Gustavo A. Ballen

Fossil freshwater fishes and the biogeography of northern South America

Peixes fósseis de água doce e biogeografia do norte da América do Sul

Original version

Thesis submitted to the Graduate Program of the Museu de Zoologia da Universidade de São Paulo in partial fulfillment of the requirements for the degree of Doctor of Science (Systematics, Animal Taxonomy and Biodiversity).

Advisor: Prof. Dr. Mário Cesar Cardoso de Pinna.

São Paulo
2020

Serviço de Bibloteca e Documentação
Museu de Zoologia da Universidade de São Paulo

Cataloging in Publication

Ballen, Gustavo A,
Fossil freshwater fishes and the biogeography of Northern South America.= Peixes fosséis de água doce biogeografia do Norte da América do Sul/ Gustavo A. Ballen; orientador Mário Cesar Cardoso de Pinna. São Paulo, 2020.
254p.

Tese Doutorado – Programa de Pós-Graduação em Sistemática, Taxonomia e Biodiversidade, Museu de Zoologia, Universidade de São Paulo, 2020.
Versão original
1. Estatistica. 2. Andes. 3. Siluriformes. 4. Anatomia. 5. Paleoictiologia. I. Pinna, Mario Cesar Cardoso de, orient. II.Título.

CDU 567

BALLEN, Gustavo A.

Fossil freshwater fishes and the biogeography of northern South America

Peixes fósseis de água doce e biogeografia do norte da América do Sul

Thesis submitted to the Graduate Program of the Museu de Zoologia da Universidade de São Paulo in partial fulfillment of the requirements for the degree of Doctor of Science (Systematics, Animal Taxonomy and Biodiversity).

Approved: _____/_____/_____

Thesis Committee:

Prof. Dr. _____ Institution: _____

Decision: _____ Signature: _____

Prof. Dr. _____ Institution: _____

Decision: _____ Signature: _____

Prof. Dr. _____ Institution: _____

Decision: _____ Signature: _____

Prof. Dr. _____ Institution: _____

Decision: _____ Signature: _____

Prof. Dr. _____ Institution: _____

Decision: _____ Signature: _____

"The... analysis demonstrates that the fish fauna of the Magdalena basin was derived in small part from the ocean and in larger part from Central America. It demonstrates beyond a peradventure that most of it had an origin in common with that of the Orinoco basin to the east of it, and that the fauna of the Magdalena was segregated from the general fauna of the Orinoco by the formation of the Cordillera de Bogotá between the two, at a time antedating the development of most of the present species. It also demonstrates that if the above conclusions are valid some species found on both sides antedate the formation of the Cordillera de Bogotá"

Carl H. Eigenmann (1920)

Acknowledgments

This thesis was supported through a doctoral scholarship (grant 2014/11558-5), and a BEPE international internship (grant 2016/02253-1), São Paulo Research Foundation (FAPESP); the opinions, hypotheses, conclusions, or statements expressed in this volume are direct responsibility of the author and do not necessarily reflect the views of or endorsement by FAPESP. Preliminary phases of this project were funded by the Smithsonian Tropical Research Institute, National Geographic, ARES Corporation, Universidad del Norte, Instituto de Ciencias Naturales-Universidad Nacional de Colombia, and Pontificia Universidad Javeriana. Workspace and facilities were provided mainly by the Museu de Zoologia da Universidade de São Paulo (Brazil), and by Pontificia Universidad Javeriana (Colombia), Universidad Nacional de Colombia (Colombia), Servicio Geológico Colombiano (Colombia), Universidad del Norte (Colombia), and Smithsonian Tropical Research Institute (Panamá). I thank the following colleagues for the kind support during visits to the collections under their care, both during my years as doctoral student, and before, when a generous amount of data were obtained: John G. Lundberg and Mark Sabaj (ANSP), Jonathan I. Bloch, Bruce MacFadden, Aldo Rincón, and Richard Hulbert (FLMNH), Saúl Prada-Pedreros, Javier A. Maldonado-Ocampo (*requiescat in pace*), Alexander Urbano-Bonilla (MPUJ), Germán Galvis, John D. Lynch, Ofelia Mejía, Edgar Larrarte, Francisco Luque (ICNMHN), Richard P. Vari (*requiescat in pace*) (NMNH), Aléssio Datovo, Osvaldo T. Oyakawa, Michel Gianeti (MZUSP), Jaime Escobar (MUN), and Herculano Alvarenga (MHNT-UT). I thank Jon Ambruster for allowing the use of his picture of *Oxydoras*the ANSP/G.W. Saul for the photograph of *Zungaro*, and Mark Sabaj for the photographs of *Hemidoras* and *Brachyplatystoma*, all the authors that provided open access to their pictures in licenses of reuse with attribution (Figure 4.18).

Special thanks go to the staff of the Museu de Zoologia da Universidade de São Paulo, my home institution during the three most academically-intense years of my life so far, whose kind support, help, and every warm word made a lot easier to adapt to a new world when moving from Bogotá to São Paulo; I want to thank specially Marta Maria Grobel and Omair Tizzot from the SPG for all their support and diligence during the processing of my academic and formal requests, and Dione Serripieri and Viviane Neves for their help with library services and personal kindness. I received enormous support and the most professional teaching from my professors at the Museu de Zoologia, Instituto de Matematica e Estatística, Instituto de Geociências and Instituto de Biociências of the Universidade de São Paulo: Mário

de Pinna, Aléssio Datovo, Mário de Vivo, Isabel Landim, Julio Singer, Taran Grant, Miguel Trefaut, Marcelo Pompeu, Paulo Prado, Luis Schiesari, Renata Pardini, Veridiana Martins, Mônica Toledo Piza Ragazzo, Ralf Britz, Fernando Dagosta, Bruno Melo, Luz Eneida Ochoa, and Fabio Roxo. I thank my qualification committee for their criticism and support during my first year of doctoral studies, Drs. Tiago Quental, Naércio Menezes, and Gustavo Bravo.

Thanks to my fellows at the Laboratorio de Ictiología at the Museu de Zoologia, Vitor Abrahão, Luiz Peixoto, Tulio Texeira, Priscilla Camelier, Manoela Marinho, Marina Loeb, Henrique Varella, Ilana FIchberg, Veronica Slobodian, Murilo Pastana, Willian Ohara, Fernando Dagosta, Arthur Lima, Paulo Presti,Guilherme Dutra, Vinicius Reis, Luz Ochoa, Vinicius Espindola, and Victor Tagliacollo, for all the company, support, laughs and *churrascos*. Thanks to Prof. Mônica Toledo Piza Ragazzo and her students, Victor Giovanetti, Kleber Mathubara, João Capretz, and Karla Soares for giving me a second how in the ichthyology lab at the IB-USP.

Thanks to the Laboratorio de Ictiología at the Pontificia Universidad Javeriana where I carried out an internship for studying the Colombian fossil specimens. My immense gratitude to Alex Urbano-Bonilla (*el propio tigrillo del Casanare*), Jhon Zamudio, Alejandro Mendez, Miguel Murcia, Guido Herrera, Edgar Collazos, Lina Camacho, Kelly Triana, and specially to Javier A. Maldonado-Ocampo (*requiescat in pace*) and Saul Prada-Pedreros. We lost Javier almost a year ago and with him, a friend, colleague, and an advisor. His place will hardly be filled by anyone as in life he managed to unite most Colombian ichthyologists for collaboration and made things that no one thought possible. Our ichthyological community hardly will see someone like him. Thanks, Javier, for supporting me from the beginning.

Although I left the amphibians lab at the ICN long ago, I want to thank my fellow students there, which became friends for a lifetime. Sandy Arroyo, Rancés Caicedo, David Sánchez, Lina Escobar, and Herón Romero, *gracias mis estimados sacajopo diversos*.

Thanks to Prof. Tiago Quental and his students, specially Gustavo Burin and Laura Alencar for the discussions, kindness, and for sharing their wisdom on the vast fields of macroecology and macroevolution. Hope to have more chances to collaborate in the future.

I thank the Brazilian Lute Community LuteBR for providing a place where my peculiar musical interests could fit perfectly well, always supportive and open to provide their kindness in academic and non-academic affairs, in sum, to let me know that I belong to that great community despite my professional background as biologist. I thank Ernesto Ett for allowing me to access the renaissance lute, a dream since my high school years; Anderson de Lima, for guiding me through the fascinating world of the early plucked strings and everything related to renaissance music from performance to theory and research; to Bruno Inácio and Alexandre Ribeiro for the beers and lively discussions on music theory and their support to me as an outsider to the professional musical community.

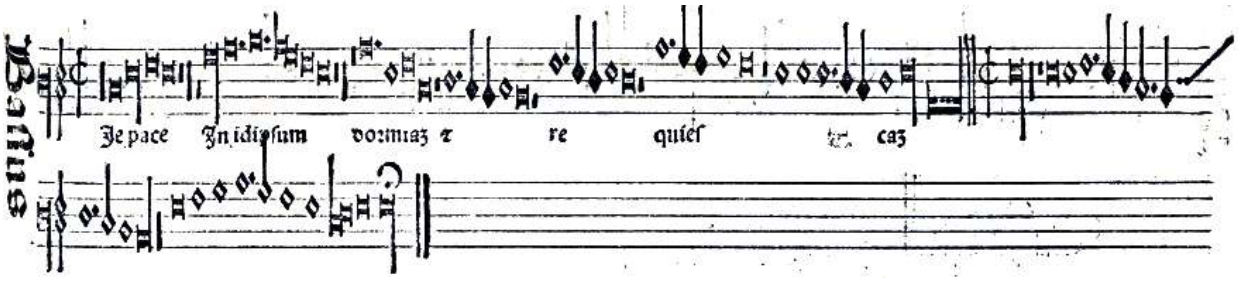
This project would not be possible without two important communities: The people who support open source technologies, and StackOverflow.com. The effect of both communities permeates even the simplest figure of this thesis. All the products of this project were

generated using open source technologies, from the very text in LaTeX to the plates edited in Inkscape and GIMP. I hope that one day, scientific institutions in Latin America will spend less money in licenses of mediocre software or pirating programs and more in the development of basic computational skills for their students such as programming, database management, GIS and graphical edition. When I started my BSc., learning to code (e.g., in R) was a plus that only tech-savvy people reached; now it is a skill as important as learning basic math. We can not stand back in this technological revolution in biological sciences.

Thanks to my advisor Mario de Pinna who believed in my ideas despite the strong shift in interests from systematics of Andean Loricariids (my former doctoral project back in 2010) to paleoichthyology and biogeography. Although not necessarily sharing my enthusiasm for numbers, he was incredibly supportive and granted me with unlimited scientific freedom. I enjoyed his friendship and shared my passion for music with him, a talented flute, violin, and piano player. Although I'm more into Ockeghem and him more into Ligeti, we always meet each other at Bach, the greatest mind in history of humanity. I am sure that with the end of this relationship between master and pupil, what is left is the friendship that I am sure, will remain strong forever. *Merci infini.*

Two important persons remain to be acknowledged. I wish to thank my wife, lover, colleague, and friend, Sandra Reinales, for all her support, patience, enthusiasm, but specially her love. Thanks for walking this way by my side, for giving me a wonderful daughter, for sharing your life with me, and for giving me the privilege of learning from you and to enjoy all your qualities. I am lucky to have married the scientist that I admire the most. Thanks to my daughter Sarita for all her patience and support, it is incredible to see you growing a better person than your parents, more intelligent, more focused, with more skills, still with a huge heart. *Las amo infinitamente.*

The very beginnings of this project started long ago. It took me nine years to finish my undergrad studies, and interestingly another (not subsequent) nine years to finish this project. It taught me a lot about people, research, nature, and the limits of our strength. As the end of this process approaches the only thing I can say to this project and myself is, as Josquin Desprez set beautifully in the tenor of the motet *Que vous madame/In pace in idipsum* ca. 1450: "*In pace in idipsum dormiam et requiescam, si dedero somnium oculis meis.*"



Resumo

Os vertebrados fósseis têm sido utilizados na literatura como fonte de informação sobre cenários paleogeográficos passados no neotrópico, embora os peixes em especial tenham sido menos estudados, apesar da disponibilidade em coleções e no campo. Restos fósseis apendiculares são uma das ocorrências mais comuns de bagres fósseis e tendem a ser abundantes localmente em faunas conhecidas de idade Cenozoico na América do Sul; no entanto, a anatomia comparada desses complexos anatômicos tem sido pouco estudada, com uma terminologia anatômica complexa e confusa, em que cada referência publicada ignora os termos anteriores propostos e, assim, cria vários sinônimos diferentes para a mesma característica osteológica. Foi realizada uma revisão completa da literatura, juntamente com o revisão de viventes da ordem Siluriformes, a fim de fornecer uma terminologia anatômica padronizada para os espinhos das nadadeiras dorsal e peitoral. Três faunas continentais de vertebrados foram encontradas no norte da Colômbia, numa área atualmente separada das drenagens da Amazônia-Orinoco pelos Andes da Cordillera Oriental na Colômbia e nos Andes de Merida na Venezuela. Os peixes fósseis de água doce da formação do meio do mioceno Castilletes são aqui descritos, juntamente com sua relevância em diferentes tópicos, desde taxonomia e anatomia comparada até reconstruções paleogeográficas e paleoambientais; essa assambleia fóssil está em conformidade com os modelos anteriores de conexões de paleodrenagens entre áreas cis- e trans-andinas na literatura. Duas faunas da idade do Plioceno das formações Sincelejo e Ware foram estudadas usando uma abordagem semelhante à da assembléia fóssil da formação Castilletes. As assembléias da idade do Plioceno implicam a presença de uma conexão hídrica entre as drenagens agora separadas pelos Andes, ou seja, elas estendem mais próximo do presente o mesmo padrão recuperado na assembléia de idade Mioceno médio e são contrarias aos modelos tectônicos clássicos dos Andes do norte que sugerem uma perda na conectividade de drenagem de 11 até 13 Ma. Um conjunto de métodos quantitativos para inferir o tempo de separação entre duas áreas biogeográficas usando dados da estimativa do tempo de divergência é aqui proposto e discutido como métodos promissores para estimativa estatística em biogeografia. Embora tenham sido projetados principalmente com padrões vicariantes em mente, esses métodos são extensíveis o suficiente para serem aplicáveis a qualquer tipo de evento que ocorra no tempo geológico deixando vestígios nos estudos de estimativa do tempo de divergência. Os diferentes métodos geralmente inferem um intervalo de separação de 2 a 5.8 Ma, consistente com as informações das assembléias fósseis que implicam uma conexão persistente para aproximadamente o mesmo intervalo; portanto, as faunas

fósseis das formações Sincelejo e Ware seriam a última evidência conexões de drenagem nos Andes. Dentro dessa estrutura temporal, vários padrões de biodiversidade, como composição da fauna, taxas de endemismo e padrões espaciais, e o momento da geração da biodiversidade em escala regional para continental devem ser reavaliados à luz dos resultados aqui fornecidos.

Palavras-chave: Estatistica, Andes, Siluriformes, Anatomia, Paleoictiologia.

Abstract

Fossil vertebrates have been used in the literature as a source of information on past paleogeographic settings in the Neotropics, although fishes in special have been less studied despite availability in collections and in the field. Appendicular fossil remains are one of the most common occurrences of fossil catfishes and tend to be locally abundant in well-known faunas of Cenozoic age in South America; however, the comparative anatomy of these anatomical complexes has been poorly studied, with a complex and confusing anatomical terminology where each published reference ignores previous terms provided and thus creates a number of different synonyms for the same osteological feature. It was carried out a thorough literature review along with direct examination of extant representatives of the order Siluriformes in order to provide a standardized anatomical terminology for the dorsal- and pectoral-fin spines in the order. Three continental vertebrate faunas have been found in northern Colombia, in an area that is currently separated from the Amazon-Orinoco drainages by the Andes of the Cordillera Oriental in Colombia and the Merida Andes in Venezuela. Fossil freshwater fishes from the middle Miocene Castilletes formation are herein described along with their bearing on different topics ranging from taxonomy and comparative anatomy to paleogeographic and paleoenvironmental reconstructions; this fossil assemblage conforms to previous models of paleodrainage connections between cis- and trans-Andean areas in the literature. Two faunas of Pliocene age from the Sincelejo and Ware formations were studied using a similar approach to that of the fossil assemblage of the Castilletes formation. The assemblages of Pliocene age imply the persistence of a hydric connection between drainages now separated by the Andes, that is, they extend towards the present the same pattern recovered in the assemblage of middle Miocene age and are against the classical tectonic models of the northern Andes that suggest a loss in drainage connectivity about 11–13 Ma. A set of quantitative methods for inferring the time of separation between two biogeographic areas using data from divergence time estimation are herein proposed and discussed as promising methods for statistical estimation in biogeography. Although primarily designed with vicariant patterns in mind, these methods are extensible enough as to be applicable to any kind of event occurring in geologic time that leaves traces in divergence time estimation studies. The different methods mostly suggest a separation interval of 2–5.8 Ma, consistent with the information from the fossil assemblages that imply a connection persisting to about the same interval, therefore, the fossil faunas of the Sincelejo and Ware formations would be the last evidence of drainage connections across the Andes. Within this temporal framework,

a number of biodiversity patterns such as faunal composition, endemism rates and spatial patterns, and the timing of generation of biodiversity at regional to continental scale should be reassessed in light of the results herein provided.
Keywords: Statistics, Andes, Siluriformes, Anatomy, Paleoichthyology.

Contents

Abbreviatures

boc	Basioccipital
CDF	Cumulative density function
CI	Confidence interval
CS	Cleared and stained specimen
Daas	Dorsal spine, Anterior articular surface
Daf	Dorsal spine, Anterior fossae
Dao, Dpo	Dorsal spine, Anterior and posterior ornaments
Dasli	Dorsal spine, Attachment surfaces for the l. interspinalia
Df	Dorsal spine, Foramen
Diar	Dorsal spine, Inflection point of the anterior longitudinal ridge
Dlas	Dorsal spine, Lateral articular surfaces
Dlc	Dorsal spine, Lateral condyles
Dlf	Dorsal spine, Lateral fossae
Dlh, Drh	Dorsal spine, Left and right hemitrichia
Dlo	Dorsal spine, Lateral ornaments
Dpf	Dorsal spine, Posterior fossae
Dpp	Dorsal spine, Posterior processes
Dps	Dorsal spine, Posterior sulcus
DS	Dry skeleton
epo	Epioccipital
ex	Extrascapular
exo	Exoccipital
fr	Frontal
HL	Head length
HPD	Interval of highest posterior density
l.	*ligamentum, ligamenti*, ligament, ligaments
m.	*musculus, musculi*, muscle, muscles
Ma	*Mega anni*, Millon years ago
masl	Meters above sea level
Pae	Pectoral spine, Anteroventral emargination
Paf	Pectoral spine, Anterior fossa

Continued on next page

Continued from previous page

Pafap	Pectoral spine, Articular facet of the anterior process
Pafdp	Pectoral spine, Articular facet of dorsal process
Pafpp	Pectoral spine, Articular facet of the posterior process
Pafs	Pectoral spine, Articular facet for the scapulocoracoid
Pafvp	Pectoral spine, Articular facet of the ventral process
Pap	Pectoral spine, Anterior process
Pbp	Pectoral spine, Basal process
Pbr	Pectoral spine, Basal recess
Pdh	Pectoral spine, Dorsal hemitrichium
PDF	Probability density function
Pdp	Pectoral spine, Dorsal process
Pfap	Pectoral spine, Flange of the anterior process
Pmp	Pectoral spine, Median process
Ppk	Pectoral spine, Posterior keel
Ppp	Pectoral spine, Posterior process
Pps	Pectoral spine, Posterior sulcus
pro	Prootic
pt	Pterotic
pts	Pterosphenoid
ptsc	Postemporo-supracleithrum
Ptvp	Pectoral spine, Trochanter of the ventral process
Pvf	Pectoral spine, Ventral fossa
Pvh	Pectoral spine, Ventral hemitrichium
Pvp	Pectoral spine, Ventral process
sco	Supraoccipital
SL	Standard length
socp	Supraoccipital process
sph	Sphenotic

List of Figures

List of Tables

Chapter 1

Praeambulus

Abstract

The main goal of the present project is to explore the biogeography of Northern South America (NSA) from a paleontological perspective and the interplay of the fossil record and the interrelationships among extant taxa as a way to infer past drainage connections and their relevance for explaining the faunal relationships among drainages in NSA. This portion of the continent present several important drainages, all of them related historically to the core Amazon/Orinoco/Guyanas, but with highly endemic and poor faunas. Such high endemism and small richness seems to be the product of tectonic events that shaped both the geography and the drainages of NSA, therefore triggering speciation and extinction events in this part of the continent. The fossil record is going to be used in the present project in order to address the question of drainage connections among cis- and trans-Andean NSA during the early Pliocene, i.e., since the last 2 Ma. This work is unique in that the fossil collections herein available are largely unstudied and are crucial for understanding the tectonic and faunal evolution of NSA during the early Pliocene. In the same way, it completes the temporal record in NSA from the middle Miocene to the Pliocene, being complementary with earlier works with important fossil faunas such as La Venta in Colombia and Urumaco in Venezuela. It is expected that this work will provide a new paleogeographic framework, as well as further information to be used in divergence time studies, that coupled with geologic information are the base for understanding the evolution of freshwater faunas in the presence of complex geological processes.

Introduction

The Neotropical freshwater fishes are by far one of the most diverse vertebrate groups in the New World (Nelson, 2006). It is not only species-rich, but also presents high endemism levels at several taxonomic and geographic scales. The Neotropics harbors ca. 320 species per km2 while Tropical Asia 165, Australasia 53, North America 42, and Europe 30 (Albert

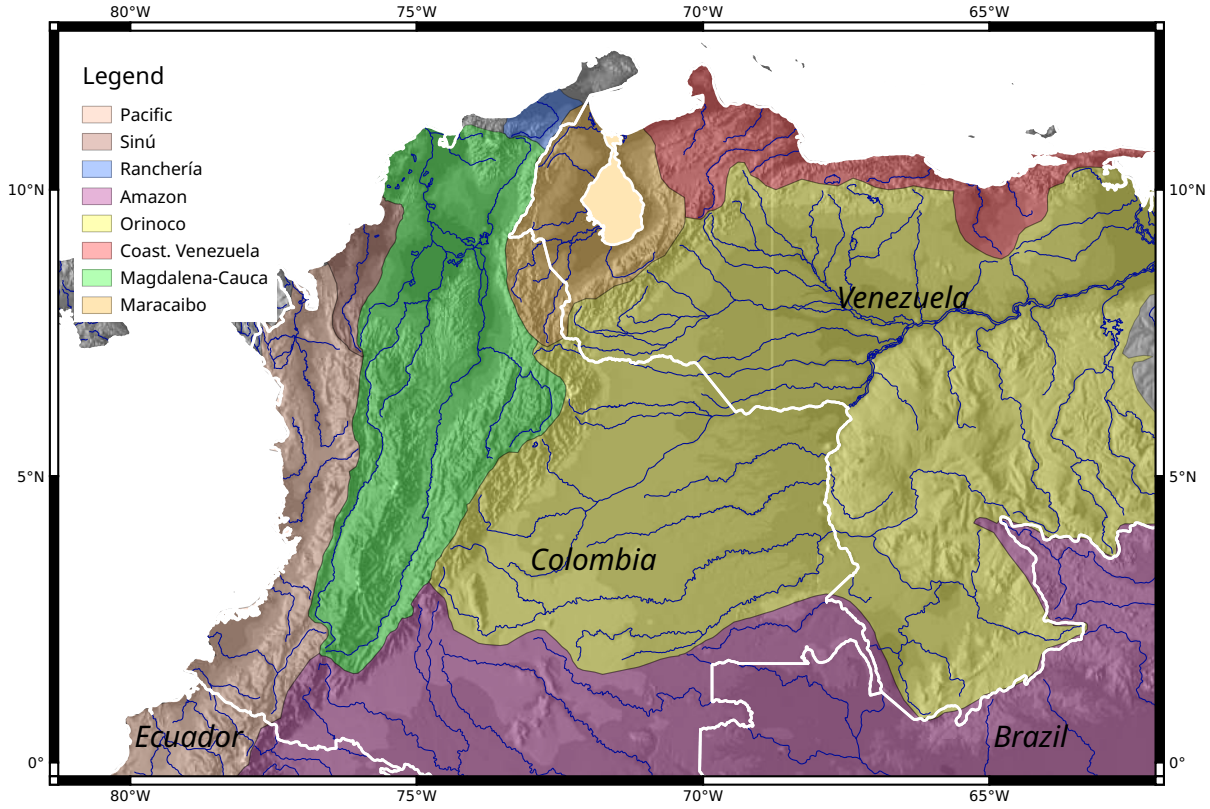


Figure 1.1: *Geographic context of major drainages in NSA. Country limits are in white and country names in italics. Drainages are colored according to the legend.*

et al., 2011). Several hypotheses have been suggested to explain the diversification and biogeographic patterns of freshwater fish biodiversity in South America, including high diversity in microhabitats, evolutionary constrains, and a complex geological and hydrological history (Albert et al., 2006, 2011; Lundberg, 1997; Lundberg et al., 1998). Such diversity arose coupled to the complex tectonic history of Northern South America (NSA hereafter) along with the consequent drainage evolution, generating several separated drainages from wider and older ones (e.g., the proto Amazonas-Orinoco; Lundberg, 1997). However, the temporal and paleogeographic context for such biotic and geological evolution remains poorly explored, with just two fossil fish faunas accounting for more than 90% of what we know about the late Cenozoic freshwater fishes of NSA (i.e., La Venta, middle Miocene of Colombia, and Urumaco, late Miocene of Venezuela). Despite such faunas have provided immense insight into the drainage evolution in NSA, much remains to be done in other lithostratigraphic units of younger age, what will provide a more precise way to look at the timing and pattern of drainage evolution and its bearing on biodiversity dynamics in geological time. Such younger units provide the unique opportunity to further constrain the orogenic history of the Northern Andes, as well as to evaluate the temporal and spatial context of the geologic history of NSA drainages.

NSA has several important river drainages (i.e., Magdalena, Maracaibo, Coastal Venezuela

and Orinoco, Figure 1.1). These drainages have changed drastically during the Neogene due to the complex tectonic and climate history of the region (Fedorov et al., 2013; Mann et al., 2006; van der Hammen et al., 1973). This area is treated as peripheral to the core Amazon region by Albert et al. (2011), and presents a poorer fish fauna when compared to the core Amazon/Orinoco/Guianas (Figure 1.2A). A large amount of its species (and even genera) is endemic to each drainage (More than half the number of species for each drainage, Figure 1.2B). The high levels of species-level and genus-level endemism of the Maracaibo and Magdalena basin suggests several speciation events during the late Neogene, and can serve as a model to study the effect of drainage history and climate change on diversification of fish lineages.

The timing of uplift of the Andes, and specially the northern Andes is still highly controversial. Some authors suggest that the Cordillera Oriental in Colombia, and the Mérida Andes and Cordillera de la Costa in Venezuela were positive by middle to late Miocene and the Perijá Range and the Mérida Andes during the late Miocene (Diaz de Gamero, 1996; Lundberg et al., 1998, and references thereir). In contrast, some authors suggest a more complex tectonic history for the Andes of NSA, arguing that the Cordillera Oriental in Colombia uplifted in pulses, with some areas being positive as early as late Paleocene (e.g., Santander Massif; Bayona et al., 2013) and with other areas uplifting from middle Eocene to middle Miocene (e.g., Perijá Range and central Cordillera Oriental; Ayala et al., 2012; Bayona et al., 2013, 2010; Caballero et al., 2010; Ochoa et al., 2012). Both scenarios would have different consequences for the freshwater fish faunas (both extant and extinct), and therefore, they could be used to better understand the evolution of the northern Andes.

A fossil vertebrate fauna was recently discovered in the Guajira Peninsula by the Smithsonian Tropical Research Institute. This fauna has been dated as spanning from 18Ma (early Miocene) to the mid Pliocene (2.7 Ma) (Moreno et al., 2015). Some of the units have yielded a continental fauna, giving a unique opportunity to study the biotic changes that took place in this area during the last 18 Ma. When analyzed along with some other fossil sites of similar age (e.g., Urumaco and San Gregorio, late Miocene to late Pliocene, both in Venezuela), one is able to explore the faunistic relationships between those areas during the Miocene to Pliocene. It is not just the geographic location of those fossil faunas but also their age what make them crucial for testing paleogeographic models, because orogenic events during late Miocene to late Pliocene are one of the conflicting aspects between the current proposals on the orogenic evolution of the Andes of NSA.

This new fossil fauna was recovered in an arid and hot region at sea level in Northern Colombia, with no present riverine connections to adjoining drainages (i.e., Magdalena and Maracaibo). The Guajira Peninsula lies to the north of the Perijá Range and presents very rich exposures of Neogene rocks, specially of late Neogene age Rollins (1965). These arid conditions along with very low precipitation (mean annual precipitation = 397 mm, mean annual temperature = 28.7 °C; Ramírez and del Valle, 2011) are strikingly contrasting with the data gathered from geology and paleontology on the environmental conditions of this

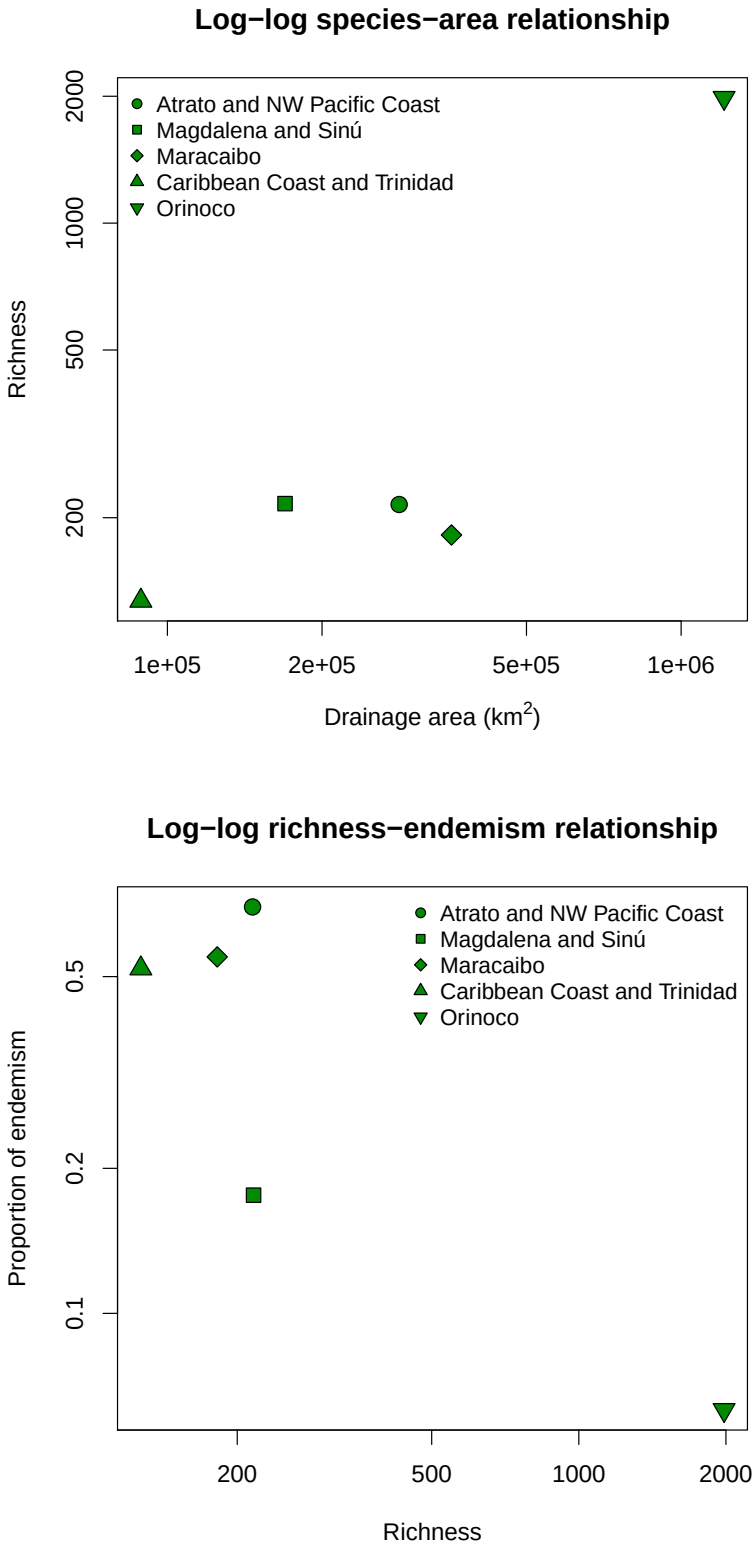


Figure 1.2: *Descriptive measures of biodiversity per drainage region. A) Log-log scatterplot of watershed area for the main drainages of NSA vs. number of species. B) Log(Number of species) vs. proportion of endemic species scatterplot for each drainage (p = endemics / total spp.). Pac = Atrato and NW Pacific; Mag = Magdalena and Sinú; Mar = Maracaibo; Car = Caribbean drainages and Trinidad; Ori = Orinoco. Note the high endemism levels for small drainages in comparison to the whole Orinoco drainage. Data taken from table 2.1 in Albert et al. (2011).*

area during the late Miocene to early Pliocene.

A preliminary paleoenvironmental reconstruction suggests that the Peninsula was a tropical forest with presence of either middle to large rivers or a deltaic area, aquatic vertebrates such as crocodylians, turtles, freshwater fishes, as well as terrestrial and amphibian mammals (Aguilera et al., 2013; Amson et al., 2016; Cadena and Jaramillo, 2015a,b; Carrillo et al., 2018; Forasiepi et al., 2014; Moreno-Bernal et al., 2016; Suarez et al., 2016). Such paleobiotic assemblage allows to explore past drainage connections with major drainages of NSA because of its intermediate location between the Magdalena and Maracaibo drainages, as well as being complementary to other fossil faunas of similar age (i.e., San Gregorio Formation in Venezuela). The main goal of this doctoral thesis is to study the freshwater fish assemblage of the Guajira Peninsula during the Miocene to Pliocene and to test different paleogeographic models by using past and present geographic distribution of the taxa recovered.

Given the state of knowledge on the tempo and pattern of drainage evolution in NSA along with the potential of the freshwater fish fossil record as a proxy for paleodrainage connections in geological time, a detailed study of the freshwater fossil fish fauna in Northern Colombia will provide a revaluation of both time and drainage separation events in NSA. The results of such study will revolutionize the tectonic models for andean orogeny in the Mérida Andes of venezuela as well as in the Cordillera Oriental of Colombia. In addition, this unique opportunity will provide a way to study the temporal component of faunal evolution, and therefore will help to explain why peripheral areas in NSA present low richness but high endemism as compared to the core Amazon/Orinoco/Guyana (Figure 2). A detailed account for the justification of this project is given below.

Justification

Recent studies on tectonic evolution of the Caribbean-South American plate dynamics have suggested a much more complex orogeny of the Andes than previously considered (Ayala et al., 2012; Bayona et al., 2010; Caballero et al., 2010; Mann et al., 2006; Ochoa et al., 2012). In spite of the large body of new geological research, geological tools are still very limited in providing information about paleotopography and landscape evolution, and consequently, there are still large controversies of when the drainages in northern South America were connected and/or separated (Bayona et al., 2013; Diaz de Gamero, 1996; Lundberg et al., 2010; Ochoa et al., 2012). Fishes on the other hand, can provide reliable information about drainage evolution, that is directly related to landscape construction and orogenic buildup. Freshwater fishes are an extremely important group when assessing paleohydrographic connections as they are directly linked to riverine environments, and both their diversification events and assemblage compositions are directly affected by drainage dynamics (Albert et al., 2006, 2011; Lundberg, 1997). Therefore, similarity in fossil fish assemblage is a powerful tool for testing drainage connections during geologic time.

Preliminary work (Aguilera et al., 2013), indicates that the early Pliocene Guajira Peninsula assemblage shows cis-Andean components (i.e., belonging to groups currently dwelling in rivers located east of the Andes as defined by the Cordillera Oriental in Colombia and the Mérida Andes and Cordillera de la Costa in venezuela). Given that all of these taxa are nowadays restricted to cis-Andean South America, their presence in the Guajira fauna suggests that by the early Pliocene a drainage connection between cis- and trans-Andean NSA was still present. On the other hand, specimens from the Sincelejo Formation in the Departamento de Sucre, northern Colombia, suggest the same kind of drainage connection between the paleo-Magdalena drainage and cis-Andean NSA. The La Venta fossil fish fauna allows the same conclusion, but the Sincelejo Formation is of early Pliocene age, extending the drainage connection between Magdalena and the proto-Amazonas-Orinoco from 12.8 Ma to around 2 Ma. This renders both the Guajira and Sincelejo faunas key to understand the final stage of drainage connection between cis- and trans-Andean NSA.

The late Miocene La Venta fauna (11–13 ma) (Lundberg, 1997) on the upper Magdalena Valley also shows a cis-Andean fauna. The San Gregorio Formation in eastern Venezuela, contains fossils of freshwater taxa nowadays extinct in Coastal Venezuela, as well as some elements currently restricted to the Maracaibo drainage (Aguilera et al., 2013) (Figure 1.1).

This geologic/paleontologic settings depicted above provide a unique opportunity to test drainage connections using the Guajira fauna in order to determine whether such paleodrainage was connected or not to the Orinoco drainage. If the fossil fauna presents components currently restricted to regions east of the Andes (cis-Andean components), then a drainage connection between the Guajira and the Orinoco for the early Pliocene needs to be advocated for explaining such faunal paleodistribution. On the other hand, if the Guajira fauna presents components currently restricted to the Magdalena drainage, it will be necessary to explore alternative models for connection with the latter drainage, but no direct connection to the paleo-Orinoco would be necessary, lending support to the idea that the Mérida Andes and Cordillera de la Costa in Venezuela were already positive areas. The hypotheses to be tested in the present project are described below.

Hypotheses and Objectives

Ho: By the late Miocene the main drainages of NSA were already formed as today, and no drainage connection was present between the Guajira Peninsula and the Orinoco drainage (Figure 1.3A). This alternative predicts a Guajira fossil fish fauna composed of taxa currently restricted to the Magdalena and/or Maracaibo drainages. As a consequence, the Mérida Andes and Cordillera de la Costa must have been uplifted by that time, restricting drainage connections between the Guajira Peninsula and the Orinoco drainage. Coupled with this consequence, the vicariant cases affected by this orogenic event are expected to correspond in timing with the late Miocene in divergence time analyses.

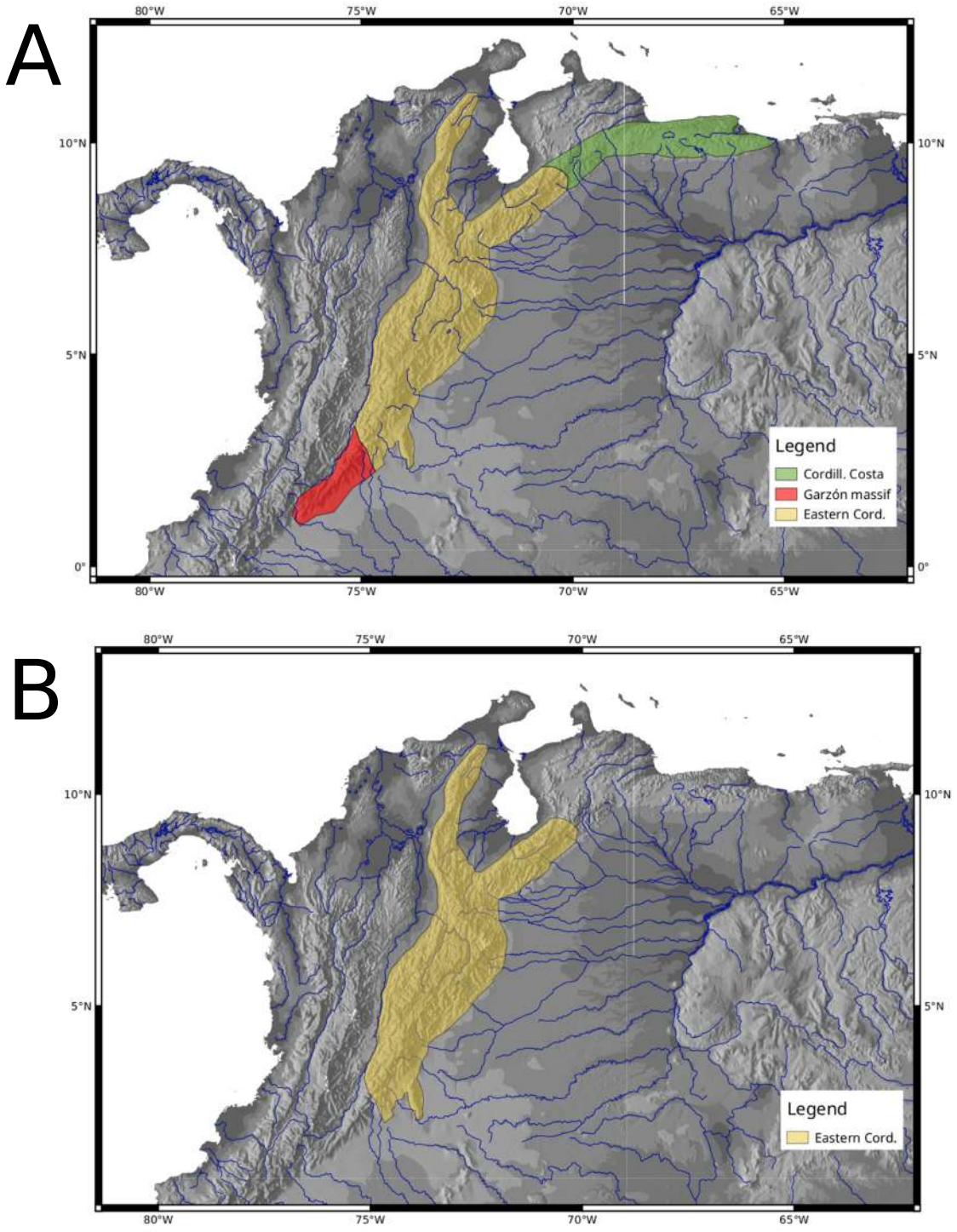


Figure 1.3: *Paleogeographic hypotheses to be tested in the present thesis. A) Null hypothesis of drainage configuration during the Pliocence. Polygons represent modern-configuration areas as water divides. B) Alternative hypothesis of drainage configuration during the Pliocene. Polygons represent the modern configuration of the Cordillera Oriental massifs while the Garzón massif and the Cordillera de la Costa in Venezuela are still not fully formed and therefore can not act as water divides. The former would allow drainage connection between the proto-Magdalena of Lundberg (1997) and the cis-Andean drainages (Amazon + Orinoco).*

Ha: The Guajira drainage was connected by the early Pliocene with the Orinoco drainage (Figure 1.3B). This alternative predicts a Guajira fossil fish fauna composed by cis-Andean taxa, suggesting drainage connections between NSA drainages and the Orinoco drainage. As a consequence, the Mérida Andes and Cordillera de la Costa were either absent or only presented partial uplift, providing space for drainage connections between the Guajira Peninsula and the Orinoco drainage. Coupled with this consequence, the vicariant cases affected by this orogenic event are expected to correspond in timing with the early Pliocene or a younger age in divergence time analyses. The objectives for the present project are as follows:

- Describe the fish assemblage present in the late Neogene sediments of the Guajira Peninsula and the Sincelejo Formation in Colombia.
- Compare the Pliocene Guajira and Sincelejo fish assemblages to with all Neogene fish faunas of NA, as well as with recent distributions.
- Test drainage connections between the Guajira Peninsula and the Orinoco based on the information provided by late Neogene fossil fish faunas as well as with time-calibrated molecular phylogenies available in the literature as well as reanalyses using the fossil information obtained in the present project.

Structure of the Thesis

The present thesis is an *opus* in five acts, as is Claudio Monteverdi's *"L'Orfeo Favola in Musica"* from 1607.

1. *Praeambulus*
2. Standardized terminology for Siluriform spines
3. A middle Miocene frehswater fish fauna from the Castilletes formation
4. Fossil fishes from the Pliocene of the Sincelejo and Ware formations
5. Statistical approaches in estimation of the time of separation between biogeographic areas

The first (and present) act settles the scene and addresses the questions to be answered and hypotheses to be tested. The second act provides the anatomical basis for the study of fossil siluriform spines (both complete and fragmentary), a prerequisite for identifying taxa from fossil spine fragments, an abundant structure often recovered from continental sediments, much neglected, and poorly understood in terms of variation and as potential sources of information for taxonomy and systematics. The third act describes the fossil assemblage in the Guajira Peninsula during the middle Miocene, a time when we expect

to evidence drainage connections across the Andes and therefore the same taxa on both sides of what we call now the Andean cordilleras. The fourth act also documents fossil assemblages but this time in more recent times, during the Pliocene in the departments of Sucre and Guajira, Colombia; this piece of information allows to mark a minimal time of persistence in drainage connection across the Andes, allowing to test initially the hypotheses of interest in the present praeambulus. Finally, the fifth act aims at testing independently the predictions of the hypothesis favored by the fossil assemblages in terms of drainage connections but this time from a statistical perspective. This last piece provides a number of analytical alternatives for using data from divergence time estimation in order to estimate a general separation time between areas as a function of individual instances recovered from these statistical estimations. Data of this nature are often found in molecular studies that frequently carry out these analysis in an almost routinely way, allowing to use reasonable samples from different biological groups for estimating general patterns in geologic time.

