacific and Atlantic Oceans, may have facilitated the vicariance and formation of two divergent lineages of crocodiles.

The rise of the CAI caused major changes on marine species composition by the separation of the two Oceans. In effect the opening of the Panama Canal, an artificial waterway, re-connected the Atlantic and the Pacific Oceans in 1914. This reconnection impacted marine and coastal faunas. For example, the coral reefs declined and coastal areas were affected after the opening of the canal (Guzman et al., 2019), there was an increase of biological invasions (Muirhead et al., 2015), fish communities changed (Sharpe et al., 2019; Smith et al., 2004), avifauna suffered local extinctions (Robinson, 2001) and population divergence and structuration has been detected in Geoffroy's Tamarins (Díaz-Muñoz, 2012). However, the effects of a secondary contact on semi aquatic species distributed on both sides of the Panama Canal has not been studied.

Our study questions are: Did the closure of the Central American Isthmus drive the emergence of two divergent lineages of crocodiles in the Caribbean and the Pacific? When did Pacific and Caribbean crocodiles diverge? What is the origin of Barro Colorado crocodile populations? Did the opening of the Panama Canal facilitate a secondary contact of crocodiles on both sides of the isthmus?

H_0: The Rise of the Central American Isthmus and complete closure of the Central American Seaway, may have interrupted the gene flow between the Caribbean and Pacific crocodile populations.

H_A: The Rise of the Central American Isthmus and complete closure of the Central American Seaway, did not affect the gene flow between the Caribbean and Pacific crocodile populations.

Methods

Tissue collection and DNA extraction

Tissue samples were acquired from 4 localities in Panama: Galeta (n=12), Barro Colorado Island (n=11), Lagartero (n=14) and Coiba island (n=17) (Figure 1). We captured a total of 54 individuals of *C. acutus*. We extracted the DNA using a phenol:chloroform:isoamyl alcohol protocol (Bardakci & Skibinski, 1994). The DNA quality was assessed by electrophoresis using 1% Ultra-Pure Agarose gel (Invitrogen). The purity and quantity of the template were measured with a Qubit Fluorometer (Thermofisher).

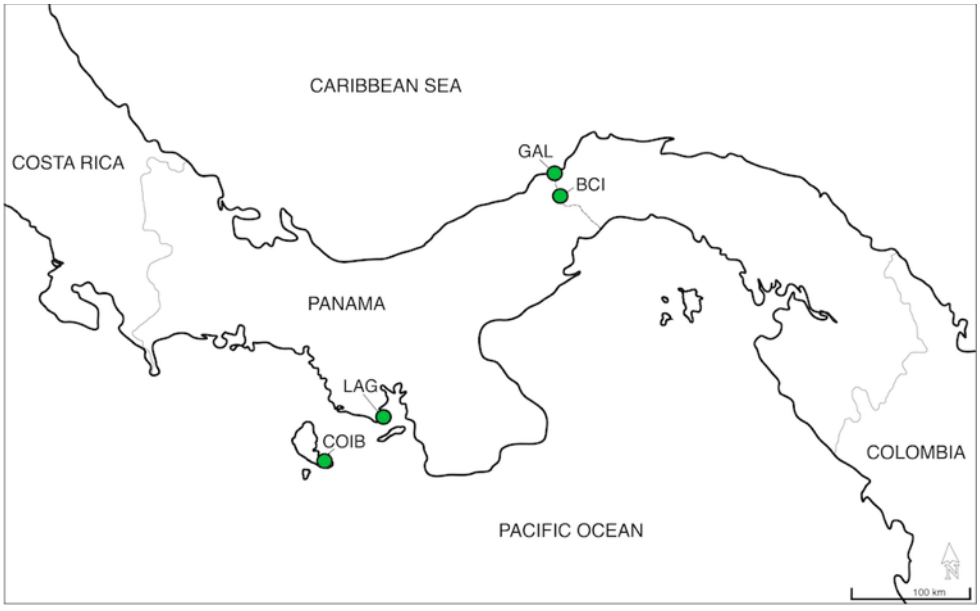

Figure 1. Localities sampled in Panama. Galeta (GAL), Barro Colorado Island (BCI), Lagartero (LAG) and Coiba (COIB)

Restriction site Associated DNA (RAD) library preparation and sequencing

We used Single digest Restriction site Associated DNA Sequencing (sRAD-Seq) (Baird et al., 2008) to create genomic scans of populations from Panama and generate Single Nucleotide Polymorphisms (SNPs). We used the restriction enzyme SbfI-HF (New England Biolabs) to digest 500 ng of DNA template of each sample. We used custom dual index adapters with unique combinations for each sample. We used a Covaris S2 for DNA fragmentation and size selected for 300 - 600 bp. Sixteen samples were pooled pre-enrichment at equimolar concentrations for each library. For each sequencing line, we pooled three post enriched libraries at equimolar concentrations. The libraries were sequenced with a HiSeq Illumina 2500 V4 paired-end 125bp at the McGill University and Génome Québec Innovation Centre, Montréal, Canada. The raw Illumina reads were demultiplexed and cleaned using *process_radtags* from the Stacks 1.46 package (Catchen et al., 2013). The RAD tags were aligned to the *C. porosus* reference genome (Green et al., 2014) using the BOWTIE2 aligner (Langmead & Salzberg, 2012). We used SAMTOOLS V.1.9 (Li et al., 2009) to sort and filter sequences with a mapping quality score under 20. The SNP's were called using the *ref_map.pl* pipeline from STACKS 1.46 requiring at least 4 reads to form a putative allele. To analyze the genotypes and generate population statistics we used the *populations* module from STACKS 1.46.

Data Analysis

Restriction site Associated DNA

We used R software (R Core Team, 2019) to calculate basic population statistics with the packages PEGAS (Paradis, 2010) and HIERFSTAT (Goudet, 2005).
The mapped SNP's were analyzed with the program FASTSTRUCTURE to infer the population structure (Raj et al., 2014). We simulated from 1 to 10 genetic clusters, with a convergence criteria of 10 e-6, under the logistic prior, which means that at a given locus, the population-specific allele frequency is generated by a logistic normal distribution, with the normal distribution having a locus-specific mean and a population-specific variance (Raj et al., 2014). We used the *chooseK.py* option to choose the number of model components that explain the structure in the data set. As an alternative method to find the number of clusters we used discriminant analysis of principal components with R software (R Core Team, 2019) and the package ADEGENET (Jombart, 2008).

Approximate Bayesian Computation

We used an Approximate Bayesian Computation implemented in DIYABC (Cornuet et al., 2014) to infer the demographic history of the populations of the Pacific and the Caribbean. For each locus, a minimum of one genotyped individual per population is required for DIYABC to run, and a maximum of 1000 SNP's can be used in the simulations (Cornuet et al., 2014). To create the dataset for this analysis we used the populations module from STACKS, to a call SNP it had to have a coverage of 12X, be present in all the populations in all the individuals. This produced a dataset of 647

polymorphic loci. We first tested six simple invasion history scenarios. The conditions were: two separation times: t2, when an ancestral population gives origin to a second population, and t1 when this ancestral population gives origin to a third population (Figure C1)

We tested a divergence model with one scenario with two populations of size N1 and N2 (30,000) diverged t generations (upper limit of 120,000) in the past from an ancestral population of size N1 (Pacific Panama) + N2 (the Caribbean and BCI) (Figure C2). To estimate the parameters N1 and N2 we used a range of 1 to 30,000, based on previous estimations of effective population size from Diffusion Approximation Demographic Inference of Chapter 2. For the number of generations, we used a range from 10 to 120,000 generations using a generation time of 20 to 25 years (Green et al., 2014). The estimated complete closure of the Central American Isthmus of 2.8 Mya (O'Dea et al., 2016) falls into this range.

Diffusion Approximation Demographic Inference

We also used the Diffusion Approximation Demographic Inference (δaδi) (Gutenkust et al, 2009), which is based on the Allele Frequency Spectrum. To create the dataset for this analysis we used the populations module from STACKS, to call a SNP the conditions were: to have a coverage of 8X be present in all the populations and in 80% of the individuals. We retrieved 17,312 SNPs and removed individuals with more than 20% missing data. The joint SFS was projected to n-1 individuals in each population to avoid missing genotypes with the script easySFS

(https://github.com/isaacovercast/easySFS). We used the folded SFS because we do not have an appropriate outgroup to determine the ancestral state of segregating sites. We considered seven demographic models of historical divergence for the pairwise comparisons: Strict Isolation (SI), Isolation-with-Migration (IM), Ancient Migration (AM), Secondary Contact (SC) and their heterogeneous migration rates versions: IM2m, AM2m and SC2m, and one demographic model with no divergence: the standard neutral model (SNM) assuming the samples were collected in an unstructured, constant-sized population.

Results

There was a higher nucleotide diversity (0.30) more polymorphic sites (>14,500) and a higher percentage of polymorphic loci (22%) in the Caribbean and Barro Colorado populations, compared to the Pacific Panama population (Table 1). Inbreeding coefficient values (F_{IS}) were close to zero for all populations. We did not detect an excess or a deficiency of heterozygotes. All populations had private allels with the highest number of private alleles present in Coiba Island (234). (Table 1)

The differentiation is high between Pacific and Caribbean populations (F_{ST} = 0.151-0.167), compared to the differentiation between Caribbean and Barro Colorado (F_{ST} = 0.039) and between Pacific populations (F_{ST} = 0.084) (Table 2).

Table 1. Restriction site Associated DNA nucleotide diversity(π), inbreeding coefficient (FIS) and polymorphic sites per population.

π	StdErr	Fis	StdErr	Polymorphic Sites	% Polymorphic Loci	Private alleles
0.3033	0.0014	0.0516	0.0056	14812	0.2284	55
0.3001	0.0014	0.0077	0.0041	14562	0.2246	48
0.2298	0.0015	0.0258	0.0047	11748	0.1812	206
0.2362	0.0015	0.0208	0.0065	12471	0.1923	234

Table 2. Restriction site Associated DNA pairwise population F_{ST}.

Population		Caribbean	Panama Canal	Pacific	
		GAL	**BCI**	**LAG**	**COI**
Caribbean	**GAL**		0.039	0.167	0.151
Panama Canal	**BCI**	0.039		0.174	0.156
Pacific	**LAG**	0.167	0.174		0.084
	COI	0.151	0.156	0.084	

FASTStructure and Discriminant Analysis of Principal Components

The FASTstructure model complexity that maximizes the likelihood for the model is the presence of 3 genetic clusters. There was one deme that included all the individuals from the Caribbean and the Panama Canal, and another cluster for Coiba island. However, the Pacific coastal population of Lagartero was conformed of two clusters, its own cluster but with some individuals assigned to Coiba Island (Figure 2 A).

The cluster assignation based on a PCA with 50 components and the Aikake information criterion (AIC) found 3 genetic clusters (AIC = 426.25). The discriminant analysis of Principal Components (DAPC) retained 45 components and 2 linear discriminants. The analysis conserved 0.916 of the variances in the data. One cluster

was again the Caribbean and Panama Canal population, one for Coiba and a third for Lagartero with some individuals assigned to Coiba (Figure 2 B and C)

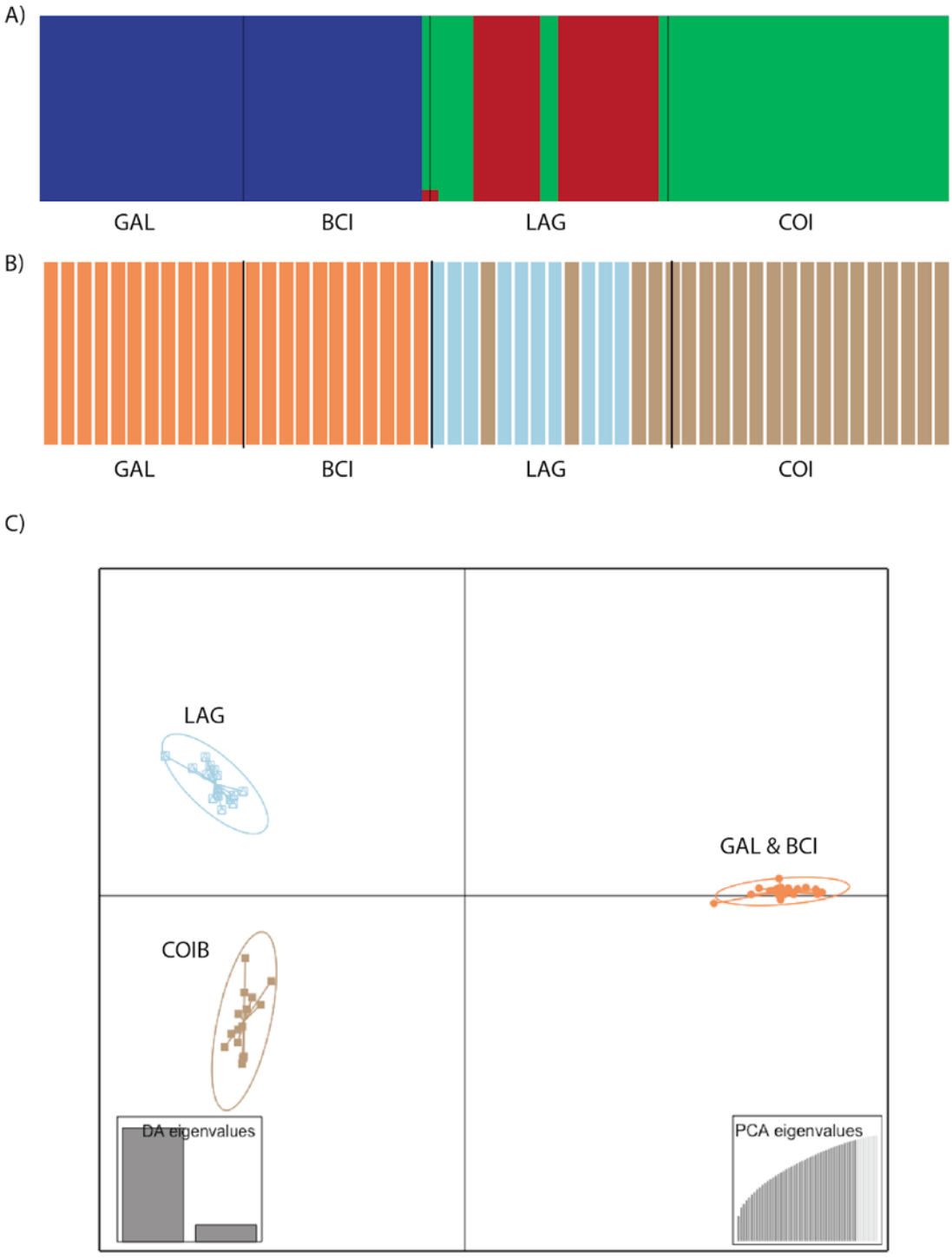

Figure 2. Restriction site Associated DNA individual assignation of A) Three genetic clusters (K=3) of FASTstructure; B) Three genetic clusters (K=3) of the DAPC C) Scatter plot based on the DAPC analysis.

DIYABC

The ABC best model from the six scenario models is scenario number 2 (Figure C3). The putative ancestral population Galeta (Caribbean) which had an estimated Ne of 14,500 individuals. It separated into Lagartero (Pacific) 1,700 generations ago (± 34 Kyr), with an effective population size of 1,680 and into BCI (Panama Canal) 543 generations ago (± 11 Kyr), (Ne = 6,920) (Figure C3). The two population divergence scenarios estimated an effective population size of 33,000 individuals for the Caribbean and 5,780 for the Pacific, and a separation time of Pacific and Caribbean populations of 4,850 generations (± 97 Kyr) (Figure 3A; C2).

Diffusion Approximation Demographic Inference

The δaδi pairwise demographic inference estimates a separation of BCI and Galeta of 407 to 507 years. The best model is isolation with the migration of 3 effective migrants per generation (EMPG) from BCI to Galeta (Table 3). Barro Colorado separated from the Pacific 37.3 - 46.7 Kyr and had a secondary contact 7.2 - 9 Kyr. The direction of the migration was from the Pacific Coast and Coiba to the Canal of Panama with one EMPG (Table 3).

Galeta (Caribbean) separated from the Pacific coast population 57.3 - 71.7 Kyr and had a secondary contact 8.5 - 10.6 Kyr with migration from the Pacific to the Caribbean (EMPG = 3) and from the Caribbean to the Pacific (EMPG = 1). The Galeta split from

Coiba was 16.5 to 20.6 Kyr with a secondary contact 14.9 - 18.6 Kyr. The direction of the flow was from the island to the Caribbean (EMPG = 1). The split of the Pacific population of the coast with the island was estimated to 2.2 - 2.7 Kyr, with the migration of 7 individuals from the island to the coast (Table 3). When we considered the populations of the Caribbean as one population and the Pacific as another, the estimated time of the split was 27.1 - 33.9 Kyr, with one effective migrant per generation from the Pacific to the Caribbean (Table 3; Figure 3B).

Table 3. Best model per population for the Demographic pairwise inference with δaδi. Akaike information criteria (AIC), the maximum-likelihood estimate, and the theta parameter for the ancestral population before the split. Following are the inferred values for the model parameters: the effective population size population 1 (N1) and population 2 (N2), the migration from population 2 to 1 (m12) and from 1 to 2 (m21), the effective migration in genomic islands from population 2 to 1 (me12) and from 1 to 2 (me21), the time of the split (Ts), the time of the secondary contact (Tsc) and the proportion of the genome evolving neutrally (P).

Population pairs	Best Model	AIC	Log-likelihood	theta	N1	N2	m12	m21	me12	me21	Ts	Tsc	P
BCI - GAL	IM	4214.129	-2101.064	4362.252	0.010	0.016	0.000	3.291	-	-	0.001	-	-
BCI - LAG	SC2M	990.820	-485.410	6540.371	0.069	0.044	18.060	0.017	49.657	26.045	0.023	0.005	0.482
BCI - COI	SC2M	1023.279	-501.640	6977.129	0.022	0.017	31.692	0.064	49.439	13.029	0.000	0.010	0.052
GAL - LAG	SC2M	1382.306	-681.153	6315.476	0.113	0.066	32.258	0.553	14.268	48.915	0.038	0.007	0.693
GAL - COI	SC2M	1129.704	-554.852	7045.266	0.024	0.018	43.318	9.982	49.782	44.526	0.001	0.010	0.779
LAG -COI	IM	2231.844	-1109.922	4123.055	0.014	0.019	0.000	7.073	-	-	0.003	-	-
Caribbean - Pacific	SC2M	4927.642	-2453.821	6589.714	0.037	0.044	18.875	8.950	49.502	0.025	0.002	0.018	0.241

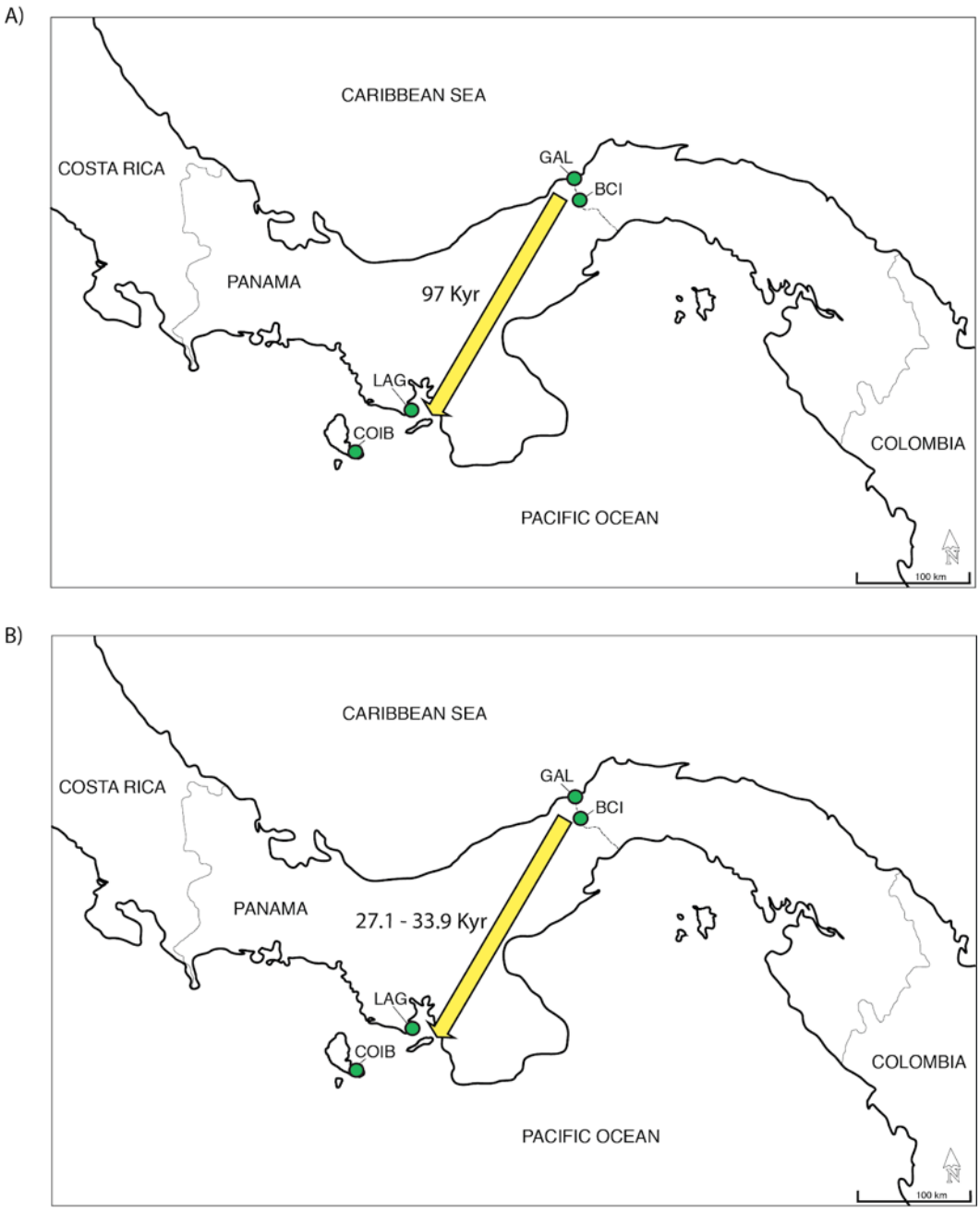

Figure 3. Demographic reconstructions for the split of the Caribbean and the Pacific A

A) DIYABC B) δaδi

Discussion

Pacific and Caribbean divergence caused by the rise of the CAI

We found three different genetic clusters: I) Caribbean and the Panama Canal, II) Pacific coast and III) Pacific island (COIBA). The differences in population structure between Caribbean and Pacific clusters can be attributed to the presence of the Isthmus, as an effective barrier for migration from one side to the other. As previously described for marine organisms (echinoids, crustaceans, fishes and molluscs) of Pacific and Atlantic species pairs through the rise of the CAI (Lessios 2008).

The species has a variety of adaptations for osmoregulation including lingual salt-secreting glands (Taplin and Grigg 1981; Taplin et al. 1982; Taplin 1988), a heavily keratinized buccal epithelium (Taplin and Grigg, 1989), and an osmoregulatory cloaca (Pidcock, Taplin, & Grigg, 1997). Additionally, it is the species with the highest tolerance to salinity, 34 ppt (Platt et al., 2013). *C. acutus* foraging environments are mangrove swamps, hypersaline lagoons, open tidal flats, turtle grass beds and reef barriers, and they can pray on insects, crustaceans and vertebrates (Platt et al., 2013).

Two main factors have caused extinction and diversification of crocodylians 1) Changes in temperature followed by aridification and 2) Changes in sea level. In the Miocene changes of crocodylian assemblages in Africa, coincided with the formation of the Sahara, and in South America with hydrographic changes driven by the Andean uplift (Mannion et al., 2015). The biology of the species makes it well adapted to all the changes caused by the rise of the CAI on the Caribbean and the Pacific and we cannot associate these changes with the emergence of two divergent lineages of crocodiles.

Timing of the Pacific and Caribbean divergence

The demographic models suggest a split between Caribbean and Pacific populations 97 Kyr. When we calibrated the results with mutation rates and generational time, the estimate is 27.1 - 33.9 Kyr. The time inferred in this study contrast with 2.5 My of divergence reported in previous studies (Pacheco et al., 2018), an explanation to this difference could be the use of pure C. acutus populations from a non-hybrid zone in our study. Two factors can explain this relatively recent divergence. 1) The Isthmus is a semipermeable barrier and 2) Changes in sea levels.

The Isthmus of panama at its narrowest point where the Panama Canal is situated is not wider than 60 kms. The Chagres drainage before the opening of the Panama Canal flowed along the Atlantic slope of the Isthmus and had four main tributaries: the Chagres, the Trinidad, la Chorrera and Gatun. While the Pacific slope was drained by small coastal streams (Meek and Hildebrand, 1916) (Figure 4). The semi-aquatic nature of crocodiles allows them to travel along small drainages (Cherkiss et al, 2014), and migrations of crocodiles are also associated with rainy seasons (Calverley & Downs, 2015). Crocodiles can also move on-land and the distance between the Pacific and Atlantic drainages could have likely been traversed by land in the rainy season. However, there are no reports of in-land migrations of crocodiles, therefore this scenario is merely speculative.

The estimated times of our calibrated population split coincide with the Last Glacial Maximum (LGM), when the ice sheets reached their maximum (10 - 50 Kyr) (Clark et al., 2009). Sea levels in the Caribbean and Mediterranean were ± 121 meters at its lowest point 26 Kya and raised gradually to the present level (8 Kya) (Peltier &

Fairbanks, 2006; Rohling et al., 2017). Changes in sea level were previously associated with the extinction and divergence of crocodyilians (Mannion et al., 2015). Successful nesting areas for *C. acutus* depend on elevated beach ridges and nearby brackish lagoons as nursery habitats (Platt & Thorbjarnarson, 2010). The availability of these areas during the LGM is unknown, but changes in sea level may have restricted the home range and availability of nesting areas in Central America, thereby isolating populations in patches where suitable habitats were available.

Other evidence of the effect of the LGM is the split of Pacific Coast and Coiba populations. Coiba was part of the continent when sea levels were 100 to 120 m below current levels in the late Pleistocene (15–18 kyr) (Cardiel et al., 1997). The complete isolation may have been reached 8 Kyr when the sea reached its current levels (Peltier & Fairbanks, 2006; Rohling et al., 2017). The island is isolated because it does not receive migrants, but migrations from the island to the coast were corroborated with the clustering analysis (Figure 2).

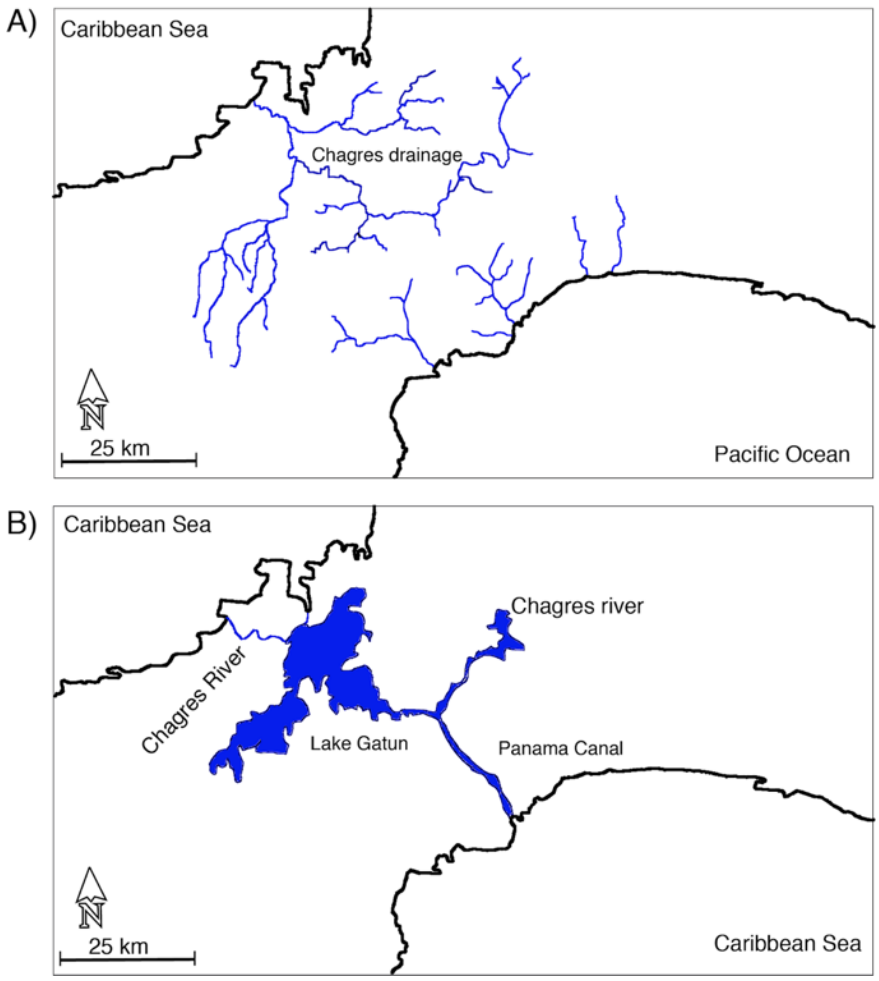

Figure 4. The Chagres drainage A) Before the Panama Canal; B) After the Panama Canal (Modified from Díaz-Muñoz, 2012).