Bibliography

Aguilera, O.A., Lundberg, J.G., Birindelli, J.L.O., Sabaj Pérez, M.H., Jaramillo, C.A., and Sánchez-Villagra, M.R. 2013. Palaeontological evidence for the last temporal occurrence of the ancient Western Amazonian river outflow into the Caribbean. *PLoS ONE*, 8:1–17. 5, 6

Albert, J.S., Lovejoy, N.R., and Crampton, W.G.R. 2006. Miocene tectonism and the separation of cis- and trans-Andean river basins: Evidence from Neotropical fishes. *Journal of South American Earth Sciences*, 21:14–27. 2, 5

Albert, J.S., Petry, P., and Reis, R.E. 2011. Major biogeographic patterns. In J.S. Albert and R.E. Reis (eds.), *Historical Biogeography of Neotropical Freshwater Fishes*, pp. 21–57. University of California Press. xiv, 1, 2, 3, 4, 5

Amson, E., Carrillo, J.D., and Jaramillo, C.A. 2016. Neogene sloth assemblages (Mammalia, Pilosa) of the Cocinetas Basin (La Guajira, Colombia): implications for the Great American Biotic Interchange. *Palaeontology*, 59:563–582. 5

Ayala, R.C., Bayona, G., Cardona, A., Ojeda, C., Montenegro, O.C., Montes, C., Valencia, V., and Jaramillo, C.A. 2012. The paleogene synorogenic succession in the northwestern Maracaibo block: Tracking intraplate uplifts and changes in sediment delivery systems. *Journal of South American Earth Sciences*, 39:93–111. 3, 5

Bayona, G., Cardona, A., Jaramillo, C.A., Mora, A., Montes, C., Caballero, V., Mahecha, H., Lamus, F., Montenegro, O., Jimenez, G., Mesa, A., and Valencia, V. 2013. Onset of fault reactivation in the Eastern Cordillera of Colombia and proximal Llanos Basin; response to Caribbean - South American convergence in early Palaeogene time. *Geological Society, London, Special Publications*. 3, 5

Bayona, G., Montenegro, O., Cardona, A., Jaramillo, C.A., Lamus, F., Morón, S., Quiroz, L., Ruiz, M.C., Valencia, V., Parra, M., and Stockli, D. 2010. Estratigrafía, procedencia, subsidencia y exhumación de las unidades Paleógenas en el Sinclinal de Usme, sur de la zona axial de la Cordillera Oriental. *Geologia Colombiana*, 35:5–35. 3, 5

Caballero, V., Parra, M., Mora, A., and Mora Bohorquez, A.R. 2010. Levantamiento de la Cordillera Oriental de Colombia durante el Eoceno tardío-Oligoceno temprano: Prove-

niencia sedimentaria en el Sinclinal de Nuevo Mundo, Cuenca Valle Medio del Magdalena. *Geologia Colombiana*, 32:45–77. 3, 5

Cadena, E. and Jaramillo, C.A. 2015a. Early to Middle Miocene Turtles from the Northernmost Tip of South America: Giant Testudinids, Chelids, and Podocnemidids from the Castilletes Formation, Colombia. *Ameghiniana*, 52:188–203. 5

Cadena, E. and Jaramillo, C.A. 2015b. The first fossil skull of *Chelus* (Pleurodira: Chelidae, Matamata turtle) from the early Miocene of Colombia. *Palaeontologica Electronica*, 18.2:1–10. 5

Carrillo, J.D., Amson, E., Jaramillo, C., Sánchez, R., Quiroz, L., Cuartas, C., Rincón, A.F., and Sánchez-Villagra, M.R. 2018. The Neogene Record of Northern South American Native Ungulates. *Smithsonian Contributions to Paleobiology*, pp. 1–67. 5

Diaz de Gamero, M.L. 1996. The changing course of the Orinoco River during the Neogene: A review. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 123:385–402. 3, 5

Fedorov, A.V., Brierley, C.M., Lawrence, K.T., Liu, Z., Dekens, P.S., and Ravelo, A.C. 2013. Patterns and mechanisms of early Pliocene warmth. *Nature*, 496:43–49. 3

Forasiepi, A.M., Soibelzon, L.H., Suárez Gomez, C., Sánchez, R., Quiroz, L.I., Jaramillo, C.A., and Sánchez-Villagra, M.R. 2014. Carnivorans at the Great American Biotic Interchange: new discoveries from the northern neotropics. *Naturwissenschaften*, 101:965–74. 5

Lundberg, J.G. 1997. Freshwater fishes and their paleobiotic implications. In R.F. Kay, R.H. Madden, R.L. Cifelli, and J.J. Flynn (eds.), *Vertebrate Paleontology in the Neotropics: The Miocene Fauna of La Venta, Colombia*, chapter 5, pp. 67–91. Smithsonian Institution Press. xiv, 2, 5, 6, 7

Lundberg, J.G., Marshall, L.G., Guerrero, J., Horton, B., Malabarba, M.C.S.L., and Wesselingh, F. 1998. The stage for Neotropical Fish diversification: A history of Tropical South American rivers. In L.R. Malabarba, R.E. Reis, R.P. Vari, Z.M. Lucena, and C.A.S. Lucena (eds.), *Phylogeny and Classification of Neotropical Fishes*, pp. 13–48. Edipucrs, Porto Alegre. 2, 3

Lundberg, J.G., Sabaj Pérez, M.H., Dahdul, W.M., and Aguilera, O.A. 2010. The Amazonian Neogene fish fauna. In C. Hoorn and F.P. Wesselingh (eds.), *Amazonia, Landscape and Species Evolution: A Look Into the Past*, pp. 281–301. Blackwell Publishing. 5

Mann, P., Escalona, A., and Castillo, M.V. 2006. Regional geologic and tectonic setting of the Maracaibo supergiant basin, western Venezuela. *AAPG Bulletin*, 90:445–477. 3, 5

Moreno, F., Hendy, A.J.W., Quiroz, L., Hoyos, N., Jones, D.S., Zapata, V., Zapata, S., Ballen, G.A., Cadena, E., Cárdenas, A.L., Carrillo-Briceño, J.D., Carrillo, J.D., Delgado-Sierra, D., Escobar, J., Martínez, J.I., Martínez, C., Montes, C., Moreno, J., Pérez, N., Sánchez, R., Suárez, C., Vallejo-Pareja, M.C., and Jaramillo, C.A. 2015. Revised stratigraphy of Neogene strata in the Cocinetas Basin, La Guajira, Colombia. *Swiss Journal of Palaeontology*, 134:5–43. 3

Moreno-Bernal, J.W., Head, J., and Jaramillo, C.A. 2016. Fossil crocodilians from the High Guajira Peninsula of Colombia: Neogene faunal change in northernmost South America. *Journal of Vertebrate Paleontology*, 36:e1110586. 5

Nelson, J.S. 2006. *Fishes of the World*. Wiley. 1

Ochoa, D., Hoorn, C., Jaramillo, C.A., Bayona, G., Parra, M., and De la Parra, F. 2012. The final phase of tropical lowland conditions in the axial zone of the Eastern Cordillera of Colombia: Evidence from three palynological records. *Journal of South American Earth Sciences*, 39:157–169. 3, 5

Ramírez, J.A. and del Valle, J.I. 2011. Paleoclima de La Guajira, Colombia; según los anillos de crecimiento de *Capparis odoratissima* (Capparidaceae). *Revista de Biología Tropical*, 59:1389–1405. 3

Rollins, J.F. 1965. Stratigraphy and structure of the Goajira Peninsula, northwestern Venezuela and northeastern Colombia. *University of Nebraska Studies New Series*, 30:1–103. 3

Suarez, C., Forasiepi, A.M., Goin, F.J., and Jaramillo, C.A. 2016. Insights into the Neotropics Prior to the Great American Biotic Interchange: New evidence of mammalian predators from the Miocene of Northern Colombia. *Journal of Vertebrate Paleontology*, 36:e1029581. 5

van der Hammen, T., Werner, J.H., and van Dommelen, H. 1973. Palynological record of the upheaval of the Northern Andes: A study of the Pliocene and lower Quaternary of the Colombian Eastern Cordillera and the early evolution of its high-Andean biota. *Review of Palaeobotany and Palynology*, 16:1–122. 3

Chapter 2

A standardized terminology of spines in the order Siluriformes (Actinopterygii: Ostariophysi)

2.1 Abstract

A standardized terminology for the anatomy of pectoral- and dorsal-fin spines in the order Siluriformes was prepared based on extensive literature review and direct examination of representatives of the order worldwide. The anatomy of the spines is described in detail and the presence of synonyms used in the literature for the same structure were noted. Most of the characters identified have already been named, while others were herein described for the first time. A quantitative approach at solution of anatomical synonyms that minimizes the difference between the general terminology proposed and the vast amount of preexistent literature is proposed, herein called the cost function. It is expected that this system will allow to solve the chaotic anatomical nomenclature of the appendicular skeleton in Siluriformes, and provide a solid basis for advances in comparative anatomy. The present terminology system is of potential application to a number of fields ranging from taxonomy to phylogenetic systematics to paleontology and archaeology.
Keywords: Morphology, Appendicular skeleton, Osteology, Anatomy, Catfishes.

2.2 Introduction

The order Siluriformes is one of the most important components of Neotropical freshwaters with more than 3800 species currently described (Eschmeyer and Fong, 2018). It is one of the most prominent component of freshwater fish biodiversity worldwide, being specially speciose in South America and Asia. Its species inhabit both freshwater and marine environments, and reach elevations above 3000 masl in the Andes. It is considerably heterogeneous in morphology across its 38 families, including miniature species of the Scoloplacidae to gi-

ants in the Pangasiidae, Pimelodidae, and Siluridae. This wide variation in morphology is reflected in its ecological, behavioral, and trophic complexity (Arratia et al., 2003).

The dorsal and pectoral spines are one of the most conspicuous features of Siluriforms, that differ from percomorph spines in that they preserve some evidence of former division, either in the plane of contact between hemitrichia, or at the apical region where growth lines are usually present for allowing growth in length through distal segment addition facilitated by the spurious ray (Kubicek et al., 2019; Reed, 1924). These stiff and pungent structures are present in most but not all of the families in the order, and are thought to play a role in mechanical defense (Bosher et al., 2006), as a poisoning mechanism (Egge and Simons, 2011; Wright, 2009), and in sound production (Kaatz et al., 2010; Parmentier et al., 2010), with several biological consequences. Most of our knowledge about its ontogeny dates back to Reed (1924) and has been greatly expanded by Kubicek et al. (2019), while its morphological relationship to the anatomy of the pectoral girdle has been studied to some extent (Alexander, 1965; Brousseau, 1976; Miano et al., 2013); on the other hand, our knowledge about the anatomy of the dorsal fin and its spine are comparatively poorer. Spines tend to be the most common Siluriform fossil remains due to its hardness and potential of preservation; this is particularly true for fossil localities showing a sedimentological setting of high energy accumulation (e.g., conglomerates and coarse sands). Given these properties, they might provide important information on extinct diversity of Siluriforms, conditional to our ability to extract information from their preserved morphological details (for a study case in the Mochokidae see Pinton et al., 2006).

Spine morphology has been demonstrated to be a useful source of characters both in systematic and taxonomic studies. For instance, interspecific variation in the anterior and posterior ornament has been used to distinguish species of *Acanthodoras* (Eigenmann, 1925), *Microglanis* (Shibatta and Benine, 2005), *Rhyacoglanis* (Shibatta and Vari, 2017), *Pimelodella* (Bockmann and Slobodian, 2014; Slobodian et al., 2017), *Synodontis* (Pinton et al., 2006), and *Pimelodus* (Costa e Silva et al., 2018). A more modest amount of characters have been recognized to be informative in phylogenetic studies, as in the Astrodoradinae (Souza, 2010), some subgroups inside Ictaluridae (i.e., the *rabida* group, *Ameiurus*, and the Ictaluridae except *Ictalurus*, Arce-H et al., 2017). However, its potential as source of characters is still to be exploited as a series of different morphologies await proper description.

Anterior and posterior ornamentations on both dorsal and pectoral spines have been repeatedly termed in several ways in different references. As an example, the following works have used a number of terms for description of antero-posterior ornaments in spines:

- Gayet and Meunier (1998): Denticles (in the Diplomystidae, p. 247)
- Eigenmann and Allen (1942): Serrae (in *Pimelodella montana*, p. 100, and *Opsodoras parallelus* [= *Nemadoras elongatus*], p. 134); serrations (in *Pimelodella peruana*, p. 101); hooks (in *Microglanis zonatus*, p. 89, *Pimelodella roccae*, p. 99, and *Ageneiosus ucayalensis*, p. 138); thorns (in *Pimelodella peruensis*, p. 98, and *Pimelodus leptus*

[= *Cheirocerus goeldii*], p. 105); teeth (in *Pimelodus jivaro*, p. 105); rugosities (in *Pimelodella peruana*, p. 101).

- Mees (1974): Teeth (in all of the Pimelodidae and Auchenipteridae studied in his work).
- Gayet and van Neer (1990): Spines (originally *épines* in french, as in *Arius*, p. 243); tubercles (originally *tubercules* in french, as in *Arius*, p. 243); denticles (originally *denticules* in french, as in *Bagrus*, p. 245).
- Pinton et al. (2006): Tubercles (for lateral ornamentations); denticles (same structures as tubercles, see description of pectoral spines in *Synodontis*, p. 24).
- Lundberg (1997): Dentations (in cf. *Pimelodus* and *Phractocephalus* sp. from La Venta Fauna, p. 78); tubercles (in Ariidae, p. 80).
- Aguilera et al. (2008): Dentations (in *Phractocephalus acreornatus*, p. 240).

From the listing above one could ask: Do these different terms refer to different morphologies? Is the term serrae in Vanscoy et al. (2015) the same condition as in Eigenmann and Allen (1942)? Do slightly-different terms such as denticles (Gayet and Meunier, 1998) and dentations (Lundberg, 1997) actually differ? These potential issues are a direct consequence of the lack of standard in anatomical description of dorsal and pectoral spines in the order Siluriformes. Several authors have noted the inconsistencies and lack of a uniform, standardized terminology for the spines in Siluriforms (*e.g.* Bisbal and Gomez, 1986; Hubbs and Hibbard, 1951). However, to date we still lack a consistent, term-rich system for describing the vast variation in spine morphology across the order. Furthermore, almost every author tends to use their own terms without consistency across works, that is, even the same author can describe a given condition with different names in different articles. Given the confusion in anatomical terms historically applied to spines and its ornaments, herein I propose a new standard terminological system to avoid future confusions. Each term is accompanied by a survey of variation across the order, and an attempt at synonymy along the numerous references examined. A quantitative approach at standardization is herein proposed for picking the optimal system from among the already proposed terms.

2.3 Materials and Methods

2.3.1 Literature review

A broad collection of sources were reviewed for spine descriptions from a wide range of fields of ichthyological research, from classic anatomical descriptions to behavioral studies. Special attention was paid when tracing the origin of wide-spread terms such as descriptors of curvature (i.e., retrorse, antrorse) given the variation in their application in the literature of the 20th century. Disambiguation of anatomical concepts was achieved by reference to

the original illustrations, and in selected cases, by direct study of osteological material. Only references with explicit reference to anatomical details either with labeled figures or detailed description of characters were taken into account, so that any misinterpretation of character definitions was minimized. Myological terminology follows Miano et al. (2013) unless for new terms herein proposed.

2.3.2 Museum specimens and anatomical preparations

A total of 197 species were examined in the present study (Appendix B). Specimens studied have been largely prepared using the method of Taylor and van Dyke (1985) for cleared-and-stained specimens, and dry skeletons cleaned with dermestid beetles for whole-body preparations. In some instances the spines were dissected and detached from the pectoral girdle, and then cleaned through exposition to about five drops of commercial preparations of sodium hypochloride (bleach) dissolved in 25 mL of tap water; sometimes more drops were added for faster results. The process was complemented by manual remotion of tissue chunks with help of insect pins and fine forceps. The spines were examined during the process under stereomicroscope until complete remotion of skin, muscle, and connective tissue; afterwards, a thorough wash of tap water was applied during several days before drying. Museum acronyms follow Sabaj Pérez (2013). Muscles and ligaments are spelled in Latin, and abbreviations *m.* and *l.* used for *musculus* and *ligamentum* respectively.

2.3.3 Cost function

Any proposal of a standardized terminology must deal with previous usage or systems. As a consequence, the more different a system to the whole body of terms already used in the literature, the higher its cost to apply or adopt it in terms of similarity and ease of comparisons with previous literature. This has a two-fold effect: 1) It is more difficult to be accepted by the scientific community, and 2) it is more difficult to compare the new system to previous terms in the literature, and therefore its amount of mismatch and informativeness would be inversely proportional.

Given the huge amount of spine terms already available in the literature on Siluriformes, the most costly alternative would be to propose a whole system with *de novo* terms for all the characters of Siluriform spines, since a completely novel terminology system would be different from the whole set of terms already available in the literature.

A particular term (if adopted) would then have a cost associated to the number of references that have used different terms for the same character/character-state, so a i^{th} term t_i would be different from a given term in a given reference t_{ref}. If different, such term would have a cost of 1 with respect to the reference, or 0 if the such term is also used in the reference.

A cost function can be defined as an expression ranking a particular combination of terms (C_{total}) consisting of particular terms t. These terms can be new or already proposed in the

literature (t_{ref}). C_{total} is calculated for a given set of terms from $i = 1$ to n.

$$C_{total} = \sum_{i=1}^{n} C(t)_i \quad \text{where} \quad C(t) = I_{\{t \neq t_{ref}\}} \tag{2.1}$$

The optimal terminology system should be the one satisfying:

$$\underset{t}{\operatorname{argmin}}\, C_{total} \tag{2.2}$$

The present approach has some advantages over a subjective selection of arbitrary terms in the definition of a given system: First, it considers all the possible combinations of terms regardless of where/when they were proposed; second, the optimality criterion is the combination of terms that minimize mismatches with the whole body of terms already proposed in the literature; third, the cost is calculated by comparison with all the available terms, one per reference, so no term has a larger *a priori* weight over others; and fourth, the terms are not biased by authoritative considerations, that is, a widely-use term not proposed by an authority has precedence over other terms.

2.3.4 Data analysis

Analyses were carried out in **R** v.3.4.1 (R Core Development Team, 2018). Implementation of cost calculations are available in Appendix E.1.

2.4 Results

2.4.1 Literature review

A total of 77 references were surveyed for Siluriform spine morphological terms and provided explicit terms that could be associated to specific spine traits, with or without explicit reference to illustrations. A total of 145 terms were found to have been used in the references examined for characters from the pectoral spine, whereas 65 terms were found to be used for description of dorsal spines. These terms can be divided in two groups: Characters from the spine shaft ornaments (62 and 37 respectively, 84 total), and those from the spine base morphology (83 and 28 respectively, 108 total). For both kinds of spines the former group is by far the largest, and constitute the most problematic part of the present attempt to standardize terminology, given that characterization of the particular ornament present in a spine must be carried out on high resolution photographs or direct examination of specimens; a common factor in the vast majority of the literature is that the reliance on illustrations produces a lack of detail in the ornamentation of spines, given that they are usually illustrated in outline and therefore lack the detail necessary to properly describe them based on illustrations. Given this challenge, it was not attempted to recognize the new

terms for the whole of literature examined and instead the focus is on providing a standard for future works and reexamination of published descriptions of Siluriform spines.

2.4.2 How to deal with chaos: Choosing a standard nomenclature with the cost function

The implementation of the cost function (Appendix E.1) reveals that several terms should take precedence over others. However, given the current literature review, some sets of terms are equally-optimal since their individual costs are equivalent (e.g., both 'median furrow' and 'deep median groove' have the same cost when choosing between them since they appear with equal frequencies in the examined references). The cost function successfully identified a least-cost combination of terms to be proposed as standard system, but failed to decide among equivalent alternatives for some cases. Further decisions needed to be done case-by-case where additional parameters are necessary to build a standardized system (Appendix A, Tables A.1,A.2).

Performance is usually of concern during implementation of procedures based on products or combinatorics. In the current implementation of the cost function a total of 6264 possible combinations were examined with convergence to the optimal solution in about 0.08 seconds since a shortcut in the algorithm was applied when noting that the best system would be the one minimizing cost individually, and since the whole function is a sum of indicator functions, then individual optimization would result in an overall optimal terminology system. Larger problems, with maybe more than three possible dimensions and much more cases could take more time, but it is expected that execution time would not differ much from a linear function relating complexity and time.

2.4.3 Proposed standardized terminology

A description of each structure along with its proposed synonyms is presented with comments whenever appropriate. The following system consists of 14 terms for the dorsal spine and 21 for the pectoral spine plus a set of three topological and 12 terms applied to surface ornaments. In general the Siluriform spines can be divided into base and shaft (Figure 2.1). The base is a complex structure formed by fusion of the basal portion of each hemitrichium along with the different elements for each kind of spine; in the dorsal spine the proximal second distal radial is fused to the base forming the anteroventral portion of it (Alexander, 1965), whereas the first proximal radial forms the proximal part of the pectoral spine base (Reed, 1924; Royero, 1999). Both spines show a blade-like shaft with or without ornaments on its anterior, posterior, dorsal, or ventral surfaces. The spines grow distally by calcification and further fusion of lepidotrichia (Kubicek et al., 2019; Vanscoy et al., 2015).

Dorsal and pectoral bases differ in striking ways, though both have the same functional role, that is, attachment and locking to the fin inner support (Alexander, 1965; Reed, 1924).

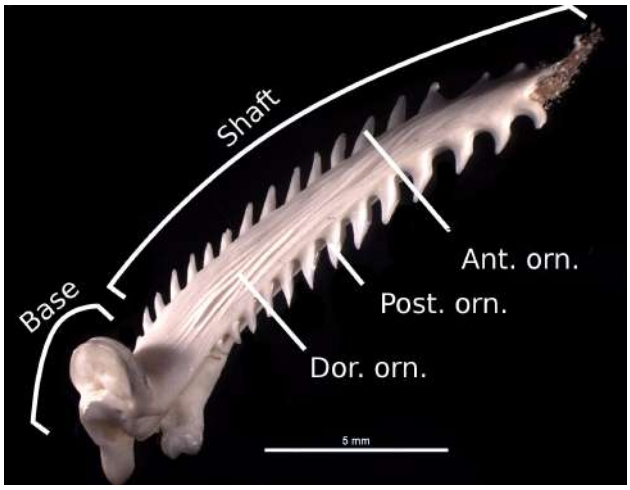


Figure 2.1: *Gross morphology of the Siluriform spine as exemplified by a pectoral spine (*Cephalosilurus fowleri*, MZUSP 73756)*

Both structures provide anchoring surfaces for muscles that erect and adpress the structure, and bony support for strong locking and sound production (Parmentier et al., 2010). In both types of spines the base is the proximal portion and it is concealed under the skin, nuchal plates, or into the cleithral cavity, whereas the defensive shaft is well exposed and comprises most of the length of the spine, where poison glands may be present along with harmful ornaments (Bosher et al., 2006; Wright, 2009). This portion undergoes extensive growth and bone deposition through the life of the animal, and is capable of regrowth and ontogenetic modification of its surface ornaments (Reed, 1924).

Ornament growth can be divided in two main processes, addition and thickening. The former process depends on the type of ornament, as those from the anterior and posterior surfaces are added through segment addition and calcification, whereas the ornaments on the dorsal, ventral and lateral surfaces are linked to the general process of thickening and dermal calcium deposition. Distal ornament addition takes place concomitant to calcification of lepidotrichia, and the orientation and shape are determined by the latter process. Little is known concerning ornament addition and its relationship to ornament diversity in shape, mostly because such variability has been overlooked in the literature.

Dorsal-fin spine

The dorsal-fin spine base consists of a compact bone mass with lateral expansions, a central rough surface for interaction with the proximal radial, and a large foramen for intrusion of the bony arc of the dorsal portion of the same radial. The dorsal-fin spine base presents rough articulating surfaces for interaction with other bones of the dorsal skeleton, and prominent surfaces for insertion of muscles and ligaments that produce abduction or adduction (sometimes called erection and depression respectively) of the spine with respect to the body. In general there are paired *m. erectores spinales* and *m. depressores spinales*, a pair of *l. interspinalia* that link the spinelet and the spine, and either a ligament (at

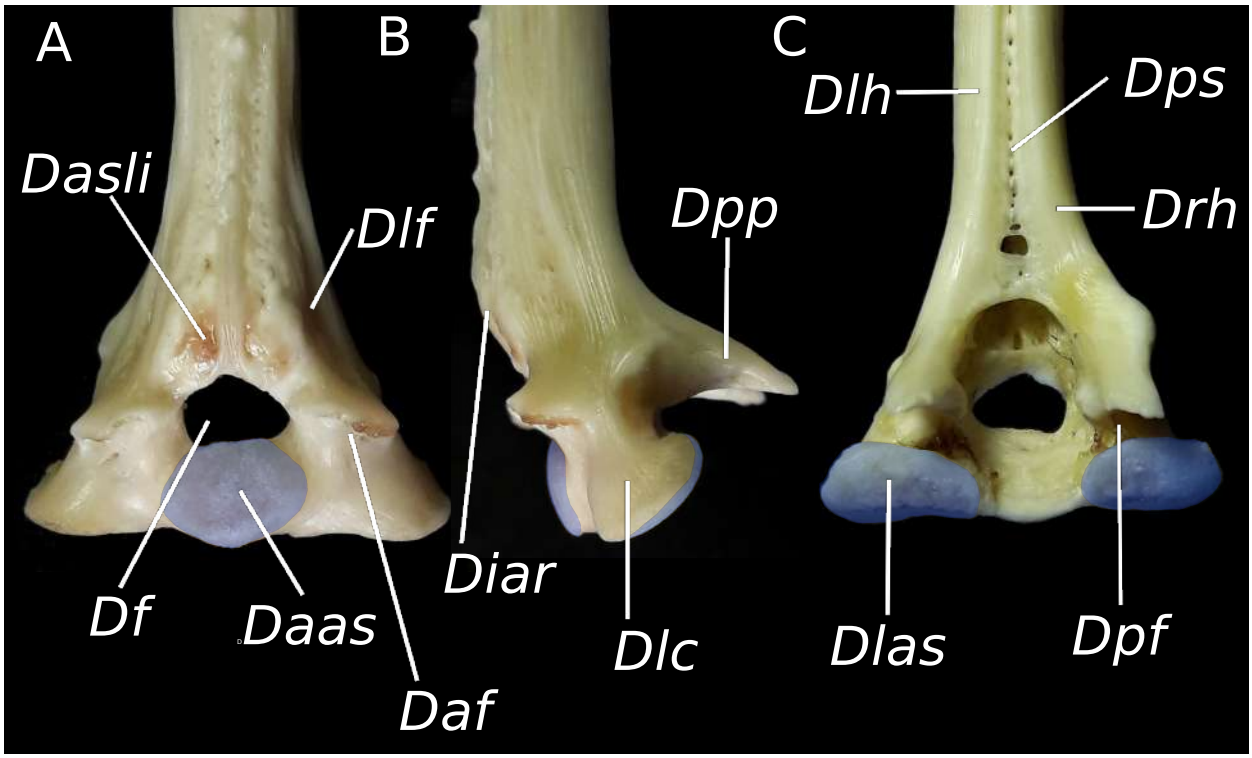


Figure 2.2: *Descriptive standard terms for the dorsal-fin spine base as exemplified by* Zungaro zungaro *(MPUJ 13213). Articular facets are shaded in blue. Abbreviations as in the text.*

least in *Pimelodus fur*) or a dense layer of connective tissue (e.g., in *Batrochoglanis villosus*) that attach the posterior face of the spine base to either the posterior nuchal plate or the pterygiophore, covering laterally the insertion of the *m. depressores spinales* (Figure 2.3); these are serially homologous to the *m. erectores dorsales*, *m. depressores dorsales*, and the *l. interradialia* present between branched rays. The following terms are illustrated in Figure 2.2, and referenced by the abbreviations between parentheses.

Anterior articular surface (*Daas*; Eigenmann, 1925): Anterior convex and rough structure located ventral to the foramen, although it can be inconspicuous in some taxa; the shape and relative size of this element is variable among Siluriforms. This structure is completely absent in *Tetranematichthys quadrifilis*, *Calophysys macropterus*, and species of *Pseudoplatystoma*, presumably as a secondary lost given its ubiquity among Siluriform taxa. The outline of this structure also varies, as some species present a dorsal notch (e.g., *Diplomystes camposensis*, *Sorubim lima*, *Zungaro zungaro*, *Liosomadoras oncinus*, *Batrochoglanis villosus*, and *Cephalosilurus fowleri*), while the dorsal outline is convex it some others (e.g., *Lophiosilurus alexandri*, *Platystomaticthys sturio*, *Schilbe sp.*, and *Synodontis schall*). The species where such structure is absent have in common a poor development of the dorsal-fin spine, and in some cases, the loss of the ability to trigger the spine-locking mechanism that requires friction between the dorsal-fin spine, spinelet, and their respective pterygiophores (Alexander, 1965). The anterior articular surface is bilaterally-divided in an aberrant individual of *Steindachneridion parahybae* from the aquaculture (MZUSP 100672). Synonyms: Articular median process (Pinton and Otero, 2010); *Condilo medial* (Royero,

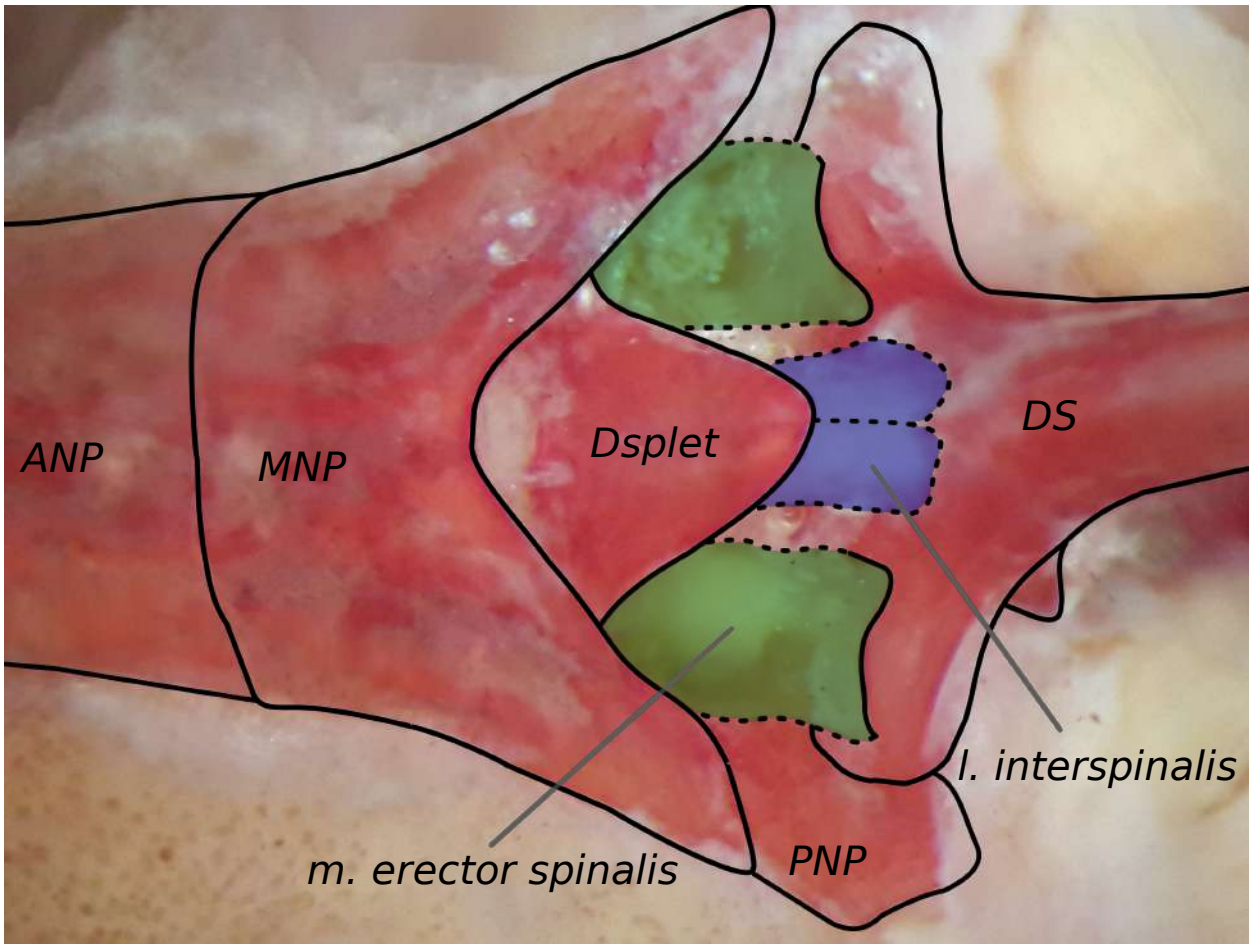


Figure 2.3: *Muscles and ligaments associated to the dorsal-fin spine as exemplified by* Hemisorubim platyrhynchos. *A) Dorsal view with spine depressed; B) Posterior view.*

unpub.); Median articular process (Otero et al., 2007); *Processus articulaire médian* (Gayet and van Neer, 1990); *Processus médian d'articulation de l'épine* (Gayet and van Neer, 1990); rd2 (radialis distalis) (Alexander, 1965).

Lateral condyles (*Dlc*; Divay and Murray, 2015, Royero, unpub): Bilateral structures that articulate with the dorsal platform of the proximal radial and contribute along with the anterior articular surface to the spine locking mechanism (Alexander, 1965). Both the lateral extension of these structures and its outline in anterior view vary across taxa. The lateral margins in anterior view are steeper in Ariids (e.g., *Arius proops*) and species of the *Pimelodus* clade (e.g., *Iheringichthys labrosus*, *Pimelodus fur*), while they are concave and projected laterally in the Auchenipteridae (*Trachycorystes trachycorystes* and *Trachelyopterichthys anduzei*). These structures seem to be homologous with the bases of contralateral hemitrichia. Synonyms: *Aile latéral* (Gayet and van Neer, 1990) Basal condyle (Lundberg, 1975); Lateral wing (Argyriou et al., 2015; Pinton et al., 2006).

Lateral articular surfaces (*Dlas*; Eigenmann, 1925; Otero et al., 2007): Rough surfaces that cap the lateral condyles on the ventral or ventro-posterior surface. These surfaces may be convex, flat, or somewhat concave, and also show striking variation in outline. Synonyms: Articular facet for posterior nuchal plate (Pinton and Otero, 2010); *Surface articulaire de l'aile latéral* (Gayet and van Neer, 1990).

Foramen (*Df*; Otero et al., 2007; Pinton et al., 2006; Pinton and Otero, 2010): Cavity with antero-posterior orientation located just dorsal to the level of the lateral condyles; with size, shape, and margins showing variation. The margins can be formed by the superposition of the anterior and posterior outlines, or be tunnel-shaped with round, complete anterior and posterior margins. When the anterior and posterior margins are not closed, the anterior face is closed by the anterior articular surface that lies just ventral to this structure. The bony ring of the proximal radial passes through the foramen so that the spine can not be disarticulated from the radial without breaking the ring (see also Alexander, 1965, p. 114). Developmental series of *Ictalurus punctatus* suggest that such annular ossification develops directly from the proximal radial (Grande and Shardo, 2002, figs. 9-10). Synonyms: Articular foramen (Argyriou et al., 2015); Basal foramen (Divay and Murray, 2015); *Foramen de la base de la espina* (Royero, unpub.); *Foramen médian* (Gayet and van Neer, 1990).

Anterior fossae (*Daf*; new name): Insertion surfaces for the *m. erectores spinales*. These surfaces are usually defined by a bony rim on the dorsal margin, a concave surface, and frequently lack a clear ventral margin. In some taxa the dorsal rim can present a central process, or the entire fossa can be similar to a process where the tendinous terminus of the muscle inserts. In some taxa the outline of the anterior fossae is round (e.g., in *Wertheimeria*), vertically oval (e.g., in *Orinocodoras eigenmanni*), horizontally oval (e.g., in *Zungaro zungaro*), or even absent where only the dorsal rim is preserved as a pointed process (e.g., in *Agamyxis albomaculatus*). Synonyms: None.

Inflection point of the anterior longitudinal ridge (*Diar*; originally "inflexion point of the median crest" in Pinton and Otero, 2010): The anterior longitudinal ridge usually

presents a swelling region just dorsal to the foramen, almost at the proximal terminus of the ridge. This region can be just the thickest proximal point of the ridge, or may present a knob with variable degree of ornamentation, or be completely absent along with the anterior longitudinal ridge. Frequently it defines the proximal limit of the anterior shaft ornament. Its knob is usually a rough, convex surface for articulation with the dorsal-fin spinelet (i.e., the first unbranched element of the dorsal fin), and contributes to the locking mechanism by means of friction with the posterior surface of the spinelet. A prominent vertically-elongate rough knob is present in the *Pimelodus* group and might prove to be a synapomorphy. Synonyms: None.

Attachment surfaces for the *l. interspinalia* (*Dasli*; new name): Anterior surfaces of the base lying just ventral to the inflection point of the anterior longitudinal ridge or any modification of it. It may be a concave surface or not, but usually consists of a faintly defined surface on the base where remains of the *l. interspinalia* are found even after cleaning the specimen. The *l. interspinalia* were described by Alexander (1965, pp. 114-115) without applying any special name to it. This feature is absent in *Plotosus lineatus*. Synonyms: None.

Lateral fossae (*Dlf*; new name): Concave regions on the lateral surface between the spine base and shaft, frequently with a spongy texture. These cavities may be sometimes very pronounced (e.g., in *Orinocodoras eigenmanni*) to even absent (e.g., *Agamyxis albomaculatus*). There are no muscles or ligaments associated to this structure. Synonyms: None.

Posterior processes (*Dpp*; Otero et al., 2007): This structure arises above the posterior fossae and lie just below the beginning of the spine shaft. This process can be hypertrophied in the Pimelodidae and some representatives of the Ariidae, where it can enter the point between the second and third proximal radials. In some catfishes this process can be strongly reduced without any clear mechanical role in spine kinematics (e.g., *Scorpiodoras heckelii*). No insertions for the *m. depressores spinales* were observed directly onto the process, though their tendinous ends can insert just below the base of this structure. Synonyms: *Processus posterieur de blocage* (Gayet and van Neer, 1990).

Posterior fossae (*Dpf*; new name): This surface serves as the insertion point for the *m. depressores spinales*, where such muscles attach tendinously. Generally this area is defined by the posterior margin of the lateral articular surfaces to the posterior process, while its lateral boundaries can be diffuse in some cases; its surface tends to be concave (e.g., in *Zungaro zungaro*) while sometimes it can be nearly flat. Synonyms: None.

Posterior sulcus (*Dps*; modified from Cione et al., 2005): Two hemitrichia fuse partially in order to form the spine; while the anterior contact between both halves is completely fused, the posterior contact leaves a discernible trace of the compound nature of the dorsal spine. The posterior sulcus is a region along the midline of the spine where the lumen of the spine opens. Posterior ornament usually develops on top of this feature. Synonyms: *Canal central* (Royero, unpub.); Deep median groove (Lundberg, 1975); Median furrow (Divay and Murray, 2015).

Left and right hemitrichia (*Dlh, Drh*): Epidermally-derived, bilateral, calcified ele-

ments that fuse to each other to form the spine. Synonyms: None.

Anterior and posterior ornaments (*Dao, Dpo*): Bony structures growing by dermal exostosis and formed by serial segment addition that determine their shape. They form on the anterior and posterior surfaces of the spine shaft respectively. They can form on a preexistent longitudinal ridge or not. A great amount of variation is present in the specific patterns of ornamentation, including variation along the shaft, so that several conditions can be present on different portions of the shaft. The addition patterns are complex and its phylogenetic structure is very poorly understood (e.g., Reed, 1924; Vanscoy et al., 2015).

Lateral ornaments (*Dlo*): Bony ornaments developing onto the lateral surfaces of the shaft by dermal exostosis.

Pectoral-fin spine

The pectoral-fin spine has a bony base for articulation with the pectoral girdle, specially with the scapulocoracoid. These spines are formed by the fusion of both the propterygium and the pectoral radial to the base of both hemitrichia, therefore reducing the number of pectoral radials to three (vs. four pectoral radials primitively in Teleosts). The spine shaft is formed by the strong contact to the eventual fusion of complementary hemitrichia, and shows the same growth patterns of the dorsal spine. Movement of the pectoral spine is controlled through a muscular system of higher complexity than the one associated to the dorsal spine (Miano et al., 2013; Winterbottom, 1973) and four unpaired muscles have been recognized to associate to the pectoral-fin spine: *m. abductor spinalis*, *m. adductor spinalis*, *m. arrector dorsalis spinalis*, and *m. arrector ventralis spinalis*. The following terms are illustrated in Figure 2.4, and referenced by the abbreviations between parentheses.

Dorsal process (*Pdp*; Hubbs and Hibbard, 1951; Kaatz et al., 2010; Miano et al., 2013; Paloumpis, 1963; Parmentier et al., 2010): Dorsal prominent expansion of bone on the base for articulation that matches a groove onto the mesial surface of the cleithrum in shape, extension, orientation, and curvature. It is formed by the base of the dorsal hemitrichium (Vanscoy et al., 2015). Alexander (1965) describes in depth the kinetic associations of this structure and also points out to its role in sound production (see also Parmentier et al., 2010, for a schematic explanation of the spine locking mechanism; fig. 8). The equivalent of this structure in *Nematogenys inermis* is double and obliquely-displaced, so it results even challenging to try to identify the homology between these two conditions. Species with poorly-developed pectoral spines tend to have a consequently ill-defined dorsal process or to lack it altogether (e.g., *Calophysus macropterus*, *Brachyplatystoma platynemum*; Vanscoy et al., 2015, pers. obs.). The degree of development of the dorsal process seems to correlate with the degree of development of the articular cavity on the scapulocoracoid, and species lacking the process usually loose the intimate the pectoral girdle encasement. Synonyms: 2 (Merriam in Hubbs and Hibbard, 1951); arched crest (Bunkenroad in Hubbs and Hibbard, 1951); Articular plateau (Otero et al., 2009; Pinton et al., 2006; Pinton and Otero, 2010);

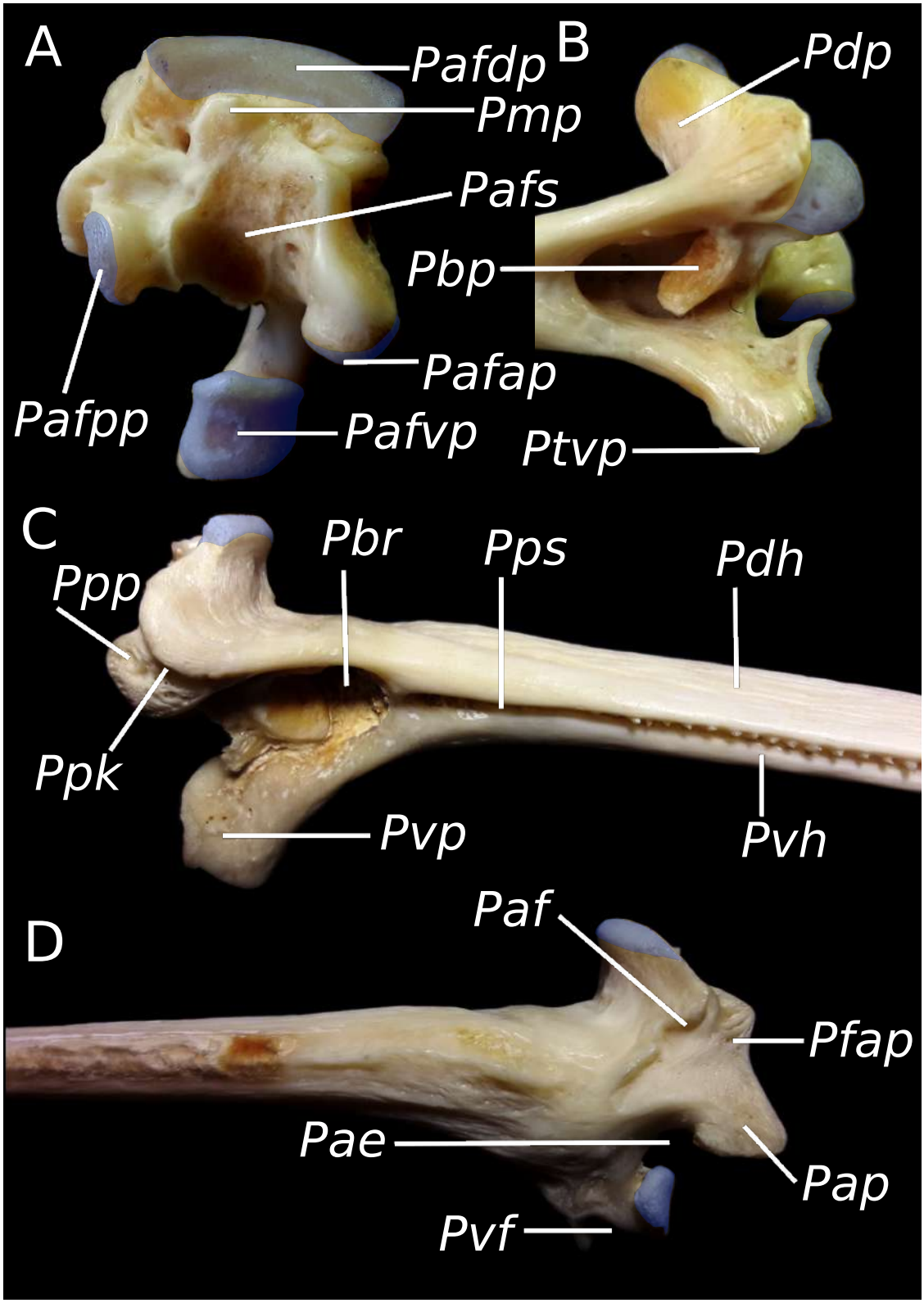


Figure 2.4: *Descriptive standard terms for the pectoral-fin spine base as exemplified by* Zungaro zungaro *(MPUJ 13213). Articular facets are shaded in blue. Abbreviations as in the text.*

Cleithral process (Argyriou et al., 2015); Curved flange (Alexander, 1965); δ (Sörensen in Hubbs and Hibbard, 1951); Dorsal articulating process (Divay and Murray, 2015; Vanscoy et al., 2015); Dorsal articulating surface (Hubbs and Hibbard, 1951); Dorsal condyle (Diogo, 2007a; Diogo et al., 2004); Dorsal process of spine (Brousseau, 1976; Cione et al., 2005); Expanded flange (Eigenmann, 1925); Helicoidal semidisk (Royero, 1999); *Surface cleithrale* (Gayet and van Neer, 1990).