Effect of the opening of the Panama Canal

The genetic clusters identified from both sides of the Isthmus show no admixture. The direction of the secondary contact from the Pacific to the Caribbean and the estimated times of the effective migrations are older than the opening of the Panama Canal. The population of Barro Colorado Island is part of the Caribbean lineage, its origin is from the Chagres drainage (Figure 4) and it is supported by our two-clustering analyses and demographic models (Figure 2; C3). The long generational time and low mutation rate of the genus *Crocodylus* (Green, 2014) may also be obscuring the

admixture of the lineages and long-term monitoring would be required to address this question.

Conclusions

The complete closure of the CAI did not have any traceable effect on the split of crocodile populations in the Pacific and the Coast. The biological adaptations of *C. acutus* to a variety of environments prevented the divergences as seen in other organisms, which were less adapted to the changes caused by the CAI. The divergence of Pacific, Atlantic and Island populations coincide with the LGM. The environmental effects of this period changed variables directly related to previous crocodilian extinctions. The population structure caused by the LGM has no detectable changes from the Anthropocene modifications, including the opening of the Panama Canal.

GENERAL DISCUSSION

Much can be gained from the genomic-scale approaches introduced in this thesis to address the evolution and ecology of *Crocodylus*. I want to approach the general discussion of the three chapters of this thesis from an eco-evolutionary point of view.

Adaptive or nonadaptive radiation in Crocodylus

What is adaptive radiation? It can be defined as the evolution of ecological and phenotypic diversity within a rapidly multiplying lineage. A single ancestor diverges into a host of species that use a variety of environments and differ on the traits used to exploit these environments (Schlutter, 2000). Four features are necessary for adaptive radiation to be detected (Schlutter, 2000): recent common ancestry, phenotype-environment correlations of adaptive traits, trait utility, and rapid speciation.

Thirteen extant species of *Crocodylus* are distributed in the tropics and subtropics of the world (Grigg & Kirshner, 2015). The genus is monophyletic, and all species are descendants of an ancestor close to *Crocodylus palaeindicus*, a crocodile from the Miocene of present day Pakistan and India (Brochu, 2000). The rapid radiation of this group all over the globe happened through transoceanic dispersal in the late Miocene and Pliocene, in a time span of three to five million years (Meredith et al., 2011; Oaks, 2011).

An important feature present in all crocodiles that matches well with their environment is their physiological tolerance to hyperosmotic conditions (Taplin, 1988). The salt tolerance of *Crocodylus* is based on lingual salt-secreting glands (Taplin and Grigg 1981; Taplin et al. 1982; Taplin 1988), a heavily keratinized buccal epithelium

(Taplin and Grigg, 1989), and an osmoregulatory cloaca (Pidcock, Taplin, & Grigg, 1997). The slender-snouted morphology of *Crocodylus* reflects ecological and functional specialization within their environments (Pooley, 1989). There are three ecomorph cranial shape categories (ESC) for crocodiles (general, blunt and slender), and most extant species fall into the general ESC, except for *C. intermedius* and *C. johnsoni,* which present a slightly slender ESC (Sadleir & Makovicky, 2008). These two species are locally restricted in their ranges, with *C. intermedius* endemic to the Orinoco Basin in the Amazonas (Balaguera-Reina et al., 2018), and *C. johnsoni* limited to the northern regions of Australia (Isberg et al., 2017). However, in these two cases, there are no reports suggesting they have different dietary habits. The utility of these traits is unquestionable. Salt tolerance allows *Crocodylus porosu*s to travel from 200 up to 590 kilometers offshore Australia in less than a month (Campbell et al., 2010). These individuals rode surface water currents for long distance travel as a low energy cost dispersal strategy (Campbell et al., 2010). All these adaptations provide the *Crocodylus* genus with a physiological phenotype that allowed the group to radiate via transoceanic dispersal (Meredith et al., 2011; Oaks, 2011), occupying ecological niches left by extinctions associated with global climatic changes of other crocodylian species in the Miocene (Brochu 2003:2006; Mannion et al., 2015).

Speciation rates in crocodiles have been studied at the chromosome level. The absence of sex chromosomes in *Crocodylus* implies a slow speciation rate (Demuth 2014). However, centric fusion/fission is a crucial evolutionary force of crocodilian diversity, with the *Crocodylus* group having karyotypic variation comprising 2n = 30, 32, and 34 (Srikulnath et al., 2012). Speciation rates might be a reflection of the rate of

chromosomal changes (i.e. gains, losses, fusions, and/or fissions) (Faria & Navarro, 2010), suggesting they have the capacity to rapidly speciate. The chronograms presented in this thesis support rapid speciation events within the species with short branches spread throughout the phylogeny. All these characteristics suggest that *Crocodylus* adaptively radiated during the late Miocene and Pleistocene.

Are speciation and divergence ecological in the genus Crocodylus?

The fact that divergent selection can cause adaptive divergence that can contribute to the evolution of reproductive isolation produces some questions, such as How often is speciation ecological? and How often may these ecological differences cause speciation? (Hendry, 2018). At the genetic level ecological, speciation requires a mechanism by which selection on genes conferring divergent adaptation is transmitted to genes causing barriers to gene flow (Nosil, 2012). Ecological speciation can be tested from a gene flow-based approach, measuring the variation among population pairs at loci not linked to the ones that are under selection, the positive correlation between adaptive phenotypic divergence and neutral genetic differentiation is called “isolation-by-adaptation”(IBA) (Nosil et al., 2008). However, another gene flow-based approach is mosaic hybrid zones (Nosil, 2012). In a literature review, Nosil et al. (2005) identified 27 mosaic hybrid zones, twenty of these were habitat associated, a pattern that is highly suggestive of a common role for ecological adaptation in hybrid zone maintenance immigrant inviability appeared to act in habitat-structured zones, and some zones showed adaptive phenotypic divergence, providing indirect evidence of ecological adaptation. Mosaic hybrid zones are formed when parental forms occupy distinct habitat

patches in a heterogeneous landscape, and when hybridization occurs, it follows a patchy distribution, caused by hybrid superiority (Barton and Hewitt, 1985).

The Yucatan peninsula hybrid zone of *C. acutus* and *C. moreletii*, based on the results of this study with restriction site associated DNA sequencing (RADseq), mitochondrial haplotypes and microsatellites suggest a mosaic hybridization pattern. An essential discovery of the thesis is the presence of hybrids in Banco Chinchorro, as opposed to previous studies that considered it a pure *C. acutus* population (Pacheco et al., 2016;2018; Cedeńo et al., 2008) and the evidence of two populations with restricted gene flow.

Crocodylus acutus population BCH1 is well differentiated from *C. moreletii* (F_{ST}= 0.54 - 0.63), and BCH2 is well differentiated from Central American *C. acutus* (F_{ST}= 0.79 - 0.81). And the differentiation between BCH1 and BCH2 (F_{ST}= 0.69) is as high as the differentiation between *C. acutus* and *C. moreletii* (F_{ST}= 0.64 - 0.77). There is a single maternal lineage in the two populations that is characteristic of *C. acutus*, but the genomic nuclear signature is of *C. moreletii*. This is a sign of introgression of a paternal *C. moreletii* to a maternal *C. acutus*. A characteristic of these hybrids is an atypical skull morphology, characterized as broader snouted resembling the one of *Crocodylus moreletii* (Labarre et al., 2017), where hybridization might be contributing to this morphotype. No postzygotic barriers are reported for the genus Crocodylus, hybrids are viable, fertile and not maladapted (Hekkala et al., 2015; Milián-García et al., 2011; Pacheco-Sierra et al., 2016; Rodriguez et al., 2008; Weaver et al., 2008). Hybrids on Banco Chinchorro also have a higher tolerance to salinity than the two parental species (Cedeño-Vázquez et al., 2008). However, the hybrids in these populations have a

smaller clutch size, egg size and egg mass (Charruau et al., 2010). The expectation should be a selection against hybrids, but these populations inhabit an environment that is different in salinity compared to the coastal parental environment, and it is isolated from the coastline by ocean currents.

In consequence, the crocodiles on the island have a high nesting success because there are no nest predators, wherein the coastline there are (Charruau et al., 2010). The hybrids also have faster growth rates and enhanced survivorship in this hyperosmotic environment (Cedeño-Vázquez et al., 2008). The cause of the improved fitness of hybrids could be explained by heterosis, arising on an additive trait (Barton, 2001). This is also called transgressive segregation, where the hybrids have higher phenotypic values than parental populations (Rieseberg et al., 1999).

Our results suggest that ecological speciation to hyperosmotic habitats was fundamental for the radiation of the genus *Crocodylus* to the tropics and subtropics of the world in the Miocene. What we found might be a speciation process of the insular crocodile populations in Banco Chinchorro. We suggest crocodiles in Banco Chinchorro have a physiological phenotype that makes them different from their parental populations, however further studies are needed to corroborate this theory.

Follow up studies from this thesis that are currently underway include generating a complete reference genome for *C. acutus*. Lack of a reference genome annotated at the chromosome level is needed to detect how introgression is acting, what is the nature of chromosome rearrangements, and to identify what genes and gene pathways are under selection from the thousands of SNPs recovered in the RADseq analyses between populations. Additional planned work is to further explore the natural history of

Banco Chinchorro crocodiles. This work will involve mark-recapture, GPS tracking, and relative fitness of the two population's fitness. This baseline work can then be compared to mainland populations to explore if these insular populations have higher fitness in these offshore atols. In turn, informed conservation recommendations can be made to protect the highest degrees of genetic variation.

What forces have shaped the distribution and divergence of extant crocodiles in the Neotropics?

The long history of crocodyliformes spans over 250 million years. During this vast time, this lineage has witnessed climate changes, sea level fluctuations, and mass extinctions. The first dramatic decline in the clade was during the marked global cooling trend during the Eocene and Plio-Pleistocene (Markwick, 1998). Although this long-term cooling had dramatic effects on crocodylian diversity, that didn't stop them from continuing to radiate throughout the tropical and subtropical coasts and waterways. More recently, our demographic models and the estimated time of the divergence time of crocodile populations in the Neotropics coincide with the Last Glacial Period and the Last Glacial Maximum. During this period, sea levels changed over 100 metres as a consequence of changes in the ice sheet cover (Clark et al., 2009; Peltier & Fairbanks, 2006; Rohling et al., 2017). Divergence of Neotropical *Crocodylus* appears to be associated with these sea level changes and may explain the high genetic diversity of the genus throughout the Neotropics. Sea level changes may be associated with the availability of nesting habitats and freshwater sources for nursing. From these observations, we can conclude that climate and hydrological changes of the Pleistocene

were a constraint for crocodile distribution, and directly affected the divergence of populations.

FINAL CONCLUSION

The three chapters of this thesis focused on the genus *Crocodylus,* a widely distributed group in the tropics and subtropics of the world, but with emphasis in Neotropical crocodiles. All the chapters are centered around genetics, genetic variation and the use of genetic variation as a measure of differentiation between species, hybrids and populations. But we used more than genetic measures to explain differentiation, we also evaluated how ocean currents, geological events and climatic events drove their diversity.

Each chapter addressed different questions from a Global to a regional and finally to a local scale. For the thesis we put to the test previous studies and described patterns.

For chapter one, I re-assessed the phylogeny of *Crocodylus* with the use of sequences of samples from only wild specimens. Our findings are generally consistent with previous studies, but with the identification of a more complex scenario of trans Atlantic crocodile radiations, suggesting that migration from Africa to the Neotropics may have occurred multiple times during the Miocene. These multiple migrations from Africa to Neotropics and back to Africa fit well with geological, marine and climatic conditions of the times.

In Chapter two, my main questions were: Are crocodiles from Banco Chinchorro a pure *C. acutus* population as previously reported? And what environmental factors are maintaining the isolation of these island crocodiles from hybridization on the coast? I used a genetic approach with different molecular markers and an extensive sampling in

the island to answer these questions. Instead we found a fine genetic structure, and the presence of two distinct lineages with limited gene flow. The arrival of each lineage to the island was from Central America but with tens of thousands of years of separating their arrival. The maternal lineage of the populations is from *C. acutus* but with a secondary contact and introgression from *C. moreletii*. The two lineages are isolated, caused by strong ocean currents but they are hybrids. The results of the chapter contradict the idea of a pure population and opens more questions about hybrid species.

For the last chapter, chapter three, the research question was: what was the effect of the rise of the Central American Isthmus (CAI) on the divergence of Pacific and Atlantic populations of crocodiles in Panama? Our results show genetic structure of the populations on both sides of the canal. But the structure is not related to the rise of the CAI. The dating of these events based on demographic models is coincident with the Last Glacial Maximum (LGM), a more recent event in geological time. These findings suggest that the biology of the species played a major role on the resilience of *C. acutus* to the changes caused by the connection of South and North America. However, the LGM may have also generated some divergence of crocodile populations in Panama.

The core of the thesis is a genetic approach, but with the integration of fossil record evidence, geologic events and the biology of the species studied. The goal of the thesis is to demonstrate the utility of using multiple genetic techniques in combination with non-genetic approaches to make inferences and describe patterns on species. Importantly, this thesis is the first to bring next-generation genomic sequencing to crocodiles and revealed important conservation issues for some populations. Finally, I

also would like to highlight the importance of collaboration for this research. The samples used in this thesis were achieved through the effort of multiple countries, institutions, researchers and students.

LITERATURE CITED

A

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. Nat Rev Genet, 17(2), 81–92.

Angst, D.; Buffetaut, E.; Lécuyer, C.; Amiot, R. (2013). "Terror Birds" (*Phorusrhacidae*) from the Eocene of Europe Imply Trans-Tethys Dispersal". PLoS ONE. 8 (11): e80357. doi: 10.1371/journal.pone.0080357

B

Bacon, C. D., Silvestro, D., Jaramillo, C., Smith, B. T., Chakrabarty, P., & Antonelli, A. (2015). Biological evidence supports an early and complex emergence of the Isthmus of Panama. Proceedings of the National Academy of Sciences. https://doi.org/10.1073/pnas.1423853112

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., … Johnson, E. A. (2008). Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. PLoS ONE, 3(10), e3376. https://doi.org/10.1371/journal.pone.0003376

Balaguera-Reina, S.A., Espinosa-Blanco, A., Antelo, R., Morales-Betancourt, M. & Seijas, A. (2018). *Crocodylus intermedius*. The IUCN Red List of Threatened

Species 2018: e.T5661A3044743. http://dx.doi.org/10.2305/IUCN.UK.2018-1.RLTS.T5661A3044743.en.

Barton, N., & Hewitt, G. (1985). Analysis of hybrid zones. Annual Review of Ecology and Systematics, 16(1), 113-148. doi:10.1146/annurev.es.16.110185.000553

Barton, N. H. (2001), The role of hybridization in evolution. Molecular Ecology, 10: 551-568. doi:10.1046/j.1365-294x.2001.01216.x

Bardakci, F., & Skibinski, D. O. F. (1994). Application of the RAPD technique in tilapia fish: species and subspecies identification. Heredity, 73, 117–123

Barrios-Quiroz, G., & Cremieux, J. (2018). Protocolo de rancheo para cocodrilo de pantano (*Crocodylus moreletii*) en México. Mexico: Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO). Mexico.

Barry, J.C., Morgan, M.E., Flynn, L.J., Pilbeam, D., Behrensmeyer, A.K., Raza, S.M., Khan, I.A., Badgley, C., Hicks, J., Kelley, J. (2002). Faunal and environmental change in the late Miocene Siwaliks of northern Pakistan. Paleobiology 28, 1–71.

Behler, J.L. (1978). Feasibility of the establishment of a captive-breeding population of the American crocodile. South Florida Research Center T-509 Report to the NP Service; Homestead, Florida.

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Ostell, J., Pruitt, K. D., & Sayers, E. W. (2017). GenBank. Nucleic acids research, 46(D1), D41–D47. doi:10.1093/nar/gkx1094

Beyer, C. (2008). Establishment of a chronostratigraphical framework for the As Sahabi sequence in northeast Libya. Garyounis Scientific Bulletin 5:59–69. (https://issuu.com/icsm/docs/gsb_5-2008-final)

Bickelmann, C., & Klein, N. (2009). The late Pleistocene horned crocodile *Voay robustus* (Grandidier & Vaillant, 1872) from Madagascar in the Museum für Naturkunde Berlin. Fossil Record, 12(1), 13–21. https://doi.org/10.1002/mmng.200800007

Bittencourt, P. S., Campos, Z., Muniz, F. de L., Marioni, B., Souza, B. C., Da Silveira, R., … Farias, I. P. (2019). Evidence of cryptic lineages within a small South American crocodilian: the Schneider's dwarf caiman *Paleosuchus trigonatus* (Alligatoridae: Caimaninae). PeerJ, accepted, 1–26. https://doi.org/10.7717/peerj.6580

Böhme, M. (2003). The Miocene Climatic Optimum: Evidence from ectothermic vertebrates of Central Europe. Palaeogeography, Palaeoclimatology,

Palaeoecology, 195(3–4), 389–401. https://doi.org/10.1016/S0031-0182(03)00367-5

Bond, M., Tejedor, M., Campbell, K., Chornogubsky, L., Novo, N., & Goin, F. (2015). Eocene primates of south america and the african origins of new world monkeys. Nature, 520(7548), 538-41. doi:10.1038/nature14120

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., … Drummond, A. J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLoS Computational Biology. https://doi.org/10.1371/journal.pcbi.1003537

Brochu, C. A. (2000). Phylogenetic Relationships and Divergence Timing of Crocodylus Based on Morphology and the Fossil Record Author (s): Christopher A. Brochu Based Phylogenetic Relationships and Divergence Timing of Crocodylus on Morphology and the Fossil Record. Copeia, 3(3), 657–673.

Brochu, C. (2004). A new late cretaceous gavialoid crocodylian from eastern north America and the phylogenetic relationships of thoracosaurs. Journal of Vertebrate Paleontology, 24(3), 610-633.

Brochu, C. (2006). Osteology and phylogenetic significance of *Eosuchus minor* (marsh, 1870) new combination, a longirostrine crocodylian from the late paleocene of North America. Journal of Paleontology, 80(1), 162-186.

Brochu, C.A. (2007). Morphology, relationships, and biogeographical significance of an extinct horned crocodile (Crocodylia, Crocodylidae) from the Quaternary of Madagascar. Zoological Journal of the Linnean Society, 150(4), 835–863. https://doi.org/10.1111/j.1096-3642.2007.00315.x

Brochu, C. A., Njau, J., Blumenschine, R. J., & Densmore, L. D. (2010). A new horned crocodile from the Plio-Pleistocene hominid sites at Olduvai Gorge, Tanzania. PLoS ONE, 5(2). https://doi.org/10.1371/journal.pone.0009333

Bronzati, M., Montefeltro, F. C., & Langer, M. C. (2015). Diversification events and the effects of mass extinctions on Crocodyliformes evolutionary history. Royal Society Open Science, 2(5), 140385. https://doi.org/10.1098/rsos.140385

Brusatte, S. L., Benton, M. J., Lloyd, G. T., Ruta, M., & Wang, S. C. (2010). Macroevolutionary patterns in the evolutionary radiation of archosaurs (Tetrapoda: Diapsida). Earth and Environmental Science Transactions of the Royal Society of Edinburgh, 101(3–4), 367–382. https://doi.org/10.1017/S1755691011020056.

Butler, R. J., Brusatte, S. L., Reich, M., Nesbitt, S. J., Schoch, R. R., & Hornung, J. J. (2011). The sail-backed reptile Ctenosauriscus from the latest Early Triassic of Germany and the timing and biogeography of the early archosaur radiation. PLoS ONE, 6(10). https://doi.org/10.1371/journal.pone.0025693.

C

Campbell, H. A., Watts, M. E., Sullivan, S., Read, M. A., Choukroun, S., Irwin, S. R., & Franklin, C. E. (2010). Estuarine crocodiles ride surface currents to facilitate long-distance travel. Journal of Animal Ecology, 79(5), 955–964. https://doi.org/10.1111/j.1365-2656.2010.01709.x

Calverley, P. M., & Downs, C. T. (2015). Movement and Home Range of Nile Crocodiles in Ndumo Game Reserve, South Africa. Koedoe, 57(1). https://doi.org/10.4102/koedoe.v57i1.1234.

Cardiel, J. M., S. Castroviejo, and M. Velayos. (1997). El Parque Nacional de Coiba. El medio físico. Pp. 11–31, in Floray fauna del Parque Nacional de Coiba (Panamá): inventario preliminar (S. Castroviejo And M. Velayos, eds.). Agencia Española de Cooperación Internacional, Madrid, 534 pp.

Carrillo, L., E.M. Johns, R.H. Smith, J.T. Lamkin, and J.L. Largier. (2015). Pathways and Hydrography in the Mesoamerican Barrier Reef System Part 1: Circulation. Continental Shelf Research 109. Elsevier: 164–76. https://doi.org/10.1016/j.csr.2015.09.014

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. Molecular Ecology, 22(11), 3124–3140. https://doi.org/10.1111/mec.12354

Cedeño-Vázquez, J.R., Platt, S.G. & Thorbjarnarson, J. (IUCN Crocodile Specialist Group) (2012). *Crocodylus moreletii*. The IUCN Red List of Threatened Species 2012: e.T5663A3045579. http://dx.doi.org/10.2305/IUCN.UK.2012.RLTS.T5663A3045579.en

Cedeño-Vázquez, J. R., Rodriguez, D., Calmé, S., Ross, J. P., Densmore, L. D., & Thorbjarnarson, J. B. (2008). Hybridization between *Crocodylus acutus* and *Crocodylus moreletii* in the Yucatan Peninsula: I. evidence from mitochondrial DNA and morphology. Journal of Experimental Zoology Part A: Ecological Genetics and Physiology, 309A(10), 661–673. https://doi.org/10.1002/jez.473

Cerón-Souza, I., Gonzalez, E. G., Schwarzbach, A. E., Salas-Leiva, D. E., Rivera-Ocasio, E., Toro-Perea, N., … Mcmillan, W. O. (2015). Contrasting demographic history and gene flow patterns of two mangrove species on either side of the Central American Isthmus. Ecology and Evolution, 5(16), 3486–3499. https://doi.org/10.1002/ece3.1569

Charruau, P. (2012). Microclimate of American crocodile nests in Banco Chinchorro biosphere reserve, Mexico: Effect on incubation length, embryos survival and

hatchlings sex. Journal of Thermal Biology, 37(1), 6-14. doi: 10.1016/j.jtherbio.2011.10.010

Charruau, P., J. R. Cedeño -Vázquez, and S. Calmé. (2005). Status and conservation of the American crocodile (*Crocodylus acutus*) in Banco Chinchorro Biosphere Reserve, Quintana Roo, Mexico. Herpetological Review 36:390–395

Charruau, P., J. R. Cedeño -Vázquez, A. Villegas, and H. González-Cortés. (2010). Tasas de crecimiento del cocodrilo americano (*Crocodylus acutus*) en estado silvestre en la Penınsula de Yucatan, Mexico. Revista Latinoamericana de Conservacion 1:63–72

Charruau, Pierre, Aníbal H. Díaz de la Vega Pérez, and Fausto R. Méndez de la Cruz. (2015). Reptiles of Banco Chinchorro: Updated List, Life History Data, and Conservation. The Southwestern Naturalist 60 (4): 299–312. https://doi.org/10.1894/0038-4909-60.4.299

Charruau, P., Thorbjarnarson, J. B., & Hénaut, Y. (2010). Tropical cyclones and reproductive ecology of *Crocodylus acutus* Cuvier, 1807 (Reptilia: Crocodilia: Crocodylidae) on a Caribbean atoll in Mexico. Journal of Natural History, 44(11–12), 741–761. https://doi.org/10.1080/00222930903490993

Cherkiss, M. S., Mazzotti, F. J., Hord, L., & Aldecoa, M. (2014). Remarkable Movements of an American Crocodile (*Crocodylus acutus*) in Florida. Southeastern Naturalist, 13(4), N52–N56. https://doi.org/10.1656/058.013.0407.

Clark, P. U., Clark, P. U., Dyke, A. S., Dyke, A. S., Shakun, J. D., Shakun, J. D., … Mccabe, A. M. (2009). The Last Glacial Maximum. Science, 325(August), 3–7. https://doi.org/10.1126/science.1172873.

Cornuet, J.-M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., … Estoup, A. (2014). DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. Bioinformatics, 30(8), 1187–1189.

Coyne, J., & Orr, H. (2004). Speciation. Sunderland, Mass.: Sinauer Associates.

D

De Queiroz, K., & Donoghue, M. J. (1988). Phylogenetic systematics and the species problem. Cladistics, 4, 317–338.

Delfino, M., Böhme, M., & Rook, L. (2007). First European evidence for transcontinental dispersal of *Crocodylus* (late Neogene of southern Italy). Zoological Journal of the

Linnean Society, 149(3), 293–307. https://doi.org/10.1111/j.1096-3642.2007.00248.x

Díaz-Muñoz, S. L. (2012). Role of recent and old riverine barriers in fine-scale population genetic structure of Geoffroy's tamarin (*Saguinus geoffroyi*) in the Panama Canal watershed. Ecology and Evolution, 2(2), 298–309. https://doi.org/10.1002/ece3.79

E

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources, 4(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Eaton, M. J., Martin, A., Thorbjarnarson, J., & Amato, G. (2009). Species-level diversification of African dwarf crocodiles (Genus Osteolaemus): A geographic and phylogenetic perspective. Molecular Phylogenetics and Evolution, 50(3), 496–506. https://doi.org/10.1016/j.ympev.2008.11.009

Ernst F. D. Ross AND C. A. Ross, C. H. (1999). *Crocodylus acutus*. Catalogue of American Amphibians and Reptiles, 700, 1–17.

Escobedo-Galvan A., Palacios-Chavez V., & Vovides-Tejera A. (2008) *Crocodylus moreletii* (Morelet's Crocodile) salinity tolerance. Herpetol. Rev. 39, 346–347.

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources, 10(3), 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

F

Faria, R., & Navarro, A. (2010). Chromosomal speciation revisited: Rearranging theory with pieces of evidence. Trends in Ecology and Evolution, 25(11), 660–669. https://doi.org/10.1016/j.tree.2010.07.008

Figueirido, B., Janis, C. M., Pérez-claros, J. A., Renzi, M. De, & Palmqvist, P. (2012). Cenozoic climate change influences mammalian evolutionary dynamics, 109(3), 722–727. https://doi.org/10.1073/pnas.1110246108/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1110246108

Finarelli, J. A., & Badgley, C. (2019). Diversity dynamics of Miocene mammals in relation to the history of tectonism and climate, (April 2010), 2721–2726. https://doi.org/10.1098/rspb.2010.0348

Fitzsimmons, N. N., Buchan, J. C., Lam, P. V., Polet, G., Hung, T. T., Thang, N. Q., & Gratten, J. (2002). Identification of purebred *Crocodylus siamensis* for reintroduction in Vietnam. Journal of Experimental Zoology, 294(4), 373–381. https://doi.org/10.1002/jez.10201

Franke, F. A., Schmidt, F., Borgwardt, C., Bernhard, D., Bleidorn, C., Engelmann, W. E., & Schlegel, M. (2013). Genetic differentiation of the African dwarf crocodile *Osteolaemus tetraspis* Cope, 1861 (Crocodylia: Crocodylidae) and consequences for European zoos. Organisms Diversity and Evolution, 13(2), 255–266. https://doi.org/10.1007/s13127-012-0107-1

G

Galván-Quesada, S., Doadrio, I., Alda, F., Perdices, A., Reina, R. G., Varela, M. G., … Domínguez-Domínguez, O. (2016). Molecular phylogeny and biogeography of the amphidromous fish genus *Dormitator gill* 1861 (Teleostei: Eleotridae). PLoS ONE, 11(4), 1–26. https://doi.org/10.1371/journal.pone.0153538

Garrick, L. D., Lang, J. W., Zoologist, A., & Winter, N. (2008). Social Signals and Behaviors of Adult Alligators and Crocodiles Social Signals and Behaviors of Adult Alligators and Crocodiles. New York, 17(1), 225–239.