Articular facet of dorsal process (*Pafdp*; new name): Rough and elaborated articular surface for friction with the cleithral groove; this structure allows both the spine locking mechanism and the sound-production system to work through stridulations. Parmentier et al. (2010) reported variation in the ultrastructure of the ornament of this surface that seems to correlate with sound-production capabilities and degree of development of the median process. Synonyms: None.

Median process (*Pmp*; modified from Diogo, 2007a,b; Diogo et al., 2003, 2004, 2006b, 2001): Generally a low, blunt ossification on the medial margin of the dorsal process. This process can be continuous with the dorsal surface of the articular facet for the scapulocoracoid (see below) so that the rotation axis of the spine presents a more firm contact to the bony vertical medial scapulocoracoid condyle with which it articulates (or scapular condyle in older literature; Brousseau, 1976). The tendinous insertion of the *m. arrector dorsalis spinalis* is generally located on the medial face of median process (Miano et al., 2013). Sometimes it can be absent at all (e.g., *Diplomystes camposensis*) or hypertrophied (e.g., in representatives of the superfamily Doradoidea, the Aspredinidae, the Pseudopimelodidae, and the Mochockidae). The median process is absent in *Plotosus lineatus*. Synonyms: Anterior process of dorsal condyle (Diogo, 2007a); Axial process (Parmentier et al., 2010; Pinton and Otero, 2010); *Processus axial* (Gayet and van Neer, 1990); Proximal tubercle (Hubbs and Hibbard, 1951); Rotator process (Royero, 1999).

Articular facet for the scapulocoracoid (*Pafs*; Diogo et al., 2006a): This semicylindrical articular surface has a cartilaginous coat for protecting the base of the spine from the extensive friction with the scapulocoracoid condyle that provides support for vertical rotation of the pectoral spine (Brousseau, 1976; Miano et al., 2013). This structure uses to be well defined with evident boundaries, although sometimes it becomes continuous with the median process for a more firm articulation with the scapulocoracoid condyle in groups with very strong spine locking system. This continuity is present in exemplars of the superfamily Doradoidea and the Mochockidae. Synonyms: Articular groove (Pinton and Otero, 2010); Articular notch (fossa) (Vanscoy et al., 2015); Central articulating surface (Hubbs and Hibbard, 1951); Cotyle of spine (Brousseau, 1976).

Anterior fossa (*Paf*; Hubbs and Hibbard, 1951): This concave surface is the insertion point of the *m. abductor spinalis* where it attaches tendinously, and lies lateral to the flange of the anterior process. Its outline is variable among groups of catfishes. Synonyms: Anterior basal recess (Divay and Murray, 2015); Insertion for arrector dorsalis dorsal division muscle (Diogo et al., 2001); Insertion surface for ventral arrector part z muscle (Brousseau, 1976).

Flange of the anterior process (*Pfap*, new name): This structure is a dorsal rim or flange onto the anterior process, presumably for maximizing the available area for muscle insertion, since the *m. abductor spinalis* is the largest of the four muscles that insert onto the pectoral spine base. Although an older name is available, we chose to propose a new name given the topographical relationship of this feature with the anterior process. Synonyms: Proximal crest (Hubbs and Hibbard, 1951).

Anterior process (*Pap*; Hubbs and Hibbard, 1951; Kaatz et al., 2010; Miano et al., 2013): This is the second largest process on the pectoral-fin spine base. This process is deflected ventrally in *Chrysichthys auratus*, *Wertheimeria maculata* and *Schilbe sp.* and straight in *Plotosus lineatus*, and *Iheringichthys labrosus*. Synonyms: Anterior articular process (Vanscoy et al., 2015); Anterior condyle (Diogo et al., 2006a); Anteroventral process (Parmentier et al., 2010); Dorsolateral process (Argyriou et al., 2015; Brousseau, 1976; Pinton et al., 2011; Royero, 1999); *Processus dorso-latéral* (Gayet and van Neer, 1990); Proximoventral process (Pinton and Otero, 2010); Ventral process (Paloumpis, 1963).

Anteroventral emargination (*Pae*; Hubbs and Hibbard, 1951; Paloumpis, 1963): The emargination is actually the concave, ventral surface of the anterior process. This is not an emargination *per se*, as the whole structure follows a curve trajectory downwards; consequently, this term may even be ignored. Brousseau (1976) called this structure a cotyle and described its spatial relationships to the scapula during spine articulation and locking; however, this structure should not be considered as a cotyle as it is not a concave articular facet. This lends further support to the alternative of removing this term from the standard terminology and is herein discussed just for sake of completeness. Synonyms: Lateral cotyle (Brousseau, 1976).

Articular facet of the anterior process (*Pafap*; new name): None of the previous terms available in the literature (all with the same cost) described properly the nature of this structure; also, this new term is consistent with them *Pap* as it is its articular facet. Synonyms: 1 (Merriam in Hubbs and Hibbard, 1951); β (Sörensen in Hubbs and Hibbard, 1951); Proximal articulating surface (Hubbs and Hibbard, 1951).

Articular facet of the posterior process (*Pafpp*, new name): This articular facet never received a name in the literature examined and is consequently herein named. Synonyms: None.

Posterior process (*Ppp*; Hubbs and Hibbard, 1951): This process is involved in the posterior surface of the articular facet for the scapulocoracoid (*Pafs*), as well as in the delimitation of the posterior keel (*Ppk*), that forms the insertion surface for the *m. adductor spinalis* (see below). The posterior process is almost absent in *Diplomystes camposensis*, *Plotosus lineatus*, *Iheringichthys labrosus*, and *Chrysichthys auratus*; it is completely absent in *Wertheimeria maculata*. Synonyms: Dorsomedian process (Diogo et al., 2006a); Proximal process (Otero et al., 2009).

Posterior keel (*Ppk*, new name): This bony vertical projection onto the posterior surface of the pectoral-spine base forms the insertion surface for the *m. adductor spinalis*, which

may insert onto the keel, or lateral to it, onto the concave surface formed lateral to this structure (*e.g.,* Miano et al., 2013, fig. 6). Two terms have been applied to this structure in the literature: Central articulating surface, and posterior fossa. However, the presence of a concavity (*i.e.*, a fossa) is variably present regardless of the presence of a bony keel, and therefore can not reliably describe this character. On the other hand, this surface is not articulating but instead may be involved in the insertion of a muscle, and therefore treating it as such surface is erroneous. The best course of action then is to create a new term for this structure. Synonyms: Posterior fossa (Hubbs and Hibbard, 1951); Central articulating surface (Paloumpis, 1963).

Trochanter of the ventral process (*Ptvp*, new name): Bony round projection on the ventral surface of the ventral process; it might be present or absent depending on the degree of development of the articular facet of the ventral process (*Pafvp*), that defines the its presence on the ventral surface of the process. Two terms have already been proposed for this structure, and the use of both is discouraged as they are misleading or ambiguous, either given that they depend on the position of another structure (a muscle insertion), or because of ambiguity (distal lobe) since the structure is said to be in a position that is not distal. Synonyms: Medial flange for insertion of superficialis abductor muscle (Brousseau, 1976); Distal lobe (Hubbs and Hibbard, 1951).

Articular facet of the ventral process (*Pafvp*, new name): This articular facet never received a name in the literature examined and is consequently herein named. Synonyms: None.

Ventral process (*Pvp*; Divay and Murray, 2015; Hubbs and Hibbard, 1951; Kaatz et al., 2010; Miano et al., 2013): This name goes in coordination with other proximal processes that occupy anterior, posterior, and dorsal positions; accordingly, this process should be termed ventral. Also, this name is already the one with the lowest cost from among a series of names that have been applied to the same structure in the literature. Synonyms: Ventrolateral process (Brousseau, 1976; Pinton et al., 2011); Distoventral process (Pinton and Otero, 2010); *Processus ventro-latéral* (Gayet and van Neer, 1990); Ventral condyle (Diogo et al., 2001); Posteroventral process (Parmentier et al., 2010); Ventromedial process (Royero, 1999); Ventral articular process (Vanscoy et al., 2015).

Ventral fossa (*Pvf*; Hubbs and Hibbard, 1951): This term was chosen based on the coordination with the anterior fossa already proposed by Hubbs and Hibbard (1951), and also because similar concavities have been called fossae in the present system. Synonyms: Depression for the arrector ventralis muscle (Pinton et al., 2011).

Basal process (*Pbp*; Hubbs and Hibbard, 1951): This process is housed inside the basal recess (*Pbr*) and articulates with the distal radials as described by Brousseau (1976). This feature has been only named once among the references examined, despite being a prominent feature of the pectoral spine base. It is absent in *Plotosus lineatus* Synonyms: None.

Basal recess (*Pbr*; Hubbs and Hibbard, 1951; Paloumpis, 1963; Vanscoy et al., 2015): The spine base is hollow in posterior view and such cavity connects to the lumen of the spine

shaft. Brousseau (1976, p. 100) indicates that "the most lateral of the four cartilaginous distal radials (RDdl) articulates by its proximal depression with the scapular process (SCps). It has a lateral process which enters the opening of the nutritive canal of the spine". This cavity can be massive relative to the shaft thickness such as in species of *Brachyplatystoma*. Synonyms: Foramen (Royero, 1999); *Fosse interne* (Gayet and van Neer, 1990); *Hendidura basal* (Bisbal and Gomez, 1986); Inner fossa (Otero et al., 2009; Pinton and Otero, 2010); Inner hole (Pinton et al., 2006); Nutritive canal (Brousseau, 1976); Posterior basal recess (Divay and Murray, 2015); Proximal foramen (Bennett, 1979; Cione et al., 2005).

Posterior sulcus (*Pps*; modified from Cione et al., 2005): Although the term "posterior groove" has the lowest cost from among alternatives, the word groove was changed to sulcus for sake of coordination with the same structure in the dorsal spine that was already named "posterior sulcus". Synonyms: Furrow (Divay and Murray, 2015); Posterior groove (Hubbs and Hibbard, 1951; Lundberg, 1975; Paloumpis, 1963).

Dorsal hemitrichium (*Pdh*): Dorsal unit composing the spine when fusing to the ventral hemitrichium. This structure should not be named otherwise as has already happened in the literature. Synonyms: *Mur supéro-interne* (Gayet and van Neer, 1990); *Rama dorsal* (Bisbal and Gomez, 1986).

Ventral hemitrichium (*Pvh*): Ventral unit composing the spine when fusing to the dorsal hemitrichium. This structure should not be named otherwise as has already happened in the literature. *Mur inféro-interne* (Gayet and van Neer, 1990); *Rama ventral* (Bisbal and Gomez, 1986).

Topological terms

Orientation terms such as antrorse and retrorse are of special importance since both are extensively used in the literature for spines and its ornaments. The proper meaning of the term is below clarified given confusion in its past use, and a further refinement proposed given the fact that orientation and curvature (herein defined as concavity) are not necessarily associated. It is proposed to describe the concavity instead as a complement to orientation.

Ornament orientation: Ornaments are said to be retrorse whenever their body is deflected towards the base of the spine; on the contrary, they are called antrorse whenever deflected towards the tip of the spine (Figure 2.5A). Mees (1974) used this term for ornaments directed in the opposite direction as those herein described, i.e., for ornamentations directed towards the tip of the spine; in the same sense, he used "directed outwards" to describe the same condition. Given that several other ichthyological references both in paleontology and neontology use the term retrorse for ornamentations curved towards the base of the spine (Boulenger, 1900; Eigenmann and Allen, 1942; Gayet and Meunier, 1998; Gayet and van Neer, 1990; Jordan, 1880; Lundberg, 1997; Pinton et al., 2006), the historical meaning of the term is adopted herein *contra* the concept of Mees and references following his work (e.g., Bisbal and Gomez, 1986). The oldest mention of ornament orientation recorded in the

present work is found in Linnaeus (1758, p.304, 10th edition). The author diagnosed his genus *Silurus* (approximately with the same composition of the modern order Siluriformes) inside his division *Pisces abdominales*, among others, with the character "*...Radius pinnarum pectoralium aut dorsalis primus spinosus, retrodentatus.*" In this sentence it is clear that the word *retrodentatus* is a compound adjective with the prefix *retro-* and the suffix *-dentatus*, literally, with backward dentations. The whole sentence reads "First radius of pectoral or dorsal fins pungent (= spiny), with backward dentations". In very special cases, the ornament can be dorsally- or ventrally-deflected, what should be referred to as oblique.

Orientation can be objectively defined as a relationship between angles (π for proximal; δ for distal) between the shaft margin and the main ornament axis; if $\pi < \delta$ the ornament is said to be retrorse because the main axis of the ornament is deflected towards the proximal region of the spine, while if $\pi > \delta$ the ornament is said to be antrorse since it is deflected towards the distal region of the spine; whenever $\pi = \delta$ it is said to be straight (Figure 2.5B).

Ornament concavity: Describes the margin of the ornament that shows concavity; it makes reference to either the proximal or distal margin (Figure 2.5C,D).

Shaft surface ornaments, other than anterior and posterior structures

Spine surfaces range from almost smooth to very elaborate ornaments, with ridges of different textures and spatial patterns. These ornaments represent further formation of bone and calcium deposition on the shaft surface. Although this process does not receive a specific name in the ichthyological literature, Trueb (1973) used the name exostosis for the same process in anurans, that present a similar sculpturing on exposed bone surfaces, particularly on the cranium. Below some terms are defined for describing the surface ornamentation on lateral surfaces of the dorsal spine, and dorsal and ventral surfaces of pectoral spines.

Smooth: No exostosis or ornament on shaft surface (Figure 2.6A,B). Spine surfaces can present some minor sculpturing in the form of longitudinal subtle sulci that make the impression of keel ornaments being; however, these instances are not elevated calcium depositions and are therefore considered as lacking ornament. Smooth shaft surfaces (at least on the proximal half of the shaft) are found on the dorsal- and pectoral-fin shafts in taxa such as *Chiloglanis swierstrai*, *Diplomystes camposensis*, *Gagata cenia*, *Pangasius micronema*, *Phyllonemus filinemus*, *Plotosus lineatus*, *Schilbe intermedius*, and *Wallago leersi*. Both *Gagata cenia* and *Plotosus lineatus* present growth lines at least on its distal one-third as described below, and therefore their spines are not completely smooth.

Ridge: Longitudinal, semiparallel to anastomosing or even reticulate bony rims. Their orientation is frequently parallel to the main spine axis. Two morphologies are frequently seen: *Sharp ridges* that present an angular and sharp midline; and *smooth ridges*, that present a soft surface devoid of an angular midline. These two morphologies tend to be present in coordination across dorsal- and pectoral-fin shafts, that is whenever smooth ridges are present on the lateral surfaces of the dorsal-fin shaft, they are also smooth on the dorsal and ventral

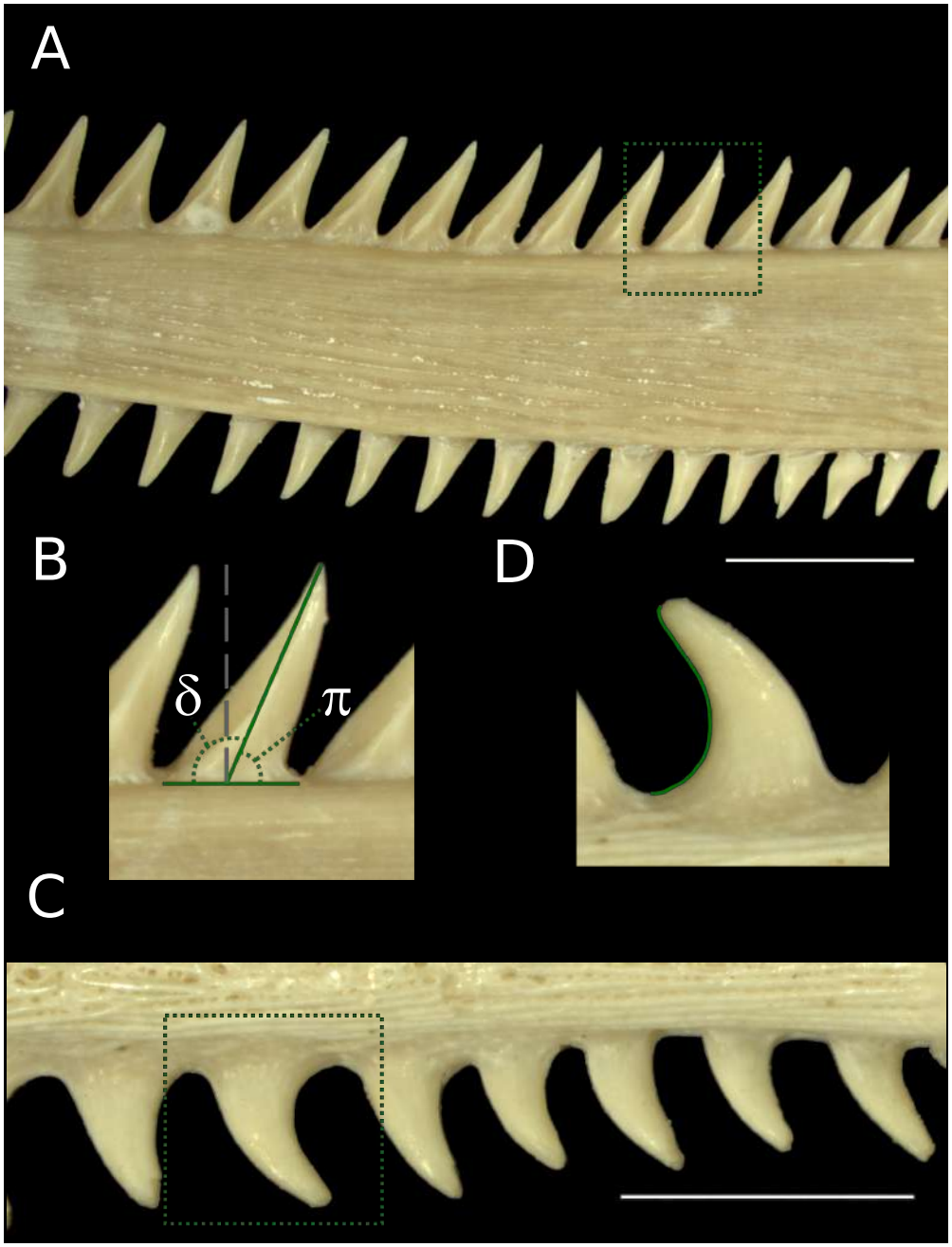


Figure 2.5: *A) Right pectoral-fin spine in dorsal view showing retrorse (posterior margin, upwards) and antrorse (anterior margin, downwards) ornaments. B) Detail of A (stippled green box) showing the definition of ornament orientation as a relationship between the proximal (π) and the distal (δ) angles with respect to the main axis of the ornament. C) Left pectoral-fin spine in dorsal view presenting concavity as a complementary property of ornaments that makes reference to the proximal or distal margin of if. D) Detail of C (stippled green box) indicating concavity on the proximal margin of the ornament which is highlighted in green outline. Scale bar in A represents 5mm, and 2mm in C; A-B* Pterodoras granulosus*, MZUSP 91655; C-D* Bunocephalus coracoideus*, MZUSP 103254.*

surfaces of the pectoral-fin shaft, as seen in *Ageneiosus pardalis*, *Bagarius* cf. *yarrelli*, *Calophysus macropterus*, and *Chrysichthys auratus*; on the contrary, sharp ridges on dorsal- and pectoral-fin shafts are present in *Anadoras grypus*, *Centrochir crocodili*, *Hypodoras forficulatus*, *Megalodoras uranoscopus*, *Rhinodoras thomersoni*, and *Scorpiodoras heckelii*. In most species the ornament on the dorsal surface of shaft is more developed than on the ventral surface (e.g., *Bagarius* cf. *yarrelli*, *Centrochir crocodili*, *Cephalosilurus apurensis*, *Hemisorubim platyrhynchos*, *Liosomadoras morrowi*, *Lophiosilurus alexandri*, *Pangasius macronema*, and *Zungaro zungaro*), whereas in few other the opposite is observed (e.g., *Acanthodoras spinosissimus*, and *Liosomadoras oncinus*). Equal degree of development of dorsal and ventral ornaments is present in several species (e.g., *Tetranematichthys quadrifilis*, *Agamyxis albomaculatus*, *Pimelodella chagresi*, *Leiarius perruno*, and *Pseudopimelodus* cf. *raninus*).

Alveolus/odontodes: Alveoli are generative cavities that produce odontodes onto bony surfaces (Bhatti, 1938). Their presence is restricted to spines of the suborder Loricarioidei (Figure 2.7A); however, its distribution vary across the families of the suborder. For instance, odontodes are restricted to the ventrolateral surface of the distal portion of the pectoral-fin shaft in *Nematogenys inermis* (Figure 2.7B), whereas small, scattered odontodes cover the spine in species of the genus *Astroblepus* (Schaefer and Buitrago-Suárez, 2002). Loricariids show more complex patterns with midline odontodes arranged in a hypertrophied row, while larger and flatter odontodes are restricted to the ventral surface of the shaft; sometimes they are also subject to sexual dimorphism and hypertrophy in males on the distal portion of the shaft (Rapp Py-Daniel and Cox-Fernandes, 2005).

Tubercle, spinule: In some cases the ridges can bear either tubercles or spinules on them, and these can be further described as blunt if they are pointed or smooth when they lack a pointed tip. As already defined for anterior and posterior shaft ornaments, tubercles are wider than tall, while spinules are taller than wider (Figure 2.6C). Tubercles on the ridges seem to be restricted to the Auchenipteridae and the Ariidae (e.g., *Trachycorystes trachycorystes*, *Liosomadoras oncinus*, *Trachelyopterichthys anduzei*, *Arius* spp.), while spinules on the shaft surface seems to be restricted to a couple genera in the Doradidae (see below).

Birindelli (2014, p. 525, character 260) mentions "serrations on the dorsal face of the pectoral-fin spine" as a character present in a few auchenipterids (*Spinipterus acsi*, *Trachelyopterichthys taeniatus*, *Tr. porosus*, *Tr. striatulus*, following Akama, 2004) and doradids (the genera *Acanthodoras* and *Agamyxis*, following Souza, 2010) defining its character states as absent (0), present on anterior portion (1), or widespread (2). Following the proposed terminology it should be described as tubercles on the dorsal surface of the pectoral-spine shaft provided that all these taxa present the same condition; however, direct examination shows that this is not the case. Given the distinction between tubercles and spinules herein where the former is never as high as wide, the dorsal shaft ornament in *Acanthodoras* and *Agamyxis* should be described as spinules, whereas species of *Trachelyopterichthys* presents tubercles. Differences in spatial arrangement of such ornaments are noteworthy, for instance, *Tr. striatulus* almost lacks such ornaments, and when present, these are restricted to the basal half

of the anterodorsal margin of the dorsal-spine shaft; on the contrary, *Tr. porosus* and *Tr. taeniatus* present stronger tubercles along the whole anterodorsal margin. *Spinipterus acsi* is a very rare auchenipterid only known from a couple collections in Peru and Brazil and could not be examined in the present study; however, Calegari et al. (2018) present detailed illustrations of the spines based on CT-scans (figs. 3-4), and these appear to be still tubercles, albeit large ones close to the threshold definition of spinules where ornament height is larger than width at the base. A better understanding of the proper term to be applied to this condition awaits direct examination and larger samples.

Growth lines: Sometimes the distal addition of segments to the spine leaves marks that persist throughout growth; this process can leave evident sculpturing on the lateral surfaces of the dorsal-fin spine or the dorsal and ventral surfaces of the pectoral-fin spine. These structures have been demonstrated to be formed through ontogenetic distal segment addition in pectoral-fin spines by Kubicek et al. (2019), who also indicates that sometimes the concave spaces between growth lines can house venom glands (e.g., *Noturus gyrinus* and *Akysis vespa*, figure 3). Prominent growth lines are present in the dorsal and pectoral spines of *Plotosus lineatus*, *Gagata cenia*, and at least in the pectoral-fin spine of *Cranoglanis bouderius* and the ictalurids *Ameiurus nebulosus* (Reed, 1924, fig. 13), *Ictalurus sawrockensis* (Bennett, 1979, fig. 7), *I. vespertinus*, *I. lavetti* (Lundberg, 1975, pl. vii, fig. 4), *Noturus flavus* (Arce-H et al., 2017, fig. 6C), and *N. miurus* (Reed, 1924, fig. 12). Given the distribution of this feature, it might be synapomorphic at some level for the Ictaluroidea and related families (e.g., Plotosidae).

Shaft ornaments on the anterior and posterior surfaces

Dorsal and pectoral spines show anterior and posterior ornaments that have been interpreted as defensive, pungent structures, sometimes associated to venom glands (Wright, 2009). Despite a few terms have been used in order to describe these ornaments, they have been used in a very inconsistent way. During the present study it was found that ornament diversity is higher than the amount of terms used to describe it, but at the same time its application has been highly inconsistent (Appendix A; Table A.2). The following terms are proposed for a finer description of anterior and posterior ornaments in pectoral and dorsal spines.

Spinule: Sharp to subsharp elongate ornament with transverse round outline, lacking cutting edges. Also, these ornaments can be flattened with oval transverse outline and are consequently named flat spinules (Figure 2.8A,B). Spinules have been called *denticulus* (plural *denticuli*) in *Ictalurus punctatus* by Kubicek et al. (2019, fig. 4). Also, flat spinules have been called serra (plural serrae) on the posterior pectoral-fin spine ornament for the same species.

Blade: These ornaments present an acute to round apex, in addition to one or more cutting edges. The orientation of the cutting edges should be also described in a cross section

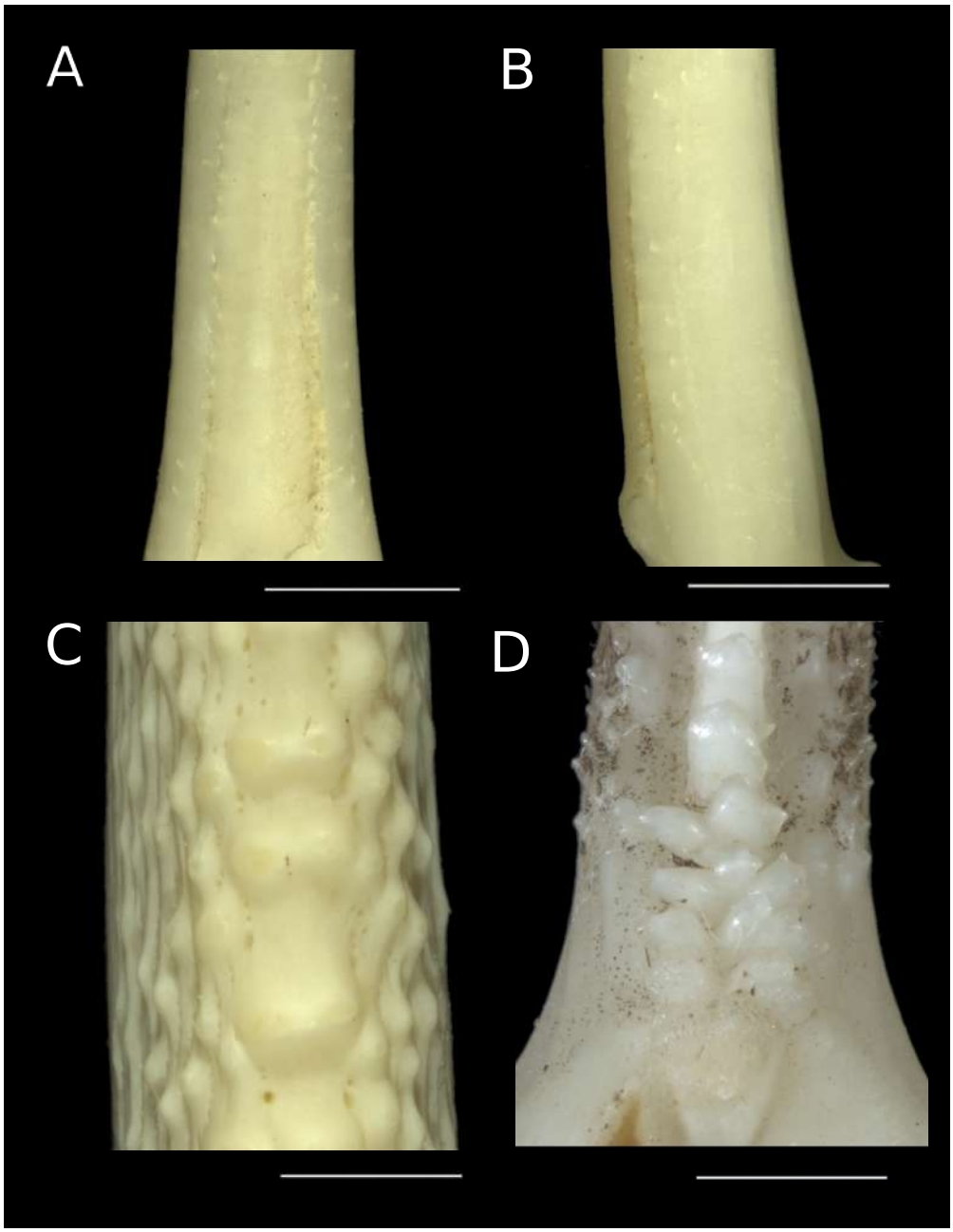


Figure 2.6: *Dorsal-fin spine of* Diplomystes camposensis *(MZUSP 88533) in A) anterior view, and B) lateral view; note the lack of ornament on the spine shaft surface. C) Tubercles in the dorsal fin of* Arius proops *(MZUSP 52842) both onto the shaft surface and the anterior margin. D) Multicuspid ornaments in the proximal region of the dorsal-fin spine in* Liosomadoras oncinus *(MZUSP 105828); they become gradually unicuspid towards the distal region of the spine.*

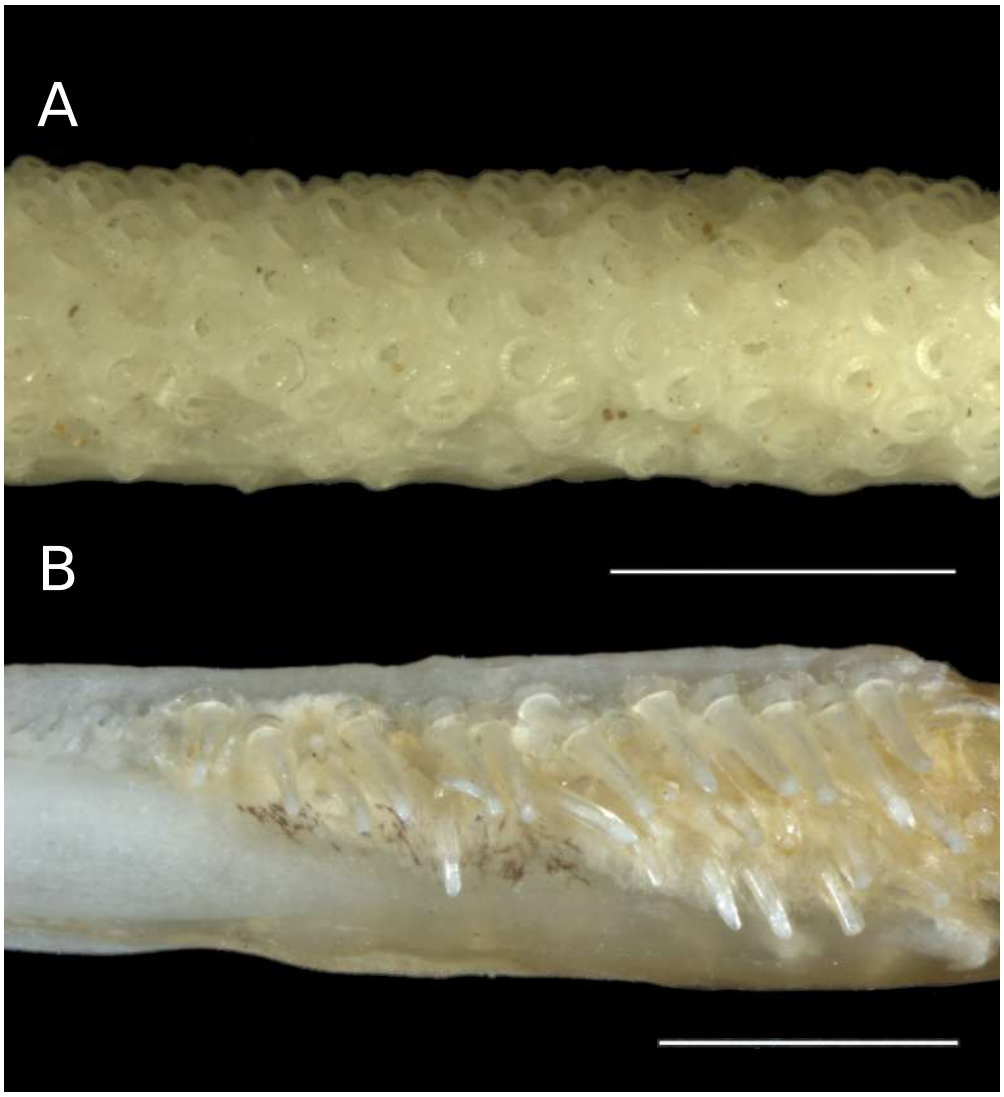


Figure 2.7: *Odontodes and alveoli in Loricarioid spines. A) Abundant odontode alveoli on the ventral surface of the pectoral-fin spine in* Hypostomus *cf.* wachereri *(MZUSP 87480). B) Odontodes on the ventral surface of the pectoral-fin spine in* Nematogenys inermis *(MZUSP 75256). All scales equal 2mm*

of the individual ornament (Figure 2.8C,D). The most common configuration of cutting edges is to show lateral ones in dorsal-fin blades, or dorsal and ventral ones in pectoral-fin blades. Given that dorsal ornaments have lateral, proximal and distal surfaces, their cutting edges should be named according to their position on these surfaces, consequently the pectoral ornaments are named dorsal, ventral, proximal, and distal. In special cases the cutting edges will not be located on the aforementioned surfaces but instead in oblique configurations, in which case they can be called with a combination of the surfaces that surround it as in dorso-distal or ventro-proximal. *Goeldiella eques*, *Chrysichthys auratus*, and *Synodontis schall* present blades on the posterior surface of the pectoral shaft with disto-dorsal and disto-ventral cutting edges; *Diplomystes camposensis* shows dorsal and ventral cutting edges on the posterior blades of the pectoral spine.

Tubercle: Ossifications that are wider than taller, usually small (Figure 2.6C). Tubercles can be smooth or sharp as in spinules. In contrast, tubercles do not tend to be flattened. Vanscoy et al. (2015, fig. 5) used the name dentations for tubercles on the anterior surface of the pectoral-fin spine of *Platynematichthys notatus*. In some instances, ontogenetic transformation from spinules to tubercles can be seen (e.g., anterior ornament of the pectoral spine in *Tocantinsia piresi*), probably through a higher calcium deposition rate on the sides of the ornament rather than longitudinally. Kubicek et al. (2019) illustrate the pectoral-fin spine of *Ictalurus punctatus* where spinules are clearly seen on the anterior shaft surface (fig. 5h); however tubercles can be seen in earlier stages of ontogenetic development (fig. 5f-g, also see fig. 4d in Vanscoy et al., 2015), what further reinforces the relationship between these two kinds of morphology through spine development. Contrary to these ontogenetic instances, some species show tubercles as the condition in the adults.

Multicuspidate complex: Most ornaments present a single tip regardless of its orientation; however, some others have complicated morphologies that present more than one tip (Figure 2.6D). So far these cases seem restricted to spinules, flat spinules, and tubercles, while instances of multicuspidate blades are unknown. The spatial arrangement and number of cusps can be further described with respect to the orientation plane. For instance, the proximal ornaments of the dorsal spine in *Liosomadoras oncinus* can be described as a tricuspid set of spinules with horizontally orientation.

Ornament fusions: Continuous structures consisting of fussions of any kind kind of individual ornaments, as in some pimelodids such as *Pseudoplatystoma magdaleniatum*. These instances differ from multicuspidate complexes in that they are formed through fusion of individual ornaments in the form of a bony ridge instead of a unit with multiple cusps.

Odontode: Dermal structures with enameloid as described by Bhatti (1938) (Figure 2.7B). A longitudinal anterior band of odontodes is present in some genera of the Callichthyidae (e.g., *Callichthys*) and the position and extend of odontode bands can be of phylogenetic relevance (G.A. Ballen, pers. obs.).

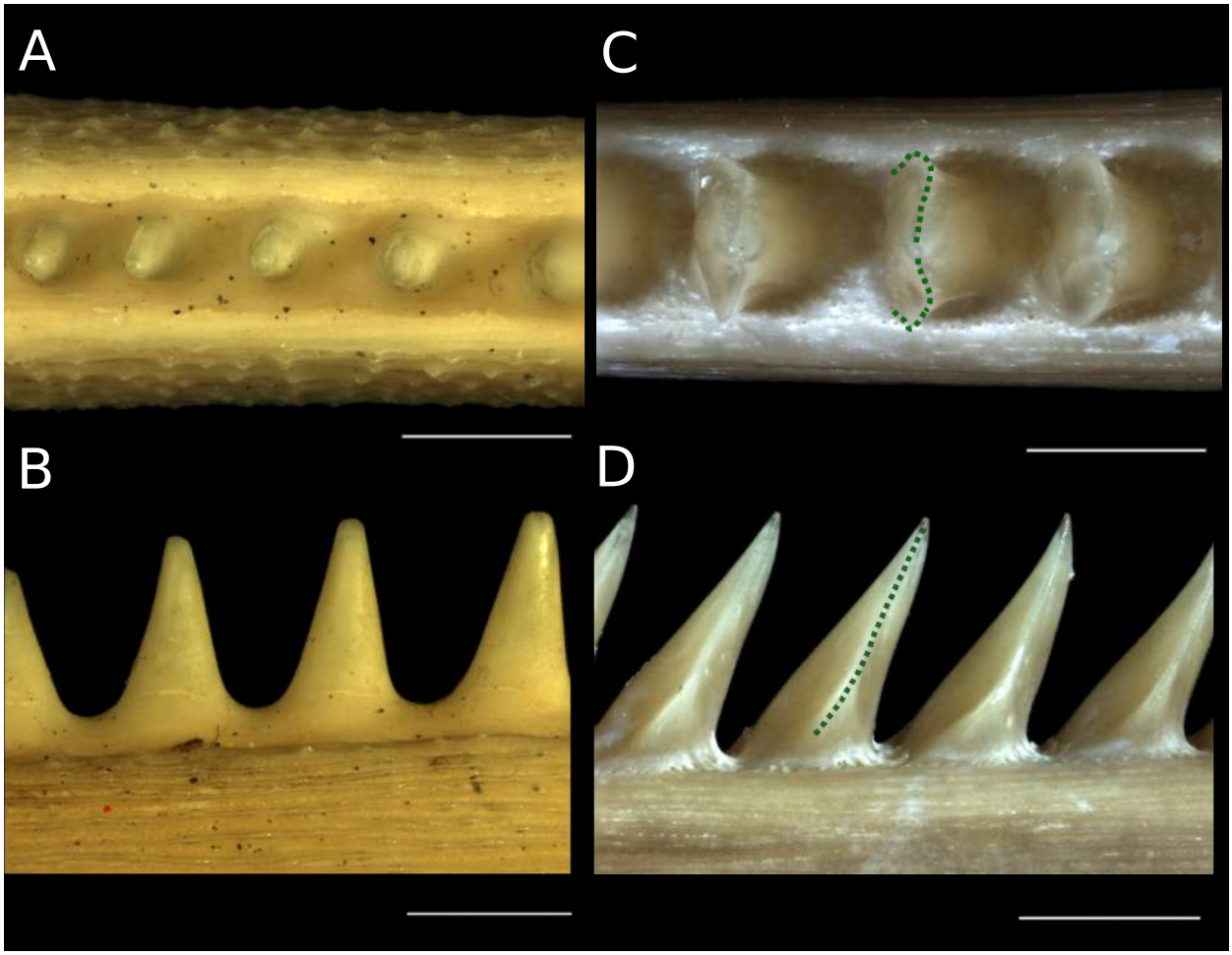


Figure 2.8: *Spine ornaments in the Siluriformes. Spinules in A) anterior view and B) dorsal view, pectoral-fin spine of* Trachycorystes *sp. (MZUSP 91659); spinules do not present cutting edges and are oval to round in cross section. Blades in C) posterior view and D) dorsal view, pectoral-fin spine of* Pterodoras granulosus *(MZUSP 91655); in contrast to spinules, blades present one or more cutting edges (stippled green lines in C and D) and their are tear-drop-shaped or have polygonal outline in cross section due to the cutting edges. All scale bars equal 2mm.*

2.5 Discussion

The present terminology system provides a standard for future descriptions and comparative studies dealing with Siluriform spines. One of the gains of its adoption is the possibility of carrying out comparisons across studies on both extinct and extant Siluriform taxa, as well as to enable the exploration of spines as a source of new morphological characters with taxonomic and systematic information. Although it was possible to produce synonym lists for characters associated to the spine base so that the same term can be traced across studies, this was not the case for most of the shaft ornament terms, where its usage and poor quality of illustrations precluded the identification of most conditions in the vast amount of literature already reviewed in the present work. The present system can be applied to taxonomic works that will further encourage researchers to assess the informativeness of spine characters in other contexts by using species descriptions and revisionary works as a baseline.

Siluriform fossil remains are common in freshwater and shallow marine environments, often found in coarse-grain sedimentary setting where they can accumulate and resist high-energy conditions. Several sources report spine remains in the fossil record; unfortunately they are often identified only to order-level as "Siluriformes indet." (e.g., Alveş et al., 2016; Antoine et al., 2007; Stirton, 1953) or in better cases to family level (e.g., Aguilera et al., 2013). This general level of taxonomic uncertainty is in contrast to the vast species richness and morphological diversity of the order. Siluriforms as a whole show more variation and at finer levels than usually acknowledged as can be seen in the variation accounts of several terms herein identified or newly proposed. The present contribution has the potential of triggering the re-study of large amounts of specimens already available in collections in order to refine their taxonomic identity, but also to direct further collection efforts to this neglected anatomical component in future field prospections. As with any anatomical complex, Siluriform spines should neither be expected to showing species-specific variation in all cases nor as being useless for documenting Siluriform occurrences because of lack of variation but instead as a useful source of information that can offer fairly detailed taxonomic identities when in conjunction with detailed study of the extant diversity. Although Siluriform occurrences has been reported back to the Turonian-early Maastrichthian (Alveş et al., 2016), morphological conservativeness is well documented so that several morphologies are comparable across several million years (Lundberg et al., 2010). For instance, occurrences assigned to the extant superfamily Doradoidea are reported as ranging from the late Holocene in Argentina (Loponte et al., 2012) to the early Paleocene in Bolivia (Gayet and Meunier, 1998). Unfortunately, most studies focusing on pre-Miocene occurrences tend to restrict comparisons with other known pre-Miocene fossil and excluding extant groups. Any work in paleontology of Siluriforms should therefore involve a generous amount of neontological work in order to better provide insights into diversification of this group in the past; here is where the current terminology system enters as the standard for comparison among fossil and extant

taxa so that more robust identifications and phylogenetic positions for fossil occurrences can be achieved. This is specially relevant for current research programs such as divergence-time estimation that lean strongly on information from the fossil occurrences and their phylogenetic position in order to put phylogenies in a temporal context (Gavryushkina et al., 2017; Heath et al., 2014; Warnock et al., 2014).

Dorsal- and pectoral-fin spines have been found associated to archaeological sites, sometimes allowing detailed reconstruction of the dietary aspects as well as a reasonable understanding of ancient fisheries (Acosta et al., 2007; Jiménez-Cano and Masson, 2016; Peña-León, 2011; Prendergast and Lane, 2010; Stewart and Cowie, 2007; Trapani, 2008; Usha et al., 2004) This demonstrates the value of pectoral-fin remains as a source of information on the interaction between humans and Siluriforms in the past. The present system can aid in the efforts of researchers in archaeoichthyology in at least two ways: 1) Allowing more detailed description of morphological diversity that can be directly used for identification purposes (e.g., Pinton and Otero, 2010), and 2) as a source of characters subject to ontogenetic variation, so that they allow to reconstruct age-dependent dynamics, specially in the context of fisheries (e.g., Jiménez-Cano and Masson, 2016; Peña-León, 2011). Although the present contribution does not render unnecessary the primary examination of preserved material, it does provide the precision needed in order to provide more accurate morphological descriptions that can be compared across studies, allowing meaningful comparisons among studies dealing with a commons set of Siluriform occurrences, and also serving as a source for spotting incorrect taxonomic identifications.

Although some authors have already recognized the value of spine characters as informative in phylogenetics (Arce-H et al., 2017), it is expected that the present work will allow to identify further instances of synapomorphies in this anatomical complex. Among the most striking instances of spine-related characters are the reticulate shaft surface ornament in the genus *Phractocephalus* that is unique among Siluriforms (Aguilera et al., 2008, ; Montes et al. *in prep.*). The rough knob on the inflection point of the anterior longitudinal ridge present in *Iheringichthys*, *Pimelodus*, *"Pimelodus" ornatus*, and *Bergiaria* is a possible synapomorphy for the *Pimelodus* group as defined by Lundberg et al. (2011). The lack of anterior articular surface of the dorsal spine might prove to be a synapomorphy for the genus *Pseudoplatystoma*, although this feature is also absent in *Calophysus macropterus* and *Tetranematichthys quadrifilis*, therefore it might prove to be homoplastic; the same applies to the hypertrophied median process of the pectoral spine in the Doradoidea, the Aspredinidae and the Mochockidae, that according to our current understanding of interrelationships (Sullivan et al., 2006) should be recovered as homoplastic. Future and more detailed study of these apparently-homoplastic instances might reveal superficially similar conditions having different evolutionary origins. The dorsal ornament of the pectoral-fin shaft in both *Acanthodoras spinosissimus* and *Agamyxis albomaculatus* consists of spinules, but these are larger and less numerous in the former and more abundant although lower in the latter. As currently defined, the spinules on the dorsal shaft of the pectoral-fin shaft seem to be a putative

synapomorphy of these two doradid genera, and suggest that the character 260 in Birindelli (2014) should be redefined and split in two different transformation series, each supporting a group in the Auchenipteridae and the Doradidae respectively. These are but a few instances of features that require more detailed study and an assessment in total-evidence settings in order to be tested as phylogenetically-informative.

The present work should not be intended as a comprehensive survey of variation in characters from the dorsal- and pectoral-fin spines of Siluriforms but instead as a starting point for future and more detailed reassessments of both its variation and meaning in several contexts. We have pointed to several promising variations that may prove to be synapomorphies or diagnostic characters at more- or less-inclusive levels within the order, but the definite tests await specific and richer sampling schemes inside families and genera of interest. What this work hopefully provides is the standardization of terminology so that meaningful comparisons can be carried out across groups of taxa under a common framework. Although no standard is by itself mandatory in zoology, we hope that the effort to compile more than a century of literature and a large body of terms in several languages can be of help in avoiding the current chaotic situation in which every author uses its own terminology, that is often incompatible with other authors' schemes.

Bibliography

Acosta, A., Loponte, D., and Musali, J. 2007. A taphonomic approach to the ichthyoarchaeological assemblage of La Bellaca Site 2, wetland of the lower Parana River, Pampean Region (Argentina). In M. Gutierrez, G. Barrientos, G. Mengoni Goñalons, L. Miotti, and M. Salamme (eds.), *Taphonomy and Zooarchaeology in Argentina*, pp. 71–88. British Archaeological Reports, International Series, Oxford, UK. 39

Aguilera, O.A., Bocquentin, J., Lundberg, J.G., and Maciente, A. 2008. A new cajaro catfish (Siluriformes: Pimelodidae: *Phractocephalus*) from the Late Miocene of southwestern Amazonia and its relationship to *Phractocephalus nassi* of the Urumaco Formation. *Palaeontologische Zeitschrift*, 82:231–245. 15, 39

Aguilera, O.A., Lundberg, J.G., Birindelli, J.L.O., Sabaj Pérez, M.H., Jaramillo, C.A., and Sánchez-Villagra, M.R. 2013. Palaeontological evidence for the last temporal occurrence of the ancient Western Amazonian river outflow into the Caribbean. *PLoS ONE*, 8:1–17. 38

Alexander, R.M. 1965. Structure and function in the catfish. *Journal of Zoology*, 148:88–152. 14, 18, 20, 22, 23, 24, 26

Alveş, Y.M., Bergqvist, L.P., and Brito, P.M. 2016. New occurrences of microvertebrate fossil accumulations in Bauru Group, Late Cretaceous of western São Paulo state, Brazil. *Journal of South American Earth Sciences*, 69:80–90. 38

Antoine, P.O., Salas-Gismondi, R., Baby, P., Benammi, M., Brusset, S., Franceschi, D.D., Espurt, N., Goillot, C., Pujos, F., Tejada, J., and Urbina, M. 2007. The middle Miocene (Laventan) Fitzcarrald fauna, Amazonian Peru. *Cuadernos del Museo Geominero*, 8:19–24. 38

Arce-H, M., Lundberg, J.G., and O'Leary, M.A. 2017. Phylogeny of the North American catfish family Ictaluridae (Teleostei: Siluriformes) combining morphology, genes and fossils. *Cladistics*, 33:406–428. 14, 33, 39

Argyriou, T., Cook, T.D., Muftah, A.M., Pavlakis, P., Boaz, N.T., and Murray, A.M. 2015. A fish assemblage from an early Miocene horizon from Jabal Zaltan, Libya. *Journal of African Earth Sciences*, 102:86–101. 22, 26, 27

Arratia, G., Kapoor, B.G., Chardon, M., and Diogo, R. 2003. *Catfishes*, volume 2 vol. Science Publishers, Inc. 14

Bennett, D.K. 1979. Late Cenozoic fish faunas from Nebraska. *Transactions of the Kansas Academy of Science*, 82:146–177. 29, 33

Bhatti, H.K. 1938. The integument and dermal skeleton of Siluroidea. *The Transactions of the Zoological Society of London*, 24:1–102. 32, 36

Birindelli, J.L.O. 2014. Phylogenetic relationships of the South American Doradoidea (Ostariophysi: Siluriformes). *Neotropical Ichthyology*, 12:451–564. 32, 40

Bisbal, G.A. and Gomez, S.E. 1986. Morfolofgía comparada de la espina pectoral de algunos Siluriformes Bonaerenses (Argentina). *Physis*, 44:81–93. 15, 29

Bockmann, F.A. and Slobodian, V. 2014. Heptapteridae. In L.J.D. Queiroz, G. Torrente-Vilara, W.M. Ohara, T.H.d.S. Pires, J. Zuanon, and C.R.d.C. Doria (eds.), *Peixes do Rio Madeira. Volume III*, pp. 15–77. Santo Antonio Energia. 14

Bosher, B.B.T., Newton, S.H.S., and Fine, M.L.M. 2006. The spines of the Channel Catfish, *Ictalurus punctatus*, as an anti-predator adaptation: an experimental study. *Ethology*, 112:188–195. 14, 19

Boulenger, G.A. 1900. On some little-known African Silurid fishes of the subfamily Doradinae. *Annals and Magazine of Natural History*, 6:520–529. 29

Brousseau, R.A. 1976. The pectoral anatomy of selected Ostariophysi II. The Cypriniformes and Siluriformes. *Journal of Morphology*, 150:79–116. 14, 26, 27, 28, 29

Calegari, B.B., Akama, A., and Reis, R.E. 2018. First record of the Driftwood Catfish *Spinipterus acsi* Akama & Ferraris, 2011 (Siluriformes, Auchenipteridae) for Brazil, Juruá River, Amazon basin. *Check List*, 14:693–697. 33

Cione, A.L., Azpelicueta, M.d.l.M., Casciotta, J.R., and Dozo, M.T. 2005. Tropical freshwater teleosts from Miocene beds of eastern Patagonia, southern Argentina. *Geobios*, 38:29–42. 23, 26, 29

Costa e Silva, T., Ribeiro, F.R.V., Lucena, C.A.S., and Lucinda, P.H.F. 2018. *Pimelodus speciosus* (Teleostei: Pimelodidae), a new catfish species from the rio Tocnantins drainage, Brazil. *Ichthyological Exploration of Freshwaters*, 28:97–106. 14

Diogo, R. 2007a. Homoplasies, consistency index and the complexity of morphological evolution: Catfishes as a case study for general discussions on phylogeny and macroevolution. *International Journal of Morphology*, 25:831–837. 26

Diogo, R. 2007b. Osteology and myology of the cephalic region and pectoral girdle of *Heptapterus mustelinus*, comparisons with other heptapterins, and discussion on the synapomorphies and phylogenetic relationships of the Heptapterinae and the Pimelodidae. *International Journal of Morphology*, 25:735–748. 26

Diogo, R., Chardon, M., and Vandewalle, P. 2003. Osteology and myology of the cephalic region and pectoral girdle of *Centromochlus heckelii*, comparison with other auchenipterids, and comments on the synapomorphies and phylogenetic relationships of the Auchenipteridae (Teleostei: Siluriformes). *Animal Biology*, 53:397–416. 26

Diogo, R., Chardon, M., and Vandewalle, P. 2004. On the osteology and myology of the cephalic region and pectoral girdle of *Chaca bankensis* Bleeker 1825, with comments on the autapomorphies and phylogenetic relationships of the Chacidae (Teleostei: Siluriformes). *Animal Biology*, 54:159–174. 26

Diogo, R., Chardon, M., and Vandewalle, P. 2006a. On the osteology and myology of the cephalic region and pectoral girdle of *Nematogenys inermis* (Ghichenot, 1848), with comments on the autapomorphies and phylogenetic relationships of the Nematogenyidae (Teleostei: Siluriformes). *Belgian Journal of Zoology*, 136:15–24. 26, 27

Diogo, R., Chardon, M., and Vandewalle, P. 2006b. Osteology and myology of the cephalic region and pectoral girdle of *Cetopsis coecutiens* Spix & Agassiz, 1829, comparison with other Cetopsids, and comments on the synapomorphies and phylogenetic position of the Cetopsidae (Teleostei: Siluriformes. *Belgian Journal of Zoology*, 136:3–13. 26

Diogo, R., Oliveira, C., and Chardon, M. 2001. On the osteology and myology of Catfish pectoral girdle, with a reflection on Catfish (Teleostei: Siluriformes) plesiomorphies. *Journal of Morphology*, 249:100 –125. 26, 28

Divay, J.D. and Murray, A.M. 2015. The late Eocene-early Oligocene ichthyofauna from the Eastend Area of the Cypress Hills Formation, Saskatchewan, Canada. *Journal of Vertebrate Paleontology*, 35:e956877. 22, 23, 26, 28, 29

Egge, J.J.D. and Simons, A.M. 2011. Evolution of venom delivery structures in Madtom Catfishes (Siluriformes: Ictaluridae). *Biological Journal of the Linnean Society*, 102:115–129. 14

Eigenmann, C.H. 1925. A review of the Doradidae, a family of South American Nematognathi, or Catfishes. *Transactions of the American Philosophical Society, New Series*, 22:279–365. 14, 20, 22, 26

Eigenmann, C.H. and Allen, W.R. 1942. *Fishes of western South America. I. The intercordilleran and Amazonian lowlands of Peru. II. The high pampas of Peru, Bolivia, and*

northern Chile. With a revision of the Peruvian Gymnotidae and of the genus Orestias. The University of Kentucky, Lexington. 14, 15, 29

Eschmeyer, W.N. and Fong, J.D. 2018. Species by Family/Subfamily. 13

Gavryushkina, A., Heath, T.A., Ksepka, D.T., Stadler, T., Welch, D., and Drummond, A.J. 2017. Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Systematic Biology*, 66:57–73. 39

Gayet, M. and Meunier, F.J. 1998. Maastrichtian to Early Late Paleocene freshwater Osteichthyes of Bolivia: Additions and comments. In L.R. Malabarba, R.E. Reis, R.P. Vari, Z.M.S. Lucena, and C.A.S. Lucena (eds.), *Phylogeny and Classification of Neotropical Fishes*, pp. 85–110. Edipucrs. 14, 15, 29, 38

Gayet, M. and van Neer, W. 1990. Caractères diagnostiques des épines de quelques silures africains. *Journal of African Zoology*, 104:241–252. 15, 22, 23, 26, 27, 28, 29

Grande, T. and Shardo, J.D. 2002. Morphology and development of the postcranial skeleton in the Channel Catfish *Ictalurus punctatus* (Ostariophysi: Siluriformes). *Fieldiana Zoology New Series*, 99. 22

Heath, T.A., Huelsenbeck, J.P., and Stadler, T. 2014. The fossilized Birth-Death Process: A coherent model of fossil calibration for divergence time estimation. *Proceedings of the National Academy of Sciences*, 111:E2957–E2966. 39

Hubbs, C.H. and Hibbard, C.W. 1951. *Ictalurus lambda*, a new Catfish, based on a pectoral spine from the lower Pliocene of Kansas. *Copeia*, 1951:8–14. 15, 24, 26, 27, 28, 29

Jiménez-Cano, N.G. and Masson, M.A. 2016. Estimation of fish size from archaeological bones of hardhead catfishes (*Ariopsis felis*): Assessing pre-Hispanic fish acquisition of two Mayan sites. *Journal of Archaeological Science: Reports*, 8:116–120. 39

Jordan, D.S. 1880. *Manual of the Vertebrates of the Northern United States, Including the District East of the Mississippi River, and North of North Carolina and Tennessee, Exclusive of Marine Species*. 3rd edition. Jansen, McClurg & Company. 29

Kaatz, I.M., Stewart, D.J., Rice, A.N., and Lobel, P.S. 2010. Differences in pectoral fin spine morphology between vocal and silent clades of catfishes (Order Siluriformes): Ecomorphological implications. *Current Zoology*, 56:73–89. 14, 24, 27, 28

Kubicek, K.M., Britz, R., and Conway, K.W. 2019. Ontogeny of the catfish pectoral-fin spine (Teleostei: Siluriformes). *Journal of Morphology*, pp. 1–21. 14, 18, 33, 36

Linnaeus, C. 1758. *Systema naturae per regna tria naturae: Secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis*. 10th edition. 30

Loponte, D., Acosta, A., and Musali, J. 2012. Allometric parameters of *Pterodoras granulosus* (Valenciennes 1833) and its application to fossil assemblages. *International Journal of Osteoarchaeology*, 22:352–360. 38

Lundberg, J.G. 1975. The fossil Catfishes of North America. *University of Michigan, Papers on Paleontology*, 11:1–51. 22, 23, 29, 33

Lundberg, J.G. 1997. Freshwater fishes and their paleobiotic implications. In R.F. Kay, R.H. Madden, R.L. Cifelli, and J.J. Flynn (eds.), *Vertebrate Paleontology in the Neotropics: The Miocene Fauna of La Venta, Colombia*, chapter 5, pp. 67–91. Smithsonian Institution Press. 15, 29

Lundberg, J.G., Sabaj Pérez, M.H., Dahdul, W.M., and Aguilera, O.A. 2010. The Amazonian Neogene fish fauna. In C. Hoorn and F.P. Wesselingh (eds.), *Amazonia, Landscape and Species Evolution: A Look Into the Past*, pp. 281–301. Blackwell Publishing. 38

Lundberg, J.G., Sullivan, J.P., and Hardman, M. 2011. Phylogenetics of the South American Catfish family Pimelodidae (Teleostei: Siluriformes) using nuclear and mitochondrial gene sequences. *Proceedings of the Academy of Natural Sciences of Philadelphia*, 161:153–189. 39

Mees, G.F. 1974. The Auchenipteridae and Pimelodidae of Suriname (Pisces, Nematognathi). *Zoologische Verhandelingen*, 132:1–246. 15, 29

Miano, J.P., Loesser-Casey, K.E., and Fine, M.L. 2013. Description and scaling of pectoral muscles in Ictalurid Catfishes. *Journal of Morphology*, 274:467–477. 14, 16, 24, 26, 27, 28

Otero, O., Likius, A., Vignaud, P., and Brunet, M. 2007. A new claroteid Catfish (Siluriformes) from the upper Miocene of Toros-Menalla, Chad: *Auchenoglanis soye*, sp.nov. *Journal of Vertebrate Paleontology*, 27:285–294. 22, 23

Otero, O., Pinton, A., Mackaye, H.T., Likius, A., Vignaud, P., and Brunet, M. 2009. First description of a Pliocene ichthyofauna from Central Africa (site KL2, Kolle area, Eastern Djurab, Chad): What do we learn? *Journal of African Earth Sciences*, 54:62–74. 24, 27, 29

Paloumpis, A.A. 1963. A key to the Illinois species of *Ictalurus* (Class Pisces) based on pectoral spines. *Transactions of te Illinois Academy of Science*, 56:129–133. 24, 27, 28, 29

Parmentier, E., Fabri, G., Kaatz, I., Decloux, N., Planes, S., and Vandewalle, P. 2010. Functional study of the pectoral spine stridulation mechanism in different mochokid catfishes. *Journal of Experimental Biology*, 213:1107–1114. 14, 19, 24, 26, 27, 28

Peña-León, G.A. 2011. Pescadores de los raudales del Río Magdalena durante el periodo Formativo Tardío. *Caldasia*, 33:295–314. 39

Pinton, A., Fara, E., and Otero, O. 2006. Spine anatomy reveals the diversity of catfish through time: a case study of *Synodontis* (Siluriformes). *Naturwissenschaften*, 93:22–26. 14, 15, 22, 24, 29

Pinton, A. and Otero, O. 2010. The bony anatomy of Chadian *Synodontis* (Osteichthyes, Teleostei, Siluriformes, Mochokidae): Interspecific variations and specific characters. *Zoosystema*, 32:173–231. 20, 22, 24, 26, 27, 28, 29, 39

Pinton, A., Otero, O., Likius, A., Mackaye, H.T., Vignaud, P., and Brunet, M. 2011. Giants in a minute Catfish genus: First description of fossil *Mochokus* (Siluriformes, Mochokidae) in the late Miocene of Chad, Including *M. gigas*, sp. nov. *Journal of Vertebrate Paleontology*, 31:22–31. 27, 28

Prendergast, M.E. and Lane, P.J. 2010. Middle Holocene fishing strategies in east Africa: Zooarchaeological analysis of Pundo, a Kansyore shell midden in northern Nyanza (Kenya). *International Journal of Osteoarchaeology*, 20:88–112. 39

R Core Development Team. 2018. R: A language and environment for statistical computing. 17

Rapp Py-Daniel, L.H. and Cox-Fernandes, C. 2005. Dimorfismo sexual em Siluriformes e Gymnotiformes (Ostariophysi) da Amazônia. *Acta Amazonica*, 35:97–110. 32

Reed, H.D. 1924. The morphology and growth of the spines of Siluroid fishes. *Journal of Morphology*, 38:431–451. 14, 18, 19, 24, 33

Royero, R. 1999. *Studies on the Systematics of the Catfish Family Auchenipteridae (Teleostei: Siluriformes)*. Ph.D. thesis, University of Bristol. 18, 26, 27, 28, 29

Sabaj Pérez, M.H. 2013. Standard Symbolic Codes for Institutional Resource Collections in Herpetology and Ichthyology - American Society of Ichthyologists and Herpetologists, Washington, DC. 16

Schaefer, S.A. and Buitrago-Suárez, U.A. 2002. Odontode morphology and skin surface features of Andean astroblepid catfishes (Siluriformes, Astroblepidae). *Journal of Morphology*, 254:139–148. 32

Shibatta, O.A. and Benine, R.C. 2005. A new species of *Microglanis* (Siluriformes: Pseudopimelodidae) from upper rio Paraná basin, Brazil. *Neotropical Ichthyology*, 3:579–585. 14

Shibatta, O.A. and Vari, R.P. 2017. A new genus of Neotropical rheophilic catfishes, with four new species (Teleostei: Siluriformes: Pseudopimelodidae). *Neotropical Ichthyology*, 15:1–30. 14

Slobodian, V., Akama, A., and Dutra, G.M. 2017. A new species of *Pimelodella* (Siluriformes: Heptapteridae) from the Guiana Shield, Brazil. *Zootaxa*, 4338:85–100. 14

Souza, L.M. 2010. *Revisão taxonômica e filogenia de Astrodoradinae (Siluriformes, Doradidae)*. Phd. thesis, Universidade de São Paulo. 14

Stewart, F.L. and Cowie, E. 2007. Dietary indications for a St. Lawrence Iroquoian site in northern New England. *Archaeology of Eastern North America*, 35:21–36. 39

Stirton, R.A. 1953. Vertebrate paleontology and continental stratigraphy in Colombia. *Geological Society of America Bulletin*, 64:603–622. 38

Sullivan, J.P., Lundberg, J.G., and Hardman, M. 2006. A phylogenetic analysis of the major groups of catfishes (Teleostei: Siluriformes) using rag1 and rag2 nuclear gene sequences. *Molecular phylogenetics and evolution*, 41:636–62. 39

Taylor, W.R. and van Dyke, G.C. 1985. Revised procedures for staining and clearing small fishes and other vertebrates for bone and cartilage study. *Cybium*, 9:107–119. 16

Trapani, J. 2008. Quaternary fossil fish from the Kibish Formation, Omo Valley, Ethiopia. *Journal of Human Evolution*, 55:521–530. 39

Trueb, L. 1973. Bones, frogs, and evolution. In J.L. Vial (ed.), *Evolutionary Biology of the Anurans: Contemporary Research on Major Problems*, pp. 65–132. University of Missouri Press, Columbia. 30

Usha, U., Ghosh, M., and Pal, T.K. 2004. Animal remains excavated from Lothal Archaeolofical Site (Gujarat) and relevance of the fauna in this Ancient Civilization. *Occasional Papers of the Zoological Survey of India*, 222:1–162. 39

Vanscoy, T., Lundberg, J.G., and Luckenbill, K.R. 2015. Bony ornamentation of the catfish pectoral-fin spine: comparative and developmental anatomy, with an example of fin-spine diversity using the Tribe Brachyplatystomini (Siluriformes, Pimelodidae). *Proceedings of the Academy of Natural Sciences of Philadelphia*, 164:177–212. 15, 18, 24, 26, 27, 28, 36

Warnock, R.C.M., Parham, J.F., Joyce, W.G., Lyson, T.R., and Donoghue, P.C.J. 2014. Calibration uncertainty in molecular dating analyses: there is no substitute for the prior evaluation of time priors. *Proceedings of the Royal Society B: Biological Sciences*, 282:20141013–20141013. 39

Winterbottom, R. 1973. A descriptive synonymy of the striated muscles of the Teleostei. *Proceedings of the Academy of Natural Sciences of Philadelphia*, 125:225–317. 24

Wright, J.J. 2009. Diversity, phylogenetic distribution, and origins of venomous catfishes. *BMC Evolutionary Biology*, 9:282–293. 14, 19, 33

Chapter 3

A fossil fish fauna from the middle Miocene of the Cocinetas basin, northern Colombia

3.1 Abstract

The sedimentary sequence of the Cocinetas basin (Guajira Peninsula, Colombia) contains a record of continental to marginal marine environments during the Neogene, which has not yet been studied in detail. Among this sequence, the Castilletes formation preserves several vertebrate assemblages including both tetrapods and marine fishes. The faunal relationships of the freshwater fish assemblage to other Miocene faunas of South America have not been properly studied or described. A novel freshwater fish fauna is herein described from a sandstone level in the locality known as Makaraipao, Guajira, Colombia. The fossil-bearing layer lies within a sequence of sandstones and conglomeratic sandstones intercalated with limestones and mudstones; the sandstone levels represent middle- to high-energy settings of rivers to floodplains. This assemblage comprises dipnoans, siluriforms, and characiforms of strictly freshwater taxa. An assessment of morphological characters permits the identification of such remains at different levels of accuracy. Morphologies were assessed across an extensive sample of living representatives of respective groups. Dental characters were found to be informative at different taxonomic levels ranging from genus- to species-level for the characiforms, pointing to definite identification of serrasalmid genera *Mylossoma* and *Piaractus*. Spine characters were found to be unambiguous for Callichthyidae fossil specimens, but not beyond family-level. Cranial fragments of the Pimelodid genus *Phractocephalus* were unambiguously diagnosed based on surface ornamentation. Pterygoid tooth plates of *Lepidosiren* differ slightly from the extant species *L. paradoxa*, but not to an extend suggesting taxonomic differentiation. A reassessment of the identity of some fossil Serrasalmid occurrences from the Paleocene to the Miocene in South America is provided based on literature data. The Makaraipao freshwater fossil fish assemblage is compared to

those of other Miocene faunas across South America, suggesting strong similarities with the Fitzcarrald and Contamana faunas, and to a lesser extent with La Venta and Rio Acre, all Miocene, implying that geographical proximity in itself does not explain faunal composition.
Keywords: Teleostei, Cenozoic, Paleoecology, South America, Morphology.

3.2 Introduction

Freshwater fishes are the richest vertebrate continental component of the Neotropics, with roughly 7000 species (Albert and Reis, 2011). They are highly diverse in ecology and morphology, and are present from brackish waters to the elevations above 3000 meters above sea level in the Andes (Schaefer, 2011). Despite this enormous diversity, freshwater fishes are not as prominent in the fossil record of South America as mammals or crocodylians, and are often recorded from bone fragments of limited diagnostic value (Lundberg et al., 2010). Our current lack of comprehensive comparative morphological analyses focused on diagnostic characters in fossil specimens further hampers the identification of new fossil specimens beyond coarse taxonomic levels such as order or family. Such drawbacks limit the potential use of Neotropical freshwater fossil fishes in paleoecological, systematic, biostratigraphic, and biogeographic studies.

Freshwater fishes component have been crucial as evidence of past drainage connections between river systems east and west of the Andes (i.e., cis- and trans-Andean respectively), thus providing biological evidence of Andean orogeny during the Neogene. The dispersal potential of freshwater fishes across mountains is very limited for lowland taxa, thus providing a strong proxy for hydrological connections through geologic time. Groups that are currently restricted to cis-Andean drainages (e.g., Orinoco and Amazon drainages) have been found in fossil assemblages of Miocene age in trans-Andean localities (Lundberg, 1997; Lundberg et al., 2010), thus indicating that they lived in a time where the Andes were not an effective hydrological barrier between these areas (Diaz de Gamero, 1996; Gregory-Wodzicki, 2000). The La Venta and Urumaco fossil faunas in Colombia and Venezuela have long been identified as evidence in favor of this hypothesis of historical drainage connections across the Andes.

During several field seasons, a multidisciplinary team led by the Smithsonian Tropical Research Institute collected fossils and reassessed the stratigraphy and geological cartography of the Cocinetas sedimentary basin and adjacent areas in northern Colombia (Moreno et al., 2015). Most of the fossil vertebrates both marine and continental have been studied and reported in publications (Aguilera et al., 2017, 2013b; Amson et al., 2016; Cadena and Jaramillo, 2015a,b; Carrillo-Briceño et al., 2019; Forasiepi et al., 2014; Moreno-Bernal et al., 2016; Suarez et al., 2016); however, the freshwater fish component remains largely unpublished (but see Aguilera et al., 2013a), and will be the subject of future contributions (Ballen in prep.).

The goals of the present work are: 1) describe the freshwater fossil fishes of the Castilletes formation in the locality Makaraipao, northern Colombia; 2) provide anatomical characters of relevance for the identification of the fossil groups recorded and then discuss their relevance for related occurrences in other faunas of South America; and 3) provide an assessment of the faunal similarity patterns among the newly recorded fauna and others of Miocene age in the continent based on freshwater fish occurrences.

3.3 Materials and methods

3.3.1 Geological setting

The Cocinetas sedimentary basin preserves a succession of continental to shallow marine strata in the northern margin of the South American Plate. It spans a time interval from the Eocene to the Pliocene. Moreno et al. (2015) redefined the stratigraphy of the basin, showing that earlier studies misinterpreted the structure, thickness, and lithological limits of several units. The Castilletes formation is one of those units, and was redefined by the latter authors as sediments of Miocene age from mostly shallow marine environments. The lower boundary of this unit is the concordant contact with the Jimol formation, and the upper boundary is defined by an unconformity with the base of the Ware formation. The Castilletes formation has been dated as lower to middle Miocene using Sr radioisotopes (Hendy et al., 2015).

Makaraipao (locality STRI 390093, 11°54'32.0"N 71°20'24.4"W) is a moderate plateau lying to the west of the Tucacas bay, municipality of Uribia, Guajira department, in northern Colombia (Figure 3.1A-C). It is contained in the section called "Long section" in Moreno et al. (2015). The locality lies about 127 stratigraphic meters above the base of the section (Figure 3.1D); however, it is located ca. 279 m above the base in the composite section of the Castilletes formation (Suarez et al., 2016). All specimens known so far come from that locality and were collected from sandstone and conglomeratic sandstone levels. Sr radioisotopic data are available from specimens of the bivalve *Anadara* from a sandstone just below the vertebrate fossil-bearing level as well as ca. 60 stratigraphic meters below, suggesting an age of roughly ca. 15 Ma.

3.3.2 Abbreviations

Institutional abbreviations are: Academy of Natural Sciences of Drexel University, Philadelphia, US (ANSP), Instituto Alexander von Humboldt, Villa de Leyva, Colombia (IAvH), Museu de Zoologia da Universidade de São Paulo, São Paulo, Brazil (MZUSP), Mapuka Museum of Universidad del Norte, Barranquilla, Colombia (MUN). Premaxilla and dentary are abbreviated PM and D respectively.

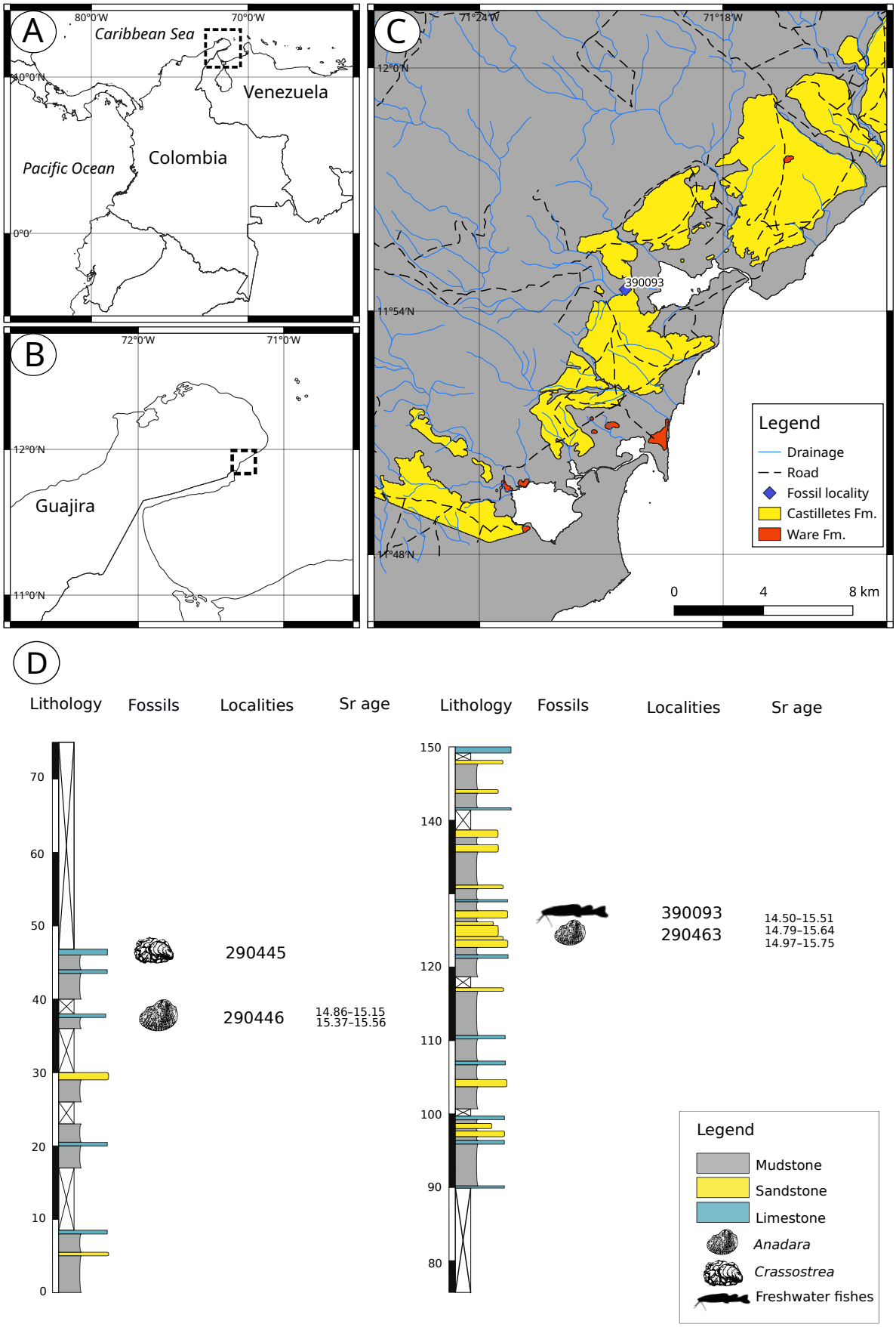


Figure 3.1: *A-C) Geographic context of the locality Makaraipao in the Cocinetas basin, Guajira Peninsula, Colombia. D) Stratigraphic column of the Long Section (loc. 170514) modified from Moreno et al. (2015). Isotopic age estimates are from Hendy et al. (2015) and the molluscan taxon from which they were taken. The vertical scale is in stratigraphic meters.*

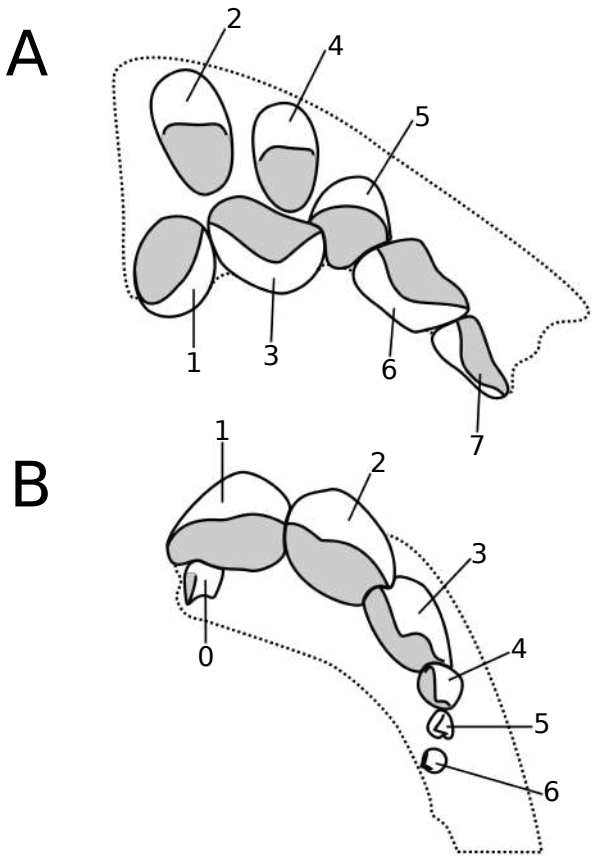


Figure 3.2: *Schematic representation of a Serrasalmid Pacu dentition based on Myloplus lucienae (Andrade et al., 2016). A) Premaxilla with teeth numbered following Cione et al. (2009). B) dentary numbering as herein proposed with the digit zero to the symphysial dentary tooth. Gray shades represent occlusal molariform surfaces for teeth where such feature is present. Cutting edge shape and position represented for each tooth.*

3.3.3 Anatomical terminology

Serrasalmid tooth nomenclature and position follows Cione et al. (2009) with modifications (Figure 3.2). Lepidosirenid anatomical nomenclature follows Criswell (2015).

3.3.4 Data analysis

Faunal composition for similarity analysis was compiled from literature data and direct observations (Aguilera et al., 2013a,b; Antoine et al., 2016; Azpelicueta and Cione, 2016; Ballen and Moreno-Bernal, 2019; Bogan et al., 2012; Cione and Azpelicueta, 2013; Cione et al., 2000, 2009; Lundberg et al., 2010; Tejada-Lara et al., 2015). An initial set of Miocene fossil fish faunas including both marine and freshwater components, and any number of fossil occurrences was considered and then reduced to a set where the number of taxa was equal or larger than the one herein described. Also, only freshwater components were included in the final results, which shows closer relationships among the fossil assemblage herein studied and Amazonian assemblages in Peru instead of the nearer Colombian and Venezuelan assemblages. A larger analysis with both marine and freshwater components and a variable

number of fossil taxa did not refute the pattern found herein (Supplementary materials). Faunal dissimilarity was measured using the Bray-Curtis coefficient as implemented in the `vegan` package v.2.5-5 (Oksanen et al., 2019) in R v.3.4.4 R Core Development Team (2018).

Extant occurrences were downloaded from the SpeciesLink and GBIF databases and specific data cleaning procedures carried out (Supplementary material). Mapping was carried out in QGIS v.3.4.12 (QGIS Development Team, 2019). Image edition and processing was carried out in GNU image manipulation program (GIMP). The complete raw data and scripts are available in the Appendix E.3.

3.4 Results

Systematic paleontology
Dipnoi
Order Lepidosireniformes
Genus *Lepidosiren* Fitzinger, 1837
***Lepidosiren* sp.**
Figures 3.3G-I,3.4A

Material examined: MUN 37667, partial left pterygoid plate preserving part of the middle and posterior pterygoid ridges, the posterior process and base of the ascending process; MUN 37693, partial pterygoid plate preserving part of the middle pterygoid ridge and all of the posterior ridge.

Description: Pterygoid tooth plates preserving the middle and posterior ridges in both specimens, although preserving the support bone in MUN 37667. Posterior pterygoid ridge somewhat sigmoid in axis in both specimens, projecting laterally from the pterygoid body. Median portion of middle pterygoid ridge approaching the contralateral ridge as observed in the preserved crown. Angle between preserved ridges about 30°, angle between posterior process of pterygoid and enamel-bearing axis of pterygoid about 130°.

Remarks: Pterygoid and prearticular tooth plates are the functional analogues of dentary and premaxilla respectively in bony fishes; however, these are not homologous structures. Criswell (2015) recovered the South American genus *Lepidosiren* as sister to the African genus *Protopterus*; these two genera can be distinguished due to the relative proportion between the posteriormost two pterygoid ridges. The second pterygoid ridge is about half the length of the first and posteriormost in *Lepidosiren*, while the second ridge is shorter than half the length of the posteriormost ridge in *Protopterus*. Pterygoid tooth plates of the Lepidosirenidae are distinguished from prearticular ones because of their more restricted amount of enameoid, a discontinuity between the ventral outline of the pterygoid ramus and the base of the enameloid in lateral view, the presence of the ascending process of the pterygoid (vs. ventral surface smooth and straight in lateral view) (Criswell, 2015). Also,

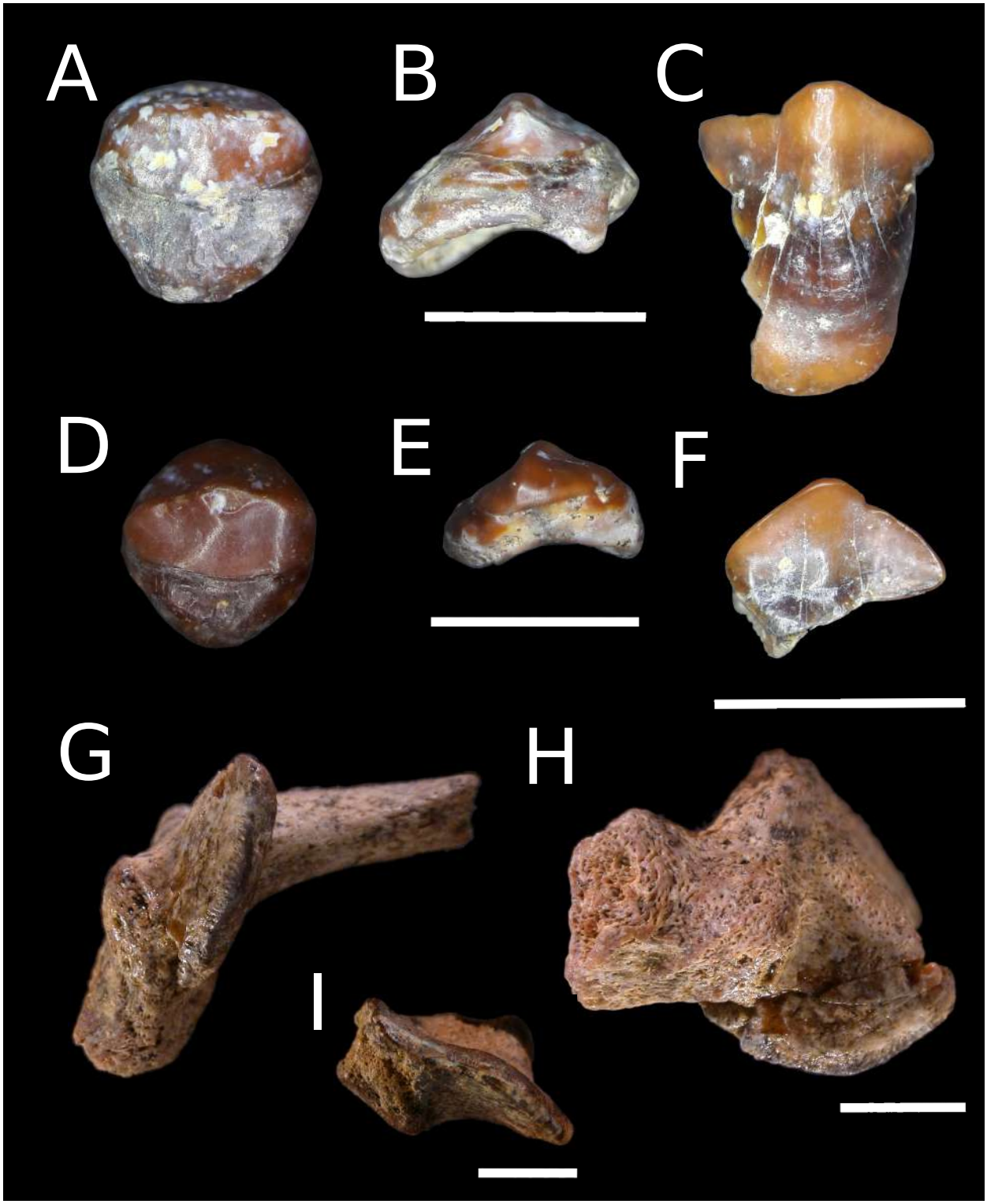


Figure 3.3: *Characiforms and Lepidosireniforms from the middle Castilletes Fm. in the locality Makaraipao. A-B,D-E)* Piaractus *aff.* brachypomus *MUN 37664, D4-6 in occlusal (A,D) and commisural (B,E) views. C,F)* Mylossoma *sp. MUN 34502 (C,F), D2 in labial (C) and occlusal (F) views. G-I)* Lepidosiren *sp. MUN 37667 (G-H) and MUN 37693 (I) in occlusal (G,I), and labial symphysial (H) views. Scale bars equal 5mm in all paired view.*

the contralateral middle ridges of the prearticular do not meet at the midline whereas they do in the middle ridges of the pterygoid.

The extant Lungfish *Lepidosiren paradoxa* is currently restricted to lenthic systems, swamps, *várzeas*, and lagoons in the Amazon and Paraná basins as well as in the Guyanas (Almeida-Val et al., 2011, Figure 3.4A). This taxon suggests that the fine-grained conditions recorded in levels adjacent to the sandy-conglomeratic lithology where the specimens were collected represent a setting of peripheral lenthic systems.

Division Ostariophysi
Order Characiformes
Family Serrasalmidae
Genus *Mylossoma* Eigenmann, 1903
***Mylossoma* sp.**
Figures 3.3C,F,3.4B

Material examined: MUN 34502, an isolated tooth.

Description: Molariform D2 tooth with asymmetric crown and flat occlusal surface. Posterior margin smoothly angular, lacking strong concavities. Anterior surface with longitudinal sulcus from cutting edge to preserved crown base. Crown unicuspid, with occlusal surface flat to slightly concave.

Remarks: Multicuspid, cutting to incisiform teeth are a well-known feature of carnivore, lepidophagous, and omnivore serrasalmids (*Serrasalmus*, *Pygocentrus*, *Pristobrycon*, *Pygopristis*, and *Catoprion*; Kolmann et al., 2018; Mirande, 2010); teeth in *Megapiranha* are still reminiscent of the ancestral, multicuspid, incisiform condition found in carnivore, lepidophagous, and onmivore genera (Cione et al., 2009). Teeth of *Acnodon*, *Mylesinus*, *Ossubtus*, and *Tometes* are multicuspid and incisiform. Contrastingly, the genera *Colossoma*, *Metynnis*, *Myleus*, *Mylossoma*, *Myloplus*, *Piaractus*, and *Utiaritichthys*, have molariform dentary teeth as in the fossil specimen reported herein. Additionally, only the recent cis-Andean species of *Mylossoma* (except the trans-Andean *M. acanthogaster*) show the vertical lingual sulcus which is diagnostic for that cluster of species, thus permitting the identification of the teeth as belonging to that genus.

Genus *Piaractus* Eigenmann, 1903
Piaractus* aff. *brachypomus
Figures 3.3A-B,D-E,3.4C

Material examined: MUN 37664, 2 isolated teeth.

Description: Commisural teeth D4-7, further serial origin uncertain. Crown well-preserved with remains of the cutting edge where only the tipo is eroded due to wear. Underlying bone preserved in both specimens, showing a central position of the cutting edge that is not lingually-deflected. Occlusal flat surface absent.

Remarks: Commisural teeth D4-7 are generally small and with deflected crowns that are usually aberrant with respect to other teeth in the same series; they tend to show a lingually-directed cutting edge that make them insuitable for crushing food items as opposed to teeth D0-4. The fossil specimens herein studied are strongly molariform and with very low cutting edge that is not lingually deflected, a condition seen in both *Colossoma* and *Piaractus* among serrasalmids. *Colossoma* still presents a unicuspid distinct cutting edge, while in *Piaractus* this feature is much less prominent, a condition also observed in the fossil specimens. Due to overall similarity supports their alignment with the genus *Piaractus*. Among the species of this genus the fossil specimens resemble more *P. brachypomus*.

Order Siluriformes
Suborder Loricarioidei
Family Callichthyidae
Gen. sp. *incertae sedis*
Figures 3.5B-G,3.4D

Material examined: MUN 37803, one dorsal and one pectoral spine fragments, the former preserving the spine base.

Description: Dorsal spine preserving the base and proximal portion of the shaft. Base triangular in outline with a small, base foramen round; anterior articular facet inverse trapezoidal in outline and finely ornamented with vertical ridges. Inflection point of the anterior longitudinal ridge poorly developed and with appearance of a tubercle; anterior longitudinal ridge absent. Anterior fossae ovoid and wide in outline. Lateral condyles poorly developed, lateral articular surfaces ovoid in outline and vertical ridges. Posterior processes present, shorter than lateral condyles, deflected ventrally. Spine shaft lacking ornaments other than odontodes alveoli covering the entire anterior surface.

Right pectoral spine preserving the shaft from the deflection point of the dorsal process to about 1/3 the shaft length. Shaft slightly depressed and oval in transverse section. Dorsal and ventral ornament consisting of parallel, fine ridges; anterior ornament consisting of fine odontode alveoli distributed along the entire anterior surface of the shaft; posterior ornament consisting of compressed blades with indication of dorsal and ventral cutting edges, spanning half of the vertical space on the posterior surface, retrorse, increasing in size from proximal to distal sections.