Glenn, T. C., Staton, J. L., Vu, A. T., Davis, L. M., Bremer, J. R. A., Rhodes, W. E., … Sawyer, R. H. (2002). Low mitochondrial DNA variation among American alligators

and a novel non-coding region in crocodilians. Journal of Experimental Zoology, 294(4), 312–324. https://doi.org/10.1002/jez.10206

Green, R. E., Braun, E. L., Armstrong, J., Earl, D., Nguyen, N., Hickey, G., … Ray, D. A. (2014). Three crocodilian genomes reveal ancestral patterns of evolution among archosaurs. Science, 346(6215). https://doi.org/10.1126/science.1254449

Grigg, G., & Kirshner, D. (2015). Biology and evolution of crocodylians. Melbourne: CSIRO Publishing. (2015)

Gompert, Z., & Buerkle, C. A. (2016). What, if anything, are hybrids: enduring truths and challenges associated with population structure and gene flow. Evolutionary Applications. https://doi.org/10.1111/eva.12380

Gompert, Z., Mandeville, E. G., & Buerkle, C. A. (2017). Analysis of Population Genomic Data from Hybrid Zones. Annual Review of Ecology, Evolution, and Systematics, 48(1), 207–229. https://doi.org/10.1146/annurev-ecolsys-110316-022652.

Goudet, J. (2005). HIERFSTAT, a package for R to compute and test hierarchical F-statistics. Molecular Ecology Notes, 5(1), 184–186. https://doi.org/10.1111/j.1471-8286.2004.00828.x

Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2009). Diffusion Approximations for Demographic Inference: ∂ a ∂ i. Nature Preceedings, 2009.

Guzman, H. M., Cipriani, R., & Jackson, J. B. C. (2019). Historical Decline in Coral Reef Growth after the Panama Canal, 37(5), 342–346.

H

Harrison, R. G., & Larson, E. L. (2014). Hybridization, Introgression, and the nature of species boundaries. Journal of Heredity, 105(S1), 795–809. https://doi.org/10.1093/jhered/esu033

Haug, G. H., & Tiedemann, R. (1998). Effect of the formation of the Isthmus of Panama on Atlantic Ocean thermohaline circulation. Nature, 393(6686), 673–676. https://doi.org/10.1038/31447

Hekkala, E. R., Platt, S. G., Thorbjarnarson, J. B., Rainwater, T. R., Tessler, M., Cunningham, S. W., … Amato, G. (2015). Integrating molecular, phenotypic and environmental data to elucidate patterns of crocodile hybridization in Belize. Royal Society Open Science, 2(9). https://doi.org/10.1098/rsos.150409

Hénaut, Y., & Charruau, P. (2012). Nest attendance and hatchling care in wild american crocodiles (*Crocodylus acutus*) in Quintana Roo, Mexico. Animal Biology, 62(1), 29-51. doi:10.1163/157075511X597629

Hendry, A. (2018). Eco-evolutionary dynamics. Princeton: Princeton University Press. (2018).

Hetem, R. S., Fuller, A., Maloney, S. K., & Mitchell, D. (2014). Responses of large mammals to climate change, (September), 115–127.

I

Isberg, S., Balaguera-Reina, S.A. & Ross, J.P. 2017. *Crocodylus johnsoni*. The IUCN Red List of Threatened Species 2017: e.T46589A3010118. http://dx.doi.org/10.2305/IUCN.UK.2017-3.RLTS.T46589A3010118.en.

IUCN 2019. The IUCN Red List of Threatened Species. Version 2019-2. http://www.iucnredlist.org

Instituto Nacional de Ecología [INE]. (2000). Programa de manejo Reserva de la Biosfera Banco Chinchorro, México. Instituto Nacional de Ecología, Ciudad de México, Distrito Federal, México.

J

Jaramillo, Carlos A. (2018). "Evolution of the Isthmus of Panama: Biological, Paleoceanographic and Paleoclimatological Implications." in Mountains, Climate

and Biodiversity, edited by Hoorn, Carina, Perrigo, Allison, and Antonelli, Alexandre., 323–338. Hoboken, New Jersey: John Wiley & Sons

Jombart, T. (2008). ADEGENET: A R package for the multivariate analysis of genetic markers. Bioinformatics, 24(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., & Ahmed, I. (2011). ADEGENET 1.3-1: New tools for the analysis of genome-wide SNP data. Bioinformatics, 27(21), 3070–3071

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics, 11(1), 94. https://doi.org/10.1186/1471-2156-11-94

K

Kalinowski, S. T. (2004). Counting Alleles with Rarefaction: Private Alleles and Hierarchical Sampling Designs. Conservation Genetics, 5(4), 539–543. https://doi.org/10.1023/B:COGE.0000041021.91777.1a

L

Labarre, Didier, P. Charruau, S. G. Platt, T. R. Rainwater, J. R. Cedeño-Vázquez, and H. González-Cortés. (2017). Morphological Diversity of the American Crocodile

(*Crocodylus Acutus*) in the Yucatán Peninsula. Zoomorphology 136 (3): 387–401. https://doi.org/10.1007/s00435-017-0356-9.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods, 9(4), 357–359. https://doi.org/10.1038/nmeth.1923

Lavocat, R. (1969). "La systématique des rongeurs hystricomorphes et la dérive des continents". C. R. Acad. Sci. Paris Sér. D., 269:1496-1497.

Lessios, H. A. (2008). The Great American Schism: Divergence of Marine Organisms After the Rise of the Central American Isthmus. Annual Review of Ecology, Evolution, and Systematics, 39(1), 63–91. https://doi.org/10.1146/annurev.ecolsys.38.091206.095815.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics, 25(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Losos, J. (Ed.). (2014). The princeton guide to evolution (Princeton reference). Princeton New Jersey: Princeton University Press. (2014).

Losos, J. B., & Ricklefs, R. E. (2009). Adaptation and diversification on islands. Nature, 457(7231), 830–836. https://doi.org/10.1038/nature07893

Losos, J., Ricklefs, R., & MacArthur, R. (2010). The theory of island biogeography revisited. Princeton: Princeton University Press. (2010).

M

Maccagno, A. M. (1948). Descrizione di una nuova specie di “Crocodilus” del giacimento di Sahabi (Sirtica). Atti della Reale Accademia Nazionale dei Lincei: Memorie della Classe di Scienze fisiche, Mathematiche e Naturale, Serie 8 1:63–96.

Machkour-M’Rabet, Salima, Yann Hénaut, Pierre Charruau, Muriel Gevrey, Peter Winterton, and Luc Legal. (2009). Between Introgression Events and Fragmentation, Islands Are the Last Refuge for the American Crocodile in Caribbean Mexico. Marine Biology 156 (6): 1321–33. https://doi.org/10.1007/s00227-009-1174-5.

Madern, P. A. A., & Hoek, L. W. Van Den. (2015). Going south: Latitudinal change in mammalian biodiversity in Miocene Eurasia. Palaeogeography, Palaeoclimatology, Palaeoecology, 424, 123–131. https://doi.org/10.1016/j.palaeo.2015.02.011

Maddison, W. P. and D.R. Maddison. (2018). Mesquite: a modular system for evolutionary analysis. Version 3.51 http://www.mesquiteproject.org

Mallet, J. (1995). A species definition for the modern synthesis. Trends in Ecology & Evolution, 10(7), 294–299. https://doi.org/10.1016/0169-5347(95)90031-4

Mannion, P. D., Benson, R. B. J., Carrano, M. T., Tennant, J. P., Judd, J., & Butler, R. J. (2015). Climate constrains the evolutionary history and biodiversity of crocodylians. Nature Communications, 6, 1–9. https://doi.org/10.1038/ncomms9438.

Marko, P. B., & Moran, A. L. (2009). Out of sight, out of mind: High cryptic diversity obscures the identities and histories of geminate species in the marine bivalve subgenus *Acar*. Journal of Biogeography, 36(10), 1861–1880. https://doi.org/10.1111/j.1365-2699.2009.02114.x

Markwick, P. J. (1998). Crocodilian diversity in space and time: The role of climate in paleoecology and its implication for understanding K/T extinctions. Paleobiology, 24(4), 470–497. https://doi.org/10.1017/S009483730002011X.

Mathews, J. C., & Samonds, K. E. (2016). A juvenile subfossil crocodylian from Anjohibe Cave, Northwestern Madagascar. PeerJ, 4, e2296. https://doi.org/10.7717/peerj.2296

Mazzotti, F. J. (1999). The American Crocodile in Florida Bay. Estuaries, 22(2), 552–561.

McAliley, L. R., Willis, R. E., Ray, D. A., White, P. S., Brochu, C. A., & Densmore Iii, L. D. (2006). Are crocodiles really monophyletic? Evidence for subdivisions from sequence and morphological data. Molecular Phylogenetics and Evolution, 39(1), 16–32. https://doi.org/http://dx.doi.org/10.1016/j.ympev.2006.01.012

McDougall, I., and C. S. Feibel. (2003). Numerical age control for the Miocene-Pliocene succession at Lothagam, a hominoid-bearing sequence in the northern Kenya Rift; pp. 43–63 in M. G. Leakey and J. M. Harris (eds.), Lothagam: The Dawn of Humanity in Eastern Africa. Columbia University Press, New York.

Meek, S. E., and S. F. Hildebrand. (1916). The fishes of the fresh waters of Panama. Field Museum of Natural History, Zoological Series, vol X, pp. 217–374, Chicago, IL.

Meredith, R. W., Hekkala, E. R., Amato, G., & Gatesy, J. (2011). A phylogenetic hypothesis for Crocodylus (Crocodylia) based on mitochondrial DNA: Evidence for a trans-Atlantic voyage from Africa to the New World. Molecular Phylogenetics and Evolution, 60(1), 183–191. https://doi.org/10.1016/j.ympev.2011.03.026

Milian-Garcia, Y., Ramos-Targarona, R., Perez-Fleitas, E., Sosa-Rodriguez, G., Guerra-Manchena, L., Alonso-Tabet, M., … Russello, M. A. (2015). Genetic evidence of hybridization between the critically endangered Cuban crocodile and the American

crocodile: implications for population history and in situ/ex situ conservation. Heredity, 114(3), 272–280. https://doi.org/10.1038/hdy.2014.96

Milián-García, Y., Russello, M. A., Castellanos-Labarcena, J., Cichon, M., Kumar, V., Espinosa, G., … Janke, A. (2018). Genetic evidence supports a distinct lineage of American crocodile (Crocodylus acutus) in the Greater Antilles Distributed under Creative Commons CC-BY 4.0. https://doi.org/10.7717/peerj.5836

Milián-García, Y., Venegas-Anaya, M., Frias-Soler, R., Crawford, A. J., Ramos-Targarona, R., Rodríguez-Soberón, R., … Bermingham, E. (2011). Evolutionary history of Cuban crocodiles *Crocodylus rhombifer* and *Crocodylus acutus* inferred from multilocus markers. Journal of Experimental Zoology Part A: Ecological Genetics and Physiology, 315 A(6), 358–375. https://doi.org/10.1002/jez.683

Muniz FL, Campos Z, Hernández Rangel SM, Martínez JG, Souza BC, De Thoisy B, Botero-Arias R, Hrbek T, Farias IP. (2018). Delimitation of evolutionary units in Cuvier's dwarf caiman, *Paleosuchus palpebrosus* (Cuvier, 1807): in- sights from conservation of a broadly distributed species. Conservation Genetics DOI 10.1007/s10592-017-1035-6.

Muirhead, J. R., Minton, M. S., & Miller, W. A. (2015). Projected effects of the Panama Canal biological invasions, 75–87. https://doi.org/10.1111/ddi.12260

N

Nesbitt SJ. (2011) The early evolution of archosaurs: relationships and the origin of major clades. Bull. Am.Mus.Nat.Hist. 352 1–292. doi:10.1206/352.

Nosil, P. (2012). Ecological speciation (Oxford series in ecology and evolution). Oxford: Oxford University Press. (2012).

Nosil, P. , Egan, S. P. and Funk, D. J. (2008), Heterogeneous Genomic Differentiation Between Walking-stick Ecotypes: “Isolation By Adaptation” And Multiple Roles For Divergent Selection. Evolution, 62: 316-336. doi:10.1111/j.1558-5646.2007.00299.x

Nosil, P., Harmon, L. J., & Seehausen, O. (2009). Ecological explanations for (incomplete) speciation. Trends in Ecology & Evolution, 24(3), 145–156.

Nosil, P., Vines, T., & Funk, D. (2005). Perspective: Reproductive isolation caused by natural selection against immigrants from divergent habitats. Evolution, 59(4), 705-719.

O

Oaks, J. R. (2011). A Time-calibrated Species Tree of Crocodylia Reveals A Recent Radiation of The True Crocodiles. Evolution, 65(11), 3285–3297. https://doi.org/10.1111/j.1558-5646.2011.01373.x

O'Dea, A. O., Jackson, J. B. C., Fortunato, H., Smith, J. T., Croz, L. D., Johnson, K. G., & Todd, J. A. (2007). Environmental change preceded Caribbean extinction by 2 million years. Geology, 104(13), 5501–5506.

O'Dea, A., Lessios, H. A., Coates, A. G., Eytan, R. I., Restrepo-Moreno, S. A., Cione, A. L., … Jackson, J. B. C. (2016). Formation of the Isthmus of Panama. Science Advances, 2(8). https://doi.org/10.1126/sciadv.1600883

P

Pacheco-Sierra, G., Gompert, Z., Domínguez-Laso, J., & Vázquez-Domínguez, E. (2016). Genetic and morphological evidence of a geographically widespread hybrid zone between two crocodile species*, Crocodylus acutus* and *Crocodylus moreletii*. Molecular Ecology, 25(14), 3484–3498. https://doi.org/10.1111/mec.13694

Pacheco-Sierra, G., Vázquez-Domínguez, E., Pérez-Alquicira, J., Suárez-Atilano, M., & Domínguez-Laso, J. (2018). Ancestral Hybridization Yields Evolutionary Distinct Hybrids Lineages and Species Boundaries in Crocodiles, Posing Unique Conservation Conundrums. Frontiers in Ecology and Evolution. https://doi.org/10.3389/fevo.2018.00138

Paradis, E. (2010). PEGAS: An R package for population genetics with an integrated-modular approach. Bioinformatics, 26(3), 419–420. https://doi.org/10.1093/bioinformatics/btp696

Paradis, E., & Schliep, K. (2019). APE 5.0: An environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics, 35(3), 526–528. https://doi.org/10.1093/bioinformatics/bty633

Payseur, B. A., & Rieseberg, L. H. (2016). A genomic perspective on hybridization and speciation. Molecular Ecology, 25(11), 2337–2360. https://doi.org/10.1111/mec.13557

Peltier, W. R., & Fairbanks, R. G. (2006). Global glacial ice volume and Last Glacial Maximum duration from an extended Barbados sea level record. Quaternary Science Reviews, 25(23–24), 3322–3337. https://doi.org/10.1016/j.quascirev.2006.04.010.

Perez-Higareda G, Rangel-Rangel A, Smith H. (1991) Maximum sizes of Morelet's and American crocodiles. Bulletin of the Maryland Herpetological Society. 27, 34–37.

Pidcock, S., Taplin, L. E., & Grigg, G. C. (1997). Differences in renal-cloacal function between *Crocodylus porosus* and *Alligator mississippiensis* have implications for crocodilian evolution. Journal of Comparative Physiology - B Biochemical,

Systemic, and Environmental Physiology, 167(2), 153–158. https://doi.org/10.1007/s003600050059

Platt S, Rainwater T. 2005A review of morphological characters useful for distinguishing Morelet's crocodile (*Crocodylus moreletii*) and American crocodile (*Crocodylus acutus*) with an emphasis on populations in the coastal zone of Belize. Bull. Chicago Herpetol. Soc. 40, 25–29.

Platt, S. G., & Thorbjarnarson, J. B. (2010). Nesting Ecology of the American Crocodile in the Coastal Zone of Belize Published by American Society of Ichthyologists and Herpetologists Nesting Ecology of the American Crocodile in the Coastal Zone of Belize. Copeia, 2000(3), 869–873.

Platt, S. G., Thorbjarnarson, J. B., & Rainwater, T. R. (1999). Distribution of Morelet's Crocodile (Crocodylus moreletii) in Southern Belize. The Southwestern Naturalist, 44(3), 395–398. Retrieved from http://www.jstor.org/stable/30055240

Platt, S. G., Thorbjarnarson, J. B., Rainwater, T. R., & Martin, D. R. (2013). Diet of the American Crocodile (*Crocodylus acutus*) in Marine Environments of Coastal Belize. Journal of Herpetology, 47(1), 1–10. https://doi.org/10.1670/12-077.

Ponce-Campos, P., Thorbjarnarson, J. & Velasco, A. (IUCN SSC Crocodile Specialist Group) (2012). *Crocodylus acutus*. The IUCN Red List of Threatened Species

2012: e.T5659A3043244.
http://dx.doi.org/10.2305/IUCN.UK.2012.RLTS.T5659A3043244.en

Pooley, A.C. (1989). Food and feeding habits. In: Crocodiles and Alligators (C.A. Ross, ed.), pp. 67–91. Facts on File, New York.

Posada D. (2008) jModelTest: Phylogenetic Model Averaging. Molecular Biology and Evolution, Volume 25(7), Pages 1253–1256, https://doi.org/10.1093/molbev/msn083

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. Genetics, 155(2), 945 LP – 959.

Putman, A. I., & Carbone, I. (2014). Challenges in analysis and interpretation of microsatellite data for population genetic studies. Ecology and Evolution, 4(22), 4399–4428. https://doi.org/10.1002/ece3.1305

R

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raj, A., Stephens, M., & Pritchard, J. K. (2014). FASTSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets. https://doi.org/10.1534/genetics.114.164350

Ray, D. A., & Densmore, L. (2002). The crocodilian mitochondrial control region: general structure, conserved sequences, and evolutionary implications. Journal of Experimental Zoology, 294(4), 334–345. https://doi.org/10.1002/jez.10198

Ribbe, J., R. D. Müller, C. J. Poulsen, M. Huber, and Y. You. (2009). “Simulation of the Middle Miocene Climate Optimum.” Geophysical Research Letters 36 (4). https://doi.org/10.1029/2008gl036571

Riff, D., P. S. R. Romano, G. R. Oliveira, and O. A. Aguilera. (2009). Neo- gene crocodile and turtle fauna in northern South America; pp. 259– 280 in C. Hoorn and F. P. Wesselingh (eds.), Amazonia: Landscape and Species Evolution. Wiley-Blackwell Publishing, Chichester, U.K.

Rieseberg, L. H., Archer, M. A., & Wayne, R. K. (1999). Transgressive segregation, adaptation and speciation. Heredity, 83(4), 363–372. https://doi.org/10.1038/sj.hdy.6886170

Robinson, W. D. (1999). Long-Term Changes in the Avifauna of Barro Colorado Island, Panama, a Tropical Forest Isolate, 13(1), 85–97.

Rodriguez, D., Cedeño-Vázquez, J. R., Forstner, M. R. J., & Densmore, L. D. (2008). Hybridization between *Crocodylus acutus* and *Crocodylus moreletii* in the Yucatan Peninsula: II. evidence from microsatellites. Journal of Experimental Zoology Part A: Ecological Genetics and Physiology, 309A (10), 674–686. https://doi.org/10.1002/jez.499

Rodriguez, D., Forstner, M. J., Moler, P., Wasilewski, J., Cherkiss, M., & Densmore Iii, L. (2011). Effect of human-mediated migration and hybridization on the recovery of the American crocodile in Florida (USA). Conservation Genetics, 12(2), 449–459. https://doi.org/10.1007/s10592-010-0153-1

Rohling, E. J., Hibbert, F. D., Williams, F. H., Grant, K. M., Marino, G., Foster, G. L., … Yokoyama, Y. (2017). Differences between the last two glacial maxima and implications for ice-sheet, δ 18 O, and sea-level reconstructions. Quaternary Science Reviews. Elsevier Ltd. https://doi.org/10.1016/j.quascirev.2017.09.009.

Ross CA, Ross FD. (1974). Caudal scalation of central American Crocodylus. Proc. Biol. Soc. Wash. 87, 231–234.

Rousset, F. (2008). GENEPOP: A complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources, 8(1), 103–106. https://doi.org/10.1111/j.1471-8286.2007.01931.x

Russello, M. A., Brazaitis, P., Gratten, J., Watkins-Colwell, G. J., & Caccone, A. (2007). Molecular assessment of the genetic integrity, distinctiveness and phylogeographic context of the Saltwater crocodile (*Crocodylus porosus*) on Palau. Conservation Genetics, 8(4), 777–787. https://doi.org/10.1007/s10592-006-9225-7

S

Sadleir, R. W., & Makovicky, P. J. (2008). Cranial shape and correlated characters in crocodilian evolution. Journal of Evolutionary Biology, 21(6), 1578–1596. https://doi.org/10.1111/j.1420-9101.2008.01602.x

Salas-Gismondi, R., Flynn, J. J., Baby, P., Tejada-Lara, J. V, Wesselingh, F. P., & Antoine, P.-O. (2015). A Miocene hyperdiverse crocodylian community reveals peculiar trophic dynamics in proto-Amazonian mega-wetlands (Vol. 282). https://doi.org/10.1098/rspb.2014.2490

Salas-Gismondi, R., Moreno-Bernal, J. W., Scheyer, T. M., Sánchez-Villagra, M. R., Jaramillo, C., Sanchez-Villagra, M. R., & Heredia, C. (2018). New Miocene Caribbean gavialoids and patterns of longirostry in crocodylians. Https://Doi.Org/10.1080/14772019.2018.1495275, 0(0), 1–27. https://doi.org/10.1080/14772019.2018.1495275.

Sánchez Herrera, O., G. López Segurajáuregui, A. G. N. O. de la H. y H. B. D. (2011). Programa de Monitoreo del Cocodrilo de Pantano (*Crocodylus moreletii*) México-Belice-Guatemala. México. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. México

Seehausen, O., Butlin, R. K., Keller, I., Wagner, C. E., Boughman, J. W., Hohenlohe, P. A., … Widmer, A. (2014). Genomics and the origin of species. Nat Rev Genet, 15(3), 176–192. https://doi.org/10.1038/nrg3644

Sepulchre, P., Arsouze, T., Donnadieu, Y., Dutay, J. C., Jaramillo, C., Le Bras, J., … Waite, A. J. (2014). Consequences of shoaling of the Central American Seaway determined from modeling Nd isotopes. Paleoceanography, 29(3), 176–189. https://doi.org/10.1002/2013PA002501

Sharpe, D. M., De León, L. F., González, R. and Torchin, M. E. (2017), Tropical fish community does not recover 45 years after predator introduction. Ecology, 98: 412-424. doi:10.1002/ecy.1648

Shaw, K.L & Mullen, S.P (2014). Speciation Continuum. Journal of Heredity, 105, 741–742. https://doi.org/10.1016/B978-0-12-800049-6.00080-9

Scheyer, T. M., Aguilera, O. A., Delfino, M., Fortier, D. C., Carlini, A. A., Sánchez, R., … Sánchez-Villagra, M. R. (2013). Crocodylian diversity peak and extinction in the

late Cenozoic of the northern Neotropics. Nature Communications, 4(May). https://doi.org/10.1038/ncomms2940

Schluter, D. (2000). The ecology of adaptive radiation (Oxford series in ecology and evolution). Oxford: Oxford University Press.

Shirley, M. H., Vliet, K. A., Carr, A. N., & Austin, J. D. (2013). Rigorous approaches to species delimitation have significant implications for African crocodilian systematics and conservation. Proceedings of the Royal Society B: Biological Sciences, 281(1776). https://doi.org/10.1098/rspb.2013.2483

Smith, S.A., G. Bell, and Bermingham, M. (2004). Cross-Cordillera exchange mediated by the Panama Canal increased the species richness of local freshwater fish assemblages. Proc. R. Soc. Lond. B 271:1889–1896.

Srikulnath, K., Thapana, W., & Muangmai, N. (2015). Role of Chromosome Changes in Crocodylus Evolution and Diversity. Genomics & Informatics, 13(4), 102–111. https://doi.org/10.5808/GI.2015.13.4.102

Stange, M., Sánchez-Villagra, M. R., Salzburger, W., & Matschiner, M. (2018). Bayesian divergence-time estimation with genome-wide single-nucleotide polymorphism data of sea catfishes (*Ariidae*) supports miocene closure of the Panamanian

Isthmus. Systematic Biology, 67(4), 681–699. https://doi.org/10.1093/sysbio/syy006

Swinton WE (1926) The geology and palaeontology of the Kaiso Bone-Beds: Fossil Reptilia. Uganda Protectorate Geological Survey Department Occasional Paper 2: 37–44.

T

Taplin, L. E. (1985). Evolution and zoogeography of crocodilians: a new look at an ancient Order; pp. 361-370 in M. Archer (ed.), Zoogeography and Evolution in Australian Animals in Space and Time. Hesperion, Perth.

Taplin, L. E. (1988). Osmoregulation in crocodilians. Biological Reviews, 63(3), 333–377. https://doi.org/10.1111/j.1469-185X.1988.tb00721.x

Taplin, L., & Grigg, G. (1981). Salt glands in the tongue of estuarine crocodile *Crocodylus porosus.* Science, 212, 1045–1048.

Taplin, L. E., and G. Grigg. (1989). Historical zoogeographie of the eusuchian Crocodilians: a physiological perspective. American Zoologist 29:885-901.

Taplin, L. E., Grigg, G. C., Harlow, P., Ellis, T. M., & Dunson, W. A. C. N.-C. (1982). Lingual salt glant in *Crocodylus acutus* and *C. johnstoni* and their absence from *Alligator mississipiens* and *Caiman crocodilus*. J Comp Physiol, 149(1982), 43–47.

Taplin, L. E., & Loveridge, J. P. (1988). Nile crocodiles, *Crocodylus niloticus*, and estuarine crocodiles, *Crocodylus porosus*, show similar osmoregulatory responses on exposure to seawater. Comparative Biochemistry and Physiology -- Part A: Physiology, 89(3), 443–448. https://doi.org/10.1016/0300-9629(88)91054-7

Targarona, R.R., Soberón, R.R., Cotayo, L., Tabet, M.A. and Thorbjarnarson, J. (IUCN Crocodile Specialist Group) (2008). *Crocodylus rhombifer* (errata version published in 2017). The IUCN Red List of Threatened Species 2008: e.T5670A112902585. http://dx.doi.org/10.2305/IUCN.UK.1996.RLTS.T5670A11516438.en

Thorbjarnarson, J. (1999). Crocodile Tears and Skins : International Trade , Economic Constraints , and Limits to the Sustainable Use of Crocodilians. Conservation Biology, 13(3), 465–470.

Thorbjarnarson, J., Mazzotti, F., Sanderson, E., Buitrago, F., Lazcano, M., Minkowski, K., … Velasco, A. (2006). Regional habitat conservation priorities for the American crocodile. Biological Conservation, 128(1), 25–36. https://doi.org/http://dx.doi.org/10.1016/j.biocon.2005.09.013

U

UNEP/IUCN. (1988). Coral Reefs of the World. Volume 1. Atlantic and Eastern Pacific.UNEP Regional Seas Directories and Bibliographies. IUCN, Gland, Switzerland and Cambridge U. K./UNEP, Nairobi, Kenya. 203-223.