Remarks: Dorsal- and pectoral-fin spines have seldom been used at a level finer than family in the Siluriform fossil record. Callichthyid remains have been previously detected in the fossil record on the basis of spines; Lundberg (1997) noted the presence of cf. *Hoplosternum* in the middle Miocene La Venta fauna in central Colombia based on specimens preserving part of the cranium and some isolated pectoral spine fragments. They are reported as bearing odontode alveoli (= bases) and strong posterior serrae, from straight to slightly retrorse. The combination of odontodes and posterior ornaments is unique to the

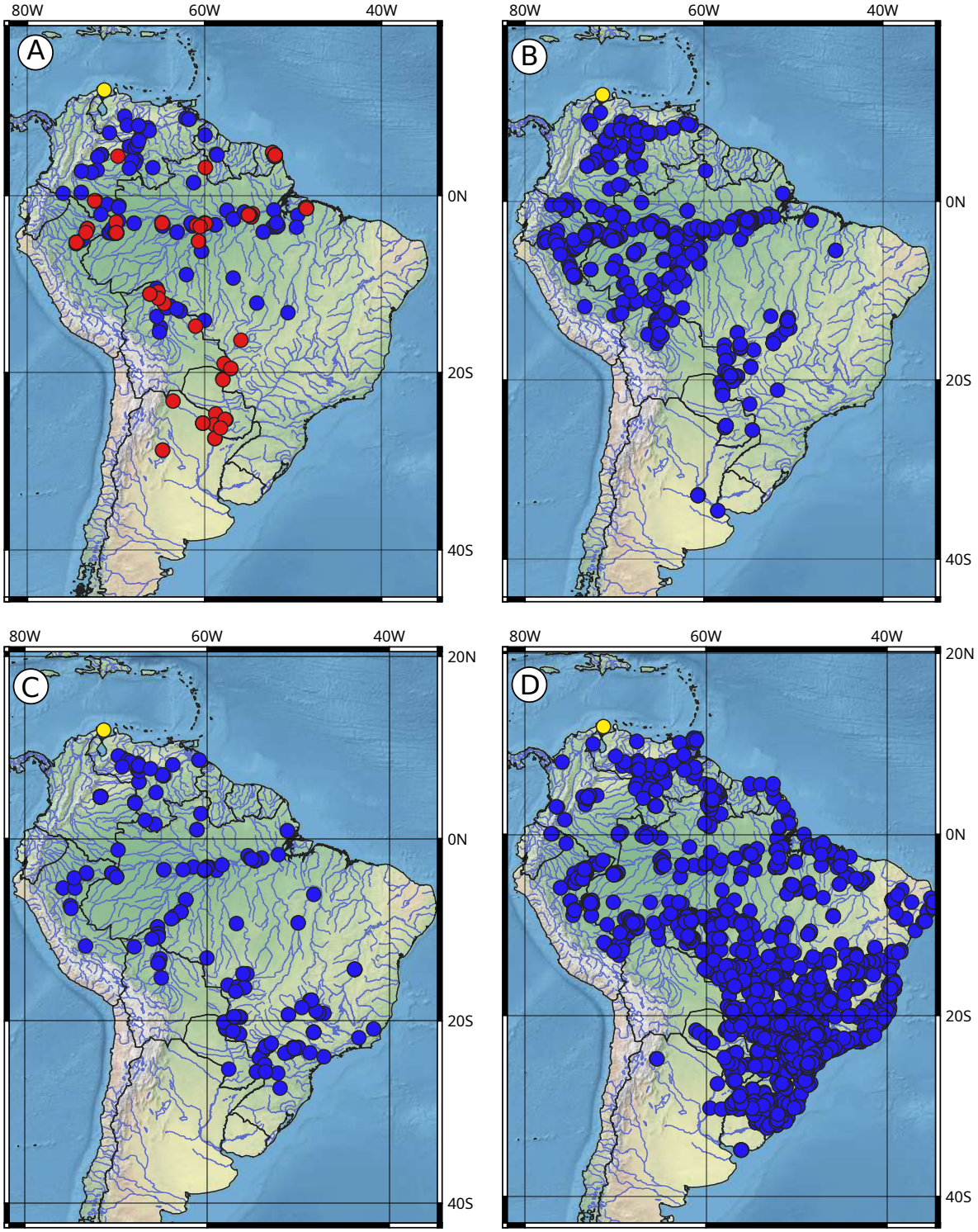


Figure 3.4: *Recent distributions of fossil taxa. A)* Lepidosiren *(red) and* Phractocephalus *(blue). B)* Mylossoma. *C)* Piaractus. *D) Callichthyidae. Yellow spots in all maps represent the fossil locality Makaraipao in northern Colombia.*

Callichthyidae among families of the suborder Loricarioidei, thus providing a useful diagnostic character for pectoral spines at family level.

Spines show strong sexual dimorphism at least in some Callichthyid genera. Pectoral spines lose their ornament due to overgrowth of the spine shaft in *Callichthys* and *Hoplosternum*, causing the loss of ornaments on the posterior surface of the pectoral spine, in contrast, females and juveniles have posterior ornaments in the form of straight to slightly retrorse spinules in *Callichthys* and flat spinules in *Hoplosternum*. Thus, the presence and absence of posterior ornament by itself cannot be used for identifying taxa; the shape of the spine, on the other hand, can be used provided that comparisons are restricted to subadults and females. It is noteworthy that spine fragments with odontode alveoli but lacking posterior ornament are extensively present in the Loricariidae, thus allowing potential ambiguity when spines combining these two features are found in the fossil record. Oftentimes Loricariids show a hypertrophied row of odontodes on the dorso-posterior angle of the spine (Ballen and Vari, 2012) that are absent in Callichthyids (pers. obs.); this feature aids in distinguishing pectoral spine fragments from those families.

The dorsal spine shaft is very wide in anterior view and antero-posteriorly compressed in the genera *Hoplosternum* and *Lepthoplosternum*; in contrast, the specimen herein studied shows a more regular cross-section outline. The anterior articular facet of the dorsal spine is inversely trapezoid in the genera *Dianema* and *Megalechis*, in contrast to the oval outline in the fossil specimen.

Suborder Siluroidei

Family Pimelodidae

Genus *Phractocephalus* Spix & Agassiz, 1829

***Phractocephalus* sp.**

Figures 3.5A,3.4A

Material examined: MUN 37660, fragment of nuchal plate.

Description: Preserved portion of nuchal plate flat with strong anastomosing ridges on dorsal surface; some isolated tubercles around pits and less frequently on ridges. Preserved bone thickness nearly uniform; margins not preserved and ventral surface somewhat eroded.

Remarks: The presence of strongly reticulate ridged ornament on the nuchal plate is diagnostic for *Phractocephalus* level among South American Siluriforms (Aguilera et al., 2008; Azpelicueta and Cione, 2016; Lundberg, 1997; Lundberg and Aguilera, 2003; Rincón et al., 2016). Although some species of the families Ariidae, Doradidae, and Andinichthyidae can have a strongly-ornamented condition, it consists of dense tubercles, but never reticulating ridges (Aguilera and De Aguilera, 2004; Birindelli, 2014; Bogan et al., 2018). Cranial bones other than the nuchal plate show the strong reticulate ornamentation; however all of these except the opercle are not mostly flat as in the preserved specimen. Besides, it does not represent an opercle given that the opercular ornament is not as reticulated as in other cranial bones and also their ridges show a radiating pattern from the proximal region of the

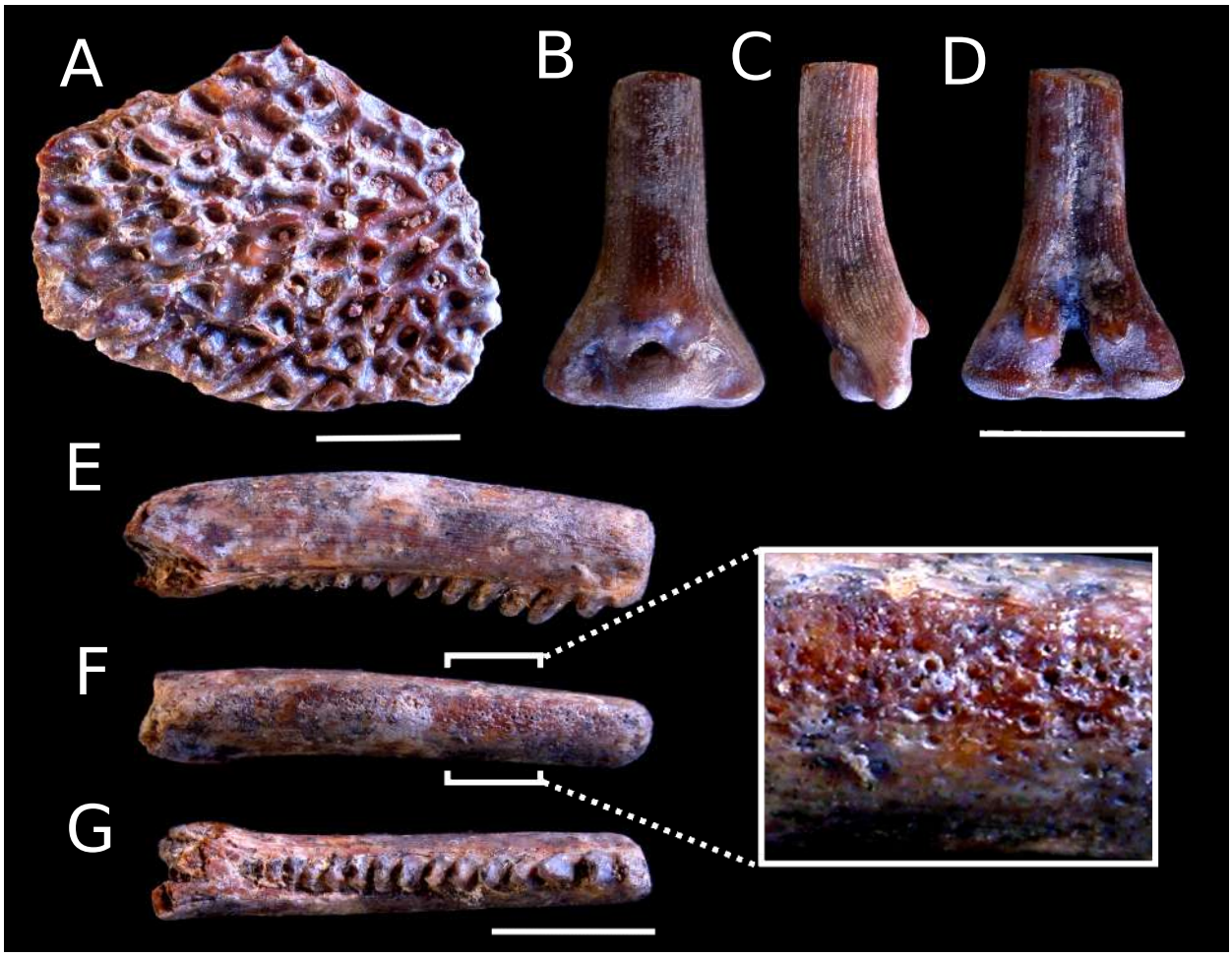


Figure 3.5: *Siluriforms from the middle Castilletes formation in the locality Makaraipao. A)* Phractocephalus *sp., nuchal plate fragment. B-D) Callichthyidae gen. et sp. indet., dorsal-fin spine fragment in anterior, lateral, and posterior views respectively. E-G) Callichthyidae gen. et sp. indet., pectoral-fin spine fragment in dorsal, anterior, and posterior views respectively. Scale bars equal 10mm in A, and 5mm in the remaining sections.*

bone. Other bones (e.g., sphenotic) can be somewhat flat but their ventral surface is not as smooth and uniform as the nuchal plate.

3.5 Faunal similarity

The freshwater fossil fish faunas herein studied vary considerably in richness of genera. The Urumaco fauna shows the richest assemblage with 13 genera while Makaraipao, Contamana and Fitzcarrald are the poorest with four genera each; Ituzaingó, Rio Acre and La Venta show intermediate values with six, eight, and eleven genera respectively. According to the Bray-Curtis dissimilarity index, the Makaraipao fish fauna is closer to the Fitzcarrald and Contamana faunas than to other units (Figure 3.6). The La Venta and Rio Acre faunas cluster together, while Ituzaingó and Urumaco are the most dissimilar of the set.

A less conservative approach including taxa of either marine or freshwater habitats, and without filtering poorly-sampled fossil faunas (e.g., Loyola-Mangan and Utuquina) does

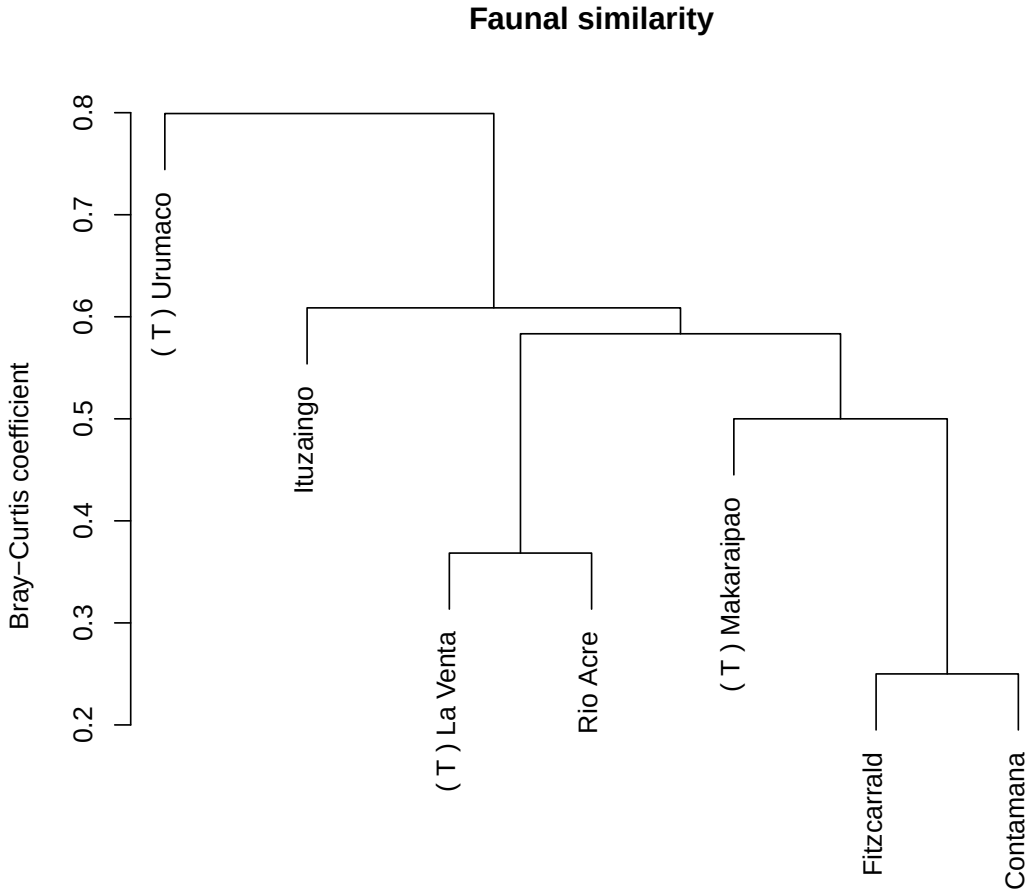


Figure 3.6: *Faunal dissimilarity for Miocene freshwater fish faunas in South America using the Bray-Curtis coefficient. The symbol (T) indicates those faunas located west of the Andes*

not break the association among Contamana, Fitzcarrald, and Makaraipao (supplementary data). Also, the relationship between La Venta and Rio Acre remains robust despite the effect of such less-stringent parameters data.

3.6 Discussion

3.6.1 The fossil record of the Lepidosirenidae in South America

Lepidosiren megalos was described by Silva Santos (1987) from the Acre fauna of Brazil based on cranial remains and isolated tooth plates. Lundberg et al. (2010) questioned the distinctiveness of that taxon from the extant *L. paradoxa*. The main difference between the two is body size. Silva Santos (1987) mentions that the coronoid process is shallower in *L. megalos* than in *L. paradoxa*; however, these differences are not visually obvious from

figures 1 and 6 in plates I and II respectively, and even if real they may correspond only to allometric size effects. More detailed statistical treatment is necessary to demonstrate significant differences between these entities. The type material of *L. megalos* was allegedly housed in the paleontology collection of MZUSP but has not been located in many years (A. Carvalho, pers. comm.). Additional specimens and proper statistical analysis are necessary to confirm validity of *Lepidosiren megalos* which is consequently considered as a junior synonym of *Lepidosiren paradoxa* Fitzinger, 1837 (Lundberg et al., 2010).

The genus *Lepidosiren* has been recorded from the Honda group in Colombia, the Solimões formation in Brazil, and the Pebas formation in both in the Contamana and Fitzcarrald areas in Peru (Antoine et al., 2016; Tejada-Lara et al., 2015). Gayet et al. (2001) reports remains of *Lepidosiren* cf. *paradoxa* from the localities Pajcha Pata and Vila Vila in Bolivia in sediments of the El Molino formation, then thought to be of Maastrichthian age. Gelfo et al. (2009) reinterpret the chronology of the Paleocene SALMAS in South America and conclude that the El Molino formation is actually Danian (early Paleocene). These remains are the oldest record of *Lepidosiren* in South America, but their specific identity remains elusive; a proper restudy of either the original or new specimens from Bolivia is necessary in order to conclude whether they are conspecific with the extant *L. paradoxa*; however, a species timespan of ca. 63 Ma seems unlikely.

The genus has been recorded only in fine-grained lithologies in the Honda group (Ballen and Moreno-Bernal, 2019; Lundberg, 1997, fig. 2) consistent with the habitat preference documented in extant representatives of *Lepidosiren*; this suggests that finer-scale environmental reconstructions in other stratigraphic units could find similar patterns in paleoecological reconstructions.

3.6.2 Paleogeography of *Phractocephalus*

The genus *Phractocephalus* has long been recorded from Neogene sedimentary units in South America. It is represented by the extant *P. hemioliopterus* (Figure 3.4B) and the extinct *P. acreornatus*, *P. ivy*, and *P. nassi* from Brazil, Argentina, and Venezuela respectively (Aguilera et al., 2008; Azpelicueta and Cione, 2016; Lundberg and Aguilera, 2003). Species of the genus are strictly freshwater and take part in large reproductive migrations along the Amazon and Orinoco drainages (Lundberg and Aguilera, 2003; Naranjo and Espinel, 2009). The extant species is currently restricted to drainages east of the Andes, although fossil species (e.g., *P. nassi*) and material of uncertain specific affinities (Lundberg, 1997, G.A. Ballen et al., *in prep.*) are known from the Neogene of the Magdalena drainage in Colombia. Given these geographic patterns in extant and extinct occurrences, the genus has long been recognized as evidence of past connections between cis- and trans-Andean drainages before separation due to the Andean orogeny (Lundberg, 1997; Lundberg et al., 1998, 2010). The occurrence of the genus in the middle Miocene of the Castilletes formation conform to this pattern of drainage connections and agrees with the occurrences of *Phractocephalus* in the

La Venta fauna in central Colombia, in the present-day Magdalena drainage.

3.6.3 Dental characters in the Serrasalmidae

The Serrasalmidae is the family with the most abundant fossil record in the Characiformes; however, it is often restricted to isolated teeth reported in the literature as indeterminate beyond the family level. The trophic ecology of serrasalmids is diverse, which in turn is reflected in morphological variability, specially concerning the oral complex. A number of isolated teeth spanning from the Paleocene to the Miocene show considerable variability, and preserve several characters herein considered to be taxonomically informative at different levels. The greatest difficulty in using dental morphology as a taxonomic character has been the lack of detailed comparative information about extant representatives of the family; there is considerable serial variation (i.e., along a given tooth series) in serrasalmids, as well as marked differences among taxa of different genera and species. The main issue is therefore the lack of a reference framework about the limits of variation at different levels, which is a necessary condition for finer taxonomic identification of isolated fossil remains.

Herein, a detailed comparative study of dentary and premaxillary teeth in serrasalmids provide for the first time a framework for the identification of taxonomically-informative variation allowing satisfactory identification of fossil serrasalmid teeth, even in the absence of preserved attachment bones (e.g., *Colossoma macropomum* from Colombia and *Megapiranha paranensis* from Argentina are both known from partial premaxillae). Our data permit identification of specimens both in the literature and directly examined.

Dahdul (2004) is one of the few efforts to attempt identification of fossil serrasalmid teeth at genus level. The author reports the genus *Mylossoma* from the Castillo formation in Venezuela. Dahdul compares here fossils with *Colossoma* and *Piaractus* among serrasalmids, but not to other taxa with molariform teeth such as *Metynnis*, *Myleus*, or *Myloplus*. Only *Mylossoma* shows two concavities in occlusal view, one labiolingual associated with the lateral cusplet, and one extensive mesiolingual, associated with the tooth D0. Other genera show a widespread conditions of having only a mesiolingual concavity (*Acnodon*, *Colossoma*, *Metynnis*, *Myleus*, *Myloplus*, *Piaractus*, and *Utiaritichthy*), two lingual concavities (*Myleus*), or lack concavities altogether on D1 in occlusal view (*Mylesinus* and *Tometes*). Given the character distribution noted above, I confirm the generic status of these remains from the Castillo formation as *Mylossoma* sp.

Gayet et al. (2001, p. 52, fig. 7c-e) describe a rich vertebrate assemblage from the Paleocene locality of Pajcha Pata in Bolivia, recording among others, isolated serrasalmid teeth. The present observations permit further taxonomic associations for that material as cf. *Acnodon*. Both teeth of the morphotype 2 (fig. 7d-e therein) are identified as premaxillary teeth due to the presence of two cutting edges, one on the labial face and one on the lingual face. Also the tooth illustrated in figure 7e can be identified as any of PM3, PM6, or PM7, due to the presence of three cusps on the lingual cutting edge and at least one on the labial cutting

edge. These characters are present in the genera *Acnodon* and *Metynnis*. Teeth with high labial and lingual cutting edges such as those illustrated by Gayet et al. in their figure 7d are present in *Acnodon* (Jégu and dos Santos, 1990, figs. 9-10).

Some published works illustrate fossil teeth with a lingual projection in occlusal view that is bounded by two concave margins (e.g., Rubilar, 1994, fig. 4c; Monsch, 1998, plate iii fig. 13); such morphology is found in PM5 across a number of genera where adjacent teeth compress the lingual margin thus creating two lingual concavities. This feature of PM5 is found in *Colossoma*, *Myloplus*, *Metynnis*, *Myleus*, *Piaractus brachypomus*, and *Utiaritichthys*; it is absent in *Acnodon*, *Mylesinus*, *Myleus setiger*, *Mylossoma*, *Piaractus mesopotamicus* and *Tometes*. The taxonomic distribution of this morphology is wider than other characters; however, a promising character that can be used in combination in order to refine the identity of fossil occurrences is the number of cusps on the labial cutting edge; the specimen illustrated by Monsch presents three cusps, while the specimen illustrated by Rubilar presents one or two cusps. These refinements should be carried out with a direct examination of the original specimens that were described in these works, since the illustrations seem inadequate for a better reassessment of their identity.

3.6.4 The fossil record of the Callichthyidae

The family has a fossil record extending from the late Paleocene to the Pleistocene. Reis (1998) reviewed the fossil record then available, reassessing the position of the iconic *Corydoras revelatus* from the Maiz Gordo formation in Argentina, the most complete fossil taxon of the family. He also discussed some additional callichthyid fossil occurrences such as *Hoplosternum* sp. from the middle Miocene Honda group in Colombia, and scattered indeterminate remains from the Acre fauna in Brazil and the Pleistocene Luján formation in Argentina. Lundberg et al. (2010) Further added an occurrence of cf. *Hoplosternum* from the Madre de Dios fauna without further comments or description of the specimens, also recording occurrences of the subfamilies Callichthyinae and Corydoradinae from the Solimões formation, Callichthyidae from Madre de Dios, and the previously known *Hoplosternum* sp. from the Honda group. Cione and Baez (2007) reviewed the Argentinian fish fossil record indicating the genera *Corydoras* and *Callichthys* from the Pleistocene of the Bahía Blanca area, also confirming an age of late Paleocene for the Maiz Gordo records of *Corydoras revelatus*. The family Callichthyidae is so far absent in the Venezuelan Urumaco formation, therefore rendering the occurrences of the family herein studied as the northernmost ones in the fossil record. Although the age of the Castilletes formation does not inform about the time of origin of the Callichthyidae, the present specimens await further and more detailed study in order to assess their phylogenetic position.

3.6.5 Paleoecology

Although the currently-known composition of the Makaraipao ichthyofauna is restricted to five taxa, each one of them is informative about the paleoenvironmental conditions present during the middle Miocene in the Guajira peninsula. Large catfishes such as *Phractocephalus* indicate large river channels with long courses permiting reproductive migrations. Present riverine conditions allow the only extant species of the genus (*P. hemioliopterus*) to migrate long distances during the reproductive season; restricted basins would not allow such events and this explains the present absence of the genus in small basins with short river courses such as the Maracaibo basin in Colombia and Venezuela, or the Coastal drainages in northern Venezuela. On the other hand, peripheral, lenthic environments such as lagoons and floodplains should have been present since *Lepidosiren* is restricted to such still and stagnant environments. On the other hand, herbivorous serrasalmids such as *Mylossoma* and *Piaractus* are known for their ability to exploit fruits and seeds in the riparian forest, often with strong preference over other possible food items; those taxa suggest that riverine environments in the Guajira peninsula had gallery forest that provided such resources to the river channel. Finally, the Callichthyidae is a very diverse Catfish family so that associated paleoecological parameters cannot at this time be determined for the still coarse taxonomic identification for the fossil remains herein studied.

These paleoenvironmental conditions indicated by those fossil fishes are similar to present-day rivers from the Amazon basin and differ starkly with current conditions in the Guajira peninsula, now arid and desertic (mean annual precipitation = 397 mm, mean annual temperature = 28.7 °C; (Ramírez and del Valle, 2011). All of the taxa found in the present study are part of the amazonian freshwater fish fauna (Figure 3.4).

3.6.6 Faunal similarity

Continental fossil faunas of Miocene age have long been recognized in South America. Their composition provides a rich amount of data relevant for paleoecological and paleogeographical reconstructions. So far, such inferences have been based mostly on their mammalian components. Although a number of works discuss the biogeographic patterns and affinities of specific vertebrate groups, Cozzuol (2006) and Carrillo et al. (2015) provide quantitative perspectives on the affinities in faunal composition among Neogene sites in South America. Cozzuol (2006) found the Acre and Urumaco faunas to be most similar to each other than to the Mesopotamian and the La Venta fauna, while the latter was found to be the most dissimilar of all. Carrillo et al. (2015), on the basis of more numerous faunal units and a richer dataset suggest complex faunal patterns in time and space. Their analyses indicate that the La Venta fauna cluster with the Fitzcarrald fauna and to a lesser degree with Collon Curá, while Urumaco is most similar to Acre and the Mesopotamian. Such conclusions agree with earlier finds by Cozzuol (2006). According to Carrillo et al. (2015) the Mesopotamian includes the fossil assemblage that I herein refer to as the Ituzaingó fish fauna in order to re-

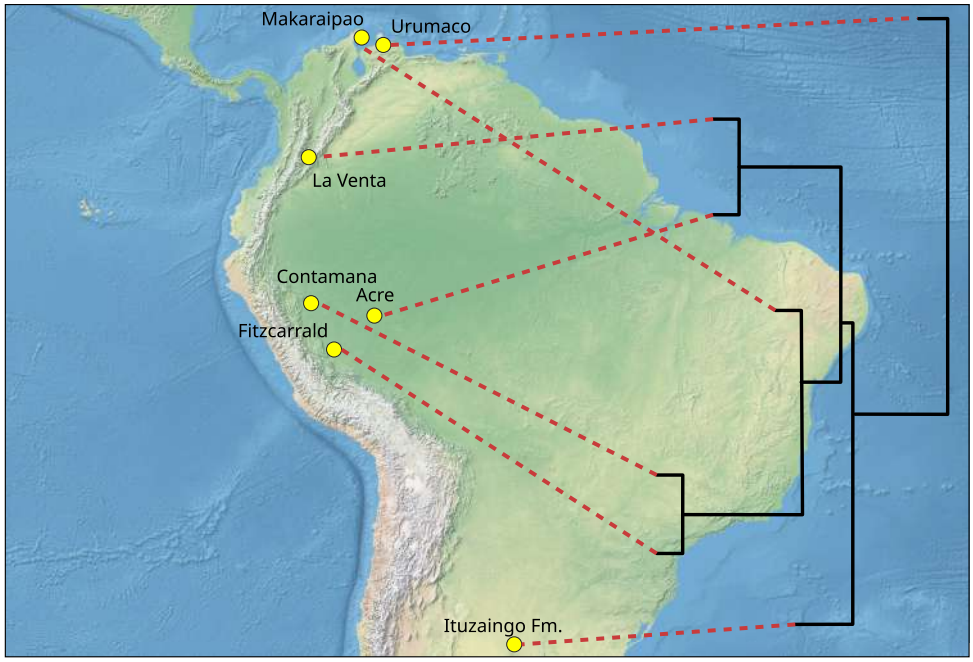


Figure 3.7: *Spatial relationships among fossil faunas of Miocene age in South America based on freshwater fossil assemblages. Dendrogram as recovered in the similarity analysis plotted in Figure 3.6*

inforce the stratigraphic and geographic provenance of the fossil association rather than the biochronological implication of the term Mesopotamian, that has been shown to be of complex application (Cione et al., 2000; Cozzuol, 2006). Several works refer to the Mesopotamian between quotation marks in order to acknowledge the problematic situation of this artificial biochronological name when referring to the faunal association instead (Brandoni, 2011; Cione et al., 2000; Cozzuol, 2006; Latrubesse et al., 2010).

Earlier studies have consistently shown a closer relationship between the Acre and Urumaco faunas, based on the fossil mammal assemblage (Carrillo et al., 2015; Cozzuol, 2006; Latrubesse et al., 2010), while the La Venta fauna has been shown to be the most distinctive unit of the fossil faunas analyzed. This pattern is due to the distinctiveness of the mammalian taxa found in La Venta in contrast to other localities in South America. The close relationship between the trans-Andean locality Makaraipao with the peruvian faunas of Contamana an Fitzcarrald based on fossil freshwater fishes is surprising given the closer geographic proximity of the former with other trans-Andean faunas such as La Venta and Urumaco (Figure 3.7). This is explained by the fact that these three faunas are almost identical, sharing components at the genus level. On the other hand, the large amount of taxa unique to the Urumaco fauna (Lundberg et al., 2010) explains its distinctiveness with respect to other Miocene faunas in South America. The components identified in both Contamana and Fitzcarrald are but a small proportion of the recovered fossil material, and several speci-

mens await more detailed study; therefore, the relationship among Makaraipao, Contamana, and Fitzcarrald may change with a better understanding of the fossil fish components of the peruvian faunas. Another possibility is that in fact the freshwater fish fauna shows faunal similarity patterns distinct from the mammalian component (i.e., Urumaco most distant to others, and Makaraipao clustering together with the peruvian faunas). This is a reasonable scenario given that fishes respond to barriers in a different way than mammals do (Dagosta and de Pinna, 2019). Further testing of these alternatives depends on more detailed study of the fossil freshwater fishes of the Contamana and Fitzcarrald.

Bibliography

Aguilera, O., Oliveira, G. A. S., Lopes, R. T., Machado, A. S., dos Santos, T. M., Marques, G., Bertucci, T., Aguiar, T., Carrillo-Briceño, J. D., Rodriguez, F., and Jaramillo, C. (2017). Neogene Proto-Caribbean Porcupinefishes (Diodontidae). *PLoS One*, 12(7):1–26. 49

Aguilera, O. A., Bocquentin, J., Lundberg, J. G., and Maciente, A. (2008). A new cajaro catfish (Siluriformes: Pimelodidae: *Phractocephalus*) from the Late Miocene of southwestern Amazonia and its relationship to *Phractocephalus nassi* of the Urumaco Formation. *Palaeontologische Zeitschrift*, 82(2):231–245. 58, 61

Aguilera, O. A. and De Aguilera, D. R. (2004). Amphi-American neogene sea catfishes (Siluriformes, Ariidae) from northern south America. *Special Papers in Palaeontology*, 71:29–48. 58

Aguilera, O. A., Lundberg, J. G., Birindelli, J. L. O., Sabaj Pérez, M. H., Jaramillo, C. A., and Sánchez-Villagra, M. R. (2013a). Palaeontological evidence for the last temporal occurrence of the ancient Western Amazonian river outflow into the Caribbean. *PLoS ONE*, 8(9):1–17. 49, 52

Aguilera, O. A., Moraes-Santos, H., Costa, S. A. R. F., Ohe, F., Jaramillo, C. A., and Nogueira, A. (2013b). Ariid sea catfishes from the coeval Pirabas (Northeastern Brazil), Cantaure, Castillo (Northwestern Venezuela), and Castilletes (North Colombia) formations (early Miocene), with description of three new species. *Swiss Journal of Palaeontology*, 132(1):45–68. 49, 52

Albert, J. S. and Reis, R. E. (2011). *Historical Biogeography of Neotropical Freshwater Fishes*. University of California Press. 49

Almeida-Val, V. M. F., Nozawa, S. R., Lopes, N. P., Aride, P. H. R., Mesquita-Saad, L. S., da Silva, M. d. N. P., Honda, R. T., Ferreira-Nozawa, M. S., and Val, A. L. (2011). Biology of the South American Lungfish, *Lepidosiren paradoxa*. In *The Biology of Lungfishes*, pages 129–147. 55

Amson, E., Carrillo, J. D., and Jaramillo, C. A. (2016). Neogene sloth assemblages (Mam-

malia, Pilosa) of the Cocinetas Basin (La Guajira, Colombia): implications for the Great American Biotic Interchange. *Palaeontology*, 59(4):563–582. 49

Andrade, M. C., Ota, R. P., Bastos, D. A., and Jégu, M. (2016). A new large *Myloplus* Gill 1896 from rio Negro basin, Brazilian Amazon (Characiformes: Serrasalmidae). *Zootaxa*, 4205(6):571–580. xvi, 52

Antoine, P.-O., Abello, M. A., Adnet, S., Altamirano Sierra, A. J., Baby, P., Billet, G., Boivin, M., Calderón, Y., Candela, A., Chabain, J., Corfu, F., Croft, D. A., Ganerød, M., Jaramillo, C. A., Klaus, S., Marivaux, L., Navarrete, R. E., Orliac, M. J., Parra, F., Pérez, M. E., Pujos, F., Rage, J. C., Ravel, A., Robinet, C., Roddaz, M., Tejada-Lara, J. V., Vélez-Juarbe, J., Wesselingh, F. P., and Salas-Gismondi, R. (2016). A 60-million-year Cenozoic history of western Amazonian ecosystems in Contamana, eastern Peru. *Gondwana Research*, 31:30–59. 52, 61

Azpelicueta, M. d. l. M. and Cione, A. L. (2016). A southern species of the tropical catfish genus *Phractocephalus* (Teleostei: Siluriformes) in the Miocene of South America. *Journal of South American Earth Sciences*, 67:221–230. 52, 58, 61

Ballen, G. A. and Moreno-Bernal, J. W. (2019). New records of the enigmatic Neotropical fossil fish *Acregoliath rancii* (Teleostei *incertae sedis*) from the middle Miocene Honda group of Colombia. *Ameghiniana*, 56(6):431–440. 52, 61

Ballen, G. A. and Vari, R. P. (2012). Review of the Andean armored catfishes of the genus *Dolichancistrus* Isbrücker (Siluriformes: Loricariidae). *Neotropical Ichthyology*, 10(3):499–518. 58

Birindelli, J. L. O. (2014). Phylogenetic relationships of the South American Doradoidea (Ostariophysi: Siluriformes). *Neotropical Ichthyology*, 12(3):451–564. 58

Bogan, S., Agnolin, F. L., and Scanferla, A. (2018). A new Andinichthyidae catfish (Ostariophysi, Siluriformes) from the Paleogene of northwestern Argentina. *Journal of Vertebrate Paleontology*, 38(3). 58

Bogan, S., Sidlauskas, B., Vari, R. P., and Agnolin, F. L. (2012). *Arrhinolemur scalabrinii* Ameghino, 1898, of the late Miocene - a taxonomic journey from the Mammalia to the Anostomidae (Ostariophysi: Characiformes). *Neotropical Ichthyology*, 10(3):555–560. 52

Brandoni, D. (2011). The Megalonychidae (Xenarthra, Tardigrada) from the late Miocene of Entre Ríos Province, Argentina, with remarks on their systematics and biogeography. *Geobios*, 44(1):33–44. 65

Cadena, E. and Jaramillo, C. A. (2015a). Early to Middle Miocene Turtles from the Northernmost Tip of South America: Giant Testudinids, Chelids, and Podocnemidids from the Castilletes Formation, Colombia. *Ameghiniana*, 52(2):188–203. 49

Cadena, E. and Jaramillo, C. A. (2015b). The first fossil skull of *Chelus* (Pleurodira: Chelidae, Matamata turtle) from the early Miocene of Colombia. *Palaeontologica Electronica*, 18.2(23A):1–10. 49

Carrillo, J. D., Forasiepi, A., Jaramillo, C., and Sánchez-Villagra, M. R. (2015). Neotropical mammal diversity and the great American biotic interchange: Spatial and temporal variation in South America's fossil record. *Frontiers in Genetics*, 5(451):1–11. 64, 65

Carrillo-Briceño, J. D., Luz, Z., Hendy, A., Kocsis, L., Aguilera, O., and Vennemann, T. (2019). Neogene Caribbean elasmobranchs: Diversity, paleoecology and paleoenvironmental significance of the Cocinetas Basin assemblage (Guajira Peninsula, Colombia). *Biogeosciences*, 16(1):33–56. 49

Cione, A. L. and Azpelicueta, M. d. l. M. (2013). The first fossil species of *Salminus*, a conspicuous South American freshwater predatory fish (Teleostei , Characiformes), found in the Miocene of Argentina. *Journal of Vertebrate Paleontology*, 33(5):1051–1060. 52

Cione, A. L., Azpelicueta, M. d. l. M., Bond, M., Carlini, A. A., Casciotta, J. R., Cozzuol, M. A., de la Fuente, M., Gasparini, Z., Goin, F. J., Noriega, J., Scillato-Yané, G. J., Soibelzon, L., Tonni, E. P., Verzi, D., and Vucetich, M. G. (2000). The Miocene vertebrates from Paraná, eastern Argentina Miocene vertebrates from Entre Ríos province, eastern Argentina. *Correlación Geológica*, 14:191–237. 52, 65

Cione, A. L. and Baez, A. M. (2007). Peces, anfibios e invertebrados cenozoicos de Argentina: los últimos cincuenta años. *Asociación Paleontológica Argentina. Publicación Especial*, 11:195–220. 63

Cione, A. L., Dahdul, W. M., Lundberg, J. G., and Machado-Allison, A. (2009). *Megapiranha paranensis*, a new genus and species of Serrasalmidae (Characiformes, Teleostei) from the upper Miocene of Argentina. *Journal of Vertebrate Paleontology*, 29(2):350–358. xvi, 52, 55

Cozzuol, M. A. (2006). The Acre vertebrate fauna: Age, diversity, and geography. *Journal of South American Earth Sciences*, 21(3):185–203. 64, 65

Criswell, K. E. (2015). The comparative osteology and phylogenetic relationships of African and South American lungfishes (Sarcopterygii: Dipnoi). *Zoological Journal of the Linnean Society*, 174(4):801–858. 52, 53

Dagosta, F. C. P. and de Pinna, M. C. C. (2019). The fishes of the Amazon: Distribution and biogeographical patterns, with a comprehensive list of species. *Bulletin of the American Museum of Natural History*, 431:1–163. 66

Dahdul, W. M. (2004). Fossil Serrasalmine fishes (Teleostei: Characiformes) from the Lower Miocene of north-eastern Venezuela. *Special Papers in Palaeontology*, 71:23–28. 62

Diaz de Gamero, M. L. (1996). The changing course of the Orinoco River during the Neogene: A review. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 123:385–402. 49

Forasiepi, A. M., Soibelzon, L. H., Suárez Gomez, C., Sánchez, R., Quiroz, L. I., Jaramillo, C. A., and Sánchez-Villagra, M. R. (2014). Carnivorans at the Great American Biotic Interchange: new discoveries from the northern neotropics. *Naturwissenschaften*, 101(11):965–74. 49

Gayet, M., Marshall, L. G., Sempere, T., Cappetta, H., and Rage, J. (2001). Middle Maastrichtian vertebrates (fishes, amphibians, dinosaurs and other reptiles, mammals) from Pajcha Pata (Bolivia). Biostratigraphic, palaeoecologic and palaeobiogeographic implications. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 169:39–68. 61, 62, 63

Gelfo, J. N., Goin, F. J., Woodburne, M. O., and De Muizon, C. (2009). Biochronological relationships of the earliest South American paleogene mammalian faunas. *Palaeontology*, 52(1):251–269. 61

Gregory-Wodzicki, K. M. (2000). Uplift history of the Central and Northern Andes: A review. *Geological Society of America Bulletin*, 112(7):1091–1105. 49

Hendy, A. J. W., Jones, D. S., Moreno, F., Zapata, V., and Jaramillo, C. A. (2015). Neogene molluscs, shallow marine paleoenvironments, and chronostratigraphy of the Guajira Peninsula, Colombia. *Swiss Journal of Palaeontology*, pages 1–31. xv, 50, 51

Jégu, M. and dos Santos, G. M. (1990). Description d'*Acnodon senai* n. sp. du Rio Jari (Brésil, Amapà) et redescription d'*A. normani* (Teleostei, Serrasalmidae). *Cybium*, 14(3):187–206. 63

Kolmann, M. A., Huie, J. M., Evans, K., and Summers, A. P. (2018). Specialized specialists and the narrow niche fallacy: A tale of scale-feeding fishes. *Royal Society Open Science*, 5(1):1–14. 55

Latrubesse, E. M., Cozzuol, M. A., da Silva-Caminha, S. A. F., Rigsby, C. A., Absy, M. L., and Jaramillo, C. A. (2010). The Late Miocene paleogeography of the Amazon Basin and the evolution of the Amazon River system. *Earth-Science Reviews*, 99(3-4):99–124. 65

Lundberg, J. G. (1997). Freshwater fishes and their paleobiotic implications. In Kay, R. F., Madden, R. H., Cifelli, R. L., and Flynn, J. J., editors, *Vertebrate Paleontology in the Neotropics: The Miocene Fauna of La Venta, Colombia*, chapter 5, pages 67–91. Smithsonian Institution Press. 49, 56, 58, 61

Lundberg, J. G. and Aguilera, O. A. (2003). The late Miocene *Phractocephalus* catfish (Siluriformes: Pimelodidae) from Urumaco, Venezuela: additional specimens and reinterpretation as a distinct species. *Neotropical Ichthyology*, 1(2):97–109. 58, 61

Lundberg, J. G., Marshall, L. G., Guerrero, J., Horton, B., Malabarba, M. C. S. L., and Wesselingh, F. (1998). The stage for Neotropical Fish diversification: A history of Tropical South American rivers. In Malabarba, L. R., Reis, R. E., Vari, R. P., Lucena, Z. M., and Lucena, C. A. S., editors, *Phylogeny and Classification of Neotropical Fishes*, pages 13–48. Edipucrs, Porto Alegre. 61

Lundberg, J. G., Sabaj Pérez, M. H., Dahdul, W. M., and Aguilera, O. A. (2010). The Amazonian Neogene fish fauna. In Hoorn, C. and Wesselingh, F. P., editors, *Amazonia, Landscape and Species Evolution: A Look Into the Past*, pages 281–301. Blackwell Publishing. 49, 52, 60, 61, 63, 65

Mirande, J. M. (2010). Phylogeny of the family Characidae (Teleostei: Characiformes): from characters to taxonomy. *Neotropical Ichthyology*, 8(3):385–568. 55

Monsch, K. A. (1998). Miocene fish faunas from the northwestern Amazonia basin (Colombia, Peru, Brazil) with evidence of marine incursions. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 143(1):31–50. 63

Moreno, F., Hendy, A. J. W., Quiroz, L., Hoyos, N., Jones, D. S., Zapata, V., Zapata, S., Ballen, G. A., Cadena, E., Cárdenas, A. L., Carrillo-Briceño, J. D., Carrillo, J. D., Delgado-Sierra, D., Escobar, J., Martínez, J. I., Martínez, C., Montes, C., Moreno, J., Pérez, N., Sánchez, R., Suárez, C., Vallejo-Pareja, M. C., and Jaramillo, C. A. (2015). Revised stratigraphy of Neogene strata in the Cocinetas Basin, La Guajira, Colombia. *Swiss Journal of Palaeontology*, 134:5–43. xv, 49, 50, 51

Moreno-Bernal, J. W., Head, J., and Jaramillo, C. A. (2016). Fossil crocodilians from the High Guajira Peninsula of Colombia: Neogene faunal change in northernmost South America. *Journal of Vertebrate Paleontology*, 36(3):e1110586. 49

Naranjo, L. G. and Espinel, J. D. A. (2009). *Plan nacional de las especies migratorias. Diagnóstico e identificacion de acciones para la conservacion y el manejo sostenible de las especies migratorias de la biodiversidad en Colombia.* Ministerio de Medio Ambiente de Colombia, Bogotá. 61

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., and Wagner, H. (2019). vegan: Community Ecology Package. 53

QGIS Development Team (2019). QGIS Geographic Information System. Open Source Geospatial Foundation Project. 53

R Core Development Team (2018). R: A language and environment for statistical computing. 53

Ramírez, J. A. and del Valle, J. I. (2011). Paleoclima de La Guajira, Colombia; según los anillos de crecimiento de *Capparis odoratissima* (Capparidaceae). *Revista de Biología Tropical*, 59(2):1389–1405. 64

Reis, R. E. (1998). Systematics, biogeography, and tyhe fossil record of the Callichthyidae: A review of the available data. In Malabarba, L. R., Reis, R. E., Vari, R. P., Lucena, Z. M. S. D., and de Lucena, C. A. S., editors, *Phylogeny and Classification of Neotropical Fishes*, pages 351–362. Edipucrs, Porto Alegre. 63

Rincón, A. D., Solórzano, A., Macsotay, O., McDonald, H. G., and Núñez-Flores, M. (2016). A new Miocene vertebrate assemblage from the Río Yuca Formation (Venezuela) and the northernmost record of typical Miocene mammals of high latitude (Patagonian) affinities in South America. *Geobios*, 49(5):395–405. 58

Rubilar, A. (1994). Diversidad ictiológica en depósitos continentales miocenos de la Formación Cura-Mallín, Chile (37-39S): Implicancias paleogeográficas. *Revista Geologica de Chile*, 21(1):3–29. 63

Schaefer, S. A. (2011). The Andes. Riding the Tectonic Uplift. In Albert, J. S. and Reis, R. E., editors, *Historical Biogeography of Neotropical Freshwater Fishes*, pages 259–278. University of California Press. 49

Silva Santos, R. (1987). *Lepidosiren megalos* n. sp. do Terciario do Estado do Acre; Brasil. *Anais da Academia Brasileira de Ciencias*, 59(4):375–384. 60

Suarez, C., Forasiepi, A. M., Goin, F. J., and Jaramillo, C. A. (2016). Insights into the Neotropics Prior to the Great American Biotic Interchange: New evidence of mammalian predators from the Miocene of Northern Colombia. *Journal of Vertebrate Paleontology*, 36(1):e1029581. 49, 50

Tejada-Lara, J. V., Salas-Gismondi, R., Pujos, F., Baby, P., Benammi, M., Brusset, S., de Franceschi, D., Espurt, N., Urbina, M., and Antoine, P.-O. (2015). Life in proto-Amazonia: Middle Miocene Mammals from the Fitzcarrald Arch (Peruvian Amazonia). *Palaeontology*, 58:341–378. 52, 61

Chapter 4

Fossil fishes from the Sincejelo and Ware formations

4.1 Abstract

Two assemblages of freshwater fossil fishes from northern Colombia are herein studied based on examination of 119 samples from the Sincelejo and Ware formations in the departments of Sucre and Guajira respectively, both of Pliocene age according to the geologic literature. A total of eleven taxa have been identified from the two stratigraphic units, comprising two orders, six families, ten genera and one instance identified to species level. Characters from dental morphology, dorsal- and pectoral-fin spines, and cranial bones are provided as taxonomic tools for present and future study of fossil fishes in the Neogene of South America. All of the taxa herein reported are members of groups currently restricted to drainages east of the Andes, suggesting that physical drainage connection was still present by the Pliocene between the Amazon-Orinoco and trans-Andean drainages such as the Magdalena-Cauca. The phylogenetic position of some taxa further reinforces their membership in cis-Andean groups. The genera *Hemidoras*, *Serrasalmus*, and *Trachelyopterichthys* are new records for the fossil fish fauna of South America. The potential presence of the genus *Rhaphiodon* would add to this list if confirmed in future studies. The genera *Hydrolycus*, *Platysilurus*, and *Zungaro* represent new records for the trans-Andean region. Most of these new occurrences also represent the youngest respective records in the fossil record (except for *Platysilurus*). The genus *Pygocentrus* is also added as a new record based on published information from the middle Miocene La Venta fauna in central Colombia. This combination of taxa suggest that the Sincelejo and Ware formations were deposited in rivers of large size that were part of a large drainage network connected to the Amazon-Orinoco despite being currently located west of the Andes. These findings challenge views where the Cordillera Oriental in Colombia and the Merida Andes in Venezuela were already dividing the drainage network in northern South America by the middle to late Miocene. Instead, they lend support to alternative geologic models as yet not considered mainstream in the literature.

Keywords: Teleostei, Cenozoic, Paleogeography, South America, Andes.

4.2 Introduction

Freshwater fishes in the Neotropics are one of the most speciose vertebrate groups in the world with thousands of species occupying a variety of habitats from level of the sea to above 4000 masl in the Andes (Albert et al., 2011). They are ecologically diverse and play a key role in aquatic habitats in the continent, as well as food resource for human populations (Carolsfeld et al., 2003). This diversity is coupled with a proportional morphological variation as well as a dramatic interval of body size and shape, from miniatures catfishes (Schaefer et al., 1989) to the giant bonytongue *Arapaima* (Castello, 2008). This high diversity at different scales has been the product of a large history of evolutionary change, diversification, and adaptation in South America (Albert et al., 2006, 2011; Lundberg, 1997; Lundberg et al., 1998).

The Andean orogeny, that is, the process on mountain build up and growth has been a key geologic process affecting the biota of South America (). Its effect on the diversification of different living groups has been shown along its entire area of influence from the limit between Chile and Argentina up to Venezuela via creation of new environments, vicariance of pre-existing wider geographic distributions, and the consequent drainage separation in the case of aquatic organisms. The timing of uplift of the Andes, and specially the northern Andes is still highly controversial. Some authors suggest that the Cordillera Oriental in Colombia, and the Mérida Andes and Coastal Cordillera in Venezuela were positive by middle to late Miocene and the Perijá Range and the Mérida Andes during the late Miocene (Diaz de Gamero, 1996; Lundberg et al., 1998, and references thereir). In contrast, some authors suggest a more complex tectonic history for the Andes of NSA, arguing that the Cordillera Oriental in Colombia uplifted in pulses, with some areas being positive as early as late Paleocene (e.g., Santander Massif; Bayona et al., 2013) and with other areas uplifting from middle Eocene to middle Miocene (e.g., Perijá Range and central Cordillera Oriental; Ayala et al., 2012; Bayona et al., 2013, 2010; Caballero et al., 2010; Ochoa et al., 2012). Both scenarios would have different consequences for the freshwater fish faunas (both extant and extinct), and therefore, they could be used to better understand the evolution of the northern Andes.

A preliminary paleoenvironmental reconstruction suggests that the Peninsula was a tropical forest with presence of either middle to large rivers or a deltaic area, aquatic vertebrates such as crocodylians, turtles, freshwater fishes, as well as terrestrial and amphibian mammals (Aguilera et al., 2013a; Amson et al., 2016; Cadena and Jaramillo, 2015a,b; Carrillo et al., 2018; Forasiepi et al., 2014; Moreno-Bernal et al., 2016; Suarez et al., 2016). Such paleobiotic assemblage allows to explore past drainage connections with major drainages of NSA because of its intermediate location between the Magdalena and Maracaibo drainages,

as well as being complementary to other fossil faunas of similar age (i.e., San Gregorio Formation in Venezuela). The main goal of this study is to study the freshwater fish assemblage of the Sincelejo and Ware formations in the departments of Sucre and Guajira during the Pliocene and to discuss their relevance for paleogeographic models by using past and present geographic distribution of the taxa recovered.

4.3 Materials and Methods

4.3.1 Abbreviations and comparative material

Preserved specimens of extant species of the orders Characiformes and Siluriformes were examined in scientific collections (Appendix B). A total of five fossil samples were examined from the Sincelejo formation, and 116 from the Ware formation. Institutional abbreviations are: Academy of Natural Sciences of Drexel University, Philadelphia, US (ANSP), Instituto Alexander von Humboldt, Villa de Leyva, Colombia (IAvH), Museu de Zoologia da Universidade de São Paulo, São Paulo, Brazil (MZUSP), Mapuka Museum of Universidad del Norte, Barranquilla, Colombia (MUN). Premaxilla and dentary are abbreviated PM and D respectively.

4.3.2 Anatomical terminology

Siluriform osteological terminology follows Lundberg and McDade (1986) with refinements by Slobodian and Pastana (2018). Siluriform appendicular terminology follows Ballen and de Pinna (in prep., see Chapter 2). Cynodontid cranial nomenclature follows Toledo-Piza (2000). Serrasalmid tooth nomenclature and position follows Cione et al. (2009) with modifications (Figure 3.2).

4.3.3 Data analysis

Fossil freshwater fish assemblages are known from Colombia and Venezuela to Argentina (Lundberg et al., 2010). In order to understand patterns of similarity, an analysis of assemblage composition was compiled from literature data and direct observations (Aguilera et al., 2013a,b; Antoine et al., 2016; Azpelicueta and Cione, 2016; Ballen and Moreno-Bernal, 2019; Bogan et al., 2012; Cione and Azpelicueta, 2013; Cione et al., 2000, 2009; Lundberg et al., 2010; Tejada-Lara et al., 2015). An initial set of Miocene fossil fish faunas including both marine and freshwater components, plus any number of fossil occurrences, was included. this initial assemblage was then reduced to a set where the number of taxa was equal or larger than the one herein described. Although the original dataset includes fossil assemblages of strictly freshwater, strictly marine, or mixed origin, only those with freshwater components were included in the final analysis. The final dataset include the Acre, Contamana, La Venta, Fitzcarrald, Ituzaingó, Makaraipao, Urumaco, and Ware fossil assemblages along with their

fossil fish occurrences. The overall results using all the data or just the subset described above did not change the conclusions of the similarity analysis. Faunal dissimilarity was measured using the Bray-Curtis coefficient as implemented in the `vegan` package v.2.5-5 (Oksanen et al., 2019) in `R` v.3.4.4 R Core Development Team (2018).

4.3.4 Phylogenetic analyses

The phylogenetic position of some taxa herein described was estimated in order to further assess their bearing on distribution patterns. Parsimony analysis was used in the case where only morphological data were available, while bayesian inference was used in when both molecular and morphological data were available. The main reason for this analytical choice is that bayesian inference allows for a more realistic and flexible model of evolution than parsimony, while the latter has the advantage of also providing information on diagnostic features in the form of synapomorphies, reinforcing the morphological evidence involved in the positioning of terminals of interest in the analysis.

Parsimony analysis

The position of fossil remains of *Hydrolycus* was assessed through parsimony analysis. A parsimony reanalysis of the morphological matrix with 72 characters published in Toledo-Piza (2000) including 11 new characters proposed herein for a total of 83 characters (Appendix D). The roestines were not included in the present analysis following Mattox and Toledo-Piza (2012) who allocated the subfamily as a tribe of the Heterocharacinae in the Characidae. A total of seven species (*Cynodon gibbus*, *C. septenarius*, *Hydrolycus armatus*, *H. scomberoides*, *H. tatauaia*, *H. wallacei*, and *Rhaphiodon vulpinus*) were included in the ingroup, while *Acestrorhynchus* and Toledo-Piza's composite outgroup (i.e., including *Acanthocharax*, *Acestrorhynchus*, *Agoniates*, *Boulengerella*, *Brycon*, *Carnegiella*, *Charax*, *Ctenolucius*, *Erythrinus*, *Galeocharax*, *Gasteropelecus*, *Gnathocharax*, *Hepsetus*, *Heterocharax*, *Hoplerythrinus*, *Hoplias*, *Hoplocharax*, *Hydrocynus*, *Lebiasina*, *Lonchogenys*, *Oligosarcus*, *Pyrrhulina*, *Roeboexodon*, *Roeboides*, and *Xenocharax*) comprise the outgroup.

Parsimony analysis was carried out using TNT v.1.5 (Goloboff et al., 2008). Nodal support values (i.e., bootstrap and jackknife values) were calculated using 10000 replications. Consistency and retention indexes were calculated using the script STATSALL.run v.1.3. TNT is freely available at http://www.lillo.org.ar/phylogeny/tnt/. Given the number of taxa we carried out an exact search with the characters unordered and rooting at Toledo-Piza's outgroup composite taxon. Characters of interest were mapped using Mesquite v.3.03, available at http://mesquiteproject.wikispaces.com/ (Maddison and Maddison, 2011). Additional tree formatting was carried out in R v.3.3 (R Core Development Team, 2018), available at http://www.r-project.org. The dataset (in tnt format) and scripts (in `R`, `bash`, and `TNT`) for reproducing the analysis are available in Appendices D and E.4.

Bayesian analysis

A phylogenetic analysis of the family Pimelodidae was carried out including the specimens of the genus *Phractocephalus* from the Sincelejo formation combining morphological and molecular data under a bayesian framework aiming at answering two basic questions: 1) Which is the phylogenetic position of these fossil remains, and 2) what information on distribution can we gather from such phylogenetic pattern. We analyzed a combined molecular and morphological dataset with 84 terminals representing 24 genera of the Pimelodidae and two genera of the Pseudopimelodidae and the Heptapteridae respectively, while the fossil taxon set was composed of the three known extinct species of *Phractocephalus* as well as the specimens herein studied, for a total of four fossil terminals. DNA sequences for the molecular partitions were taken from Lundberg et al. (2011), who included the single-copy nuclear recombination activating genes (*rag1* and *rag2*), a mitochondrial region comprising the *12S* rRNA, tRNA-val and *16S* rRNA genes, and the cytochrome-b *cytb* gene (including Threonine tRNA and partial Proline tRNA regions), for a total of four molecular partitions. All available sequences were retrieved from GenBank and aligned with `MAFFT` v.7.271 (Katoh and Standley, 2013) using the `G-INS-i` algorithm with 1000 iterations except for *rag1* where we used the `L-INS-i` algorithm due to the unequal size of sequences in this partition. Minor manual adjustments were carried out for the *rag1* and *rag2* alignments. Additionally, we used `Gblocks` v.0.91 (Castresana, 2000; Talavera and Castresana, 2007) for reproducible exclusion of ambiguous and hypervariable positions of the *12S* alignment with the following settings: `-t=d -b=a -d=y`; a total of 84 (4%) positions were removed from the original alignment. Accession numbers are available in Section C.2.2; missing data were coded as "`?`". A total dataset with 84 terminals and 7578 characters was assembled for analysis.

Phylogenetic analysis was performed using Bayesian Inference as implemented in `MrBayes` v.3.2.6 (Ronquist et al., 2012). Substitution models for each partition were selected using `JModeltest` 2.1.10 v.20160303 (Darriba et al., 2012) and `PhyML` 3.0 v.20131022 (Guindon et al., 2010) using the Bayesian Information Criterion (BIC) following Darriba et al. (2012). The best-fit substitution models were GTR+I+Γ (*rag1*); K80+I+Γ (*rag2*); GTR+I+Γ (*12S*), and HKY+I+Γ (*cytb*). Parameters other than topology were unlinked across partitions. Two runs with eight independent Markov chains were run in parallel for 2.000.000 generations and sampling every 2.000 generations. Convergence of runs was determined based on the average standard deviation of the split frequencies (ASDSF) < 0.01, and the effective sample size (ESS), calculated using `Tracer` v1.6.0 (Drummond et al., 2012), that was > 300 for all parameters. Additionally, the potential scale reduction factor (PSRF) approached 1.0, suggesting convergence in the estimation of the posterior probabilities of nodes and branch length parameters. The posterior density graphs of the two independent runs were also examined in `Tracer`, and not visual differences were found between them. The 25% of trees were discarded as burn-in and a 50% majority-rule tree and posterior probabilities (PP) for

node support were calculated using the remaining trees. The dataset (in nexus format) and scripts (in `R`, `bash`, and `MrBayes`) for reproducing the analysis are available in Appendices C and E.2.

4.3.5 Geological Setting

The fossil localities from the Ware and Sincelejo formation herein studied belong to two geologic regions of northern Colombia respectively: the Cocinetas sedimentary basin and the San Jacinto tectonic belt. Both regions show an important sedimentary sequence of spanning the Cenozoic, although cretacic stratigraphic units are sometimes preserved. Their paleoenvironments are diverse and record important events in the geologic evolution of northern South America.

The Cocinetas sedimentary basin

The Cocinetas sedimentary basin is composed of units recording the environmental dynamics during the Eocene to Pliocene timespan (Renz, 1960; Rollins, 1965). It is bounded by the Macuira Fault to the northeast, the Cuisa Fault to the southwest, and the Serranía de Jarara to the northwest. The basin genesis has been associated to regional tectonics involving the migration of tectonic blocks and the subsequent formation of pull-apart basins along the northern margin of South America (Moreno et al., 2015). It is located to the east of the Guajira Peninsula in northern Colombia, along the international border with Venezuela.

Historically it has been composed of five stratigraphic units that have suffered strong redefinitions in the literature. Renz (1960) recognized six formations, the Guasare formation of Paleocene consisting of drown limestones, an unnamed unit of late Eocene age consisting of conglomerate and limestone, three units spanning the middle to late Oligocene (Uitpa, Siamana, and Jimol formations), and finally the Tucacas formation of early Miocene age. A preliminary assessment of their fossil content (mostly mollusks and foraminifera) along with their stratigraphic position was used in order to infer the chronology of the sedimentary succession. Later on Rollins (1965) redefined the units recognizing the Guasare formation and the unnamed late Eocene succession of Renz as synonyms of the Macarao formation and restricting the age of the succession to the Eocene. The same author suggested the age of the Siamana formation to be early to middle Oligocene, refined the age of the Uitpa formation as late Oligocene to early Miocene, and that of the Jimol formation as middle Miocene, finally renaming the Tucacas formation as Castilletes formation and assigning it an age of late Miocene to Pliocene. Moreno et al. (2015) restudied the stratigraphy of the Cocinetas basin with an emphasis on the Neogene units (the Jimol and Castilletes formations of Rollins), further refined the age of all units, and documented the rich fossil assemblages other than marine mollusks of these Neogene units. The age of the Siamana formation was further constrained to the late-middle Oligocene to the Oligocene-Miocene boundary; the Uitpa and Jimol formations were restricted to the early Miocene, while the Castilletes formation

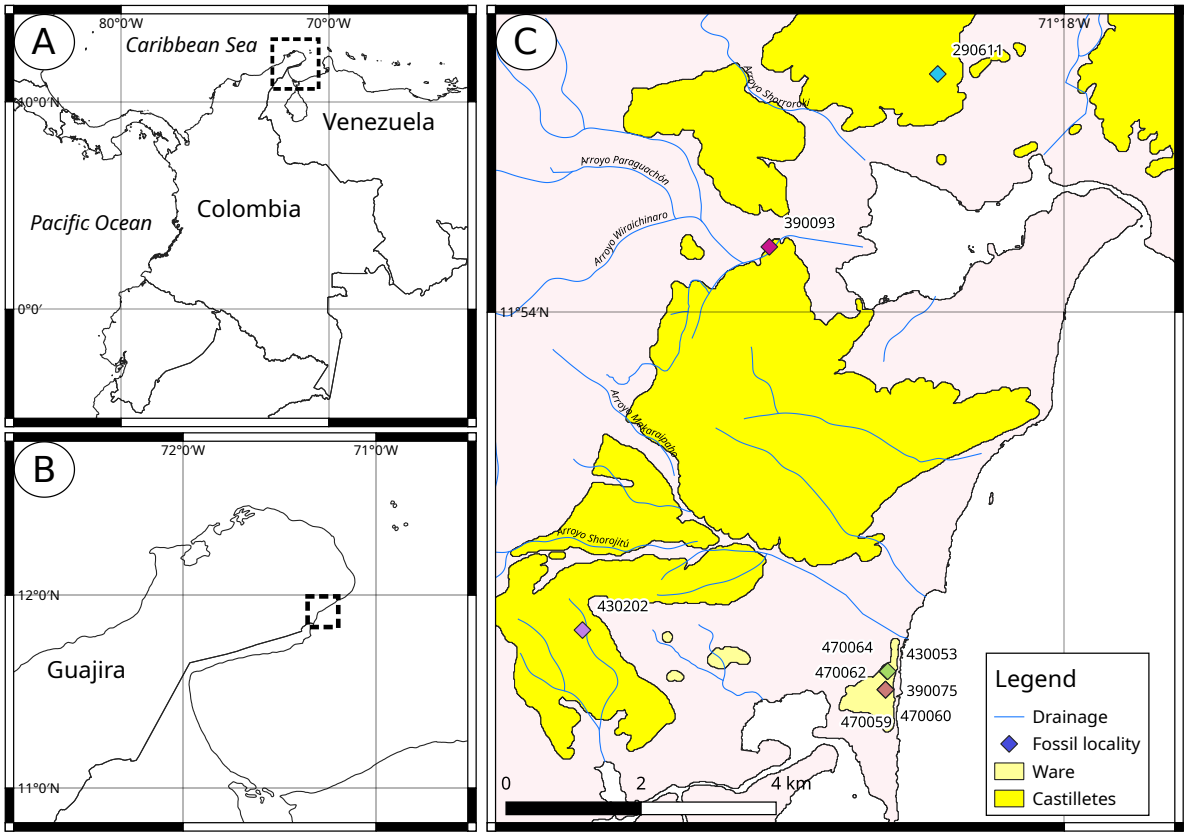


Figure 4.1: *Cartography of the Guajira Peninsula.*

was restricted to the middle Miocene. The Ware formation was described for the uppermost portion of the Castilletes formation of Rollins, and a Pliocene age proposed for the new unit.

As currently defined, the Macarao formation comprises a succession of limestone, dark sandstone, and gray clays, it contains abundant mollusks and was deposited in shallow marine conditions (Hendy et al., 2015); the Siamana formation preserves coral reefs in a marine succession (Flórez et al., 2018, 2019); the Uitpa formation contains clays and shales, with shark fossil remains (Carrillo-Briceño et al., 2016). The Jimol formation is a marine succession where sandstone packages are more prominent and competent than in the overlying sediments; it records shallow marine conditions with marginal environments, as suggested by the presence of both continental and marine fossils (Carrillo et al., 2018; Flórez et al., 2018, 2019; Hendy et al., 2015; Moreno-Bernal et al., 2016). The Castilletes formation is a succession of mostly mudstone interleaved with some levels of limestone and conglomerates; oyster banks are present in the middle segment levels of coquina of local extent and lateral continuity are present, a rich assemblage of fossil vertebrates have been documented from both marine and continental communities (Aguilera et al., 2017, 2013b; Amson et al., 2016; Cadena and Jaramillo, 2015a,b; Carrillo-Briceño et al., 2019; Forasiepi et al., 2014; Moreno-Bernal et al., 2016; Suarez et al., 2016, Ballen and de Pinna, in prep., see Chapter 3).

The Ware formation is a continental succession bounded by an uncomformity with the

Castilletes formation to the base and a coquina of regional extent to the top; the preserved succession is generally thin (ca. 25 m) but can be followed regionally to the east of the Cocinetas basin. The Ware formation is a succession of medium to coarse sandstone intercalated with fine levels of mudstone and some prominent levels of conglomerate and conglomeratic sandstone, a strong coquina marks finishes the sedimentary record at the top; the Ware formation preserves a rich vertebrate assemblage (Amson et al., 2016; Carrillo et al., 2018; Carrillo-Briceño et al., 2019; Forasiepi et al., 2014; Moreno-Bernal et al., 2016) (Table 4.1; Figure 4.3). The Ware formation has been dated as Pliocene based on radiometric data from Sr isotopes (Hendy et al., 2015). Although Aguilera et al. (2013a) reported Catfishes of the families Doradidae and Pimelodidae from the upper Castilletes formation (later named Ware formation), most of the freshwater fishes are for the first time herein studied.

The San Jacinto tectonic belt

The San Jacinto tectonic belt is an accretionary feature that forms the Cordillera de San Jacinto, a mountain range to the west of the Magdalena-Cauca drainage. It is bound along its east and west margins by the Sinú and Romeral lineaments (Duque-Caro, 1984). This structure and the associated Sinú belt comprise the Sinú–San Jacinto terrane that was formed by compressional stress between the Caribbean Oceanic and the South American crusts, forming north–south folds. It has been subject to a sequence of uplift events during the Cenozoic, with an alternation of unconformities and mostly marine sedimentary units, while the uppermost one is composed of terrestrial sediments. Duque-Caro (1984) calls "Sincelejan Stage" to the association of continental sediments of Pliocene age around Sincelejo between the pre-Late Pliocene and Pleistocene tectonic emergences affecting the San Jacinto tectonic belt. These two events are evidenced by the uncomformities bounding the continental sediments of the Sincelejo formation. The geology of this region has a complex history since the first field surveys in the region by Beck (1921) and Werenfelds (1926). Posterior refinements from both the petroleum industry and the Colombian geological survey (de Porta, 1962; Dueñas and Duque-Caro, 1981; Duque-Caro, 1966, 1967, 1984) have generated the bulk of information on this region and its stratigraphy, leaving a publication record that despite being at first confusing, has been reviewed in the past decades and now allows a better understanding of the stratigraphy in the area of study (Alfaro and Holz, 2015; Bacca et al., 2010; Bermúdez et al., 2009; Duque-Caro, 1984). Ongoing efforts to revise the geology and refine the context of the fossil record by researchers at the Universidad del Norte anticipate further contributions concerning the continental sequence of the San Jacinto tectonic belt in wider structural, paleogeographic, and stratigraphic contexts (F. Lamus, com. pers.).

Stratigraphy of the continental units—Werenfelds (1926) proposed the name Sincelejo sandstones for a sequence of sandstone and conglomeratic sandstone on the upper part of this sequence, resting on the Savana sandstone. Later, Kassem et al. (1964) proposed the name Morroa for the uppermost conglomeratic portion with massive sandstones below fine-grained

sediments of the Betulia formation. Dueñas and Duque-Caro (1981) raised the Sincelejo formation to the rank of group and included in it the Sincelejo, Morroa, and Betulia formations; although these were not distinguishable in their study area (the cartographic quadrat F-8 in p.9, from about Planetarica to the south to Ciénaga de Oro to the north, and the Serranía de San Jerónimo to the west to the alluvial plains of the Rio San Jorge to the east), according to the authors, the units are thicker and better differentiated to the north, in the vicinity of Sincelejo (the topographic cell E-8, p.9). Clavijo-Torres and Barrera-Olmo (2001) revised extensively the stratigraphy and structural geology of the cartographic quadrats 44 and 52 in the San Jacinto belt. The Sincelejo formation is said to rest unconformably on the Carmen and San Jacinto formations, while it is apparently planar or in maybe angular unconformity over the Cerrito formation. From the mapped units in cartographic quadrat 52, the Sincelejo rests on the Cerrito formation in the vicinity of the Sincelejo-Corozal area of the fossil localities herein studied (Figure 4.2). The Sincelejo formation was subdivided by Clavijo-Torres and Barrera-Olmo in three subunits: The lower Sincelejo, the upper Sincelejo, and the Morroa members. Bacca et al. (2010) reiterates in a more succinct form the description of the Sincelejo formation already presented by Clavijo-Torres and Barrera-Olmo (2001). According to the author, the Sincelejo formation correlates to the Corpa formation of the Sinú basin (See figure 3 in Bacca et al., 2010). Bermúdez et al. (2009) described the stratigraphy and micropaleontological content of its units, along with paleoenvironmental reconstructions for the marine-related units. They recognize the Sincelejo at formation rank and assign it an age of Pliocene. The lower contact is said to be discordant on rocks of the Ciénaga de Oro–El Floral formations as evidenced from well logs, none of which recovered sediments of the Cerrito and San Jacinto formations between the Sincelejo and the Ciénaga de Oro–El Floral formations, older than the sequence Cerrito to San Jacinto. The Carmen formation of Clavijo-Torres and Barrera-Olmo (2001) is a synonym of the El Floral formation (Bermúdez et al., 2009, figure 2).

Fossil assemblage of the Sincelejo formation—The Sincelejo formation has long been known to preserve fossil vertebrates; however, very few taxa have since been reported in press, and they were restricted to mammals until the present contribution. The first mention of fossil vertebrates from the Sincelejo formation was the report of fossil remains from the Corozal–Sincelejo area by Royo y Gomez (1946, p.499) of reptile bone fragments and a tooth of the Rodent later described as *?Gyriabrus royoi* by Stirton and R. (1947) from the Sierra Peñata locality of the (?) San Antonio sandstone. Stirton (1953) acknowledges the complex stratigraphic of the area, where the names Savana, Sincelejo, and San Antonio have been applied to sandstone facies cropping out in the vicinity of Sincelejo, all with a poor understanding of lateral extent, age, and regional correlation. Stirton however, argues that these sandstone beds might be rather recent based on evolutionary trends in tooth morphology of the Dinomyidae. de Porta (1962), Duque-Caro (1966), and Duque-Caro (1967) focused on the marine facies underlying the Sincelejo formation and studied extensively the molluscs and microfossils, mentioning only succinctly the presence of vertebrates in the

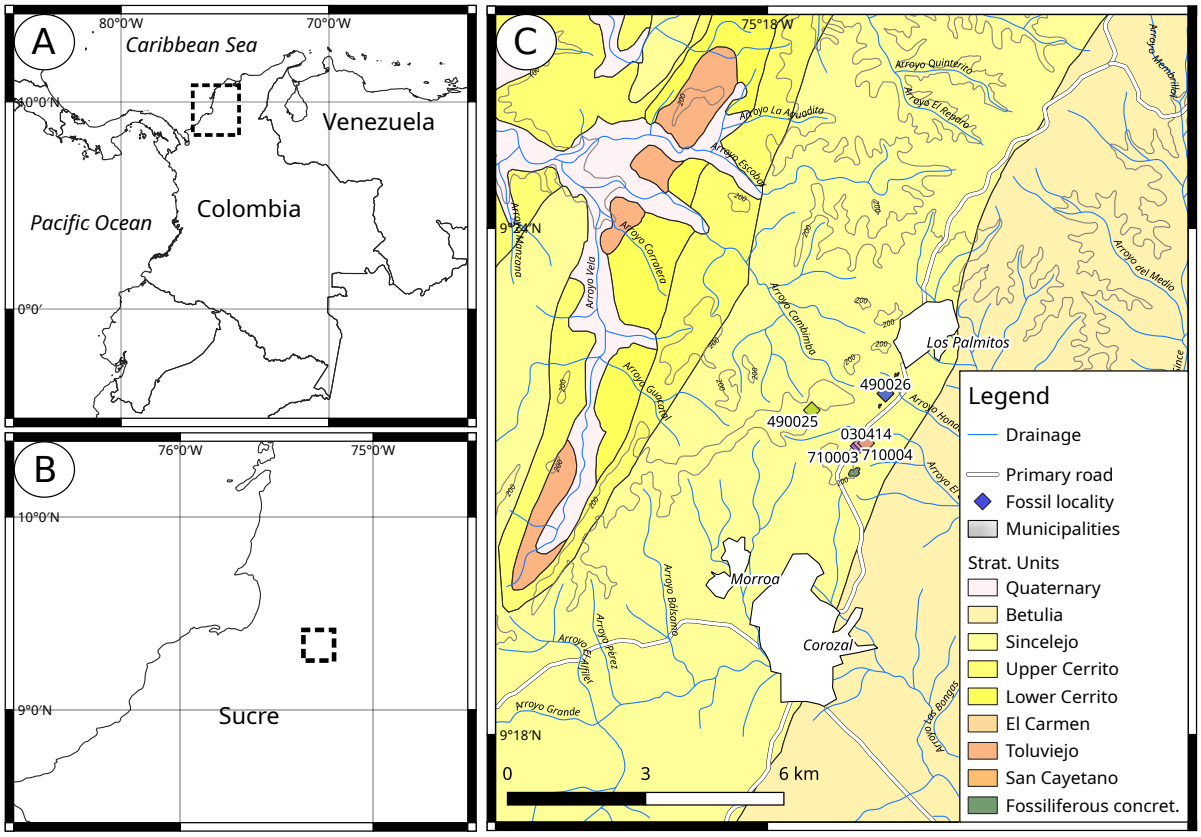


Figure 4.2: *Cartography of the Corozal area.*

continental facies already reported in publications; their descriptions of geographic features make it an important reference for referencing geological features and localities in this area, where the official cartography seems to lack some detail. All these references consistently avoid the nomenclatural problem of the stratigraphy in the region by referring to the sediments above the Cerrito formation with the general term "continental series". Marshall et al. (1983) mention the fossil record of the Sincelejo area as the Sierra Peñata fauna, and reports the presence of a Toxodont citing de Porta (1961) as the source. The latest review of the fossil assemblages of northern Colombia was carried out by Villarroel and Clavijo (2005), who also described the Glyptodont *Neoglyptatelus sincelejanus* from a locality 6 km south of Sincejelo in the Calle Fría–Segovia road. Villarroel and Clavijo also reported indeterminate remains of a Toxodont from 2 km to the north of Los Palmitos, in the vicinity of Corozal, and the Toxodont cf. *Trigodonops* from sandstone beds where the town of Corozal was built. These authors recognize with some uncertainty the Sincelejo formation as the source of the indeterminate Toxodont and cf. *Trigodonops* mentioned above, and an uncertain unit below the Sincelejo formation as the source of *Neoglyptatelus sincelejanus* and *?Gyriabrus royoi* (see figure 3, p.354), also challenging the local stratigraphy of previous works on the basis of field observations. Villarroel and Clavijo however, did not cite Clavijo-Torres and Barrera-Olmo (2001) and therefore can not be correlated directly to the latter work that has been followed by later authors in recognizing the Betulia, Sincelejo, and Cerrito as distinct

formations with more or less clear stratigraphic boundaries.

Fossil localities

Most of the vertebrate localities from the Ware formation are very spatially restricted, occurring in a small area in the stratotype of the unit. The main difference between localities is the fine-scale stratigraphic provenance, sometimes to fractions of stratigraphic meters. Several localities have been named differently when in fact they are synonyms as the main criterion for locality designation in the field was stratigraphic position. Most of the explicit, textual information about stratigraphic position of the different STRI localities come from Amson et al. (2016) and Carrillo et al. (2018), while other references have either cited any of these sources, the original paper by Moreno et al. (2015), or stated the STRI sample online database as online repository of provenance information.

A compendium of published localities in the stratotype of the Ware formation is provided in Table 4.1, along with a synonymy of the stratigraphic section (Figure 4.3). As the original catalog information sometimes suffers from synonymous locality labeling and the ultimate criterion for locality naming is stratigraphic position, some of the original localities from the same fossil level bear two locality names. I retain the original locality labels as they are expected to remain attached to the associated provenance data and field notes from different collectors (e.g., Figure 4.18). Since online repositories may or may not be available in the future, the present compendium of locality information is expected to serve as permanent future reference.

Table 4.1: *STRI localities from the Ware formation referenced in different sources since 2013. Several localities refer to the same level and spatial point, being therefore synonyms.*

Locality	Strat (m)	Taxon	Source	Comments
290045	NA	NA	Moreno et al. (2015)	From supplementary spreadsheet
290466	NA	NA	Moreno et al. (2015)	From supplementary spreadsheet
390017	5	*Pliomegatherium lelongi*	Amson et al. (2016); Carrillo et al. (2018)	In referred material of Amson et al. without strat position; strat position said to be in the stri database. In table 13 of Carrillo et al.
390018	4.5	Proterotheriidae indet.	Carrillo et al. (2018)	In table 13

Continued on next page

Table 4.1 – *Continued from previous page*

Locality	**Strat (m)**	**Taxon**	**Source**	**Comments**
390020	5	Proterotheriidae indet., Toxodontinae indet.	Carrillo et al. (2018)	In table 13
390022	NA	Tardigrada indet.	Amson et al. (2016)	In referred material without strat position. Strat position said to be in the stri database
390023	3.5	Lestodontini indet.	Amson et al. (2016); Carrillo et al. (2018)	In Amson et al. in referred material without strat position; strat position said to be in the stri database. In table 13 of Carrillo et al.
390024	4.5	cf. *Nothotherium*, *Pliomegatherium lelongi*	Amson et al. (2016); Carrillo et al. (2018)	In referred material in Amson et al. without strat position; strat position said to be in the stri database. In table 13 of Carrillo et al.
390025	4.5	Scelidotheriinae indet.	Amson et al. (2016); Carrillo et al. (2018)	In Amson et al. in referred material without strat position; strat position said to be in the stri database. In table 13 of Carrillo et al.
390026	5	*Pliomegatherium lelongi*	Amson et al. (2016); Carrillo et al. (2018)	In referred material in Amson et al. without strat position; strat position said to be in the stri database. In table 13 of Carrillo et al.
390075	4	Elasmobranchii	Carrillo-Briceño et al. (2019)	Locality and position previously discussed in Ballen et al. (in prep). Supplementary table S1 of Carrillo-Briceño, no strat position

Continued on next page

Table 4.1 – *Continued from previous page*

Locality	**Strat (m)**	**Taxon**	**Source**	**Comments**
390077	1.7	Elasmobranchii, Megalonychidae sp. nov.	Amson et al. (2016); Carrillo et al. (2018); Carrillo-Briceño et al. (2019)	In referred material of Amson et al. without strat position; said to be in the stri database. In table 13 of Carrillo et al.. In supplementary table S1 without strat position of Carrillo-Briceño et al.
390080	5	Elasmobranchii	Carrillo-Briceño et al. (2019)	Supplementary table S1 of Carrillo-Briceño et al., no strat position
390081	21	NA	Moreno et al. (2015)	From supplementary spreadsheet
390083	NA	Elasmobranchii	Carrillo-Briceño et al. (2019)	Supplementary table S1, no strat position
430052	5	Proterotheriidae indet., Elasmobranchii	Carrillo et al. (2018); Carrillo-Briceño et al. (2019)	In table 13 of Carrillo et al.. Supplementary table S1 in Carrillo-Briceño et al., no strat position
470059	1,7	NA	Moreno et al. (2015)	From supplementary spreadsheet
470059	2	Elasmobranchii, Toxodontinae indet.	Carrillo et al. (2018); Carrillo-Briceño et al. (2019)	In table 13 of Carrillo et al. Supplementary table S1 of Carrillo-Briceño et al., no strat position
470060	4	Camelidae indet., Lestodontini indet., Megalonychidae sp. nov., Proterotheriidae indet., Toxodontinae indet.	Amson et al. (2016); Carrillo et al. (2018)	In referred material of Amson et al. without strat position; strat position said to be in the stri database. In table 13 in Carrillo et al.

Continued on next page

Table 4.1 – *Continued from previous page*

Locality	**Strat (m)**	**Taxon**	**Source**	**Comments**
470061	4.3	*Chapalmalania* sp., Proterotheriidae indet., Toxodontinae indet.	Carrillo et al. (2018); Forasiepi et al. (2014)	In table 13 of Carrillo et al. Lower Ware formation in the text, 4.3 in geological setting of Forasiepi et al.
470062	4	NA	Moreno et al. (2015)	From supplementary spreadsheet
470062	5	Crocodylidae indet., *Crocodylus* sp., Elasmobranchii, Mylodontidae indet., Proterotheriidae indet.	Amson et al. (2016); Carrillo et al. (2018); Carrillo-Briceño et al. (2019); Moreno-Bernal et al. (2016)	In referred material of Amson et al. without strat position, strat position said to be in the stri database. In table 13 of Carrillo et al. Supp table S1 of Carrillo-Briceño et al. In Moreno-Bernal et al. said to be Police station, no explicit strat position, although the illustrated position in figure 1C points to just below the third conglomerate of the stratotype
470064	NA	Elasmobranchii	Carrillo-Briceño et al. (2019)	Supplementary table S1, no strat position
600001	NA	NA	Moreno et al. (2015)	From supplementary spreadsheet
NA	ca. 4.8	*Brachyplatystoma* cf. *vaillantii*	Aguilera et al. (2013a)	Said to be from the upper Castillets = Ware. Without strat position data, without STRI locality data
NA	ca. 4.8	Doradidae indet.	Aguilera et al. (2013a)	Said to be from the upper Castillets = Ware. Without strat position data, without STRI locality data

4.4 Results

4.4.1 Systematic Paleontology

Sincelejo Formation

Teleostei
Ostariophysi
Characiformes
Anostomidae
Genera *Leporinus* or *Hypomasticus*
Gen. et. sp. indet.
Figure 4.4.

Material examined—Locality STRI 710004, GAB-P 415, isolated symphysial dentary tooth.

Description—Isolated tooth lacking strong asymmetry, unicuspid, without strong distal projection. Crown elongate, not wider than long. Occlusal surface striate somewhat inverted triangular in outline.

Remarks—The isolated tooth can be assigned with confidence to family and has been further identified as a member of either *Leporinus* or *Hypomasticus* (as defined by Sidlauskas and Vari, 2008). Those genera consist of strictly cis-Andean species, so this record further reinforces a relationship between the Magdalena-Cauca drainage with cis-Andean drainages by the Pliocene. The genus *Megaleporinus* was recently separated from *Leporinus* to accommodate ten species formerly considered as either *Leporinus* (9 spp.) or *Hypomasticus* (1 sp.) (Ramirez et al., 2017). Species of *Megaleporinus* are mostly cis-Andean with one member found in trans-Andean drainages (*Megaleporinus muyscorum*), and have dentary teeth that are very wide, semispherical, and spoon-shaped (e.g., fig. 2, Ramirez et al., 2017) in contrast to the fossil specimen which is clearly elongate and not spoon-shaped. *Hypomasticus* is a problematic genus whose delimitation is still confusing, and *Leporinus* is a large genus whose alpha-taxonomy has improved considerably in the last few decades. Further study of recent material of those genera as well as a better understanding of their respective delimitation are needed in order to evaluate the relationships of the specimen.

Teleostei
Ostariophysi
Siluriformes
Pimelodidae
***Phractocephalus* sp.**
Figure 4.6.

Material examined—Locality STRI 710004, MUN 41058, MUN 43679, pectoral spine fragments.

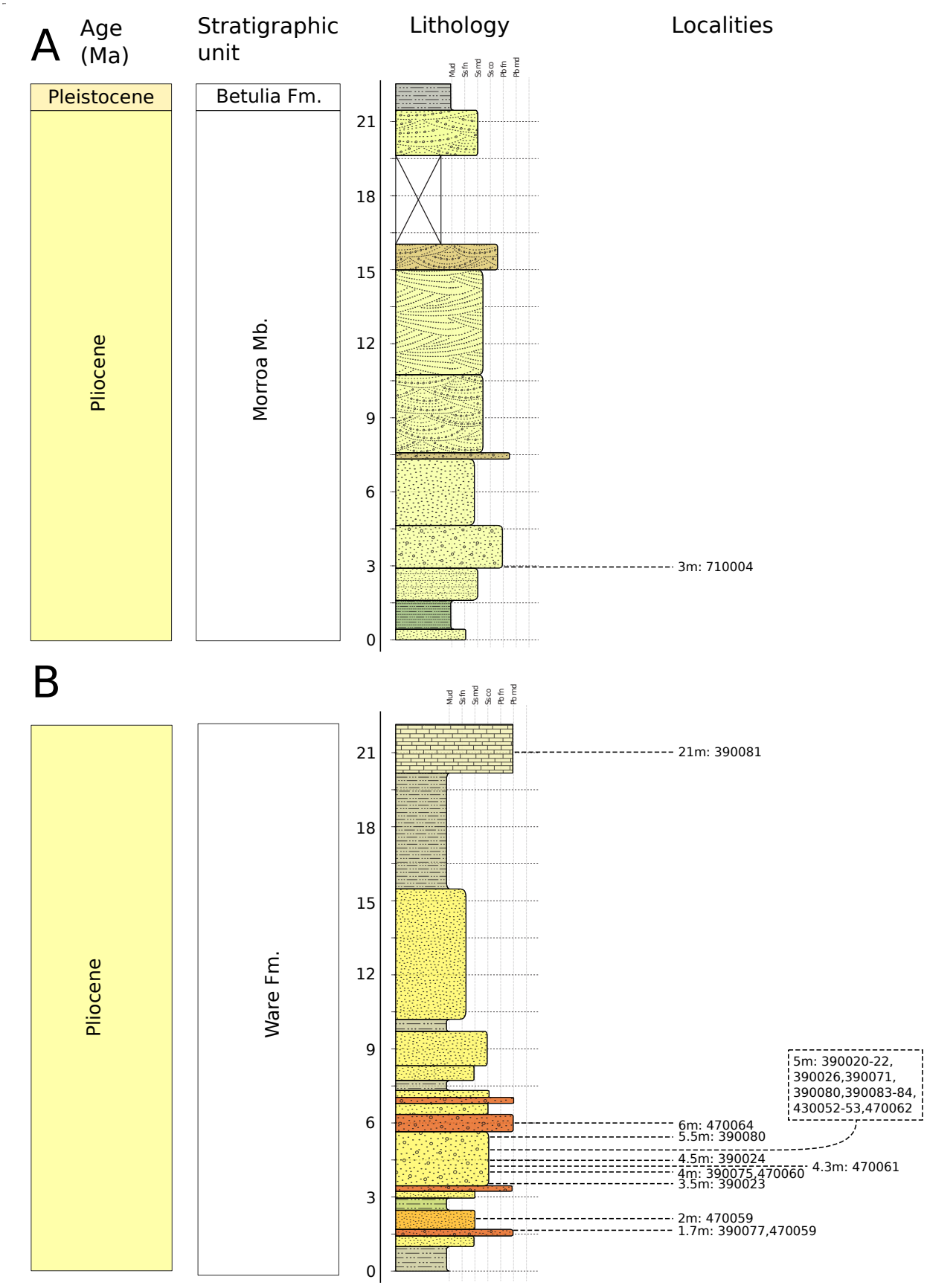


Figure 4.3: *Stratigraphic position of the fossil specimens herein studied. A) San Francisco farm section. B) Ware formation stratotype. Staff scales in stratigraphic meters. Vertical guides are granulometry; Mud = mud, Ss = sand, Pb = pebbles, fn = fine, md = medium, co = coarse. Localites are labeled as with the stratigraphic position in meters from base of the sequence and then the list locality numbers. Color reflects the one recorded in the outcrop. Stratigraphic columns modified from Moreno et al. (2015) and Montes et al. (in prep.)*

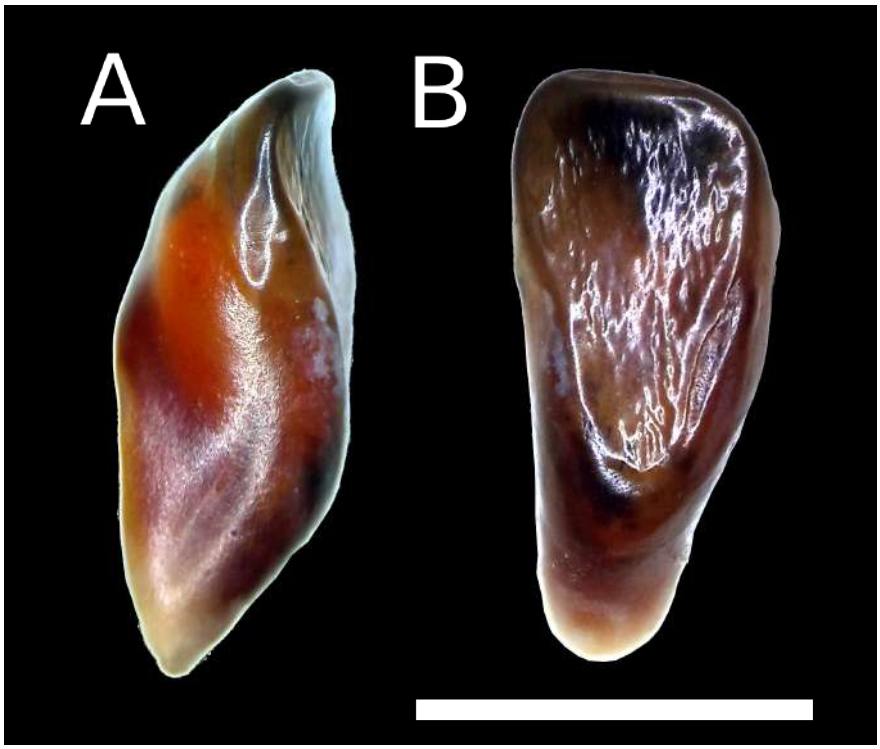


Figure 4.4: *Anostomid tooth from the Sincelejo formation. A) Lateral view. B) Lingual view. Scale bar equals 5 mm.*

Description—MUN 41058 preserves the proximal portion of the right pectoral-spine shaft with dorsal, ventral, anterior and posterior ornament preserved, with proximal exposition of the medullar region. Dorsal and ventral ornament consisting of reticulate ridges with ovoid pits; dorsal and ventral ornament appear to be equally well developed. Anterior ornament consisting of straight, vertically-elongate ornament with some tips worn. Posterior ornament consisting of retrorse, oblique spinules directed somewhat ventrally; the ornament disposition is typical of that of ornaments on the basal portion of the spine where the fin membrane tends to insert dorsal to the ornaments, making them to direct slightly to the ventral portion of the spine shaft. Even when the posterior ornament is retrorse in curvature, they tend to be displaced forward, a condition typical of basal ornaments in catfishes of the Pimelodidae. Medullar region in cross-section spongy with a small and oval lumen. MUN 43679 preserves only the anterior portion of the middle region of the right pectoral-spine shaft with dorsal and ventral reticulate ridges; anterodorsal and anteroventral margins smooth; anterior ornament consisting of straight, vertically-elongate ornament, with tips eroded due to transport.