V

Valen, L. Van. (1976). Ecological Species, Multispecies and Oaks. Taxon, 25(2), 233–239

Venegas-Anaya, M., Crawford, A. J., Escobedo Galván, A. H., Sanjur, O. I., Densmore, L. D., & Bermingham, E. (2008). Mitochondrial DNA phylogeography of *Caiman crocodilus* in Mesoamerica and South America. Journal of Experimental Zoology Part A: Ecological Genetics and Physiology, 309A(10), 614–627. https://doi.org/10.1002/jez.502

W

Waples, R. S., & Allendorf, F. (2015). Testing for Hardy-Weinberg proportions: Have we lost the plot? Journal of Heredity, 106(1), 1–19. https://doi.org/10.1093/jhered/esu062

Weaver, J. P., Rodriguez, D., Venegas-Anaya, M., Cedeño-Vázquez, J. R., Forstner, M. R. J. J., & Densmore, L. D. (2008). Genetic characterization of captive cuban

crocodiles (*Crocodylus rhombifer*) and evidence of hybridization with the American crocodile (*Crocodylus acutus*). Journal of Experimental Zoology Part A: Ecological Genetics and Physiology, 309(10), 649–660. https://doi.org/10.1002/jez.471

Wichura, H., Jacobs, L. L., Lin, A., Polcyn, M. J., Manthi, F. K., Winkler, D. A., … Clemens, M. (2015). A 17-My-old whale constraints onset of uplift and climate change in East Africa, 112(13), 3910–3915. https://doi.org/10.1073/pnas.1421502112

Wiley, E. O. (1978). The Evolutionary Species Concept Reconsidered. Systematic Zoology, 27, 17–26. https://doi.org/10.2307/2412809

Wilson, A., Arcese, P., Keller, L. F., Pruett, C. L., Winker, K., Patten, M. A., & Chan, Y. (2009). The contribution of island populations to in situ genetic conservation. Conservation Genetics, 10(2), 419–430. https://doi.org/10.1007/s10592-008-9612-3

Wirtz, P. (1999). Mother species-father species: Unidirectional hybridization in animals with female choice. Animal Behaviour, 58(1), 1–12. https://doi.org/10.1006/anbe.1999.1144

APPENDICES

Appendix A: Chapter 1

Table A1. Samples used in the study and accession numbers in Genbank

Species	Genbank Accesion Number	Origin	Reference	Locality	Latitude	Longitude
C.acutus	160wgs_BCH	Wild	This study	Banco Chinchorro	18.594084°	-87.317156°
C.acutus	BC3019_PacificMex	Wild	This study	Chiapas	15.872824°	-93.663355°
C.acutus	COZHA143_Cozumel	Wild	This study	Cozumel	20.315231°	-86.986280°
C.acutus	QRCHA103_Yucatan	Wild	This study	Xcalak	18.274250°	-87.840748°
C.acutus	RL82_RiaLagartos	Wild	This study	Rial Lagartos	21.472408°	-88.131728°
C.intermedius	HM636895	Unknown	Zhang et al., 2011			
C.intermedius	NC015648	Unknown	Zhang et al., 2011			
C.johnsoni	HM488008	Captive	Meganathan et al., 2011			
C.johnsoni	NC015238	Captive	Meganathan et al., 2011			
C.mindorensis	GU144287	University Collection	Feng et al., 2010			
C.mindorensis	NC014670	University Collection	Feng et al., 2010			
C.moreletii	ALT16_Tamaulipas	Wild	This study	Altamira	22.376960°	-98.009175°
C.moreletii	ALT20_Tamaulipas	Wild	This study	Altamira	22.376960°	-98.009175°
C.moreletii	CIB2527_Tamaulipas	Wild	This study	Cabezas	21.843127°	-99.306016°
C.moreletii	PA2521_Nmex	Wild	This study	Altamira	22.376960°	-99.306016°
C.niloticus	JF502243	Wild	Meredith et al., 2011	Gambia		
C.niloticus	JF502246	Wild	Meredith et al., 2011	Madagascar		
C.niloticus	JF502244	Wild	Meredith et al., 2011	Mauritania		
C.niloticus	JF502245	Wild	Meredith et al., 2011	Zimbabwe		
C.novaeguineae	HM636896	University Collection	Zhang et al., 2011			
C.novaeguineae	NC015651	University Collection	Zhang et al., 2011			
C.palustris	HM488007	Captive	Meganathan et al., 2011			

C.palustris	GU144286	University Collection	Feng et al., 2010			
C.palustris	HM488007	Captive	Meganathan et al., 2011			
C.porosus	AF542533	Wild	Fitzsimmons et al., 2002			
C.porosus	AF542535	Wild	Fitzsimmons et al., 2002			
C.porosus	AF542536	Wild	Fitzsimmons et al., 2002			
C.porosus	AF542538	Wild	Fitzsimmons et al., 2002			
C.porosus	AJ810453	Unknown	Janke et al., 2005			
C.porosus	DQ273698	Unknown	Li et al., 2007			
C.porosus	JQ237683	Wild	Luck et al., 2012			
C.porosus	JQ237684	Wild	Luck et al., 2012			
C.porosus	JQ237685	Wild	Luck et al., 2012			
C.rhombifer	JX292787	Unknown	Unpublished			
C.rhombifer	NC024513	Unknown	Unpublished			
C.siamensis	AF542540	Wild	Fitzsimmons et al., 2002			
C.siamensis	DQ353946	University Collection	Ji et al., 2008			
C.siamensis	EF581859	WIld	Srikulnath et al., 2012			
C.siamensis	NC008795	University Collection	Ji et al., 2008			
G.gangeticus	AB079596	Unknown	Unpublished			
G.gangeticus	AJ810454	Unknown	Janke et al., 2005			
G.gangeticus	NC008241	Unknown	Unpublished			
M.cataphractus	EF551000	Unknown	Unpublished			
M.cataphractus	NC010639	Unknown	Unpublished			
O.tetraspis	NC009728	Unknown	Roos et al., 2007			
A.sinensis	AF511507	Captive	Wu, 2003			
A.mississippiensis	Y13113	Wild	Janke & Arnason, 1997			
C.crocodylus	AJ404872	Captive	Janke et al., 2001			
P.trogonatus	NC009732	Unknown	Janke et al., 2005			
T.schlegelli	NC011074	Unknown	Janke et al., 2005			

Table A2. Best three models of the Jmodel test for the AIC and BIC criterions.

	AIC		BIC	
Crocodylus	TPM3uf+I+G	3491.791	HKY+I+G	3878.847
	TIM3+I+G	3493.578	TPM3uf+I+G	3880.25
	TPM1uf+I+G	3494.217	TPM1uf+I+G	3882.676
Crocodylus+Gavialis	TPM3uf+I+G	4540.899	TPM3uf+I+G	4991.439
	TIM3+I+G	4542.678	HKY+I+G	4992.694
	TVM+I+G	4543.984	TPM1uf+I+G	4995.372

Table A3. Character matrix with the geographic models tested in Mezquite

Character	Model 1	Model 2	Model 3	Model 4	Model 5
0	Indopacific	Indopacific	Indopacific	Indopacific	Indopacific
1	Africa	West Africa	West Africa	Africa	West Africa
2	Netropics	East Africa	East Africa	Neotropics 1	East Africa
3		Neotropics 1	Neotropics	Neotropics 2	Neotropics 1
4		Neotropics 2			Neotropics 2

Sample	Model1	Model 2	Model 3	Model 4	Model 5
C. acutus BCH160	2	3/4	3	3	4
C. acutus BCH273	2	3/4	3	3	4
C. acutus BCH285	2	3/4	3	3	4
C. acutus Boca del Cielo BC3001	2	3/4	3	3	4
C. acutus Boca del Cielo BC3020	2	3/4	3	3	4
C. acutus Cozumel COZ196	2	3/4	3	3	4
C. acutus Cozumel COZHA135	2	3/4	3	3	4
C. acutus RBSK QRCHA89	2	3/4	3	3	4
C. acutus Xcalak QRCHA101	2	3/4	3	3	4
C. intermedius HM636895.1	2	3/4	3	3	4
C. intermedius NC 015648.1	2	3/4	3	3	4
C. johnsoni HM488008.2	0	0	0	0	0
C. johnsoni NC 015238.2	0	0	0	0	0
C. mindorensis GU144287.1	0	0	0	0	0
C. mindorensis NC 014670.1	0	0	0	0	0
C. moreletii Altamira ALT16	2	3/4	3	2	3
C. moreletii Altamira ALT20	2	3/4	3	2	3
C. moreletii Altamira ALT2530	2	3/4	3	2	3
C. moreletii Altamira ALT2531	2	3/4	3	2	3
C. moreletii Altamira PA2539	2	3/4	3	2	3
C. moreletii Cabezas CIB2524	2	3/4	3	2	3
C. moreletii Cabezas CIB351	2	3/4	3	2	3
C. acutus Ria Lagartos RL82	2	3/4	3	3	4
C. niloticus Gambia JF502243.1	1	1/2	1	1	1
C. niloticus Madagascar JF502246.1	1	1/2	2	1	2
C. niloticus Mauritania JF502244.1	1	1/2	1	1	1
C. niloticus Zimbabwe JF502245.1	1	1/2	2	1	2
C. novaeguineae HM636896.1	0	0	0	0	0
C. novaeguineae NC 015651.1	0	0	0	0	0
C. palustris HM488007.1	0	0	0	0	0
C. porosus AF542533.1	0	0	0	0	0

C. porosus AF542535.1	0	0	0	0	0
C. porosus AF542536.1	0	0	0	0	0
C. porosus AF542538.1	0	0	0	0	0
C. porosus AJ810453.1	0	0	0	0	0
C. porosus DQ273698.1	0	0	0	0	0
C. rhombifer JX292787.1	2	3/4	3	2	3
C. rhombifer NC 024513.1	2	3/4	0	2	3
C. siamensis AF542540.1	0	0	0	0	0
C. siamensis DQ353946.1	0	0	0	0	0
C. siamensis EF581859.1	0	0	0	0	0
C. siamensis NC 008795.1	0	0	0	0	0
M. cataphractus EF551000.1	1	1/2	1	1	1
M. cataphractus NC 010639.1	1	1/2	1	1	1
O. tetraspis NC 009728.1	1	1/2	1	1	1

Table A4. BioGeoBears probabilities of the different models for the nodes in Figure 2.

Model	Biogeographic origin							
DEC + J Stratified	**None**	**Africa**	**Neotropics**	**Indopacific**	**Africa-Neotropics**	**Africa-Indopacific**	**Neotropics - Indopacific**	**A+N+I**
1	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
3	0.00	0.98	0.02	0.00	0.00	0.00	0.00	0.00
4	0.00	0.52	0.48	0.00	0.00	0.00	0.00	0.00
5	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
6	0.00	0.81	0.00	0.14	0.00	0.05	0.00	0.00
7	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
8	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
9	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
11	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
12	0.00	0.06	0.00	0.87	0.00	0.07	0.00	0.00
13	0.00	0.40	0.00	0.22	0.00	0.38	0.00	0.00
14	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
DEC Stratified	**None**	**Africa**	**Neotropics**	**Indopacific**	**Africa-Neotropics**	**Africa-Indopacific**	**Neotropics - Indopacific**	**A+N+I**
1	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	0.91	0.00	0.09	0.00	0.00	0.00
3	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
4	0.00	0.00	0.33	0.00	0.67	0.00	0.00	0.00
5	0.00	0.99	0.00	0.00	0.01	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00
7	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
8	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
9	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
11	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
12	0.00	0.00	0.00	0.18	0.00	0.82	0.00	0.00
13	0.00	0.45	0.00	0.00	0.00	0.55	0.00	0.00
14	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
DEC + J Unstratified	**None**	**Africa**	**Neotropics**	**Indopacific**	**Africa-Neotropics**	**Africa-Indopacific**	**Neotropics - Indopacific**	**A+N+I**
1	0	0.00	1.00	0.00	0.00	0.00	0.00	0.00
2	0	0.04	0.95	0.00	0.01	0.00	0.00	0.00
3	0	0.13	0.87	0.00	0.00	0.00	0.00	0.00

4	0	0.05	0.95	0.00	0.01	0.00	0.00	0.00
5	0	0.54	0.41	0.00	0.04	0.00	0.00	0.00
6	0	0.13	0.07	0.60	0.01	0.09	0.05	0.04
7	0	0.00	0.00	1.00	0.00	0.00	0.00	0.00
8	0	0.00	0.00	1.00	0.00	0.00	0.00	0.00
9	0	0.00	0.00	1.00	0.00	0.00	0.00	0.00
10	0	0.00	0.00	1.00	0.00	0.00	0.00	0.00
11	0	0.00	0.00	1.00	0.00	0.00	0.00	0.00
12	0	0.07	0.03	0.66	0.00	0.13	0.07	0.04
13	0	0.22	0.01	0.15	0.02	0.50	0.01	0.10
14	0	1.00	0.00	0.00	0.00	0.00	0.00	0.00

DEC Unstratified	None	Africa	Neotropics	Indopacific	Africa-Neotropics	Africa-Indopacific	Neotropics - Indopacific	A+N+I
1	0	0.00	1.00	0.00	0.00	0.00	0.00	0.00
2	0	0.00	0.00	0.00	1.00	0.00	0.00	0.00
3	0	0.00	0.00	0.00	1.00	0.00	0.00	0.00
4	0	0.00	0.00	0.00	1.00	0.00	0.00	0.00
5	0	0.01	0.00	0.00	0.99	0.00	0.00	0.00
6	0	0.00	0.00	0.00	0.00	0.02	0.00	0.98
7	0	0.00	0.00	1.00	0.00	0.00	0.00	0.00
8	0	0.00	0.00	1.00	0.00	0.00	0.00	0.00
9	0	0.00	0.00	1.00	0.00	0.00	0.00	0.00
10	0	0.00	0.00	1.00	0.00	0.00	0.00	0.00
11	0	0.00	0.00	1.00	0.00	0.00	0.00	0.00
12	0	0.00	0.00	0.00	0.00	0.02	0.00	0.98
13	0	0.00	0.00	0.00	0.01	0.02	0.00	0.97
14	0	1.00	0.00	0.00	0.00	0.00	0.00	0.00

Table A5. Mesquite maximum likelihood probabilities of the different models for the nodes in Figure 2.

Node	TPM3uf + I + G			HKY + I + G		
	Indopacific	Africa	Netropics	Indopacific	Africa	Netropics
1	0	0	1	0	0	1
2	0.01	0.02	0.97	0	0.02	0.98
3	0	0.03	0.97	0	0.03	0.97
4	0	0.02	0.98	0	0.03	0.97
5	0.41	0.14	0.45	0.43	0.14	0.43
6	0.97	0.015	0.015	0.98	0.01	0.01
7	1	0	0	1	0	0
8	0.99	0.005	0.005	0.99	0.005	0.005
9	1	0	0	1	0	0
10	0.99	0.005	0.005	1	0	0
11	0.99	0.005	0.005	0.99	0.005	0.005
12	0.98	0.01	0.01	0.98	0.01	0.01
13	0.42	0.5	0.08	0.42	0.5	0.08
14	0.33	0.59	0.08	0.33	0.59	0.08

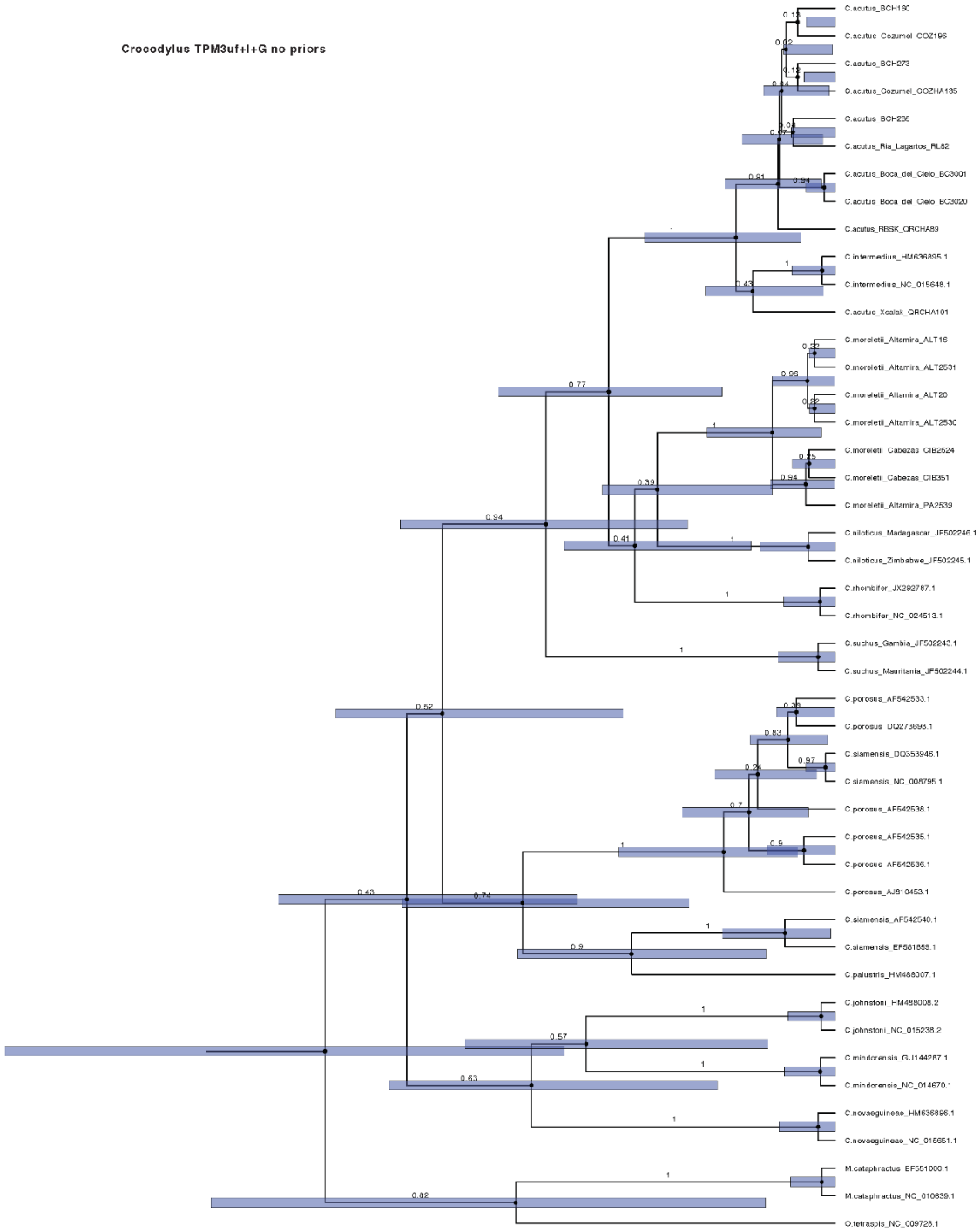

Figure A1. Maximum clade credibility tree for *Crocodylus* TPM3uf + I + G with no priors.

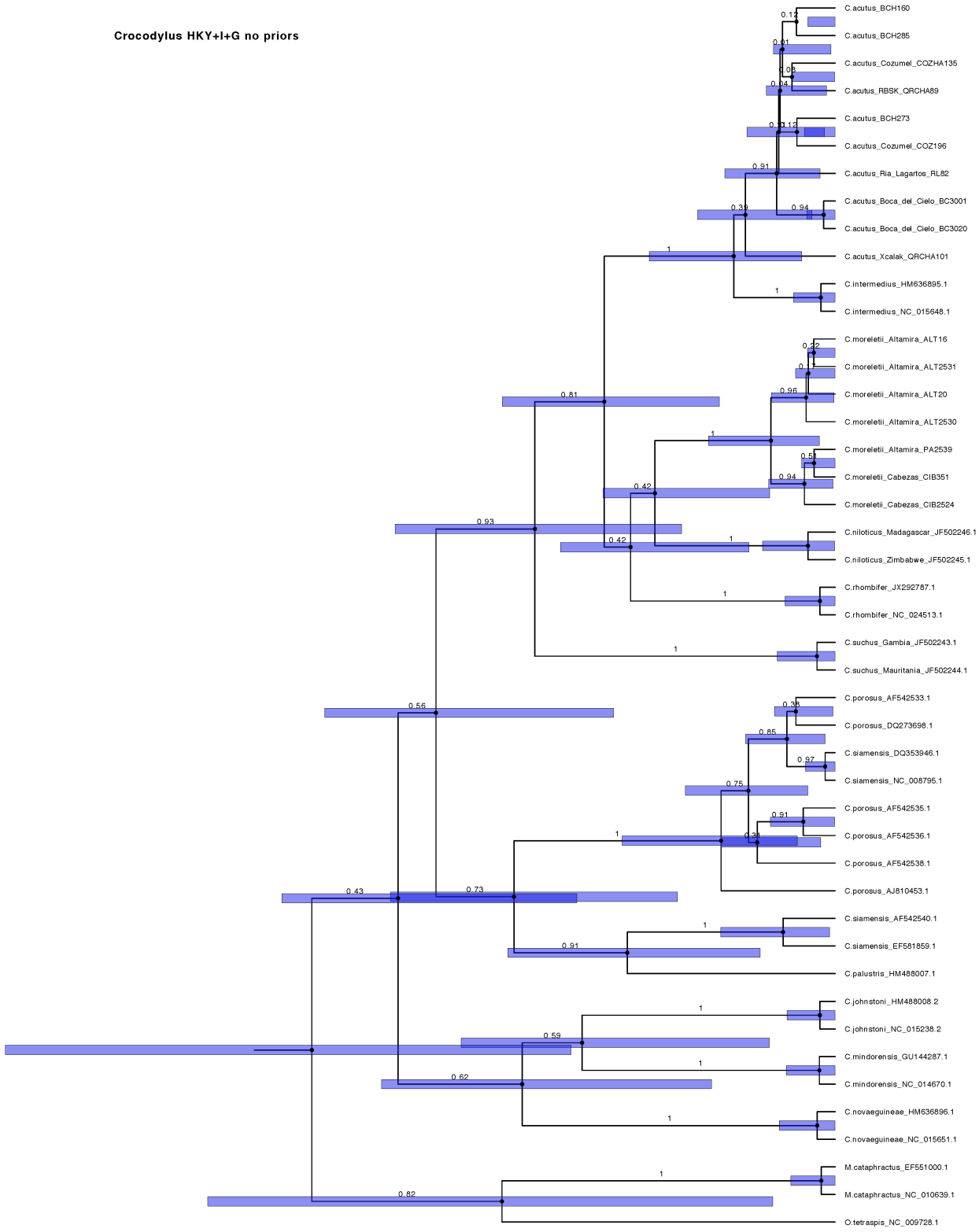

Figure A2. . Maximum clade credibility tree for *Crocodylus* HKY + I + G with no priors.

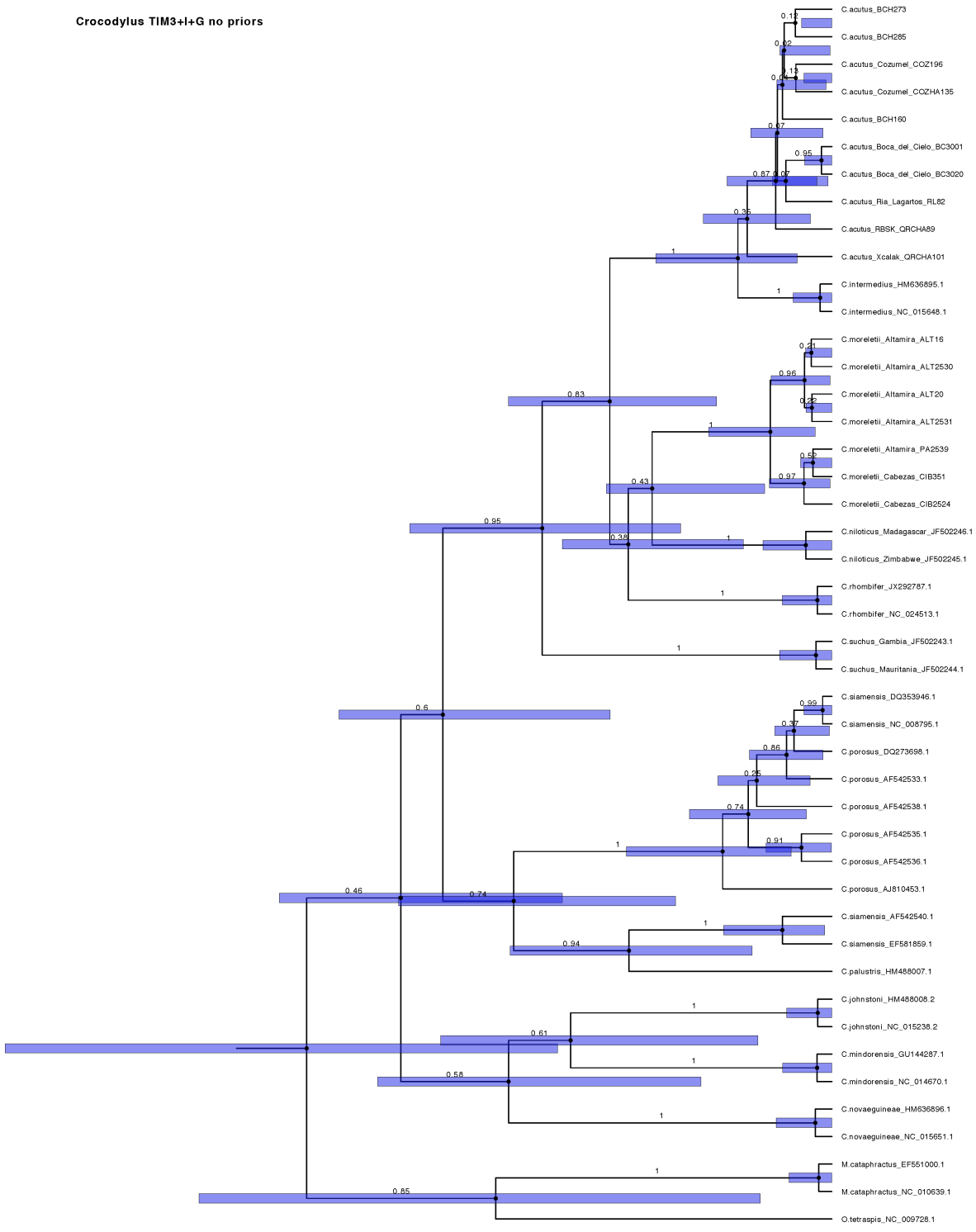

Figure A3. Maximum clade credibility tree for *Crocodylus* TIM3 + I + G with no priors.

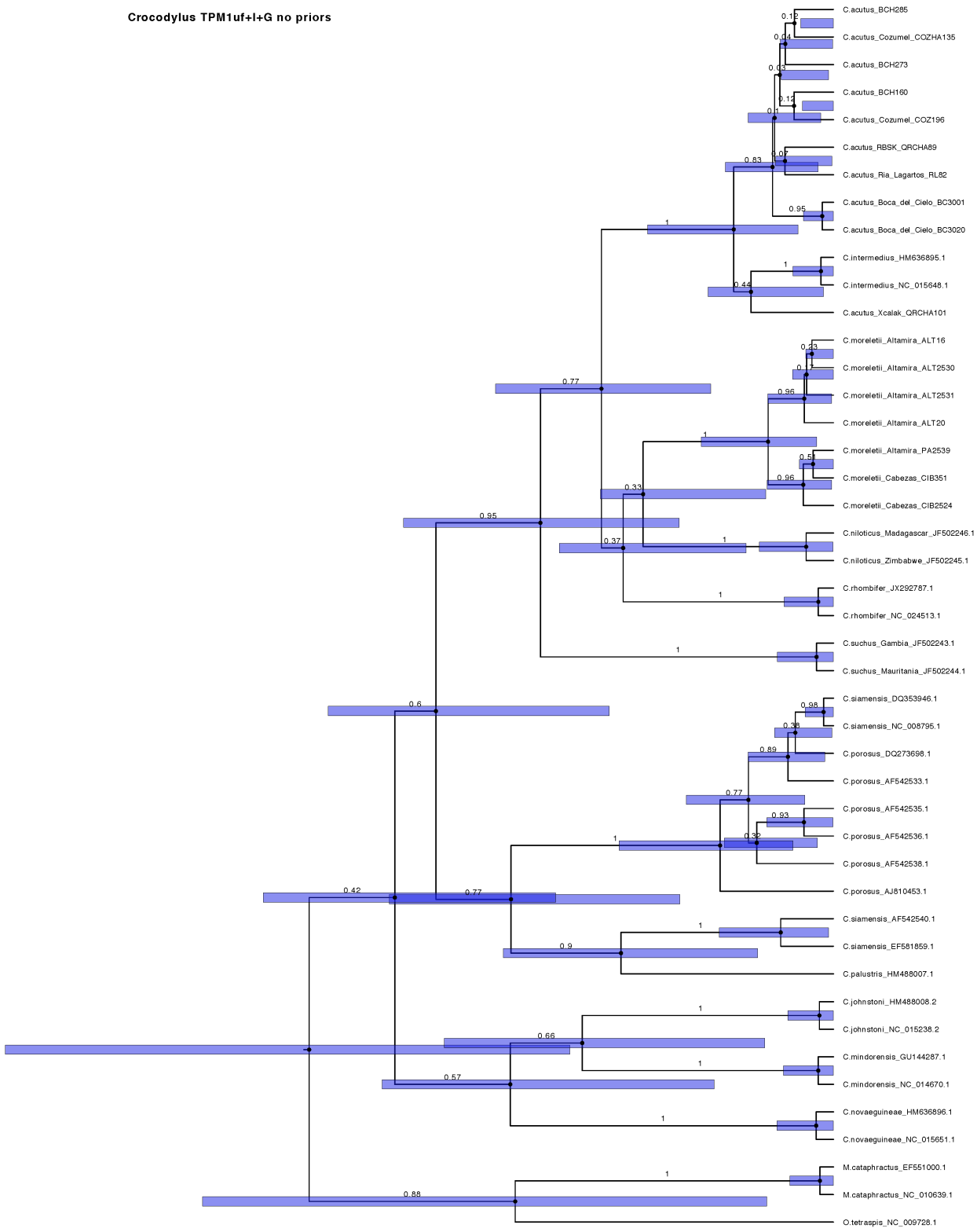

Figure A4. Maximum clade credibility tree for *Crocodylus* TPM1uf + I + G with no priors.

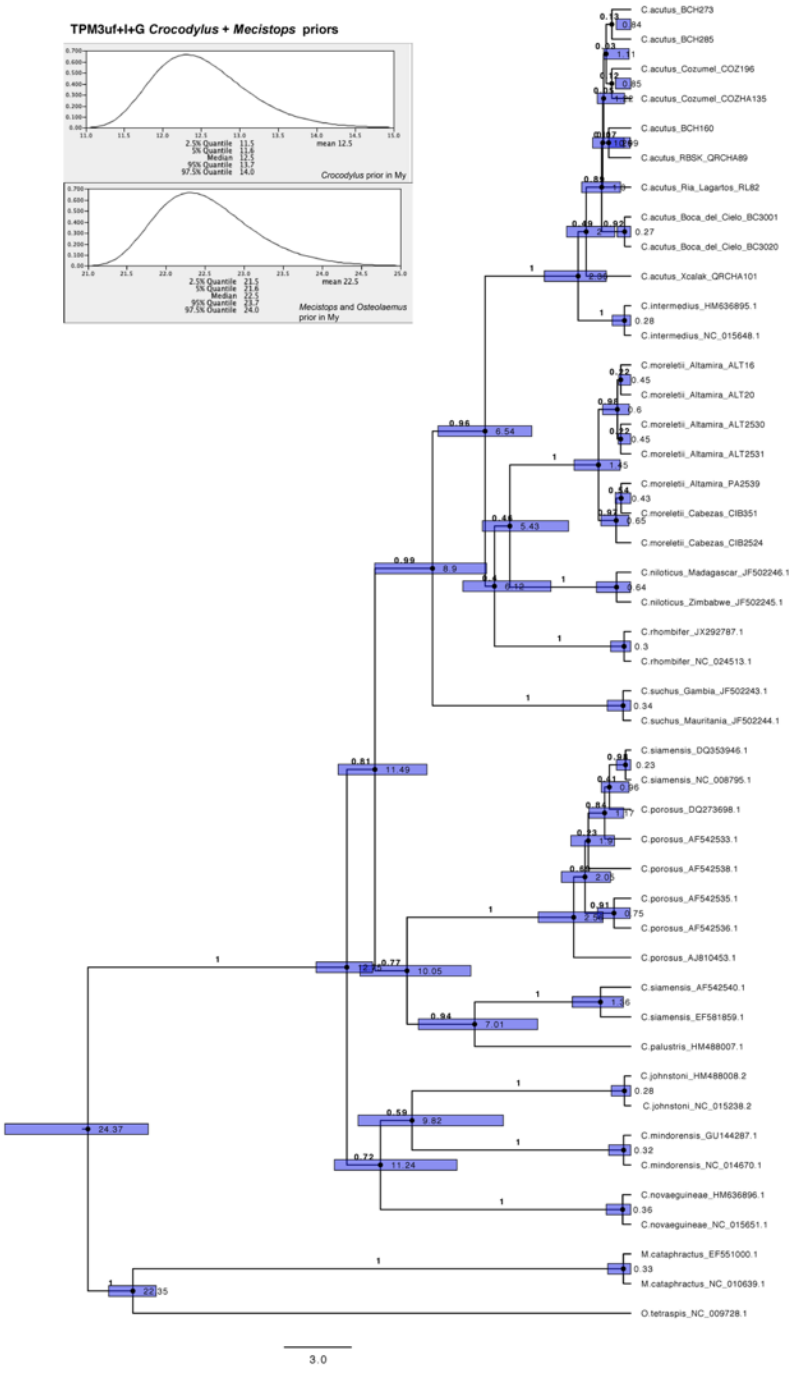

Figure A5. Maximum clade credibility tree for TPM3uf + I + G with Crocodylus divergence time of 10 - 14 Ma, and *Mecistops* and *Osteolaemus* divergence time of 20 - 24 Ma.

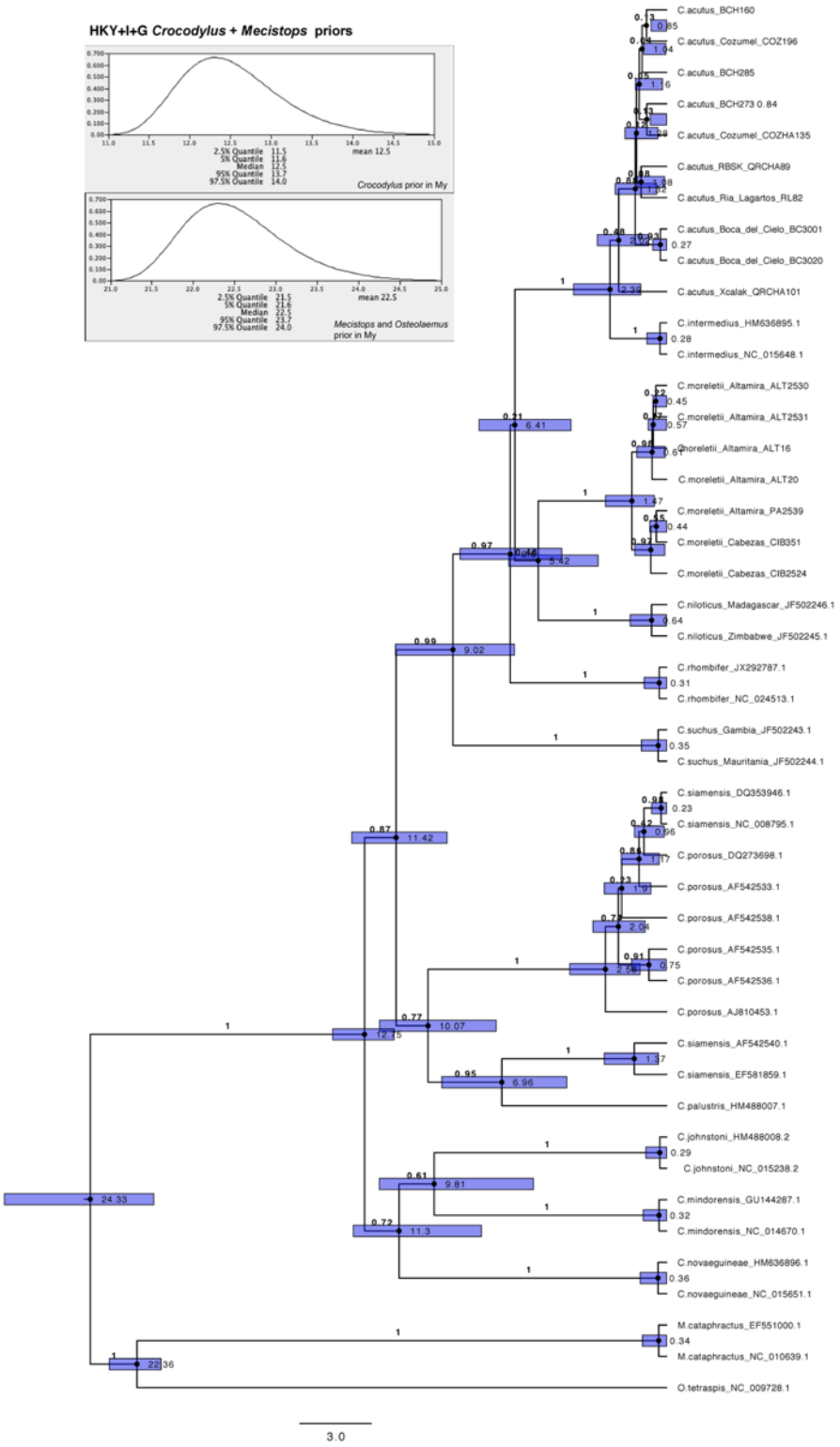

Figure A6. Maximum clade credibility tree for HKY + I + G with Crocodylus divergence time of 10 - 14 Ma, and *Mecistops* and *Osteolaemus* divergence time of 20 - 24 Ma

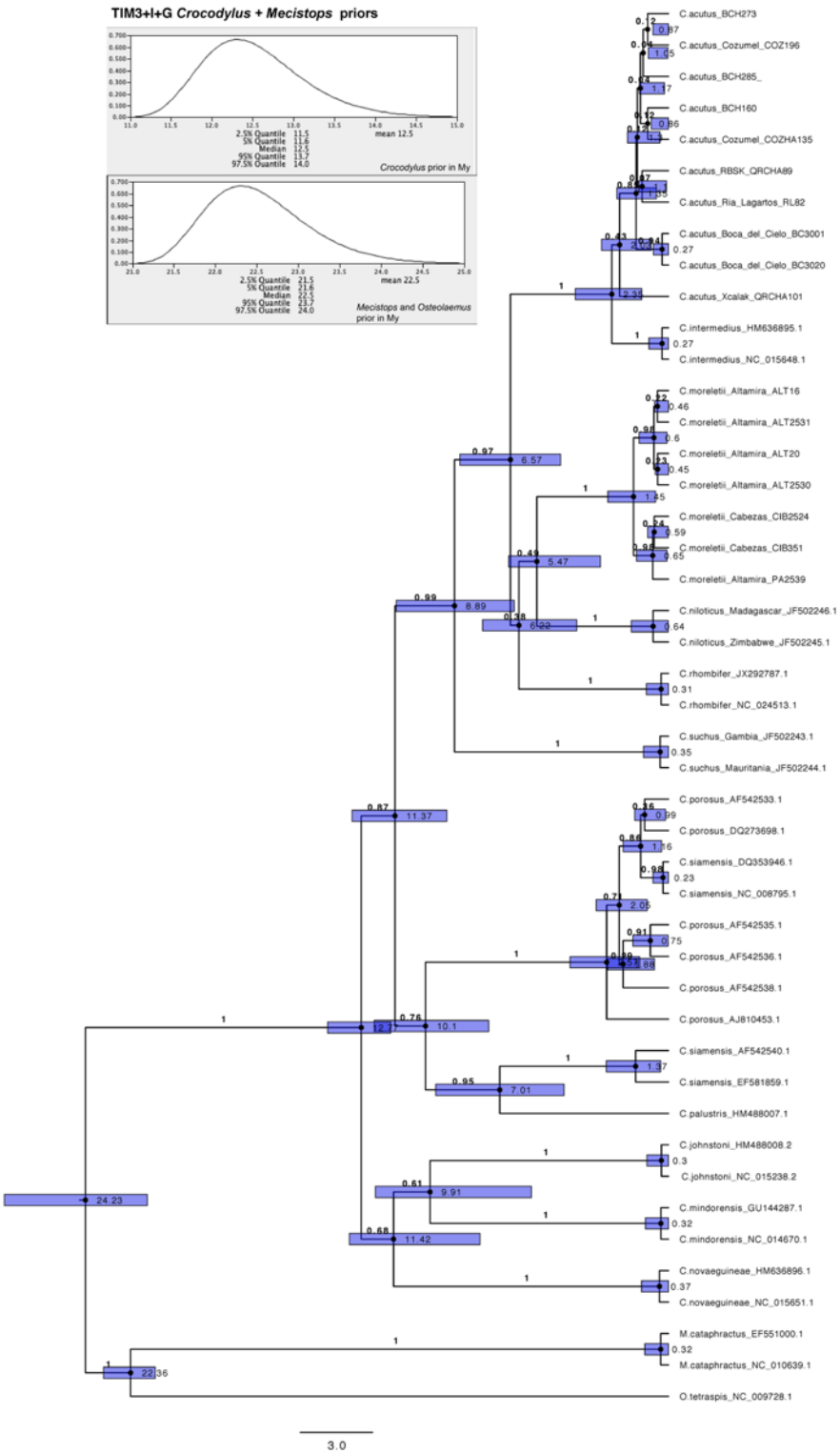

Figure A7. Maximum clade credibility tree for TIM3 + I + G with Crocodylus divergence time of 10 - 14 Ma, and *Mecistops* and *Osteolaemus* divergence time of 20 - 24 Ma.

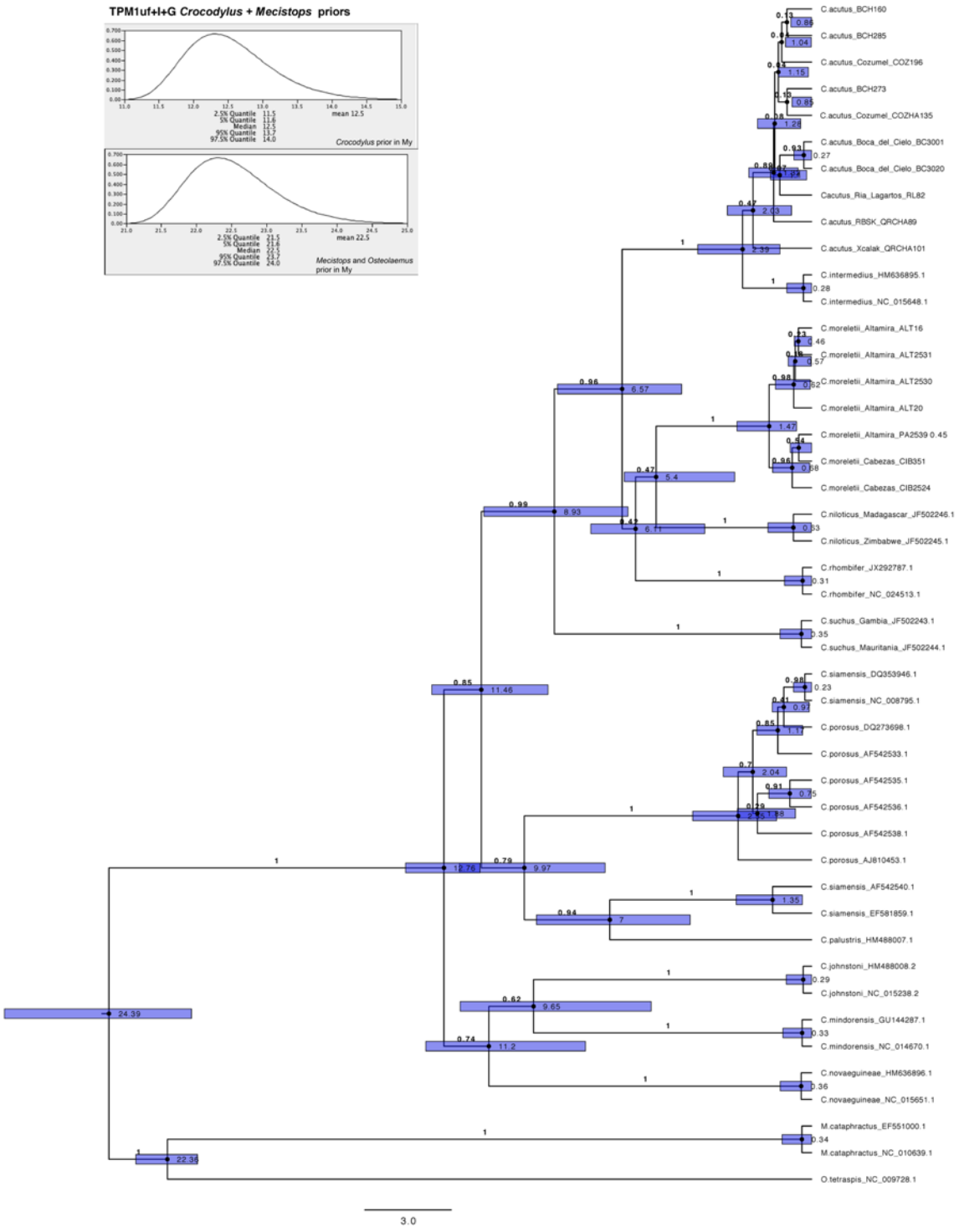

Figure A8. Maximum clade credibility tree for TPM1uf + I + G with Crocodylus divergence time of 10 - 14 Ma, and *Mecistops* and *Osteolaemus* divergence time of 20 - 24 Ma.

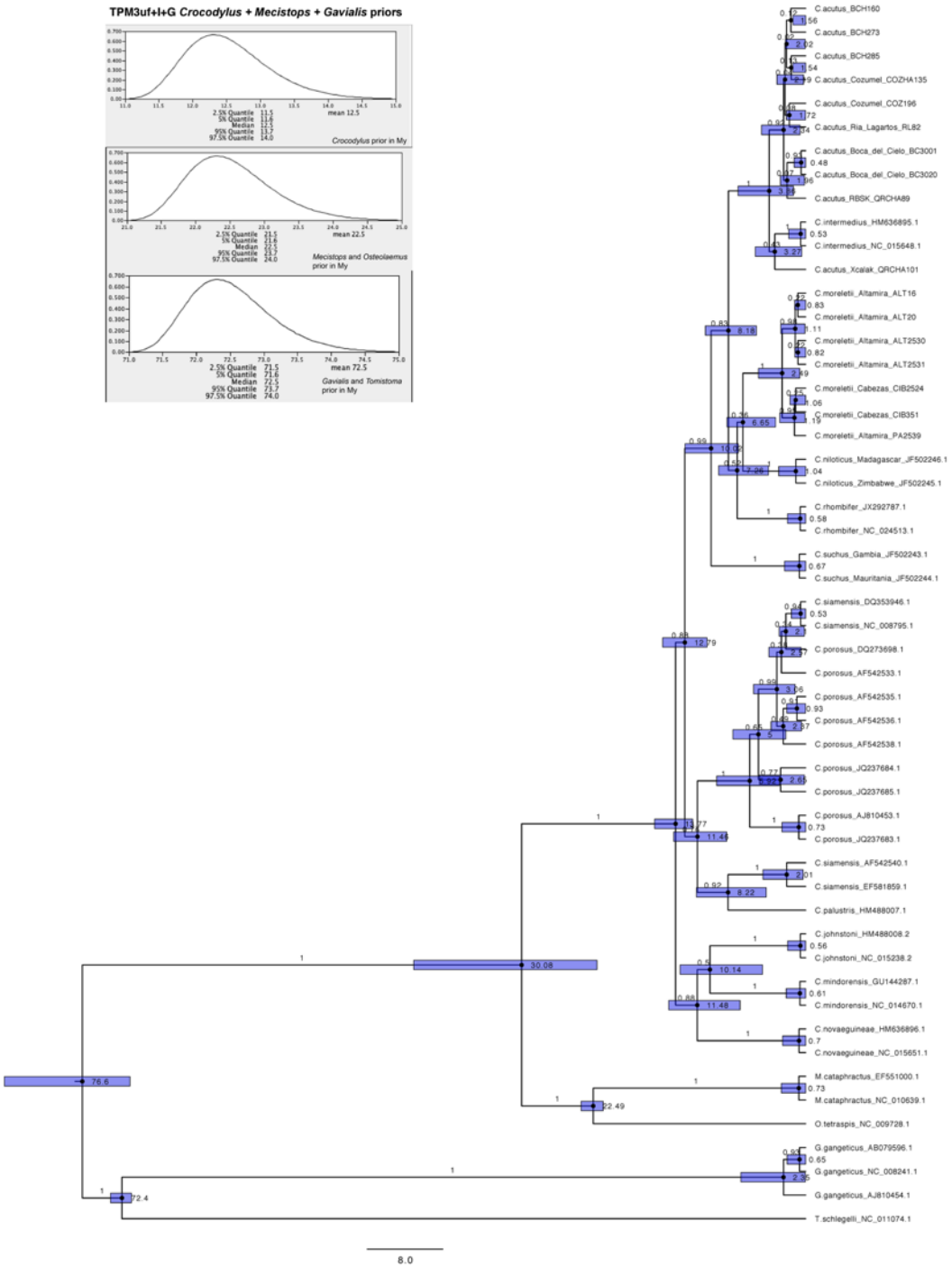

Figure A9. Maximum clade credibility tree for TPM3uf + I + G with Crocodylus divergence time of 10 - 14 Ma, *Mecistops* and *Osteolaemus* divergence time of 20 - 24 Ma and *Gavialis* divergence time of 70 – 100 Ma.

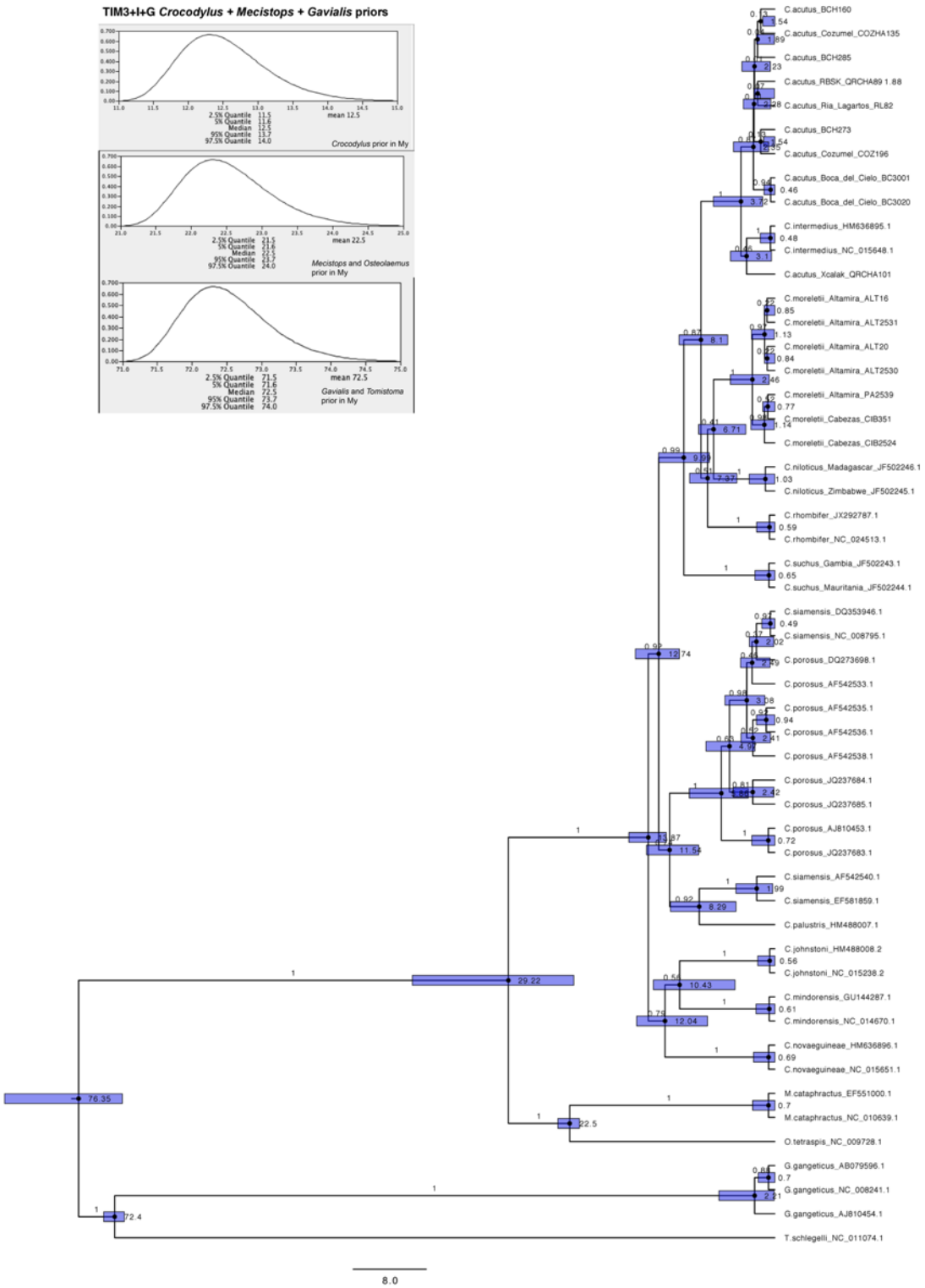

Figure A10. Maximum clade credibility tree for TIM3 + I + G with Crocodylus divergence time of 10 - 14 Ma, *Mecistops* and *Osteolaemus* divergence time of 20 - 24 Ma and *Gavialis* divergence time of 70 – 100 Ma.

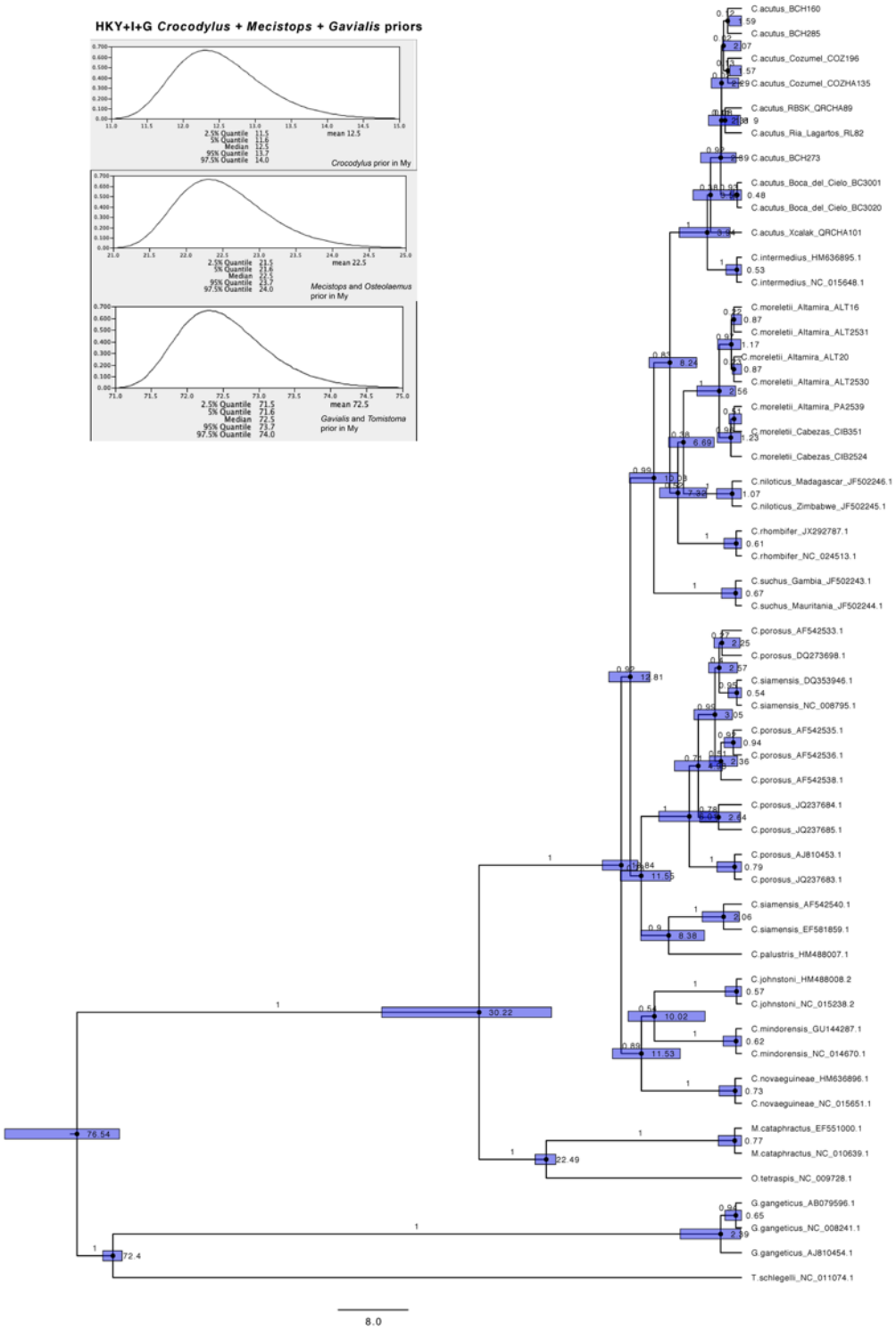

Figure A11. Maximum clade credibility tree for HKY + I + G with Crocodylus divergence time of 10 - 14 Ma, *Mecistops* and *Osteolaemus* divergence time of 20 - 24 Ma and *Gavialis* divergence time of 70 – 100 Ma.