Remarks—The phylogenetic analysis of combined molecular and morphological data for the family Pimelodidae indicates that the specimens of *Phractocephalus* sp. herein described lie in a clade along with *P. nassi* and *P. ivy*, the two extinct species of *Phractocephalus* (0.75), while *P. acreornatus* is recovered as basal to that trichotomy (0.96); the extant species *P. hemioliopterus* is to the most basal member of the genus (Figure 4.5). Two large traditional groups recognized informally for the family were recovered as monophyletic in our

analysis (the so-called "sorubimines" and the "OCP clade" of Lundberg et al., 2011), while *Steindachneridion*, *Phractocephalus*, and *Leiarius* were found to be successive groups to the base of the Pimelodidae. Overall the phylogenetic analysis shows high nodal support values with posterior probabilities above 0.95 in almost all cases; nine out of 80 nodes showed posterior probabilities between 0.5 and 0.9. These results place the specimens form the Sincelejo within *Phractocephalus* with high confidence and corroborate earlier hypotheses of the interrelationships of Pimelodid Catfishes (Lundberg et al., 2011).

The pectoral spines of *Phractocephalus* differ from all other neotropical genera by the presence of reticulated dorsal and ventral surface ornament (vs. spines lacking ornament or with ornament present but never reticulated). Also, pectoral spines of *Phractocephalus* differ from pectoral spines in other pimelodid genera by the presence of vertically-bifid anterior ornament (vs. anterior ornament consisting of vertically-unicuspid ornament or absent), and from other pimelodids except *Leiarius* by the presence of unicuspid and smooth spinules (vs. anterior ornament consisting of tubercles or absent).

Ware Formation

Teleostei
Division Ostariophysi
Order Characiformes
Family Cynodontidae
Genus *Hydrolycus* Müller & Troschel, 1844
Hydrolycus scomberoides
Figure 4.8A-B.

Material examined—Locality STRI 430062, MUN 16211, anterior fragment of both dentaries preserving the right leading canine.

Description—Specimen with complete symphysis, right leading canine, left dentary ramus up to level of ninth tooth and small portion of right mandibular ramus (Figure 4.8A-B). Tooth bases still preserved, although crowns are almost completely eroded. Dentary ramii are preserved at an angle of about 70° in dorsal view (Figure 4.8B,D), corresponding to the highest angle attained by the dentaries when the mouth is completely open. Base of symphysial teeth distinguished in dorsal view; two symphysial teeth present.

Only right leading canine complete. Crown length 21.0 mm (Figure 4.8B), fissured at level of base, enameloid heavily worn, but preserved enough as to show hypertrophy and presence of commissural, symphysial, and lingual cutting edges. No serrations evident on cutting edges, due to wear of enameloid. Base of canines curved and the tooth is inserted oblique with respect to the main horizontal plane of the dentary; symphysial teeth mesial to leading canines. No tooth remains preserved behind right leading canine. Only portion of leading canine peduncle preserved on left mandibular ramus. Small, oblique crack followed posteriorly by base of small tooth and second canine base, about one-third the diameter

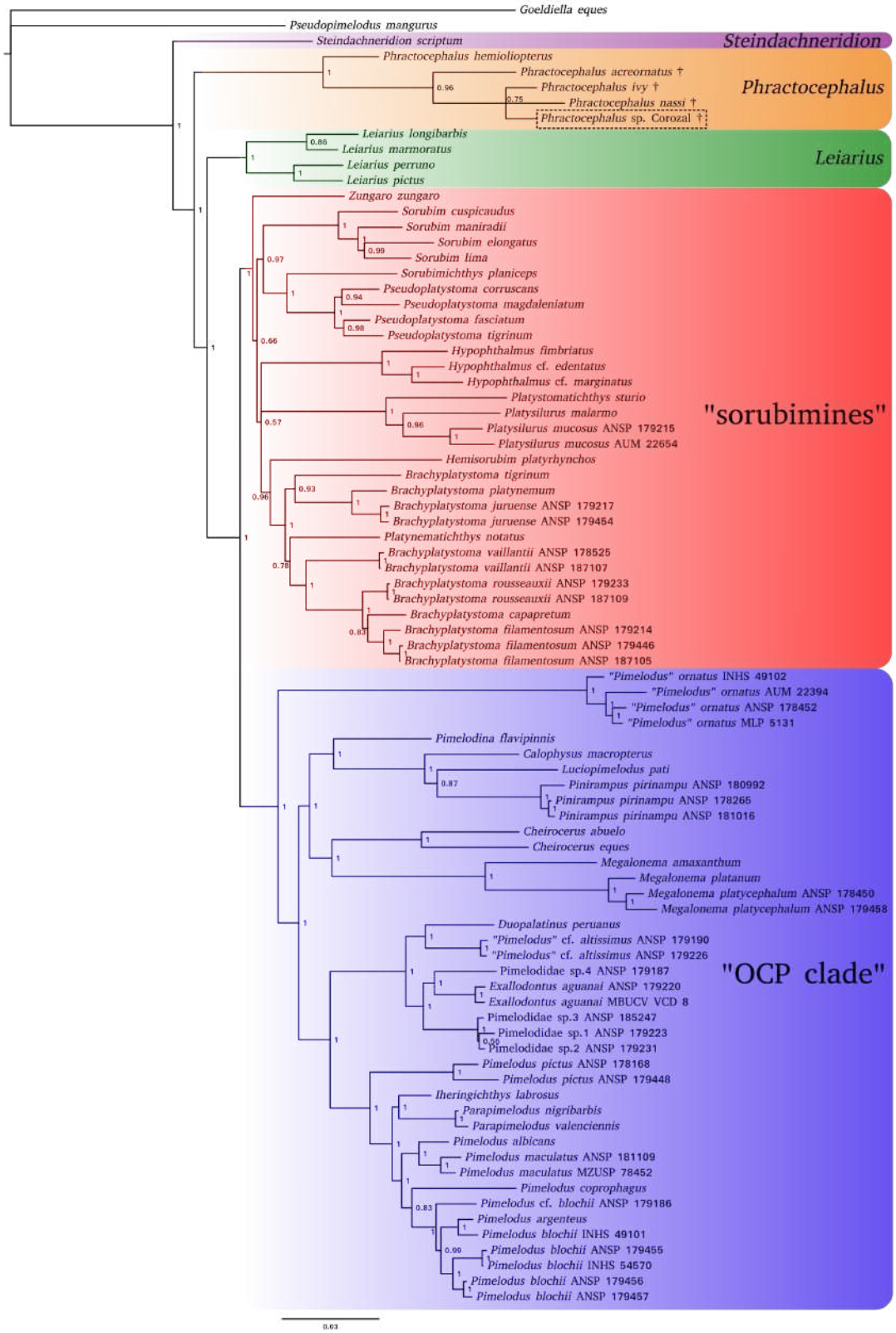


Figure 4.5: *Bayesian phylogenetic analysis of the Pimelodidae including fossil representatives of the genus* Phractocephalus. *Nodal support values are bayesian posterior probabilities. Dagger (†) indicates extinct lineages while the fossil occurrence from the Sincelejo Formation is enclosed in the dashed box. Relevant monophyletic groups are highlighted in colors.*

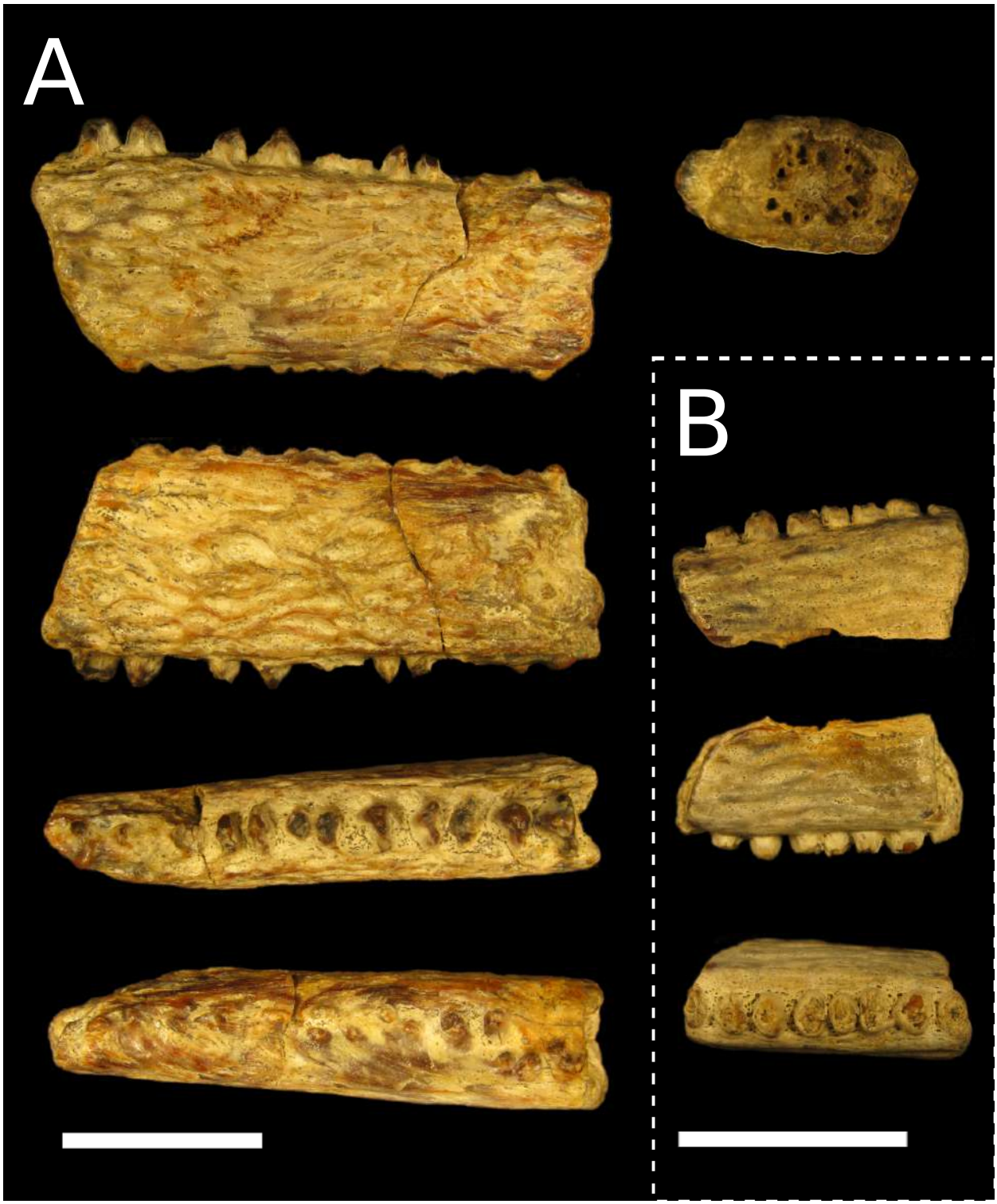


Figure 4.6: *Pectoral-fin spines of* Phractocephalus *sp. A) Fossil pectoral spine fragment, MUN 41058, specimen in dorsal, ventral, anterior, posterior, and proximal cross-section views. B) Fossil spine fragment, MUN 43679, specimen in dorsal, ventral, and anterior views. Scale bar equals 20 mm in all cases.*

of first one. Series of four small tooth bases separates second and third canines, about half diameter of largest canine. Obscured trace suggests an additional tooth base preserved posterior to third canine. Replacement trench present on lingual surface of each mandibular ramus. Unerupted leading canine developing inside left replacement trench; developing tooth horizontally oriented behind base of left leading canine (Figure 4.8).

Remarks—The parsimony analysis (Figure 4.7, Appendices D and E.4) reveals characters 73, 74, 75, 79 and 83 as informative for the position of this fossil among cynodontids. Their distribution in the most parsimonious cladogram indicates that the derived conditions of characters 73, 74, and 79 are synapomorphic for *Hydrolycus*, the derived condition of character 75 is synapomorphic for *Hydrolycus* + *Rhaphiodon*, and derived state of character 83 as autapomorphic for *H. scomberoides*. Therefore, the position of the fossil is well defined not only as a member of the genus *Hydrolycus* but also as part of the extant species *H. scomberoides*. *Hydrolycus scomberoides* is an apex predator restricted to freshwater systems of the Amazon drainage (Mamoré, Guaporé, Beni-Madre de Dios, middle-lower Madeira, Madeira Shield Tributaries, Purus, Tefé, Juruá, Javari, Ucayali, Marañon-Nanay, Napo-Ambyiacu, Putumayo, Japurá, Negro, Branco, Urubu-Uatumã, Trombetas, Amazonas main channel, Dagosta and de Pinna, 2019). Literature records of the species in the Apure basin (Orinoco drainage) have been found to be actually congeners *H. tatauaia* and *H. armatus* (Toledo-Piza et al., 1999). These two widespread species are present both in the Amazon and Orinoco drainages, and compose the sister clade to *H. scomberoides*. Detailed biogeographical and phylogeographical studies might shed light onto the specific patterns of historical relationship between these two cis-Andean areas.

***Hydrolycus* sp.**

Figure 4.8C.

Material examined—Locality STRI 430062, MUN 16230, left canine, very abraded; Locality STRI 390080, MUN 16340 left canine tooth, almost complete; Locality STRI 390077, MUN 16591, complete right canine; Locality STRI 390084, MUN 16540, right canine tooth, almost complete; Locality STRI 470060, MUN 34399 complete left canine, lacking base; MUN 34426, incomplete left canine, without tip and very abraded; MUN 34444, one incomplete left canine with intact enameloid and an abraded right canine.

Description—Teeth can be recognized as cynodontid leading canines based on both shape and structure (Figure 4.8C). Most can be assigned to *Hydrolycus* since base of enameloid is expanded with base of root oblique to horizontal plane. Such configuration suggests that teeth were inserted in an oblique plane on dentary surface, (derived condition of character 74; Appendix D). Teeth elongated, pointed, curved along anterior margin and gently sigmoid along the posterior margin. Divided in distinct base and enameloid crown (Figure 4.8C); sections delimited by constricted neck. Longitudinal ridges present on basal portion of crown, extend some distance towards tip. Symphysial and commisural cutting edges present in all specimens. Third cutting edge can be seen on lingual surface of all specimens, near

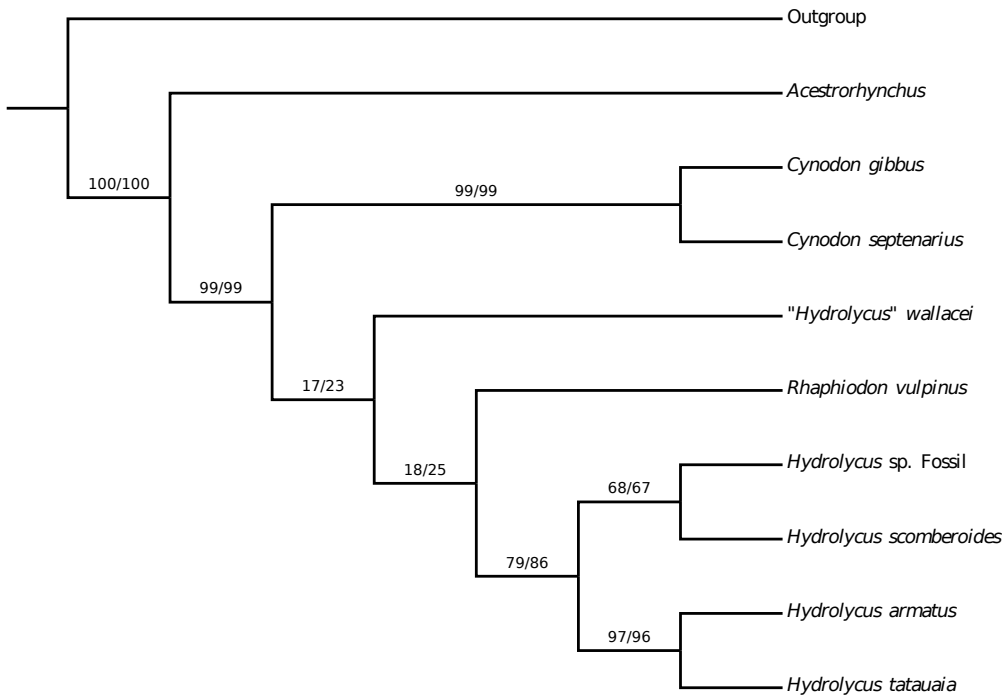


Figure 4.7: *Most parsimonious tree found for the systematics of the Cynodontidae. The genera* Cynodon *and* Hydrolycus *exclusive of* "H." wallacei *present high nodal support whereas the relationship of the latter species and* Hydrolycus + Rhaphiodon *is not well supported. Values above branches are bootstrap/jackknife. Nodes are numbered to the left starting from base to top.*

distal main edge. Enameloid almost completely worn in MUN 16230, which obscures cutting edges. Distal edges of MUN 16340, 16591 and 16540 show more or less worn serrations.

Remarks—These records can not be assigned to *H. scomberoides* given that there is evidence of sympatry among species of the genus (Toledo-Piza et al., 1999). Consequently, these remains may not be assigned to definitely to *H. scomberoides* on the basis of the presence of such species in the Ware formation until diagnostic features at the species level are found among species of *Hydrolycus*.

Genus *Rhaphiodon* Agassiz, 1829
cf. *Rhaphiodon* sp.

Figure 4.8D.

Material examined—Locality 390075, MUN 37734, incomplete tooth preserving only the base of the enameloid and longitudinal half of the tooth.

Description—Single partial tooth resembling enameloid base of *Rhaphiodon* leading canines with straight base, possibly indicating straight insertion on dentary (Figure 4.8D). Preserved dimensions 7.9 mm length and 3.0 mm base width. Only half of tooth preserved in which appears to be straighter than teeth of *Hydrolycus*, thus reinforcing resemblance with *Rhaphiodon*.

Remarks—This specimen it is herein tentatively assigned to the genus *Rhaphiodon* due to the presence of a straight insertion plane of the leading canine.

Figure 4.8: *Fossil representatives of the Cynodontidae. A-B* Hydrolycus scomberoides*, partial dentary, MUN 16211, right lateral and dorsal views respectively. C-D* Hydrolycus *sp. isolated teeth, MUN 34444-5, MUN 34399-3, and MUN 34444-2. D cf.* Rhaphiodon *sp., MUN 37734, incomplete canine. Scale bars equal 20 mm in A, 10 mm in B, and 5 mm in C-D.*

Family Serrasalmidae
Genus *Serrasalmus* Lacépède, 1803
***Serrasalmus* sp.**
Figure 4.9.

Material examined—Locality STRI 470060, MUN 34443, one isolated symphysial dentary (D1) tooth. Locality STRI 470059, MUN 37712, one dentary tooth D3-D4, almost complete.

Description: Left symphysial dentary tooth labio-lingually compressed, crown slightly curved lingually, tricuspid with main cusp largest and lateral cusps about same size, cutting edge very sharp. Root bifid in commisuro-symphysial view, root halves in angle of ca. 40°, basal outline smoothly straight to slightly concave in middle section. Right dentary tooth D3-D4 labio-lingually compressed, crown almost completely triangular occlusal outline in labial view and slightly curved lingually, bicuspid, with main cusp largest and very small commisural cusp somewhat deflected comisurally. Root bifid in commisuro-symphysial view, root halves at angle of ca. 55°, basal outline smoothly convex in labial view.

Remarks—It is possible to distinguish between premaxillary teeth from dentary teeth in carnivore serrasalmids given that the premaxillary ones have either a lingual projection that forms an occlusal platform similar to the molariform platform in herbivorous pacus or are much more asymmetric than dentary teeth; oftentimes, the premaxillary teeth have very wide base and low crown, with the commisural cusplet more developed than in dentary teeth. Also, premaxillary teeth have a less-developed inter-dental anchoring system when compared to dentary teeth where such anchoring mechanism is less developed.

Dentary symphysial teeth in carnivore piranhas (i.e., *Pygocentrus* and *Serrasalmus*) are tri- to penta-cuspidate and lack the commisural socket present in subsequent dentary teeth for anchoring of the anterior adjacent tooth. In *Pygocentrus* the cusplets around the main one are strongly heterogeneous in size, being the commisural cusplet larger than the symphysial one; on the contrary, these cusplets are about the same size in *Serrasalmus*. Other Serrasalmids with multicuspidate teeth present either mamilliform (*Catoprion*) or multicuspidate with four or more cusplets (*Pristobrycon* and *Pygopristis*).

The identification of the specimens herein studied as dentary teeth of the genus *Serrasalmus* are based on the fact that they lack the occlusal platform seen in premaxillary teeth, the presence of cusplets on both sides of the main cusplet which is found in symphysial teeth, and the the presence of cusplet of about the same size that allows to identify the genus *Serrasalmus*.

The tricuspidate condition in piranha teeth is also seen in the fifth premaxillary tooth of the genus *Pygocentrus* (*vs.* fifth premaxillary tooth bicuspid in *Serrasalmus*). As already mentioned the symphysial teeth of either the premaxilla or the dentary can also be tricuspid, have narrow base, and have a nearly symmetrical main cusplet; in contrast, the fifth premaxillary tooth in *Pygocentrus* shows a very wide base and asymmetrical main cusplet which is

commisurally-oriented. Lundberg (1997) illustrated this morphology from an isolated tooth IGM 251277 from the middle Miocene La Vent fauna in central Colombia; although the author was uncertain about the affinities of such specimen and thus identified it as "*Serrasalmus*, *Pygocentrus*, or *Pristobrycon* sp.", it is possible to narrow its identity as a fifth premaxillary tooth of the genus *Pygocentrus* sp.

Order Siluriformes
Family Auchenipteridae
Genus *Trachelyopterichthys* Bleeker, 1862
Trachelyopterichthys sp.

Figure 4.10.

Material examined—Locality 470060, MUN 34401, one pectoral-spine fragment.

Description—Left pectoral-spine fragment preserving the middle portion of the spine shaft along with the anterior, posterior, dorsal, and ventral ornaments. Shaft depressed and robust, with antero-dorsal, antero-ventral, postero-dorsal, and postero-ventral margins smooth. Preserved shaft showing some curvature along the axis posteriorly. Dorsal and ventral ornament consisting of undulating, subparallel ridges, decreasing in thickness from anterior to posterior across shaft. Abundant tubercles onto the dorsal ridges throughout the shaft, ventral tubercles present, yet restricted to the anterior 1/4 of ventral surface. Anterior ornament consisting of straight to slightly antrorse spinules directly implanted onto anterior shaft surface. Posterior ornament consisting of irregular blades poorly preserved but showing extensive fusion between units, straight to slightly retrorse in orientation. Lumen comprising ca. 1/2 of shaft area in cross-section; lumen outline very irregular.

Remarks—The anterior ornament of the pectoral-fin spine in *Trachycorystes* consists of straight spinules with fused bases, so that the whole ornament series consists of a continuous ridge emerging from the anterior surface in dorsal view. In contrast, the anterior spinules are sessile and implanted directly on the anterior shaft surface without fusion at the bases in *Trachelyopterichthys*. This is the first fossil record of the family Auchenipteridae anywhere. The family has a Neotropical distribution in drainages on both sides of the Andes, with a greater diversity in cis-Andean basins. The genus *Trachelyopterichthys* is composed of only two species: *T. anduzei* from the Orinoco drainage, and *T. taeniatus* from the Amazon drainage (Birindelli, 2014; Calegari et al., 2019). Further differences between species are not known, but further study of extant specimens of the genus may provide further information on the identity of the fossil as belonging to either of the extant species or even an extinct, undescribed one. At present it was not possible to reach at a conclusion as to the specific identity of the fossil remains.

Family Doradidae
Genus *Hemidoras* Bleeker, 1858
Hemidoras sp.

Figure 4.11.

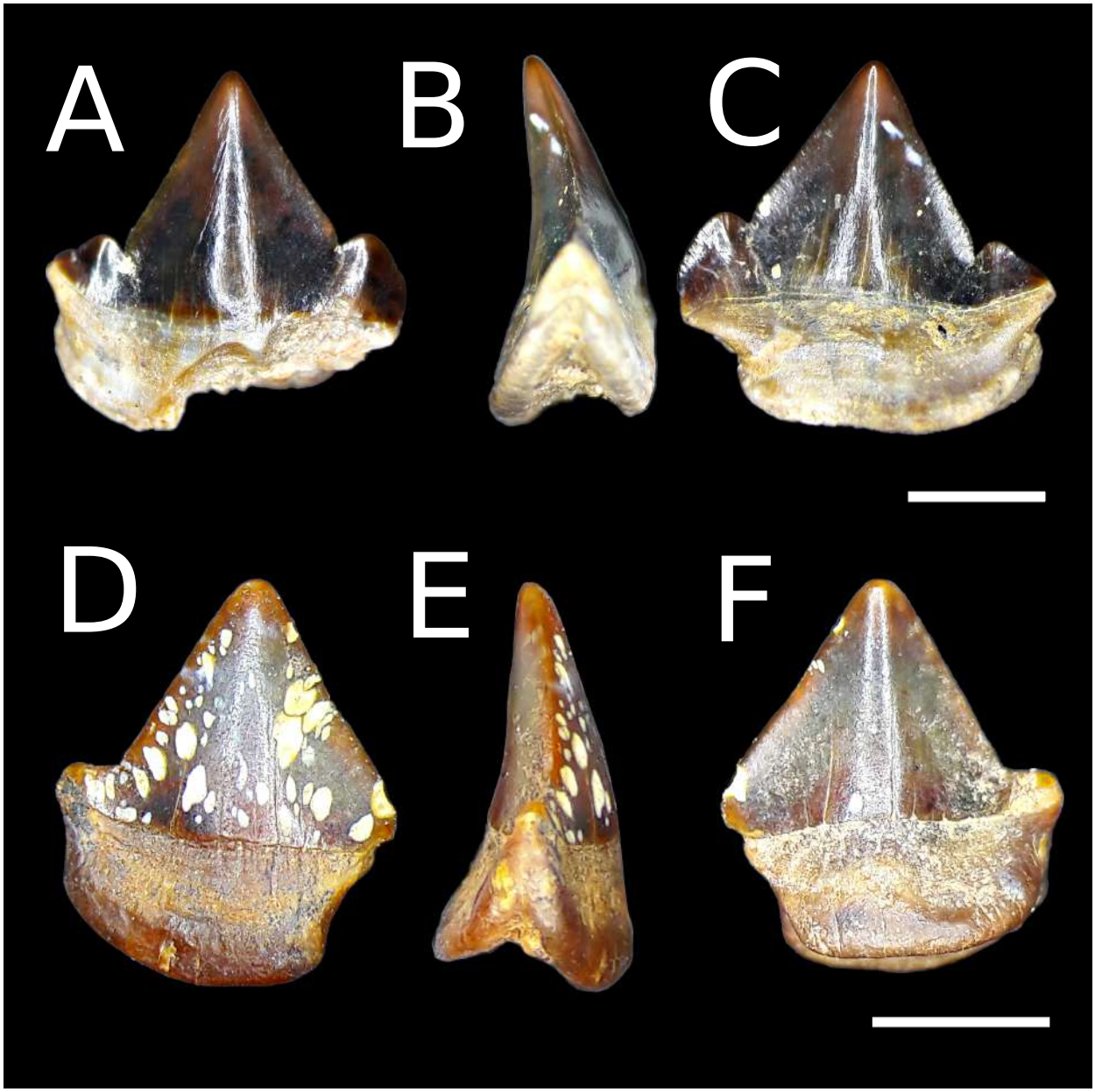


Figure 4.9: *Fossil isolated teeth of* Serrasalmus *sp. A-C) Right D1 tooth in labial, symphysial, and lingual views respectively. D-F) Left D4-D5 tooth in labial, commisural, and lingual views respectively. Scale bars equal 2 mm.*

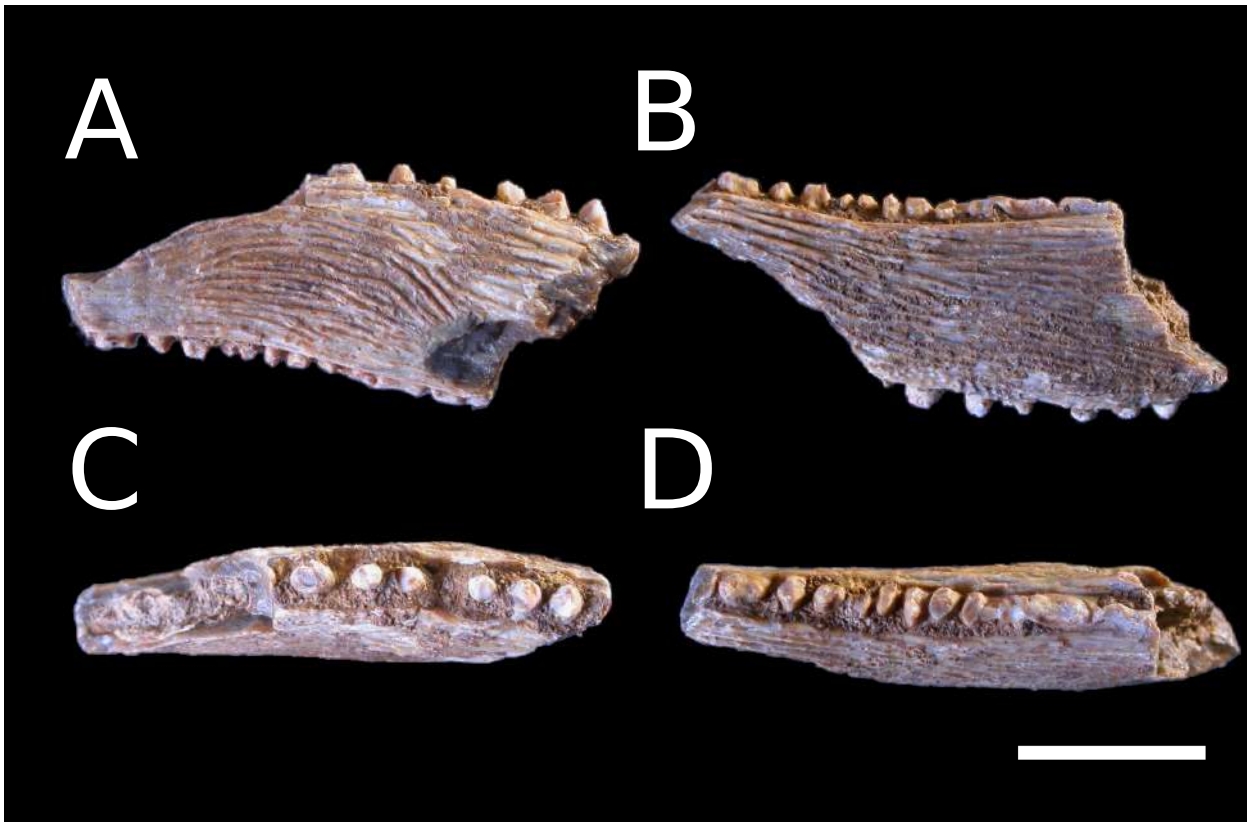


Figure 4.10: *Left pectoral-fin spine of* Trachelyopterichthys *sp., MUN 34401. A-D Dorsal, ventral, anterior, and posterior views respectively. Scale bar equals 10 mm.*

Material examined—Locality 470060, MUN 34455, one partial dorsal-fin spine preserving the a left portion of the base and about 1/3–1/2 of the shaft.

Description—Base preserving half inflection point of anterior longitudinal ridge, and left lateral condyle but without lateral articular surface; posterior process not preserved. Shaft preserving anterior, posterior, and lateral ornaments, as well as fenestrae communicating interior of shaft lumen with epidermal tissue. Anterior ornament consisting of antrorse spinules from inflection point of anterior longitudinal ridge up to preserved dorsalmost portion of shaft; spinules closely set, leaving space inbetween less than spinule diameter. Posterior ornament absent on basal third of preserved spine length, then consisting of antrorse spinules with space inbetween measuring equaling spinule diameter or more. Lateral ornament consisting of subparallel ridges oriented somewhat oblique to shaft main axis; ridges with angular surface and flat sides. Shaft outline quadrangular in transversal view; lumen maximum axis comprising 1/4 of shaft width in anteroposterior axis.

Remarks—The presence of both anterior and posterior antrorse dorsal-fin spine ornaments is a rare feature in the Doradidae, found only in the genera *Hemidoras*, one species of *Leptodoras* (*L. acipenserinus*), one species of *Nemadoras* (*N. humeralis*), *Ossancora*, *Oxydoras*, *Petalodoras*, one species of *Platydoras* (*P. armatulus*), and *Trachydoras*. In contrast, most other taxa show anterior antrorse and posterior retrorse or straight ornaments. Among those taxa, posterior ornaments are separated by a large space of more than 1.5 of the adja-

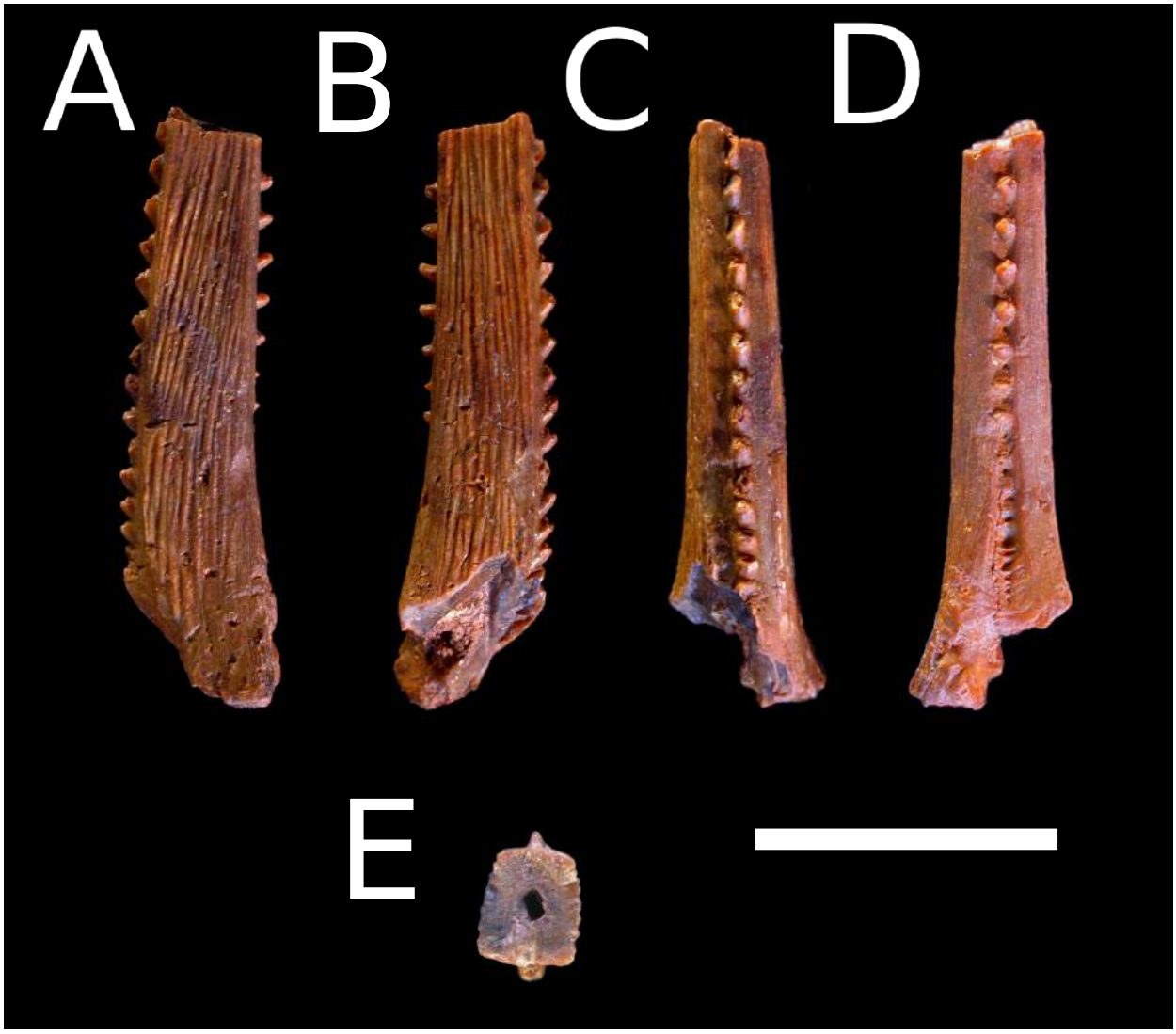


Figure 4.11: *Dorsal-fin spine fragment of* Hemidoras *sp., MUN 34455. A-D) Left, right, anterior, and posterior views respectively. E) Shaft cross section. Scale bar equals 10 mm.*

cent ornament base width in *L. acipenserinus*, *Ossancora*, *Oxydoras*, *Petalodoras*, *Platydoras*, and *Tenellus*. *Nemadoras humeralis* shows very closely-set posterior ornaments separated by a space of ca. 1.2 of ornament base width. This is in contrast to *Hemidoras* and the fossil specimen where it is between 1.0 and 1.5. Species of *Trachydoras* have extensive posterior ornament developed down to the level of the dorsal-fin spine base, while in the fossil and *Hemidoras* there is at least an extension of 1/4 of the spine shaft length devoid of ornaments basally. This combination of characters allow positive identification of dorsal-fin spines to *Hemidoras* among doradids.

Birindelli (2014) considered *Opsodoras* as a synonym of *Hemidoras*, comprising now five species distributed through the Amazon region in the Amazonas main channel as well as in the basins Beni-Madre de Dios, Branco, Guaporé, Japurá, Middle-lower Madeira, Mamoré, Marañon-Nanay, Negro, Purus, Putumayo, Tefé, Trombetas (Dagosta and de Pinna, 2019), all in cis-Andean South America. Phylogenetic analyses consistently recover *Hemidoras* as part of a group including *Anduzedoras*, *Doras*, *Ossancora*, *Nemadoras*, *Hassar*, *Trachydoras*, and *Leptodoras*; however, the interrelationships among these genera sometimes referred to as the "*Doras* clade" or the "fimbriate-barbel doradids" are variable depending on both the data source (either morphology or sequence data) and the analytical strategy employed (ML, bayesian inference, or parsimony). Arce H et al. (2013, fig. 2) considered *Hemidoras* monophyletic only with the inclusion of *Opsodoras*. The two genera were later synonymized by Birindelli (2014, fig. 70). Birindelli also recovered *Hemidoras* as monophyletic (after inclusion of *Opsodoras*) and sister to a clade comprising *Nemadoras*, *Tenellus*, *Hassar*, *Anduzedoras*, and *Leptodoras*.

Genus *Oxydoras* Kner, 1855
cf. *Oxydoras* sp.

Figure 4.12.

Material examined—Locality 470060, MUN 34409-2, distal portion of a dorsal-fin spine shaft.

Description—Shaft preserving anterior, posterior, and lateral ornaments. Anterior ornament consisting of antrorse spinules along preserved shaft surface; spinules irregularly spaced, leaving space inbetween less than spinule diameter, fused basally in a longitudinal ridge. Posterior ornament consisting of straight to retrorse spinules with large space inbetween larger than 1.0–1.5 times spinule diameter. Lateral ornament almost absent, with very smooth surface and just faint indication of longitudinal small ridges. Shaft outline hexagonal in transversal view; lumen maximum axis comprising 1/2 of shaft width in anteroposterior axis.

Remarks—The genus *Oxydoras* is peculiar among doradids in having a very slender dorsal-fin spine with very spaced out anterior and posterior ornaments. Species of *Oxydoras* have almost smooth lateral shaft surface lacking prominent ridges, unlike all the remaining

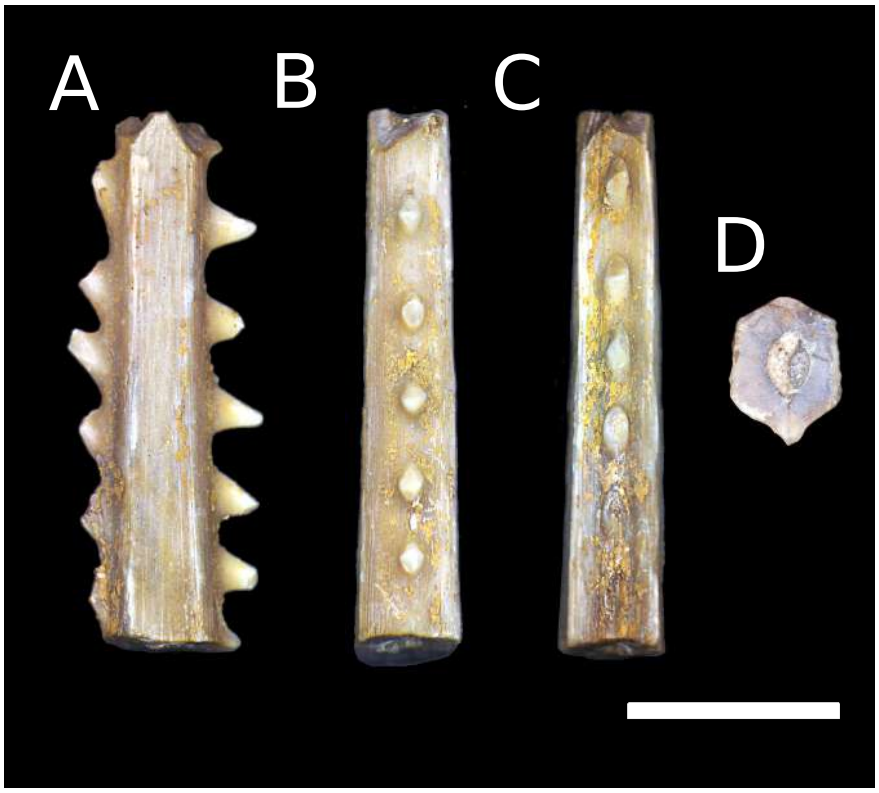


Figure 4.12: *Dorsal-fin spine fragment of cf.* Oxydoras *sp., MUN 34409-2. A-C) Left, posterior, and anterior views respectively. D) Shaft cross section. Scale bar equals 5 mm.*

genera in the family that show lateral shaft surfaces with ridges of varying degrees of development, from fine and numerous as in *Hemidoras* to few in number but hypertrophied as in *Petalodoras* and members of the Astrodoradinae. The fossil specimen herein described matches representatives of the genus *Oxydoras* in most details, except for the orientation of the posterior dorsal-fin spine ornament which is antrorse to straight in most species of *Oxydoras* and retrorse in the fossil specimen. *Oxydoras sifontesi* differs further from the fossil specimen due to the presence of small tubercles on the lateral shaft ornament restricted to the anteriormost third on the spine, while such tubercles are absent in other representatives and the fossil specimen. Due to the latter incongruent character the fossil cannot be allocated with certainty to the genus *Oxydoras*.

Family Pimelodidae

Genus *Brachyplatystoma* Bleeker, 1862

Brachyplatystoma* cf. *vaillantii

Figure 4.13.

Material examined—Locality 470059, MUN 37567, partial left sphenotic.

Description—Preserving most of the original bone, including the articular facet for the hyomandibula and the surface ornament. Preserved outline renoid in dorsal view. Most of the ventral surface eroded and exposing the internal trabecular structure. Dorsal ornament consisting of parallel ridges on both the lateral surface and the anterior half of the bone in

a radiating pattern with focus in the posterior half. Tubercles concentrated in the region of the ridge focus, somewhat aligned to the ridges. Articular facet for the hyomandibula smooth and concave, becoming shallower towards the sides of the bone and suggesting that the posterior portion on the pterotic should have been reduced as most of the functional surface is restricted to the sphenotic.

Remarks—*Brachyplatystoma* is a genus with seven extant and one extinct species (Lundberg, 2005; Lundberg and Akama, 2005). Fossil occurrences of this group are known from several localities of Neogene age in northern South America (Aguilera et al., 2013a; Lundberg et al., 2010), mostly identified as *Brachyplatystoma* cf. *vaillantii* and *B. promagdalena*. The former species is distinctive among congeners by the pattern of ornamentation on exposed cranial bones, with ridges and tubercles aligned in radiating patterns with focus in the sphenotics. In contrast, other species of the genus have very smooth cranial bones. The only specimen known of *B. promagdalena* is a weberian apparatus, and consequently we lack further anatomical details to compare with congeners. By the distribution of character states, *B. promagdalena* should also show smooth cranial bones, quadrangular opercle, and a well-defined sulcus on the anterior surface of the pectoral spine, the latter two characters being morphological synapomorphies for the genus (Lundberg and Akama, 2005). The numerous Colombian and Venezuelan Neogene occurrences historically associated to *B. vaillantii* must be restudied in order to better assess their specific identity, as they appear to be a different, extinct species sharing the pattern of cranial ornament with the extant *B. vaillantii* (J.G. Lundberg, *pers. comm.*).

Genus *Phractocephalus* Spix & Agassiz, 1829

Phractocephalus sp.

Figure 4.14.

Material examined—Locality 470060, MUN 34425, partial nuchal plate.