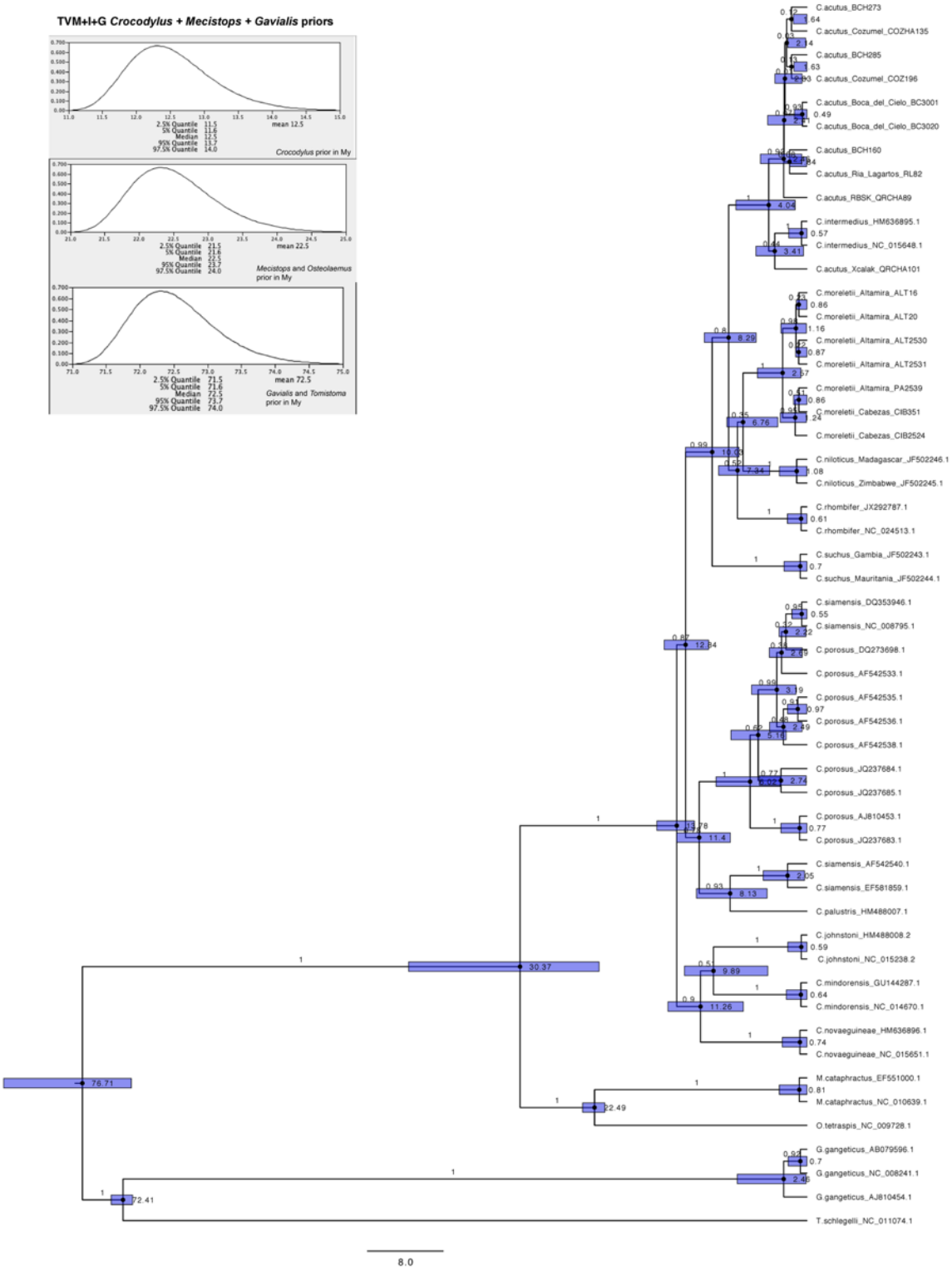

Figure A12. Maximum clade credibility tree for TVM + I + G with Crocodylus divergence time of 10 - 14 Ma, *Mecistops* and *Osteolaemus* divergence time of 20 - 24 Ma and *Gavialis* divergence time of 70 – 100 Ma.

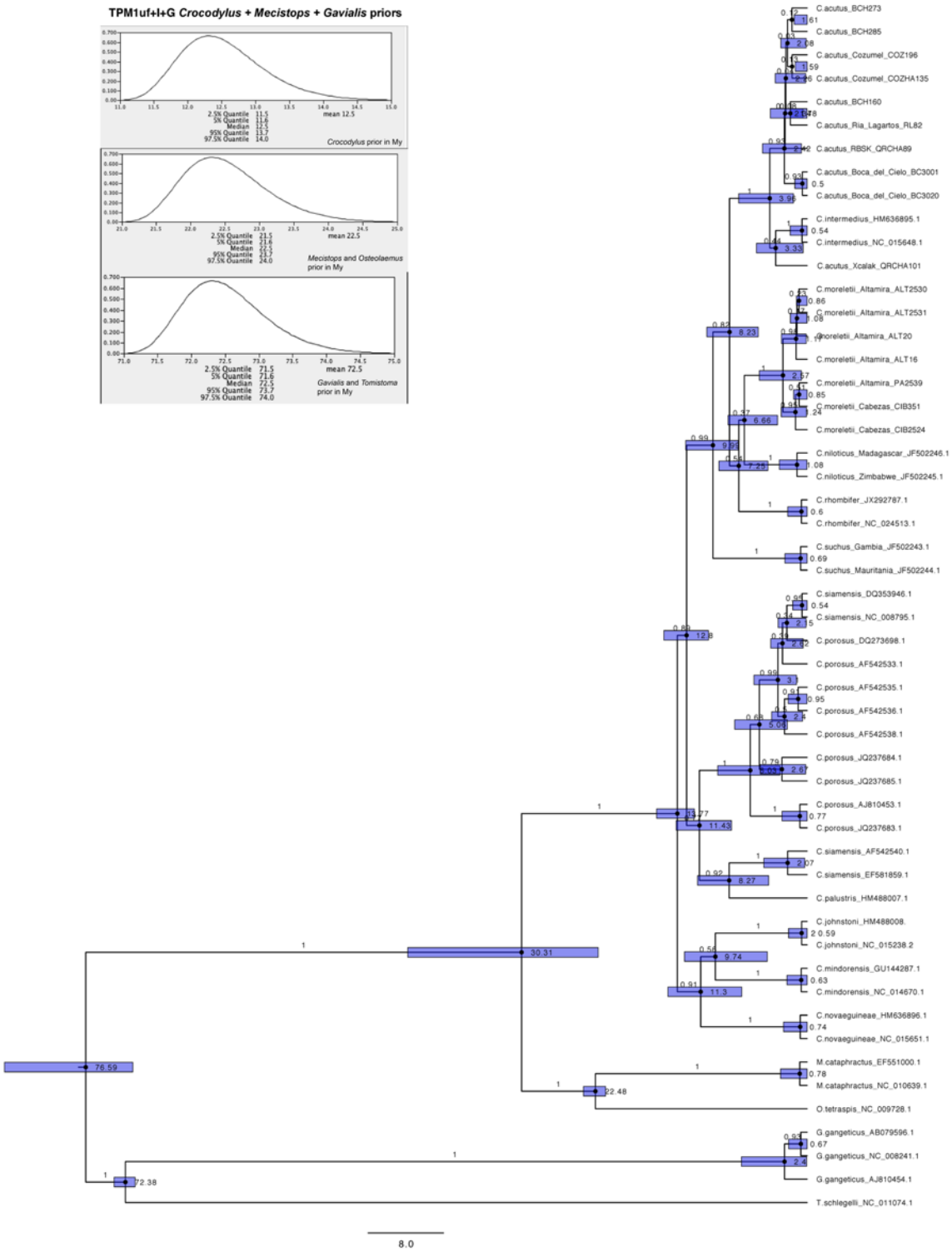

Figure A13. Maximum clade credibility tree for TPM1uf+ I + G with Crocodylus divergence time of 10 - 14 Ma, *Mecistops* and *Osteolaemus* divergence time of 20 - 24 Ma and *Gavialis* divergence time of 70 – 100 Ma.

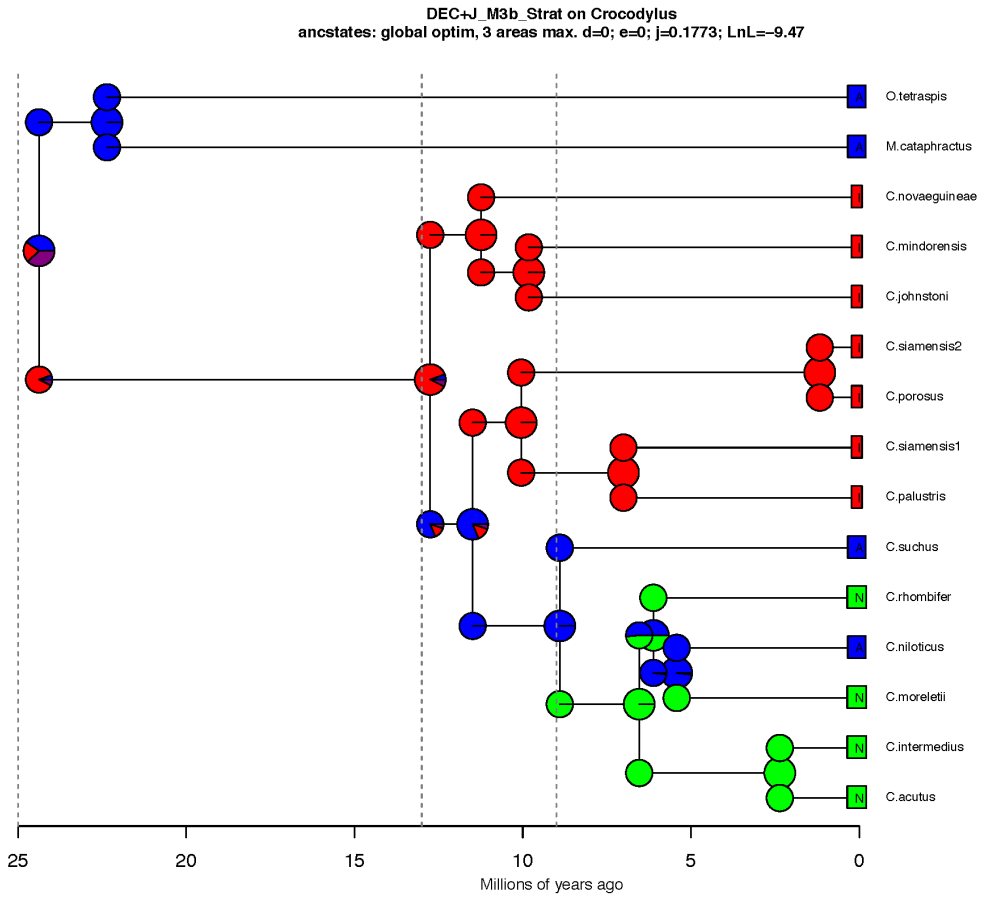

Figure A14. Ancestral state reconstruction of BioGeoBears for the TPM3uf + I + G tree with stratified analysis with DEC + J. Probabilities of the nodes are represented in the circles. The red color corresponds to Indopacific, the blue color to Africa and the green to the Neotropics.

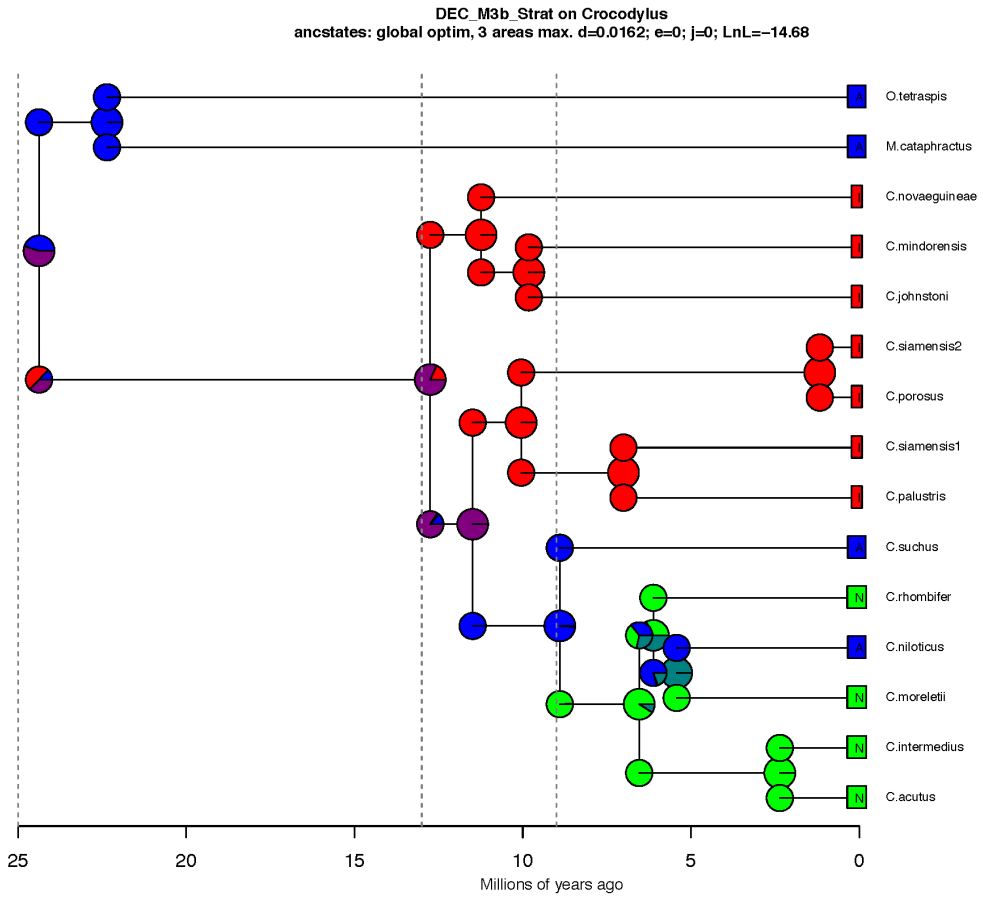

Figure A15. Ancestral state reconstruction of BioGeoBears for the TPM3uf + I + G tree with stratified analysis with DEC. Probabilities of the nodes are represented in the circles. The red color corresponds to Indopacific, the blue color to Africa and the green to the Neotropics.

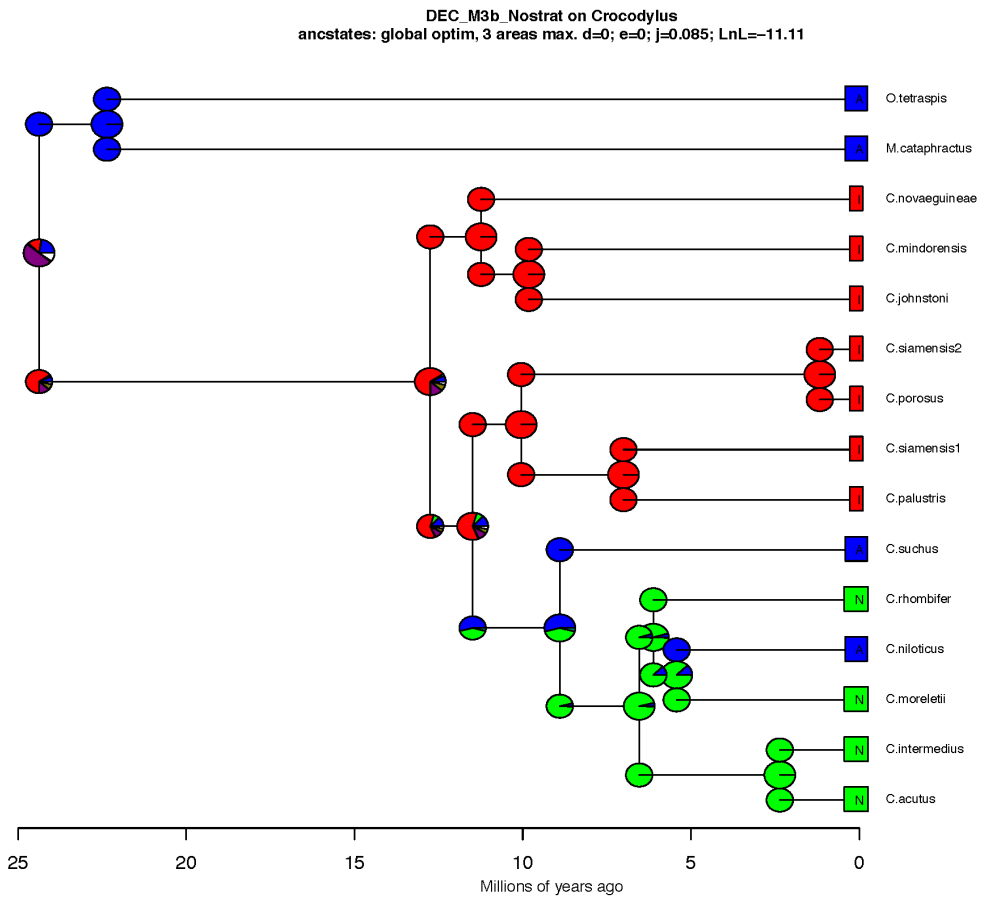

Figure A16. Ancestral state reconstruction of BioGeoBears for the TPM3uf + I + G tree with non-stratified analysis with DEC+J. Probabilities of the nodes are represented in the circles. The red color corresponds to Indopacific, the blue color to Africa and the green to the Neotropics.

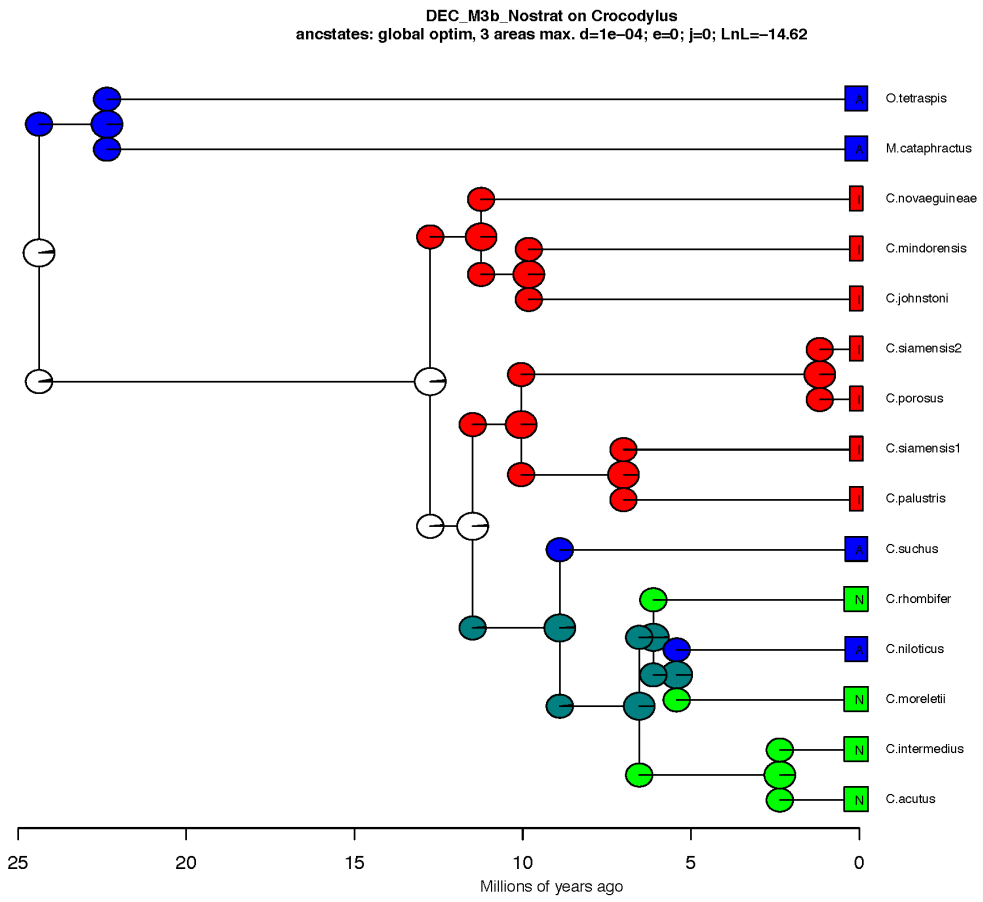

Figure A17. Ancestral state reconstruction of BioGeoBears for the TPM3uf + I + G tree with non-stratified analysis with DEC. Probabilities of the nodes are represented in the circles. The red color corresponds to Indopacific, the blue color to Africa and the green to the Neotropics.

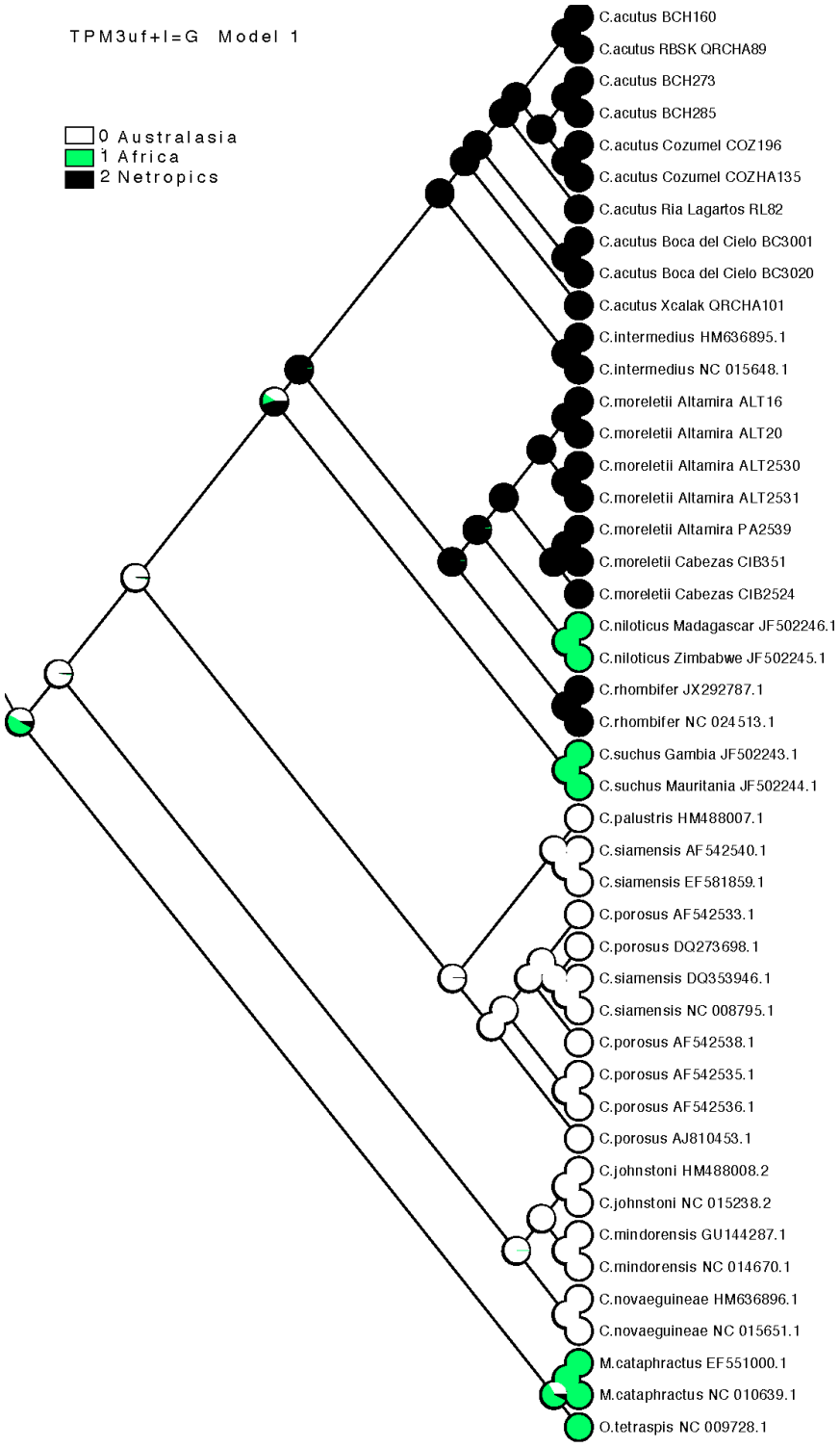

Figure A18. Mesquite Maximum likelihood tree for the TPM3uf + I + G model 1. Probabilities of the nodes are represented in the circles.

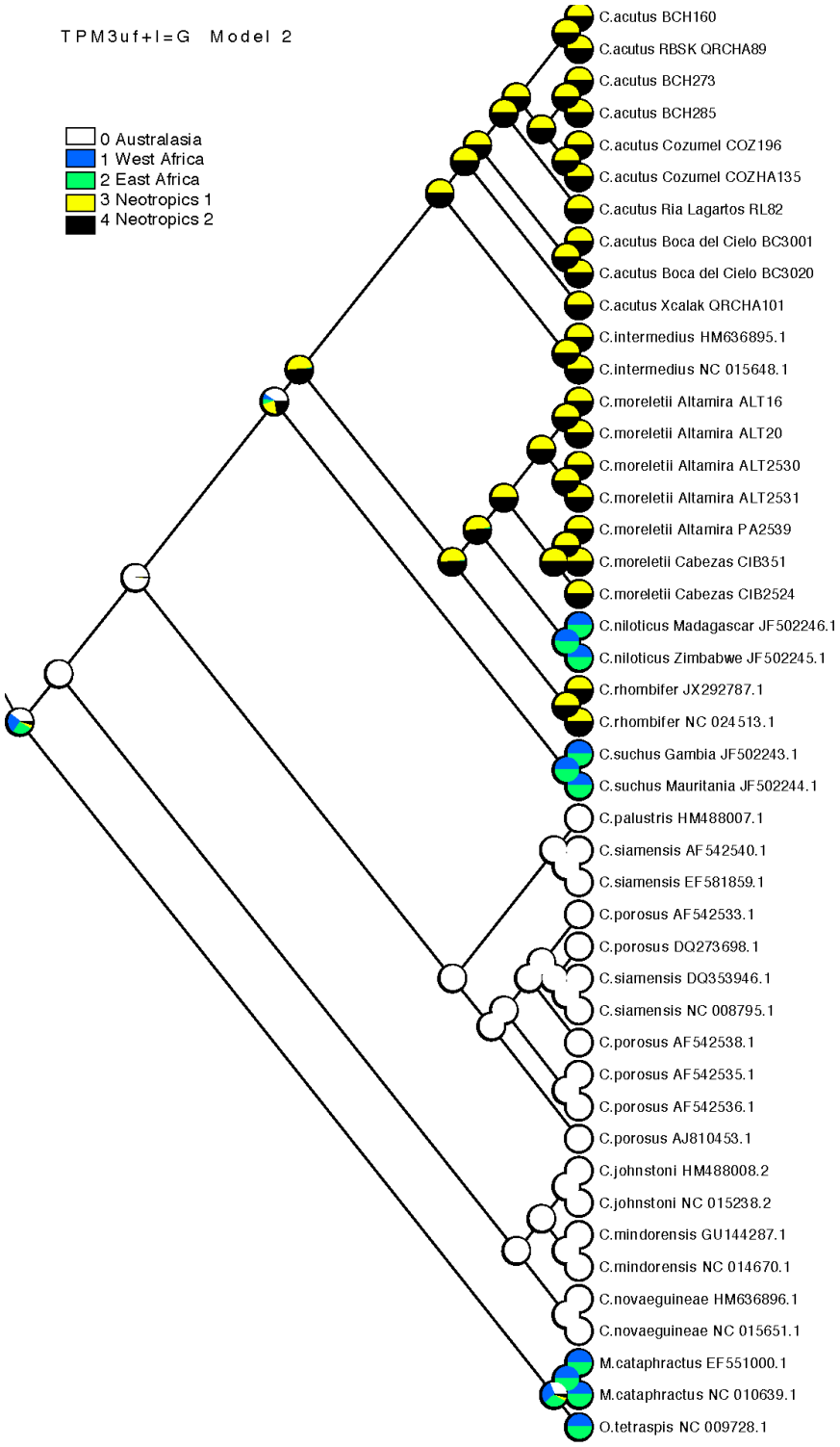

Figure A19. Mesquite Maximum likelihood tree for the TPM3uf + I + G model 2. Probabilities of the nodes are represented in the circles.

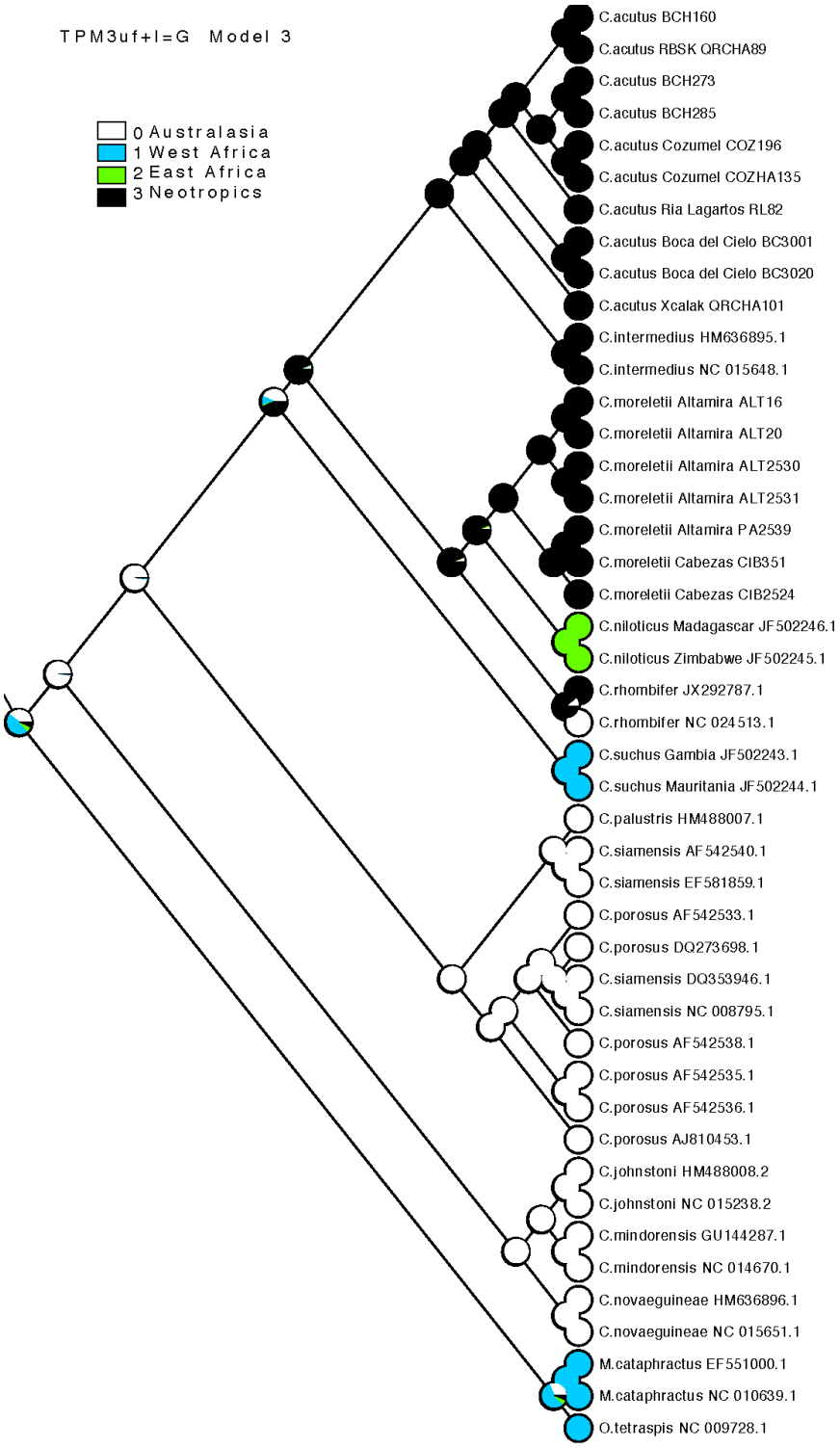

Figure A20. Mesquite Maximum likelihood tree for the TPM3uf + I + G model 3. Probabilities of the nodes are represented in the circles.

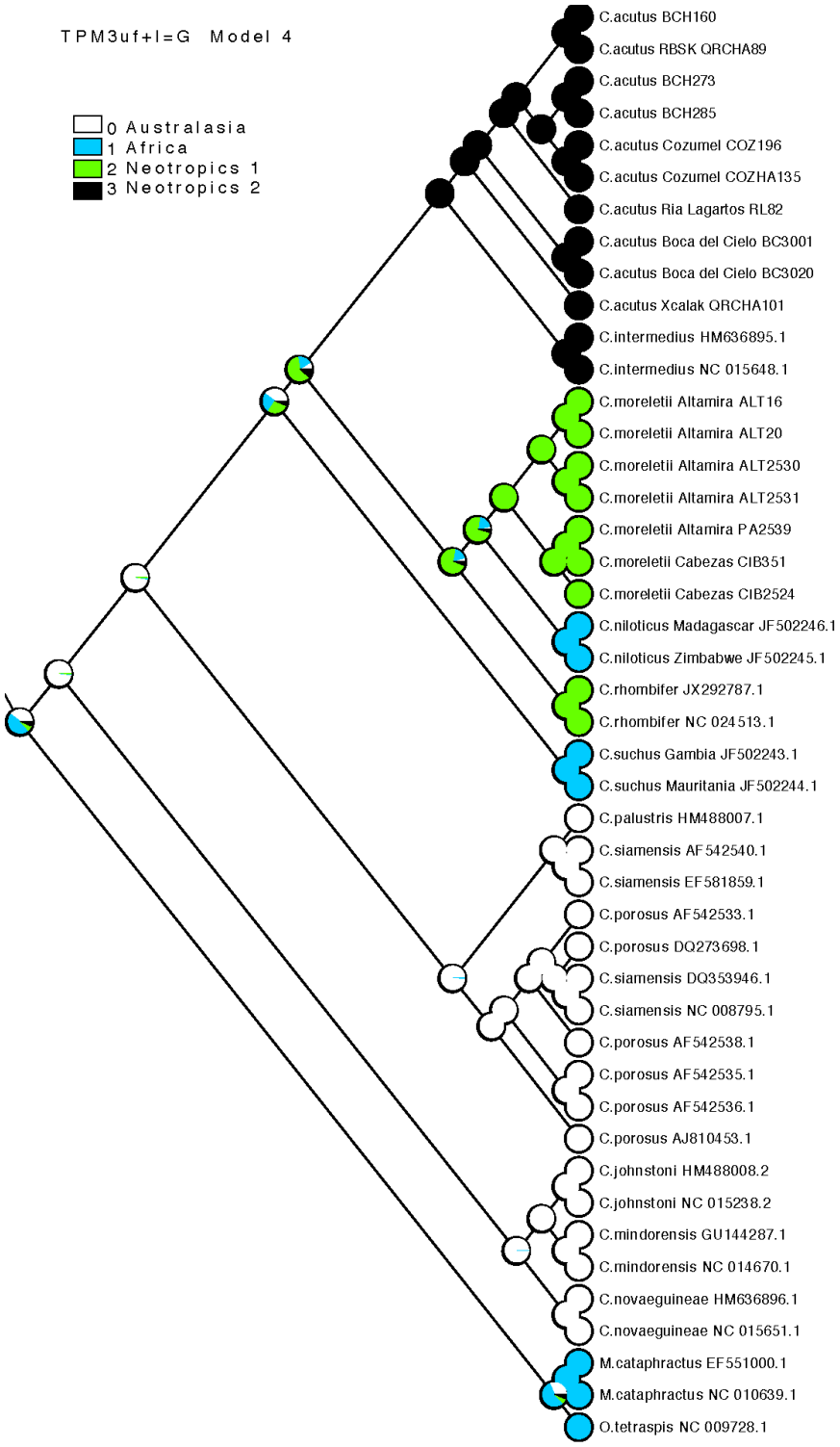

Figure A21. Mesquite Maximum likelihood tree for the TPM3uf + I + G model 4. Probabilities of the nodes are represented in the circles.

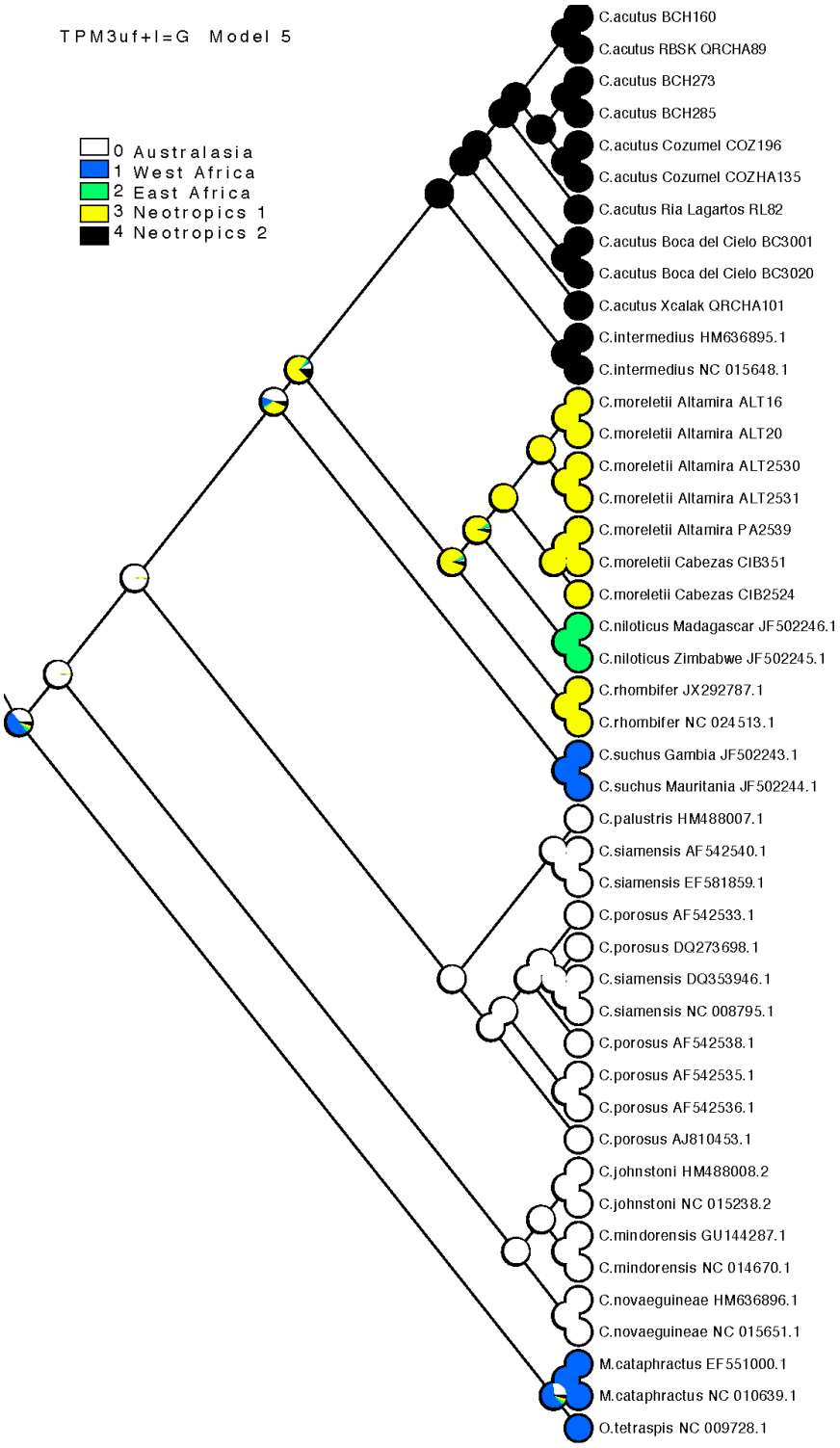

Figure A22. Mesquite Maximum likelihood tree for the TPM3uf + I + G model 5. Probabilities of the nodes are represented in the circles.

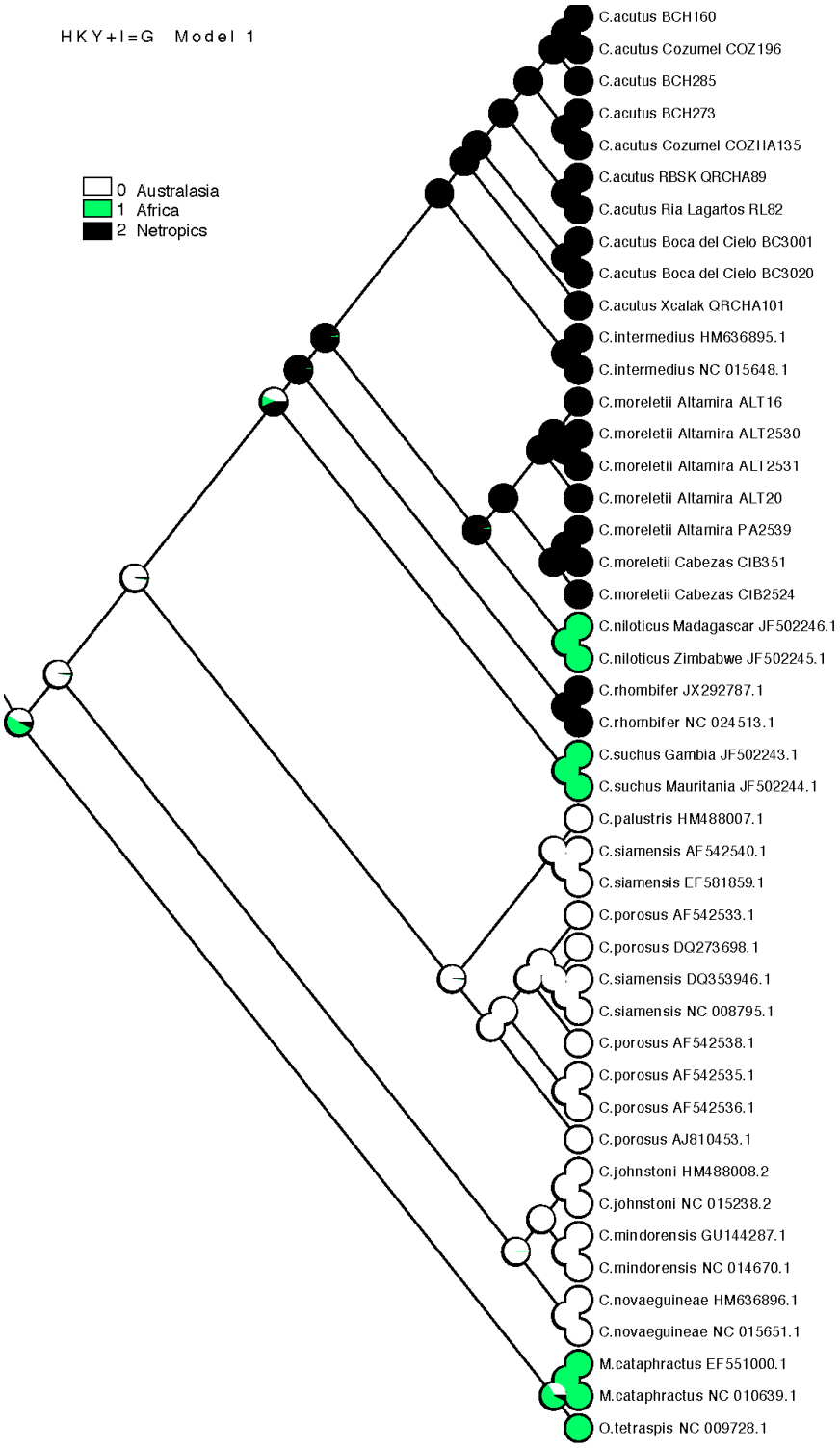

Figure A23. Mesquite Maximum likelihood tree for the HKY + I + G model 1. Probabilities of the nodes are represented in the circles.

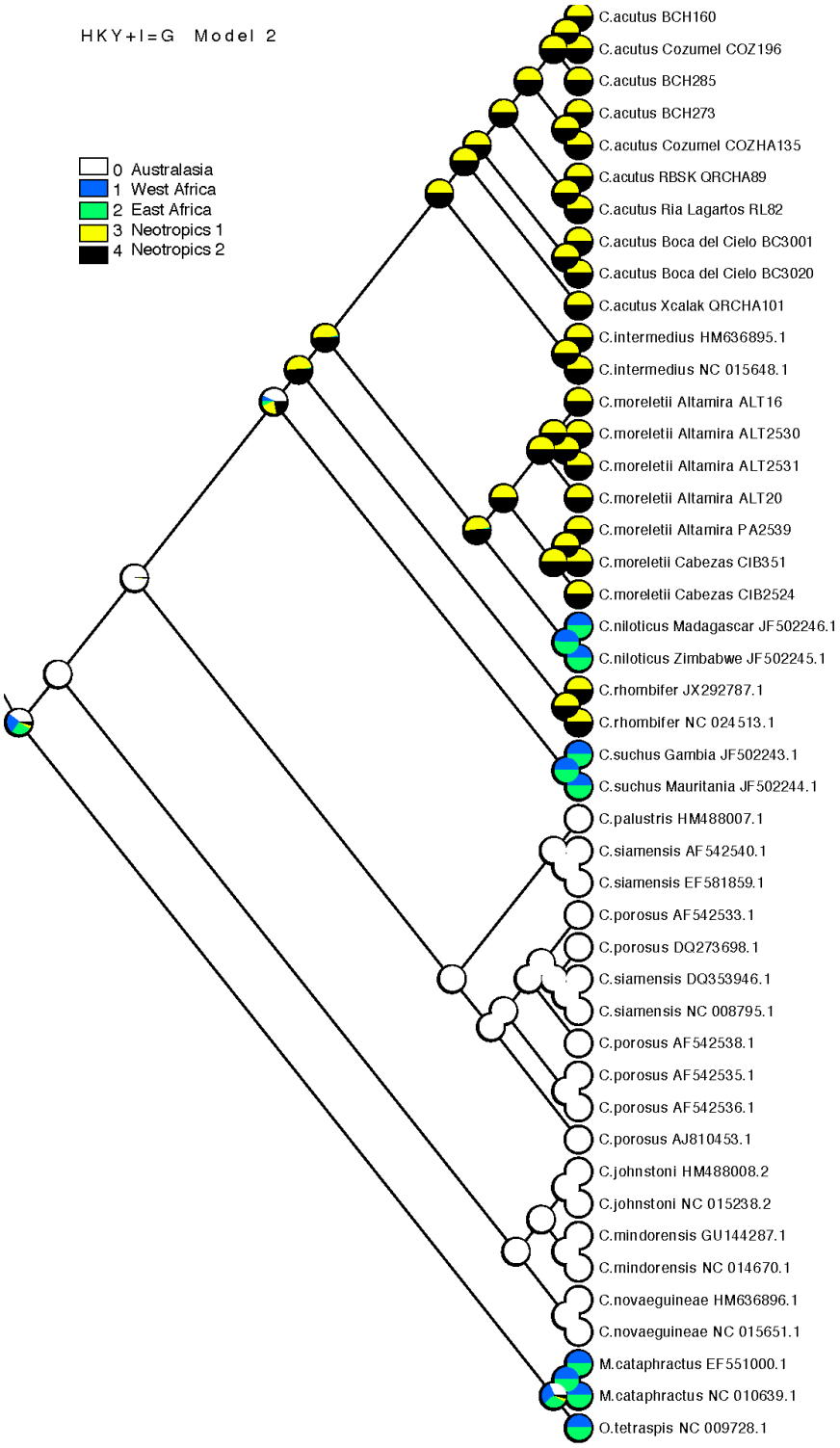

Figure A24. Mesquite Maximum likelihood tree for the HKY + I + G model 2. Probabilities of the nodes are represented in the circles.

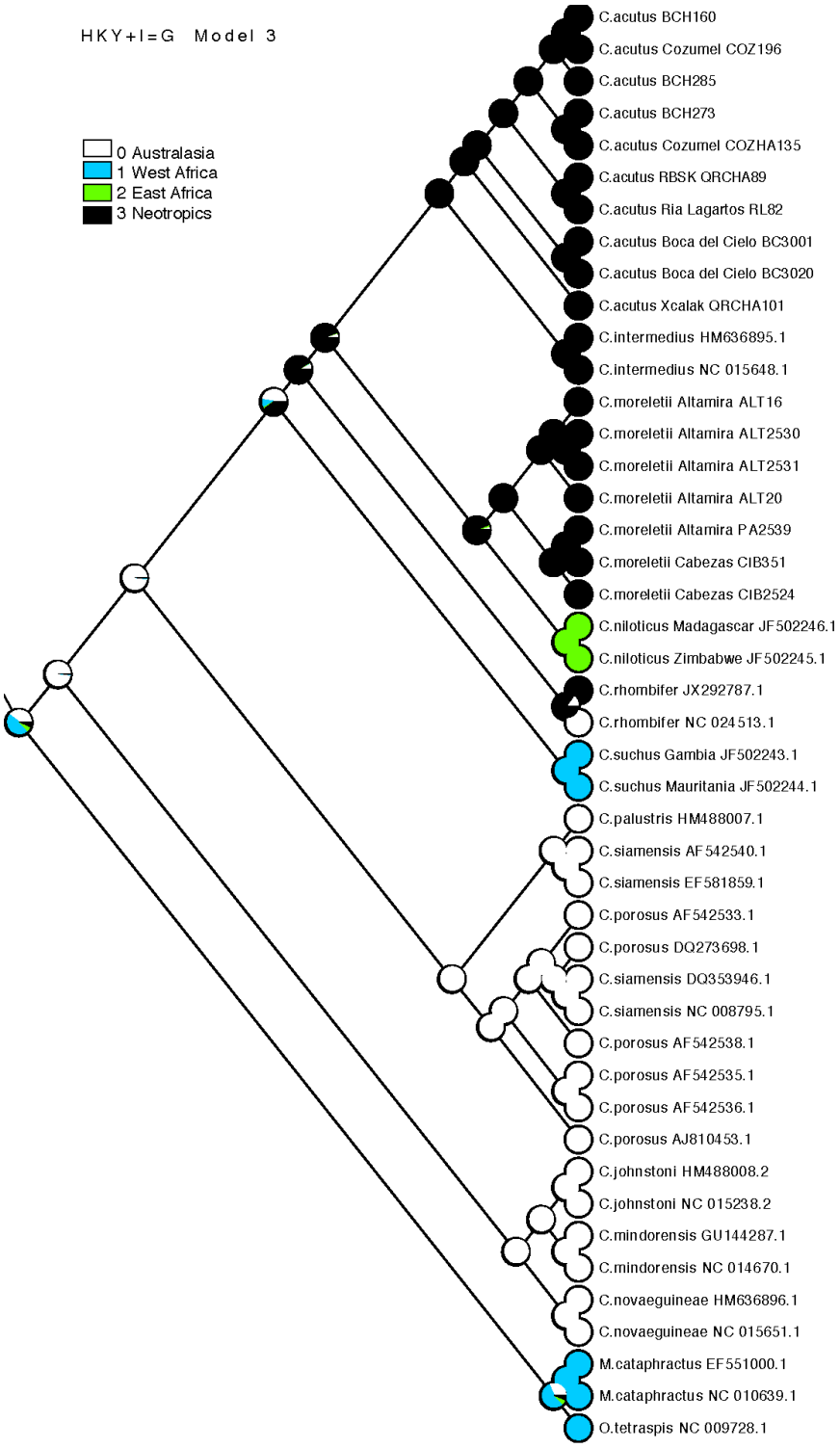

Figure A25. Mesquite Maximum likelihood tree for the HKY + I + G model 3. Probabilities of the nodes are represented in the circles.

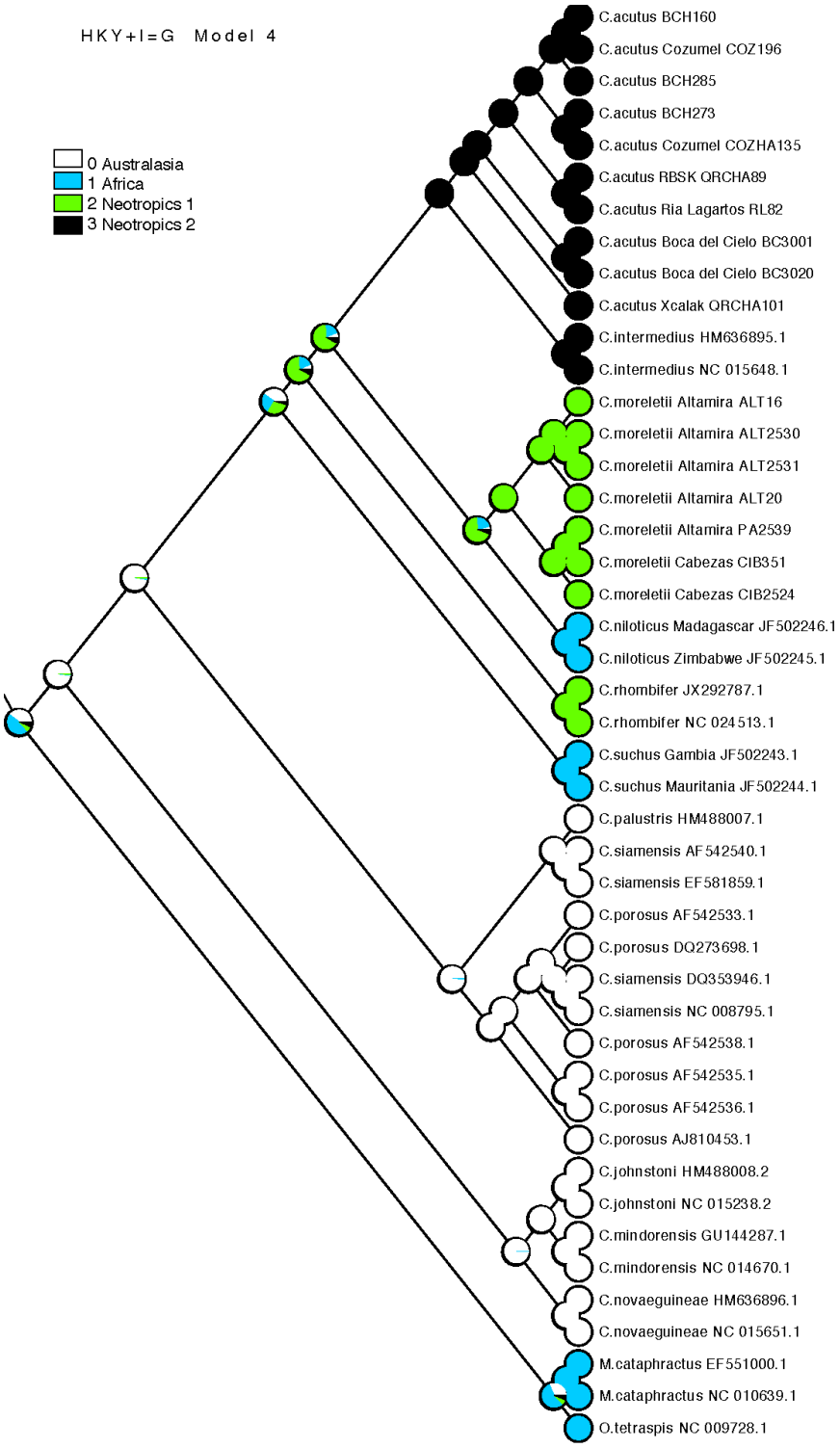

Figure A26. Mesquite Maximum likelihood tree for the HKY+ I + G model 4. Probabilities of the nodes are represented in the circles.

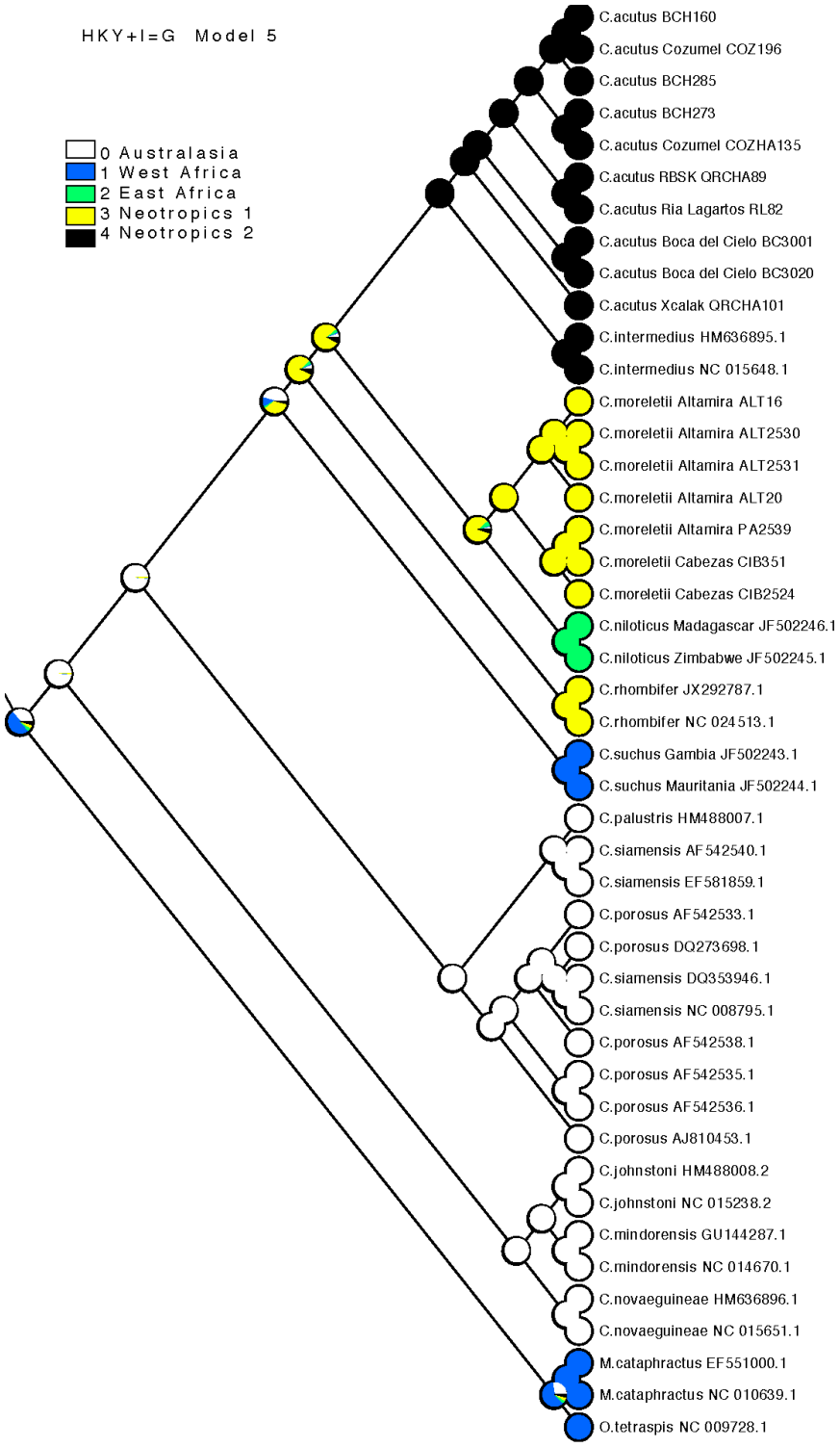

Figure A27. Mesquite Maximum likelihood tree for the HKY + I + G model 5. Probabilities of the nodes are represented in the circles.

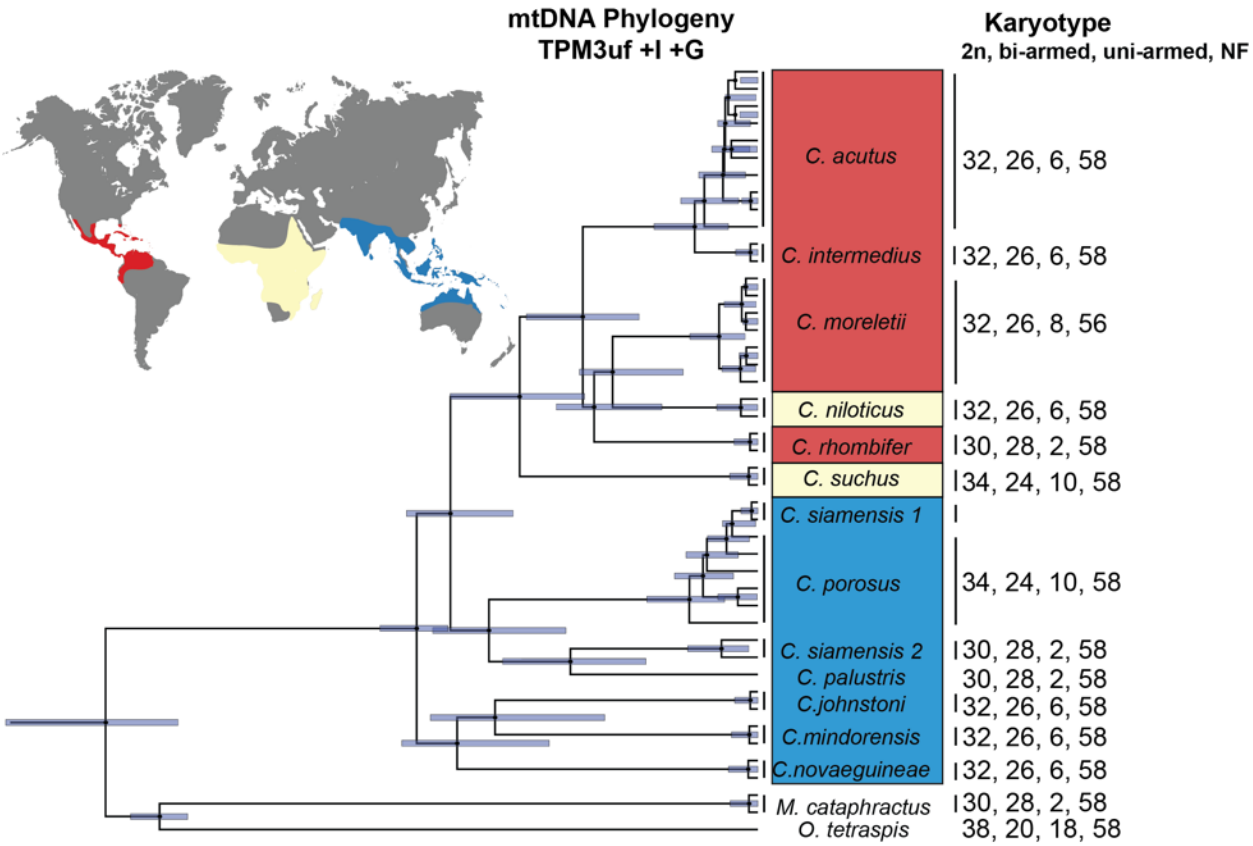

Figure A28. Mitochondrial DNA phylogeny for the TPM3uf + I + G model, with the information of the karyotypes for each species from Srikulnath et al (2015).

Appendix B: Chapter 2

Table B1. Fluorescent labeled primers for the microsatellites used from (Fitzsimmons et al., 2000) and multiplex reactions.

Primer name	Primer Sequence 5'-3'	DYE	Color	MultiplexReaction
Cj18	F: ATCCAAATCCCATGAACCTGAGAG R: CCGAGTGCTTACAAGAGGCTGG	VIC	Green	1
CUD68	F: GCTTCAGCAGGGGCTACC R: TGGGGAAACTGCACTTTAGG	PET	Red	
C391	F: ATGAGTCAGGTGGCAGGTTC R: CATAAATACACTTTTGAGCAGCAG	NED	Yellow	
CU5-123	F:GGGAAGATGACTGGAAT R:AAGTGATTAACTAAGCGAGAC	6-FAM	Blue	
Cj131	F: GTTTGTCTTCTTCCTCCTGTCCCTC R: AAATGCTGACTCCTACGGATGG	6-FAM	Blue	2
Cj128	F: ATTGGGGCAGATAAGTGGACTC R:GTTTCTGCTTCTCTTCCCTACCTGG	VIC	Green	
Cj127	F: CCCATAGTTTCCTGTTACCTG R: GTTTCCCTCTCTGACTTCAGTGTTG	VIC	Green	
Cj119	F: GTTTGCTGTGGAATGTTTCTAC R: CGCTATATGAAACGGTGGCTG	NED	Yellow	
Cj16	F: CATGCAGATTGTTATTCCTGATG R: TGTCATGGTGTCAATTAAACTC	NED	Yellow	3
CUC20	F: GATCTGCAGTGCAAGAAAG R: GGTTTAGCGGTCACAGTAAC	PET	Red	
CUJ131	F:GTCCCTTCCAGCCCAAATG R:CGTCTGGCCAGAAAACCTGT	VIC	Green	

Table B2. Allelic richness after rarefaction for the microsatellite markers.

Species	Population	Alleic richness of all Alleles		Allelic richness of Private Alleles	
		Average over Loci	Sdev	Average over loci	Sdev
C. moreletii	Altamira	3.860	1.241	0.640	0.547
	Cabezas	1.990	1.139	0.120	0.265
	Ria Lagartos	3.820	0.999	0.620	0.715
	Cozumel	3.790	1.364	0.560	0.880
C. acutus	Banco Chinchorro 1	3.010	1.361	0.360	0.522
	Banco Chinchorro 2	2.620	1.057	0.220	0.367
	RBSK	3.690	1.024	0.210	0.419
	Huach	4.990	0.936	0.540	0.322
	Xcalak	3.880	0.585	0.290	0.306
	Sumidero	2.570	0.949	0.210	0.332
	Boca del Cielo	3.030	1.123	0.180	0.215

Table B3. Restriction site Associated *DNA* nucleotide diversity(π), inbreeding coefficient (F_{IS}) and polymorphic sites per population.

Species	Population	π	StdErr	Fis	StdErr	Polymorphic Sites	% Polymorphic Loci
C. moreletii	Altamira	0.169	0.001	0.393	0.004	8694	0.337
	Panuco	0.091	0.001	0.010	0.003	4053	0.157
	Carpintero	0.097	0.002	0.027	0.006	4228	0.164
	Cabezas	0.309	0.002	0.499	0.006	9454	0.366
Hybrid zone	Ria lagartos	0.275	0.002	0.271	0.006	9317	0.361
Yucatan *C. acutus*	Cozumel	0.179	0.002	-0.008	0.013	7439	0.286
	Banco Chinchorro 1	0.152	0.002	0.001	0.006	5178	0.199
	Banco Chinchorro 2	0.048	0.001	0.033	0.005	2900	0.112
	Huach	0.267	0.002	0.443	0.007	9397	0.361
	Xcalak	0.127	0.002	0.022	0.009	5112	0.196
Central America *C. acutus*	Galeta	0.099	0.002	0.015	0.005	3496	0.135
	Barro Colorado Island	0.097	0.002	0.000	0.003	3400	0.132
	Lagartero	0.071	0.001	0.005	0.004	2474	0.096
	Coiba	0.073	0.001	0.005	0.006	2635	0.102
Pacific *C. acutus*	Sumidero	0.125	0.001	0.521	0.002	8312	0.322
	Boca del Cielo	0.057	0.001	-0.008	0.003	1843	0.071

Appendix C: Chapter 3

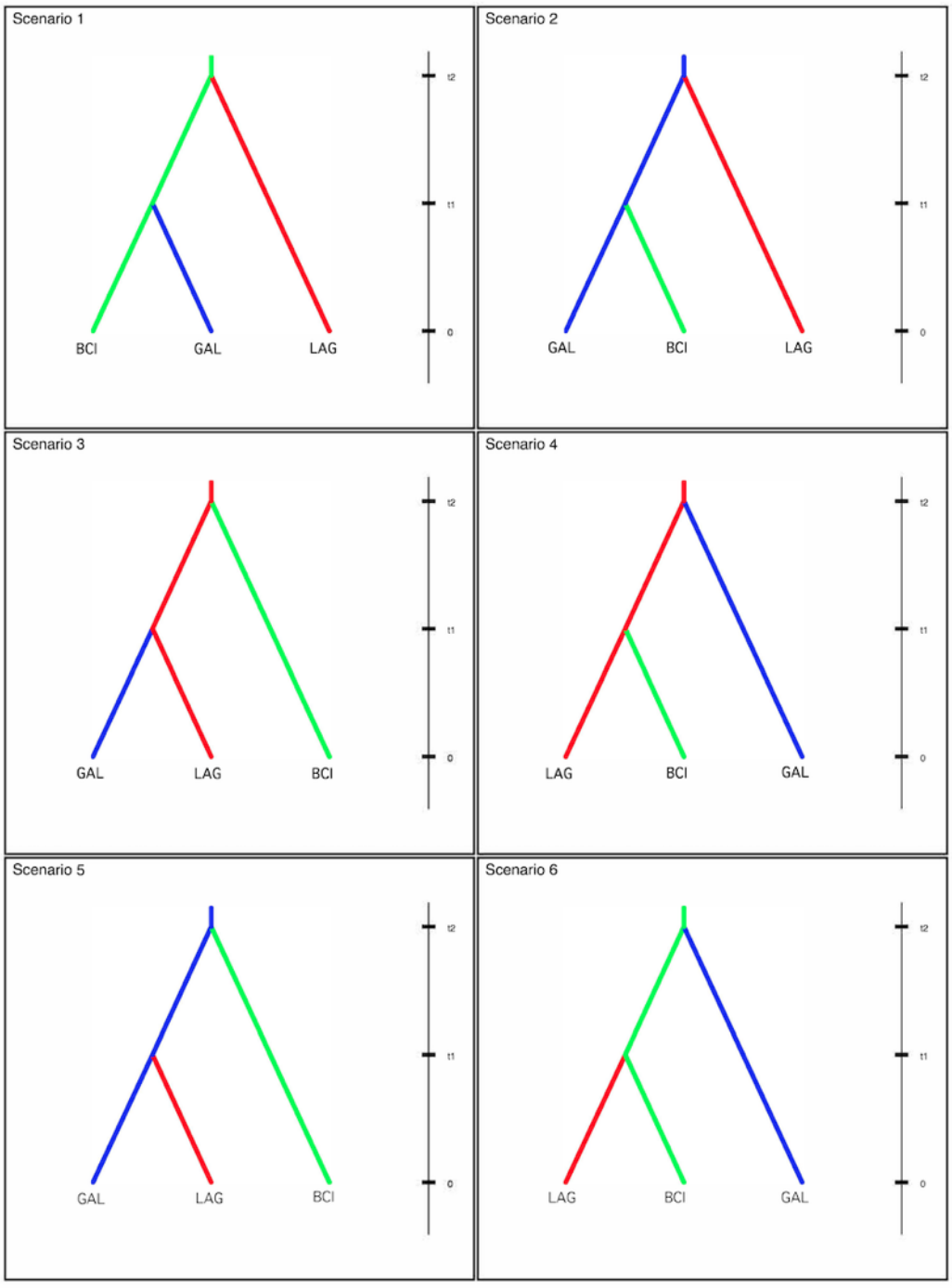

Figure C1. Historical model scenario used in DIYABC to infer population history.

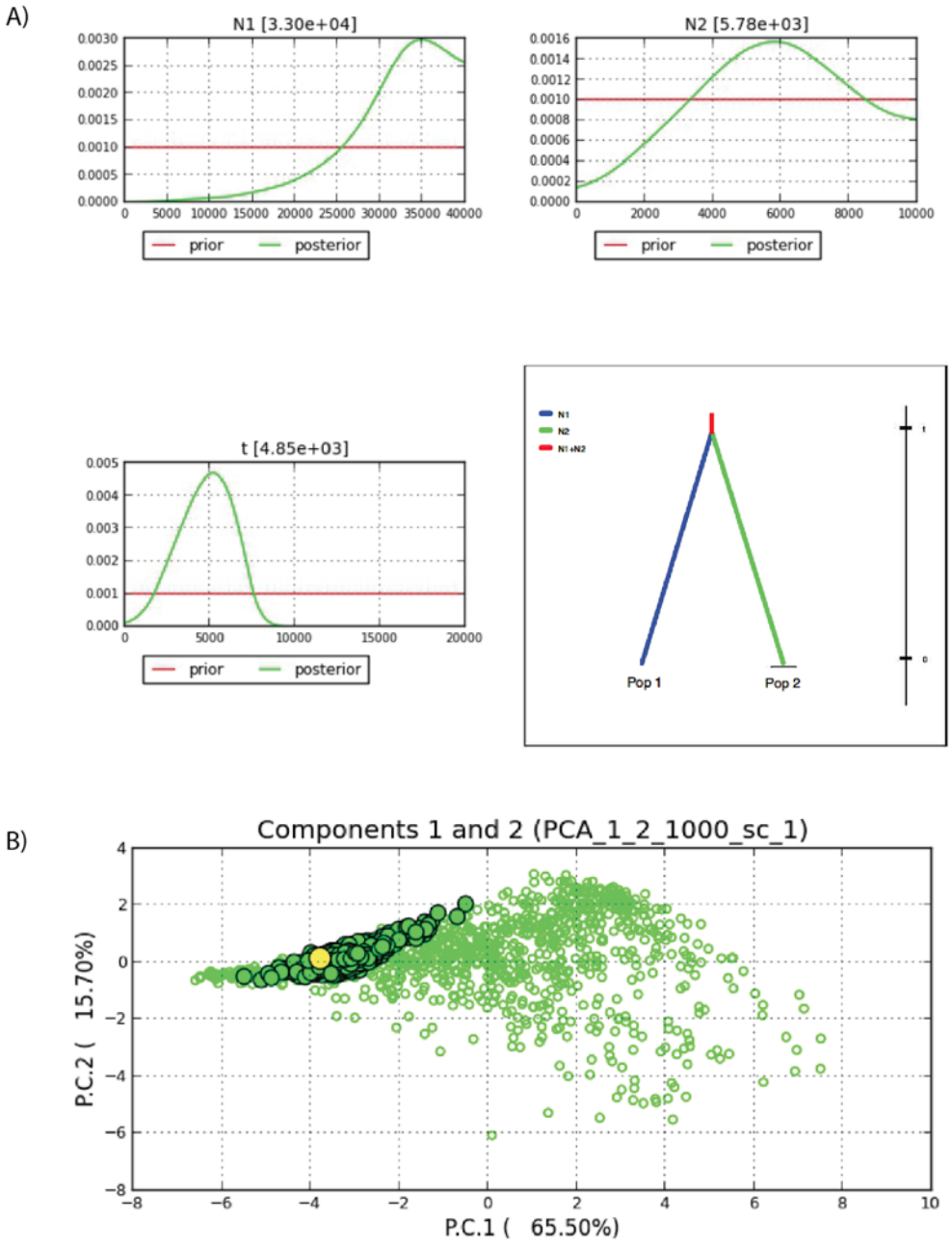

Figure C2. DIYABC results A) Results and scenario used for the split of the Caribbean and Pacific populations B) Plot with the prior and posterior distributions for the split of Pacific and Caribbean. The dense presence of dots around the data set (yellow) indicate a good fit of the scenario.

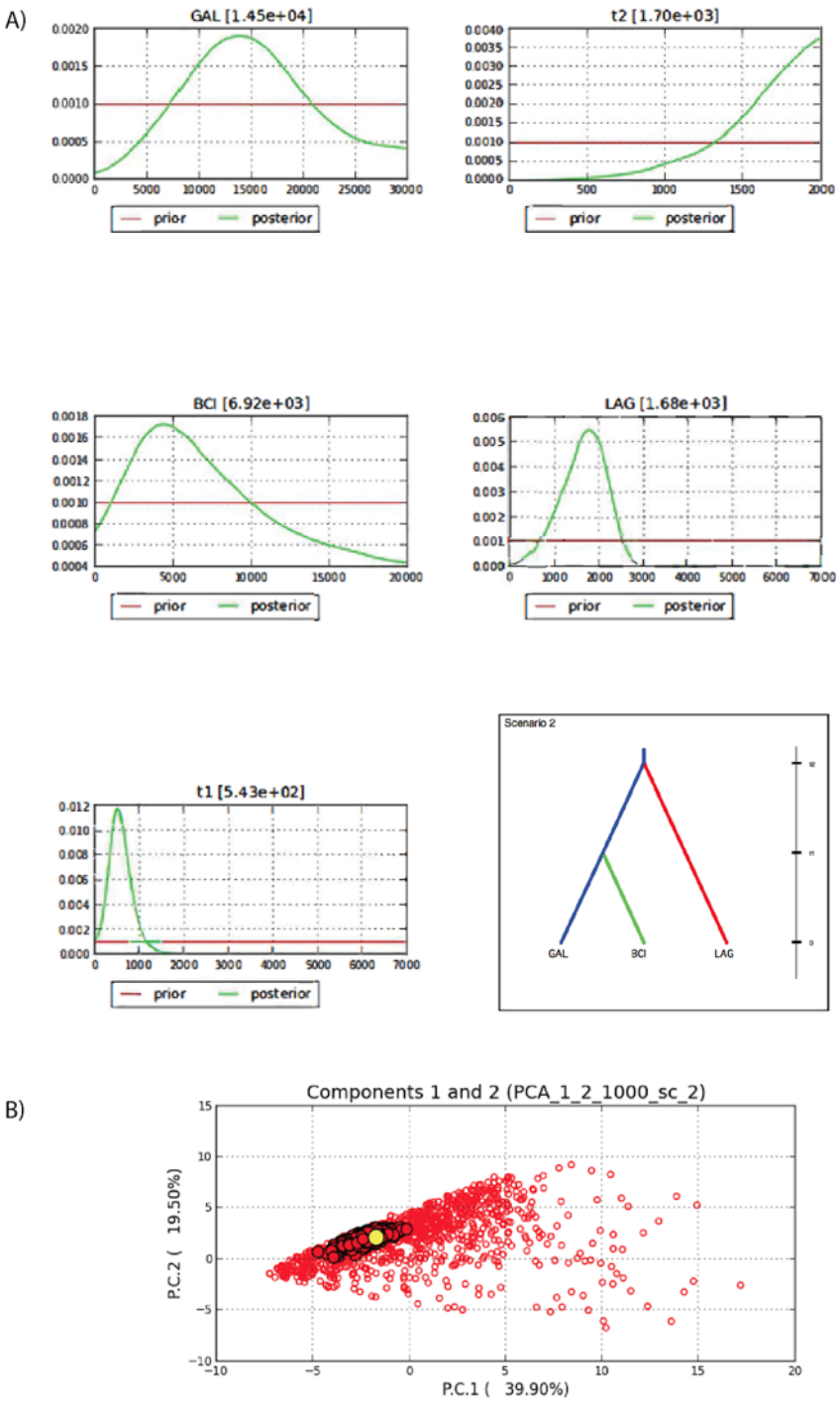

Figure C3. DIYABC results for A) The best scenario of Caribbean, Pacific and Panama Canal split. B) Plot with the prior and posterior distributions. The dense presence of dots around the data set (yellow) indicate a good fit of the scenario.

Appendix D: External examiner Dr. Adalgisa Caccone questions.

Chapter 1

1- Captive vs wild caught. I consider that using samples from wild populations is more informative and better suited to the purpose of the chapter. The majority of the samples come from wild populations except for the ones that no other information was available. However the main purpose of the chapter was to reassess the phylogeographic history from previous authors. The difference is that we combine the genetic data with the fossil record, paleo climatic and paleo bathymetric and paleo ocean currents to reconstruct the possible scenarios of dispersal of Crocodylus from Africa to America and back. If we consider how relevant is in terms of the broad geographic context, we do not change the biogeographic history, but we are seeing more dispersal events, not reflected in previous multi-gene phylogenies. This multiple dispersals have been suggested based on fossil record, morphology and paleogeography.

2- Using mtDNA markers in a group known to hybridize. One of the main reasons to use a single marker was the lack of funds to sequence more genes. But now this is something that can be expanded for the publication. I decided to use the mitochondrial because I am trying to track the lineages in time and space, not trying to find the hybrids. However, the presence of undetected hybrids in the reconstruction of the phylogeny could result in a topology not reflecting the complete history of the Crocodylus. But this could be the case for previous studies too.

3- Comparison with previous studies that includes mtDNA and nuclear markers. A limiting factor of the study is the use of a single marker for the phylogenetic reconstruction, compared to what was done by previous author that used more genes from the mitochondria and nuclear genes. The studies are not comparable in the genetic level. But our study integrates variables that put in context the nature of the dispersal of Crocodylus around the globe.

4- Choice of D-loop: I chose the Dloop because it is a highly variable region, compared to slower evolving mitochondrial genes. And the purpose was to track recent changes for the reconstruction.

A general conclusion of this chapter is that needs more support from many genes, we are working on using next generation sequencing on samples of all extant species to give a stronger support to our results for publication.

Chapter 2

1- Unbalance of sampling. I consider that having the same number of individuals per species could be informative. But I think the Bias here would be having uneven numbers for each population. A way to correct it would be having the same number of individuals per population from a randomization.

2- DAPC vs PCA. The PCA is used to simplify the complexity of data reducing it to a few dimensions called principal components The DAPC is a two-step method, first a PCA is performed to transform the data and then the genetic clusters are identified by the use of a Discriminant Analysis. I used both approaches for the data and the results were consistent. However, I consider that the DAPC in fact detects a higher number of genetic clusters than the PCA and might overestimate the number of populations.

3- A priori BCH two clusters. For the RadSeq analysis I first did an approach considering it as a single population. The result was that the F_{ST} range for the population was considerably wider than for the other populations. I also did the Clustering methods (Structure, DAPC, PCA) assuming one population and the results were two genetic clusters in Banco Chinchorro. The differences that the microsatellites and the RadSeq can show is first caused by the number of markers. For microsatellites is can be in the hundreds compared to the RadSeq that its usually an order of magnitude higher (thousands). This increase in the markers also increases the resolution. The microsatellites could also not be polymorphic or fixed for some populations that have gone through bottlenecks or have limited genetic diversity.

Chapter 3

1- Run of ABC analysis with bottleneck events. The use of ABC for the thesis only considered a simple scenario of divergence between Pacific and Atlantic populations. A way to make the model more realistic would be incorporating bottlenecks, migration events and admixture. I will run these models for the publication.

2- Genetic Isolation of Coiba island. The reason why Coiba island has a unique signature might be also related to the Last Glacial Maximum. With a decrease in sea level of 150 meters Coiba would not be an island but part of the continent, and the populations already present in the island while these events happened would stay isolated by the lack of water bodies to facilitate dispersal. There are two possible reasons to explain that crocodiles are migrating in only one direction. The first one is the ocean currents flow from the island to the coast facilitating the dispersal of crocodiles in that direction. And a second one could be a sample bias, I do not have as many samples from the coast as I have from the island.

3- Migrants and Introgression history. The pure individuals would be the ones from the parental species (*C. acutus* and *C. moreletii*). The F1 Hybrids could be of two types:

   a. Maternal *C. acutus* X Paternal *C. moreletii*
   b. Maternal C. moreletii X Paternal *C. acutus*

   The hybrids we detected in Banco Chinchorro based on the Maternal mtDNA haplotype and the RadSeq seem to have Maternal *C. acutus* X Paternal *C. moreletii*.

   The backcross will be a Hybrid F1 with the parental species, detecting in which category the sampled individuals of the populations falls is fundamental to understand the history and demography of the events that led to the formation of the populations. An approach could be using the R package INTROGRESS, with

this would be easier to polarize and create a hybrid index. Then assign the hybrids to hybrid classes to understand the direction of the introgression.