Description—Fragment of nuchal plate, oval in outline. No margins preserved, only dorsal and ventral surfaces. Dorsal ornament consisting of strong reticulating ridges surrounding round to ovoid pits. Ventral surface smooth with traces of abrasion and some exposure of spongy bone internal structure.

Remarks—The genus *Phractocephalus* has been discussed in the respective entry above (Sincelejo formation), and the same general remarks apply here. The occurrence from the Ware formation is a cranial element, the nuchal plate. The characteristic surface ornament consisting of strong reticulating ridges around round to oval pits is a feature that easily distinguishes the genus *Phractocephalus* from other Neotropical Siluriforms; this feature has already been used in the literature when assessing the generic position of bone elements in this genus in Neotropical fossil fish faunas (Aguilera et al., 2008; Azpelicueta and Cione, 2016; Lundberg and Aguilera, 2003). Although catfishes of the families Auchenipteridae, Ariidae, and Doradidae display very developed cranial ornamentation through exostosis,

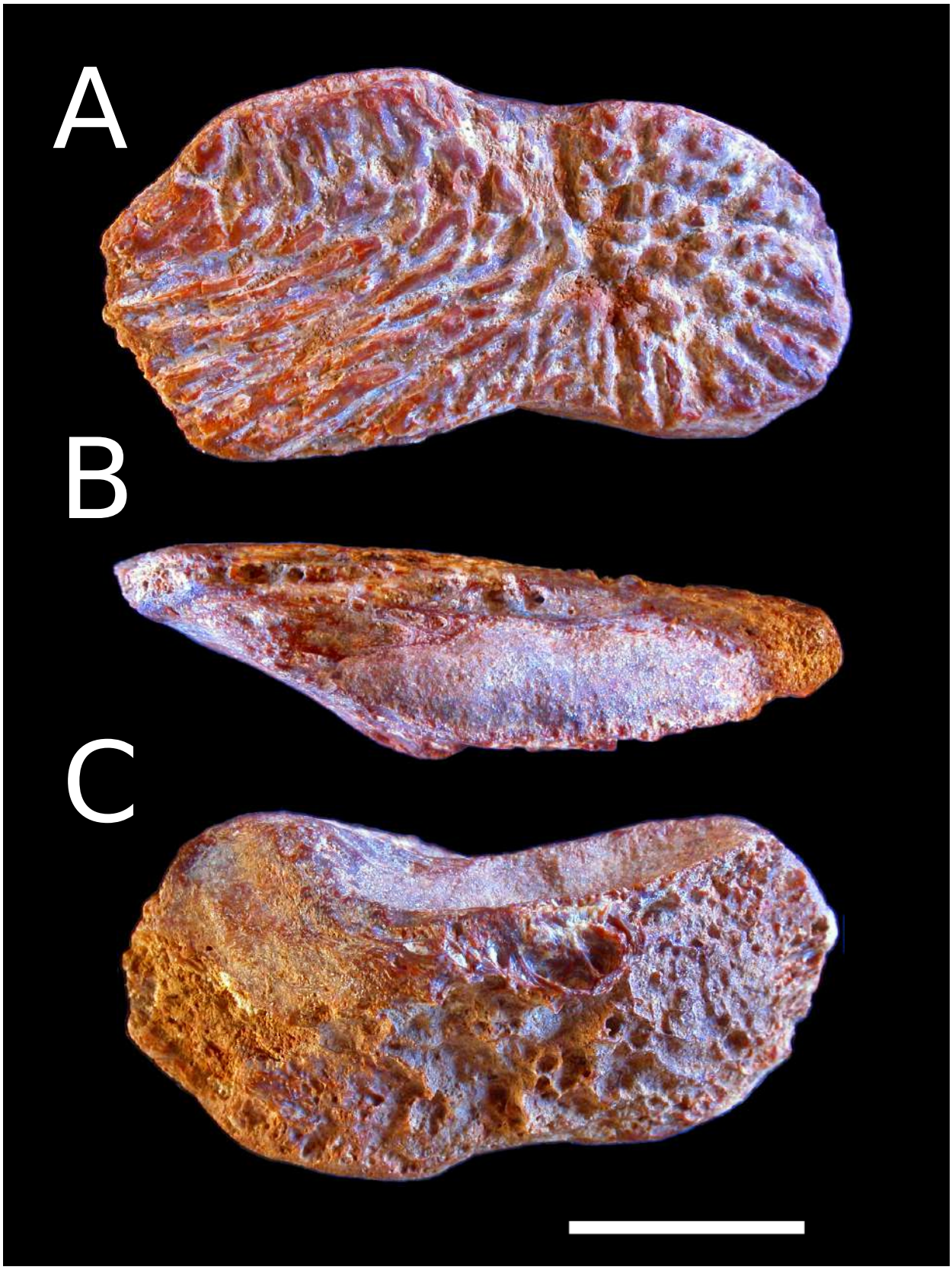


Figure 4.13: *Left sphenotic of* Brachyplatystoma *cf.* vaillantii*, MUN 37567. A-C) Dorsal, lateral, and ventral views respectively. Scale bar equals 10 mm.*

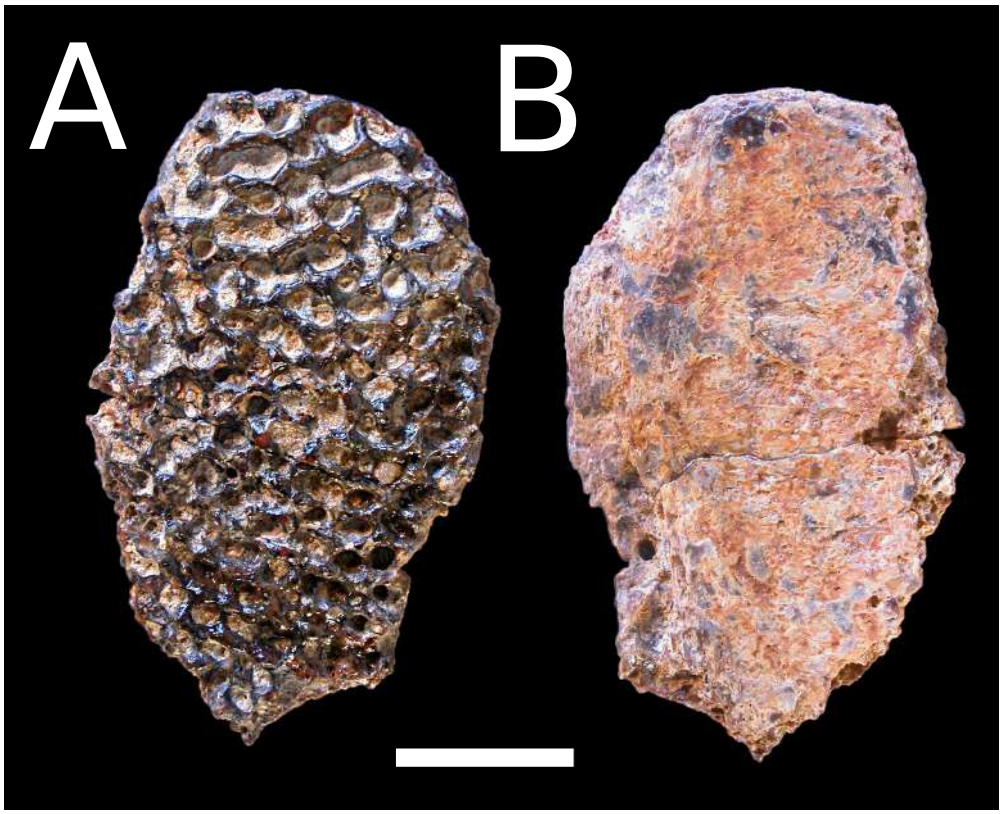


Figure 4.14: *Nuchal plate fragment of* Phractocephalus *sp., MUN 34425. A-B Dorsal and ventral views respectively. Scale bar equals 10 mm.*

those cases often consist of either subparallel ridges, series of tubercles, or a combination of both, but never arranged in a reticulate pattern.

Genus *Platysilurus* Haseman, 1911

Platysilurus sp.

Figure 4.15.

Material examined—Locality 470062, MUN 37605, partial neurocranium.

Description—Specimen preserving several cranial bones and surface ornament. Tridimensional structure compromised by squashing due to diagenesis, although most bones remain spatially in place. Sediment and iron cover tend to obscure most sutures between bones, making difficult to trace specific bone outlines. Dorsal bones (frontal, sphenotic, pterotic, supraoccipital, extrascapular) with ornament consisting of parallel and radiating ridges with origin foci in supraoccipital and pterotics; ridges sometimes undulating, often interleaved with scattered tubercles. Presence of evident depression on posteriormost region of frontals, anterior to suture with supraoccipital. Dorsal surface noticeably concave between level of pterotics to anteriormost preserved frontals, then strongly convex from supraoccipital ornament focus to preserved posterior limit of supraoccipital process. Absence of median longitudinal sulcus along supraoccipital process.

Frontal incomplete and preserving only about posterior half of original extent, showing medial depression posterior to region of anterior cranial fontanelle, surface covered with

parallel, divergent, straight ridges; frontal in contact with supraoccipital and sphenotic, suture between frontal and supraoccipital strongly interdigitate, suture with sphenotic smooth and clean to slightly curve in some regions, suture with contralateral frontal mainly smooth with some degree of interdigitation at level of frontal depression posterior to anterior cranial fontanelle. Sphenotic elongate and renoind in outline, almost completely preserved in dorsal view; sphenotic in contact with frontal, supraoccipital, pterotic, pterosphenoid, and prootic; mesial suture with frontal as described above, posterolateral suture with pterotic nearly straight and smooth dorsally and interdigitate ventrolaterally, mesial suture with supraoccipital finely interdigitate, ventral suture with pterosphenoid concealed under sediment and bone fragments, ventral suture with prootic obscured due to lateral displacement of ventral wing of sphenotic during diagenesis, apparently smooth and straight. Lateral surface present yet squashed, although preserving spatial relationship to neighboring bones of otic region and some sutures. Hyomandibular articular facet elongate, concave, smooth, spanning sphenotic and anterior half of pterotic. Pterotic ovoid in dorsal view with posteromedian concavity for contact with exstrascapular; pterotic in contact with sphenotic, supraoccipital, extrascapular, prootic, postemporo-supracleithrum, and exoccipital; anteromesial suture with sphenotic as already described above, mesial suture with supraoccipital finely interdigitate, posterolateral suture with postemporo-supracleithrum and exoccipital obscured by fractures and sediment, ventral suture with prootic interdigitate although region of suture is fractured. Extrascapular barely visible in posterior view, strongly covered by sediment. Supraoccipital well preserved, polygonal in outline with posterior prominent supraoccipital process. Sutures as already described with adjacent bones.

Pterosphenoid poorly preserved, only posterior portion in contact with sphenotic and prootic present although very fractured and covered with sediment. Prootic in contact with pterosphenoid, sphenotic, pterotic, postemporo-supracleithrum, basioccipital, and exoccipital. Sutures with adjacent bones mostly obscured by fractures, however, preserved outlines suggest interdigitate sutures at least with basioccipital, exoccipital and sphenotic. Exoccipital in contact with prootic and basioccipital. Outlines poorly preserved, although interdigitate sutures are present with adjacent bones. Basioccipital poorly preserved, fractured, missing basal half, preserving articular facet for weberian apparatus in posterior view. Sutures with adjacent bones as already described above. Epioccipital poorly preserved, only part of bone exposed out of strong sediment cover, therefore obscuring sutures.

Remarks—Most of the fossil occurrences of the Ware formation so far are represented by very fragmentary specimens. Two notable exceptions are the weberian apparatus of *Brachyplatystoma* cf. *vaillantii* described by Aguilera et al. (2013a) and the neurocranium of *Platysilurus* sp. herein described. The fossil record of the genus was until now restricted to the Urumaco and Rio Yuca formations in Venezuela (Lundberg et al., 2010; Rincón et al., 2016; Sabaj Pérez et al., 2007). The latter records occurrences are interesting because they are represented by the same anatomical region of the head as the specimen herein described; all of these specimens are also only identified to genus level and therefore its specific identity

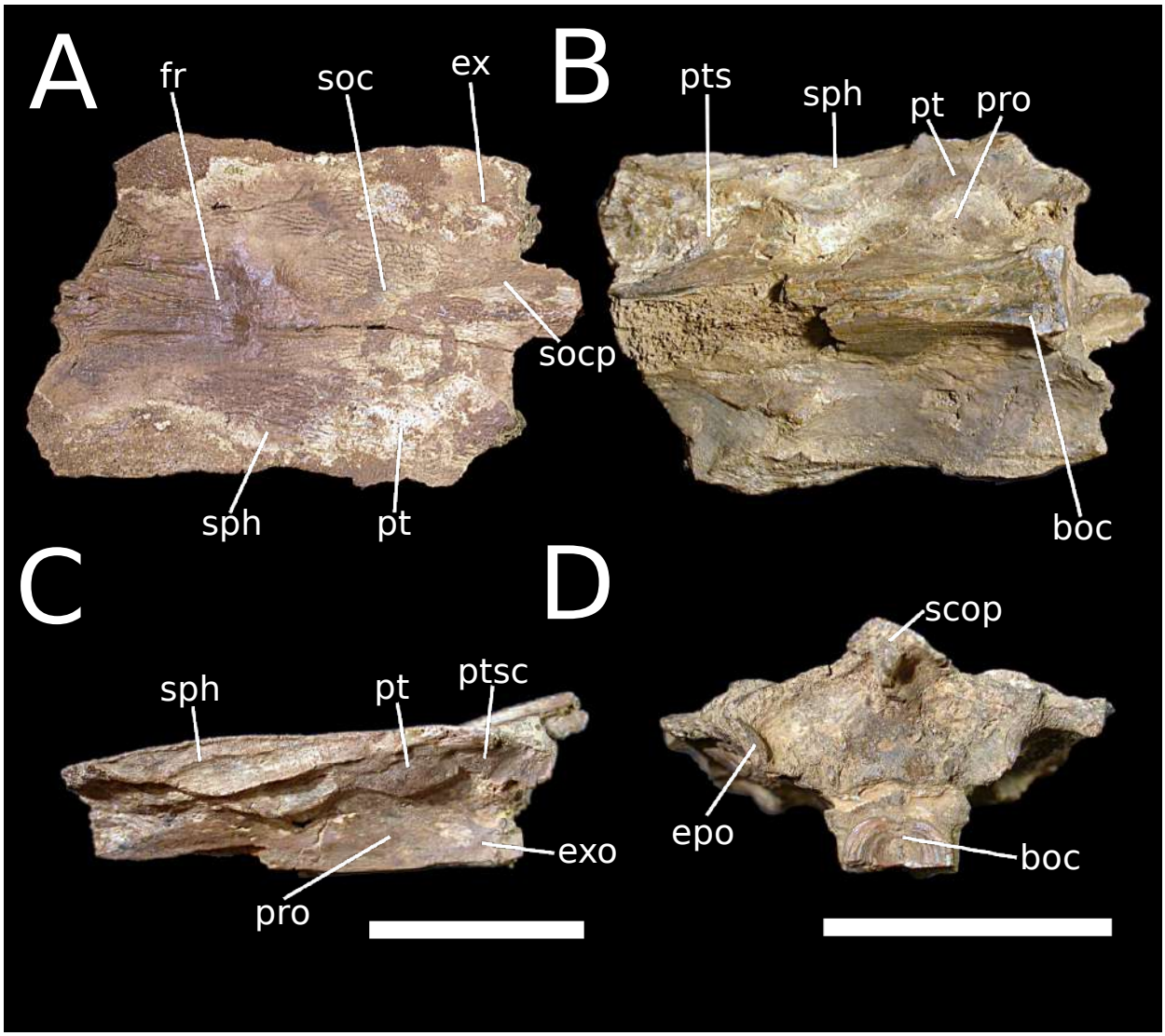


Figure 4.15: *Partial neurocranium of* Platysilurus *sp., MUN 37605. A-D) Dorsal, ventral, lateral, and posterior views respectively. Scale bars equal 20 mm; the scale bar in C applies to A-C while the bar in D applies only to such section.*

and affinities await re-study. The genus *Platysilurus* includes small to medium-sized piscivorous catfishes restricted to the Maracaibo drainage (*P. malarmo*), and Orinoco-Amazonas drainages (*P. mucosus*); it is readily distinguishable from other pimelodid genera by the highly raised supraoccipital process, a feature herein verified in museum specimens for the extant species, evident in the figures of Venezuelan fossil specimens (Rincón et al., 2016; Sabaj Pérez et al., 2007), and already suggested by Sabaj Pérez et al. (2007) as diagnostic for the genus. Despite the completeness of the material, family-scale comparative anatomy is still needed in order to determine whether the preserved characters are taxonomically informative. Both extant species show a prominent longitudinal sulcus along the supraoccipital process. *Platysilurus mucosus* shows continuity between interfrontal concatity and the longitudinal supraoccipital sulcus, in contrast with the specimen herein described where such connection is missing as well as the supraoccipital sulcus. *Platysilurus malarmo* has a nearly flat surface between the eyes anteriorly to before supraoccipital process, while the fossil specimen herein studied has a concave surface. Our fossil specimen is overall more similar to *P. malarmo* than to *P. mucosus*, despite the clear differences with either species already mentioned; this suggests that the fossil specimen might represent an extinct, undescribed species. Further detailed anatomical study of the congener *P. malarmo* is however needed before claiming any concise specific assignment for the fossil *Platysilurus* of the Ware formation. Specimens from the Urumaco and Rio Yuca formations are overall similar to the *Platysilurus* from the Ware formation; however, it was impossible to examine any of these specimens in order to decide whether they represent only one or different taxa. Given the cis- and trans-Andean distributions of both extant species, this occurrence may either support the presence or absence of drainage connections until its phylogenetic position in the family is better understood.

Genus *Zungaro* Bleeker, 1858

Zungaro sp.

Figure 4.16.

Material examined—Locality 470060, MUN 34483, left partial pectoral spine preserving part of the spine base and proximal 1/4 of the shaft.

Description—Spine base preserving dorsal process with wide articular facet, otherwise strongly eroded, base of spine not preserving both proximal articular facet and proximal process. Anterior ornament absent; posterior ornament consisting of straight spinules; dorsal ornament consisting of smooth ridges, widest restricted to lateral 2/3 of shaft surface, additional narrow ones restricted to 1/3 mesial surface of shaft surface, oblique sulcus separating both kinds of surface ornaments.

Remarks—The genus *Zungaro* comprises *Z. jahu* and *Z. zungaro*, the former distributed in the Paraná-Paraguay drainage, and the latter in the Amazon and Orinoco drainages (Boni et al., 2011; Dagosta and de Pinna, 2019), therefore being one more representative of the Siluriformes restricted to cis-Andean South America. The fossil record of the genus includes

Figure 4.16: *Left pectoral-fin spine of* Zungaro *sp., MUN 34483. A-C Dorsal, anterior, and posterior views respectively. Scale bar equals 5 mm.*

a partial mesethmoid from the Rio Acre fossil fauna of Madre de Dios, Perú (LACM 128395, Lundberg et al., 2010). The present occurrence in the Guajira Peninsula represents the first trans-Andean record of the genus, a region outside of its current distribution. The combination of absent anterior ornamentation, and two regions with differential ridge thickness on the dorsal ornament of the pectoral-fin spine are diagnostic for the genus among Neotropical Siluriforms. *Zungaro* is a migratory species that occupies deep portions of lotic environments (Agostinho et al., 2003).

4.4.2 Faunal similarity

The freshwater fossil fish faunas herein studied vary considerably in richness of genera. The Urumaco fauna is the richest, with 13 genera while Makaraipao, Contamana and Fitzcarrald are the poorest with four genera each; Ituzaingó, Rio Acre and La Venta show intermediate values with six, eight, and eleven genera respectively.The Ware fauna has a comparatively large number of taxa (10 genera). According to the Bray-Curtis dissimilarity index, the Makaraipao fish fauna is closer to the Fitzcarrald and Contamana faunas than to other units (Figure 4.17). This high degree of dissimilarity between the Ware fauna and other Neogene faunas of South America, either cis- or trans-Andean, is expected given that several genera are only known from the this fossil locality (e.g., *Hemidoras*, *Serrasalmus*, *Trachelyopterichthys*) while other taxa were shared with few other fossil localities (e.g., *Brachyplatystoma* and *Platysilurus*). The La Venta and Rio Acre faunas cluster together, while Ware is the most dissimilar of the set, followed by Ituzaingó and Urumaco. The pattern of similarity does not seem to correlate with geographic proximity or chronostratigraphic closeness.

4.5 Discussion

Table 4.2: *STRI localities from the Sincelejo and Ware formations referenced in the present study. Several localities refer to the same level and spatial point, being therefore synonyms, see Figure 4.3.*

Taxon	Catalog number	Strat. unit	Locality	Strat. pos. (m)
Brachyplatystoma cf. *vaillantii*	MUN 37567	Ware	470059	1.7
Hemidoras sp.	MUN 34455	Ware	470060	4
Hydrolycus scomberoides	MUN 16211	Ware	430062	5
Hydrolycus sp.	MUN 16230	Ware	430062	5
Hydrolycus sp.	MUN 16340	Ware	390080	5
Hydrolycus sp.	MUN 16540	Ware	390084	5
Hydrolycus sp.	MUN 16591	Ware	390077	1.7

Continued on next page

Table 4.2 – *Continued from previous page*

Taxon	Catalog number	Strat. unit	Locality	Strat. pos. (m)
Hydrolycus sp.	MUN 34399, MUN 34426, MUN 34444	Ware	470060	4
Leporinus or *Hypomasticus*	GAB-P 415	Sincelejo	710004	3
cf. *Oxydoras* sp.	MUN 34409-2	Ware	470060	4
Phractocephalus sp.	MUN 34425	Ware	470060	4
Phractocephalus sp.	MUN 41058, MUN 43679	Sincelejo	710004	3
Platysilurus sp.	MUN 37605	Ware	470062	5
cf. *Rhaphiodon* sp.	MUN 37734	Ware	390075	4
Serrasalmus sp.	MUN 34443	Ware	470060	4
Serrasalmus sp.	MUN 37712	Ware	470059	1.7
Trachelyopterichthys sp.	MUN 34401	Ware	470060	4
Zungaro sp.	MUN 34483	Ware	470060	4

4.5.1 Informativeness of fragmentary specimens and taxonomic accuracy

Further field prospection will be necessary in order to expand our knowledge on the freshwater fishes of the Sincelejo formation, specially focusing on the facies with concretions where fossil vertebrates seem to be better represented (Figure 4.2C). Despite several field seasons and collection efforts throughout several years, few specimens of any vertebrate group have been recovered in the vicinity of Corozal when compared to rich fossil faunas such as La Venta and Ware, which suggests that recovery rate per rock volume tends to be low. Although original low concentration of remains in the sedimentary record is a possibility, other reasons such as low rock are exposure, fewer exposures with horizontal continuity, and vegetation cover consisting of grasses and dry-forest cover might account for the poor recovery of specimens from Corozal when compared to the fossil record of the Ware and La Venta.

Exposures of the Ware formation, although regionally restricted, are locally well developed and devoid of vegetation cover as is the rule in the desertic Guajira Peninsula. Although numerous fossil specimens have been recovered in this unit, specially in the locality of the stratotype section, they suffer from excessive erosion and fragmentation, greatly reducing preservation of diagnostic features in fossil specimens. This explains why despite having numerous bone fragments available for study, only a portion have shown diagnostic features that allow to identify them to genus level.

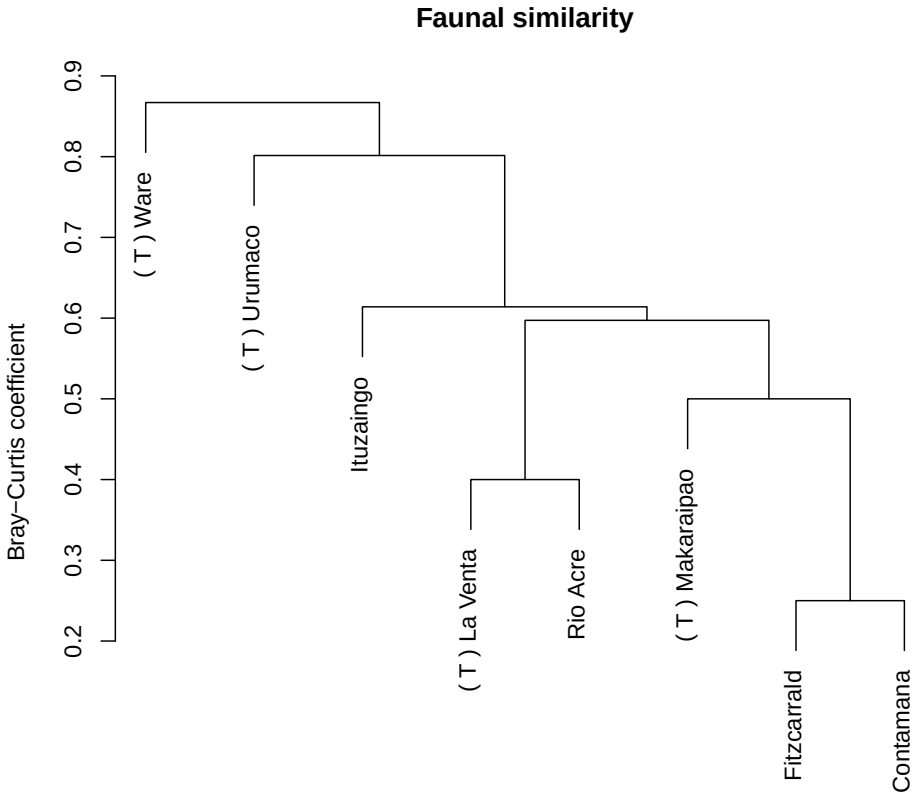


Figure 4.17: *Faunal dissimilarity for Neogene freshwater fish faunas in South America using the Bray-Curtis coefficient. The symbol (T) indicates those faunas located west of the Andes*

4.5.2 New records

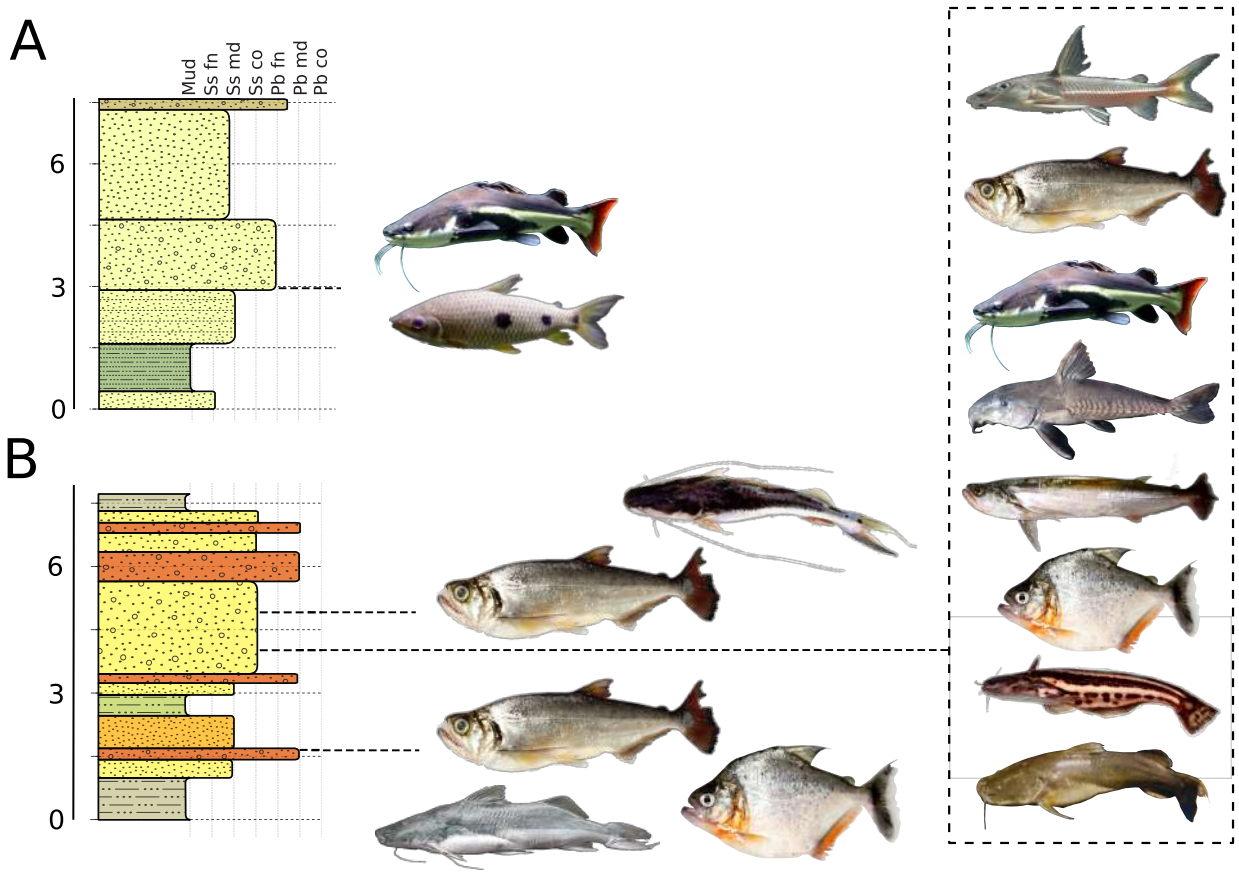
The present occurrences from the Sincelejo formation represent the first record of freshwater fishes from the Pliocene of the Magdalena-Cauca drainage (Figure 4.18). This component of the extinct continental diversity in trans-Andean Northern South America has long been neglected despite being known from a number of stratigraphic units in the Magdalena-Cauca drainage since the first half of the XX century (e.g., the Coyaima, La Venta, Carmen de Apicalá faunas; Stirton, 1953, pp. 610, 612, 616). Among upper Magdalena faunas historically known to contain freshwater fishes, only the La Venta fauna has been extensively studied (Ballen and Moreno-Bernal, 2019; Lundberg, 1997, 2005; Lundberg and Chernoff, 1992; Lundberg et al., 1986, 2010), with other records restricted to to order-level identification or simply indeterminate fish remains. Herein, one remain was identified to genus level and another to family.

The Ware fauna is by far the most rich of both faunas herein studied with ten recognized taxa, mostly to genus level (Table 4.2, Figure 4.18). As mentioned, the genera *Hemidoras*, *Serrasalmus*, *Trachelyopterichthys*, and possibly *Rhaphiodon* are entirely new records for the fossil freshwater fishes of South America. The genera *Hydrolycus*, *Platysilurus*, *Phractocephalus*, and *Zungaro* are already known from other fossil faunas of South America; however, all records herein reported are new for the trans-Andean region, and almost all being also the youngest of the fossil record of their respective groups (Lundberg et al., 2010). The only exception is the genus *Platysilurus*, whose occurrence in the Rio Yuca formation of Venezuela is also of Pliocene age (Bermúdez et al., 2015). The genus *Pygocentrus* in the La Venta fauna is the first reliable occurrence of the genus in the fossil record.

4.5.3 Paleoenvironments

Fossil-bearing facies of both the Sincelejo and Ware formations share granulometric and sedimentary features typical of high-energy environments such as rivers, subject to high rates of transport and erosion. Although transport should have taken place in such a high-energy depositional environment, this process appears to have been only local in scale, as judging from the size class mixture of the collected remains (small fragments at the millimeter order of magnitude are mixed with centimeter-scale particles). Large-scale transport, on the other hand, would have generated a stronger size sorting that is not present in the fossiliferous facies of the Ware stratopype, where large-sized remains such as crocodylian bones are intermingled with small-sized fragments such as catfish spine fragments and piranha teeth, roughly spanning the cm to mm orders of magnitude. On the other hand, if large-scale transport was present and mixed with local supply of bone fragments, differential degrees of erosion would be present with small specimens better preserved than large-sized remains, the latter having traveled a longer erosive distance. Detailed taphonomic studies could shed light on the likely degree of transport and sorting in this fossil assemblage.

Some of the taxa recovered show certain degree of ecological specialization. For instance,



,

Figure 4.18: *Stratigraphic distribution of the taxa recovered in this study. A) San Francisco farm section with* Leporinus *or* Hypomasticus *and* Phractocephalus *sp. B) Ware formation stratotype with* Brachyplatystoma *cf.* vaillantii*,* Hydrolycus *spp., and* Serrasalmus *at 1.7m;* Hemidoras *sp.,* Hydrolycus *spp.,* Phractocephalus *sp., cf.* Oxydoras *sp., cf.* Rhaphiodon *sp.,* Serrasalmus *sp.,* Trachelyopterichthys *sp., and* Zungaro *sp. at 4m;* Hydrolycus *spp. and* Platysilurus *sp. at 5m. Staff scales in stratigraphic meters. Vertical guides are granulometry; Mud = mud, Ss = sand, Pb = pebbles, fn = fine, md = medium, co = coarse. Stratigraphic columns as in Figure 4.3. Sources of the pictures: Cliff on wikicommons (*Phractocephalus hemioliopterus*), ANSP/G.W. Saul (*Zungaro zungaro*), Jonathan Armbruster (*Oxydoras niger*), Citron on wikicommons (*Leporinus friderici*), Clinton & Charles Robertson (*Hydrolycus tatauaia*, Serrasalmus *sp.*, Rhaphiodon vulpinus*,* Trachelyopterichthys taeniatus*), Mark Sabaj (*Brachyplatystoma vaillantii*,* Hemidoras morrisi*), and Galvis et al. (1997,* Platysilurus malarmo*).*

large catfish species today (e.g., *Brachyplatystoma*, *Phractocephalus*, *Platysilurus*, and *Zungaro*) are known to undergo seasonal reproductive migrations in large rivers, therefore requiring large drainages (Carolsfeld et al., 2003). This combination of taxa strongly suggest that the Ware formation is composed by sediments of a medium to large-sized river that was part of a large drainage where these taxa could migrate. Cynodontid fishes (*Hydrolycus* and *Rhaphiodon*) apparently also take part in migrations, although not to the same distances as pimelodid catfishes.

Most of the taxa preserved in the Ware and Sincelejo formations are of the piscivorous–carnivorous trophic guild, with a notable absence of herbivorous especialists such as pacus and common omnivores such as characids; this can be explained by the fact that these high-order nodes in trophic networks tend to be the largest in body size in Neotropical environments, and facilitates its recovery in fossil samples while more fragile and smaller fishes tend to be less represented or even absent at all. Pacu teeth (see e.g. Chapter 3) are very resistant and their remains are expected to be recovered in future collection efforts as they are a prominent component in most South American Neogene faunas. A minor proportion tends to be more directed towards detritivory (e.g., the Doradidae).

Although the Characidae is a prominent, very rich, abundant and important component of the biomass in fish communities in the Neotropics, they are rarely represented in fossil samples and are generally restricted to isolated teeth.Abrasion easily destroys the laminar and fragile bones of these fishes. This also explains the fact that piranhas are only known from isolated teeth since their bones are as fragile as those of most Characiformes. Such preservation bias explains the general lack of representativeness of this order in the fossil record in South America (Lundberg et al., 2010).

The dissimilarity analysis recovered a pattern consistent with previous analyses prior to the study of the Ware fauna (Figure 4.17). The main difference relative to previous analyses (Figure 3.6 in Chapter 3) is the position of the Ware fauna, at the base of the other Neogene fossil faunas. As was found in Chapter 3, this pattern is expected given the amount of taxa that are unique to the Ware fauna, as well as the few shared components with other fossil faunas. The lack of correlation with geographic proximity is a pattern already recovered in previous studies using mammal faunas (Carrillo et al., 2015), and freshwater fishes (Ballen and de Pinna, in prep., Chapter 3). This pattern may be affected by an insufficient knowledge of the Peruvian faunas as already pointed out in Chapter 3, or perhaps highlight ecological heterogeneity among sites. Until the peruvian fish faunas are studied in detail, we might not be able to distinguish between the two alternatives.

4.5.4 Paleogeography

The rich fossil assemblage herein studied shows a consistent pattern of association with cis-Andean groups. As already mentioned, both taxa from the Sincelejo fauna are currently restricted to drainages west to the Andes (*Leporinus* or *Hypomasticus*, and *Phractocephalus*),

while the phylogenetic position of the latter reinforces its paleogeographic relevance. Although the Sincelejo fauna is small, its two fossil representatives suggest that there was a drainage connection between the Magdalena-Cauca drainage and the Amazon-Orinoco drainages by the Pliocene.

The Ware fauna is composed of taxa that are consistently part of groups restricted to the drainages east to the Andes (*Brachyplatystoma*, *Hemidoras*, *Hydrolycus*, *Phractocephalus*, *Serrasalmus*, *Trachelyopterichthys*, and *Zungaro*). The doubtful occurrence of *Rhaphiodon* also adds to this pattern, because their family is entirely restricted to cis-Andean drainages. The genus *Platysilurus* provides ambiguous paleogeographic information given that one of its extant species is cis-Andean while the other is found in the Maracaibo drainage (trans-Andean). *Platysilurus* from Ware and potentially those from the Rio Yuca and Urumaco formations in Venezuela do not seem conspecific with *Platysilurus malarmo*, the trans-Andean species. The only occurrence confidently identified to species level, *Hydrolycus scomberoides*, is a component restricted to the Amazon drainage, today absent in the Orinoco drainage, therefore reinforcing the association suggested by other taxa; it is intriguing that this species is currently absent in the Orinoco drainage while present in northern Colombia by the Pliocene. This indicates a case of regional extirpation.

The strong resemblance between the Sincelejo and Ware faunas on the one hand, and cis-Andean freshwater fish communities on the other, is not an isolated pattern., It conforms with the same composition found in the La Venta fauna, where almost all of the taxa in the fossil record are nowadays restricted to the Amazon drainage (Ballen and Moreno-Bernal, 2019; Lundberg, 1997, 2005). Present results would support the model of drainage connection to the mid- to late Pliocene by Lundberg for the La Venta fauna. The present data, however, extend that connection to the Pliocene, suggesting that some kind of drainage connection was still present at that period across the Cordillera Oriental of Colombia and/or the Merida Andes in Venezuela. A possible geological pattern explaining this persistence is the late uplift of the Garzón massif in the southern Cordillera Oriental during the Pliocene, creating room for drainage connection between the upper Magdalena and the Amazon-Orinoco drainages (Saeid et al., 2017). This kind of connection would explain both the fish composition and the required drainage properties such as a long course and connection to cis-Andean drainages. Alternative mechanisms such as connections across the Merida Andes in Venezuela have also been suggested in the literature, although they seem less likely as they require extensive uplift in a large area during the past million years (Audemard, 2003; Audemard and Audemard, 2002). The latter model has also been challenged in the literature because the present-day elevation of the Merida Andes is thought to have been reached during the late Miocene to Pliocene (Bermúdez et al., 2015). Anyway, the biotic component is pointing to a consistent pattern of drainage connection that calls for a reassessment of our geological models of Andean orogeny in the Cordillera Oriental. Further combination of fossil information, molecular and morphological data, tectonic models, and complex probabilistic modeling has the potential to provide a better understanding of the process of mountain

building, its associated drainage evolution, and ultimately the evolution of the neotropical biota.

Bibliography

Agostinho, A.A., Gomes, L.C., Suzuki, H.I., and Júlio, H. 2003. Migratory Fishes of the Upper Paraná River Basin. In: (, eds.). World Fisherie. In J. Carolsfeld, B. Harvey, C. Ross, and A. Baer (eds.), *Migratory Fishes of South America: Biology, Fisheries and Conservation Status*, pp. 19–99. International Development Research Centre, Ottawa. 110

Aguilera, O., Oliveira, G.A.S., Lopes, R.T., Machado, A.S., dos Santos, T.M., Marques, G., Bertucci, T., Aguiar, T., Carrillo-Briceño, J.D., Rodriguez, F., and Jaramillo, C. 2017. Neogene Proto-Caribbean Porcupinefishes (Diodontidae). *PLoS One*, 12:1–26. 79

Aguilera, O.A., Bocquentin, J., Lundberg, J.G., and Maciente, A. 2008. A new cajaro catfish (Siluriformes: Pimelodidae: *Phractocephalus*) from the Late Miocene of southwestern Amazonia and its relationship to *Phractocephalus nassi* of the Urumaco Formation. *Palaeontologische Zeitschrift*, 82:231–245. 103

Aguilera, O.A., Lundberg, J.G., Birindelli, J.L.O., Sabaj Pérez, M.H., Jaramillo, C.A., and Sánchez-Villagra, M.R. 2013a. Palaeontological evidence for the last temporal occurrence of the ancient Western Amazonian river outflow into the Caribbean. *PLoS ONE*, 8:1–17. 74, 75, 80, 86, 103, 106

Aguilera, O.A., Moraes-Santos, H., Costa, S.A.R.F., Ohe, F., Jaramillo, C.A., and Nogueira, A. 2013b. Ariid sea catfishes from the coeval Pirabas (Northeastern Brazil), Cantaure, Castillo (Northwestern Venezuela), and Castilletes (North Colombia) formations (early Miocene), with description of three new species. *Swiss Journal of Palaeontology*, 132:45–68. 75, 79

Albert, J.S., Lovejoy, N.R., and Crampton, W.G.R. 2006. Miocene tectonism and the separation of cis- and trans-Andean river basins: Evidence from Neotropical fishes. *Journal of South American Earth Sciences*, 21:14–27. 74

Albert, J.S., Petry, P., and Reis, R.E. 2011. Major biogeographic patterns. In J.S. Albert and R.E. Reis (eds.), *Historical Biogeography of Neotropical Freshwater Fishes*, pp. 21–57. University of California Press. 74

Alfaro, E. and Holz, M. 2015. Stratigraphic relationships between the Colombian, Sinú

Offshore and Sinú-San Jacinto basins based on seismic stratigraphy. *Brazilian Journal of Geology*, 44:607–625. 80

Amson, E., Carrillo, J.D., and Jaramillo, C.A. 2016. Neogene sloth assemblages (Mammalia, Pilosa) of the Cocinetas Basin (La Guajira, Colombia): implications for the Great American Biotic Interchange. *Palaeontology*, 59:563–582. 74, 79, 80, 83, 84, 85, 86

Antoine, P.O., Abello, M.A., Adnet, S., Altamirano Sierra, A.J., Baby, P., Billet, G., Boivin, M., Calderón, Y., Candela, A., Chabain, J., Corfu, F., Croft, D.A., Ganerød, M., Jaramillo, C.A., Klaus, S., Marivaux, L., Navarrete, R.E., Orliac, M.J., Parra, F., Pérez, M.E., Pujos, F., Rage, J.C., Ravel, A., Robinet, C., Roddaz, M., Tejada-Lara, J.V., Vélez-Juarbe, J., Wesselingh, F.P., and Salas-Gismondi, R. 2016. A 60-million-year Cenozoic history of western Amazonian ecosystems in Contamana, eastern Peru. *Gondwana Research*, 31:30–59. 75

Arce H, M., Reis, R.E., Geneva, A.J., and Sabaj Pérez, M.H. 2013. Molecular Phylogenetics and Evolution Molecular phylogeny of thorny catfishes (Siluriformes : Doradidae). *Molecular Phylogenetics and Evolution*, 67:560–577. 101

Audemard, F.A. 2003. Geomorphic and geologic evidence of ongoing uplift and deformation in the Mérida Andes, Venezuela. *Quaternary International*, 101-102:43–65. 116

Audemard, F.E. and Audemard, F.A. 2002. Structure of the Mérida Andes, Venezuela: Relations with the South America-Caribbean geodynamic interaction. *Tectonophysics*, 345:299–327. 116

Ayala, R.C., Bayona, G., Cardona, A., Ojeda, C., Montenegro, O.C., Montes, C., Valencia, V., and Jaramillo, C.A. 2012. The paleogene synorogenic succession in the northwestern Maracaibo block: Tracking intraplate uplifts and changes in sediment delivery systems. *Journal of South American Earth Sciences*, 39:93–111. 74

Azpelicueta, M.d.l.M. and Cione, A.L. 2016. A southern species of the tropical catfish genus *Phractocephalus* (Teleostei: Siluriformes) in the Miocene of South America. *Journal of South American Earth Sciences*, 67:221–230. 75, 103

Bacca, J.M., Pardo, O.H., and Vasquez, L.E. 2010. Determinación de la geometría del Acuífero de Morrosquillo y geología detallada del Acuífero de Toluviejo, Sucre. *Geología colombiana*, 35. 80, 81

Ballen, G.A. and Moreno-Bernal, J.W. 2019. New records of the enigmatic Neotropical fossil fish *Acregoliath rancii* (Teleostei *incertae sedis*) from the middle Miocene Honda group of Colombia. *Ameghiniana*, 56:431–440. 75, 113, 116

Bayona, G., Cardona, A., Jaramillo, C.A., Mora, A., Montes, C., Caballero, V., Mahecha, H., Lamus, F., Montenegro, O., Jimenez, G., Mesa, A., and Valencia, V. 2013. Onset of fault

reactivation in the Eastern Cordillera of Colombia and proximal Llanos Basin; response to Caribbean - South American convergence in early Palaeogene time. *Geological Society, London, Special Publications*. 74

Bayona, G., Montenegro, O., Cardona, A., Jaramillo, C.A., Lamus, F., Morón, S., Quiroz, L., Ruiz, M.C., Valencia, V., Parra, M., and Stockli, D. 2010. Estratigrafía, procedencia, subsidencia y exhumación de las unidades Paleógenas en el Sinclinal de Usme, sur de la zona axial de la Cordillera Oriental. *Geologia Colombiana*, 35:5–35. 74

Beck, E. 1921. Geology and oil resources of Colombia, the coastal plain. *Economic Geology*, 16:457–473. 80

Bermúdez, H.D., Alvarán, M., Grajales, J.A., Clemencia, L., Rosero, J.S., Guzmán, C., Ruiz, E.C., Esther, R., Jaramillo, C., and Osorno, F. 2009. Estratigrafía y evolución geológica de la secuencia sedimentaria del Cinturón Plegado de San Jacinto. *Memorias del XII Congreso Colombiano de Geología*, 98:1–27. 80, 81

Bermúdez, M.A., Hoorn, C., Bernet, M., Carrillo, E., van der Beek, P.A., Garver, J.I., Mora, J.L., and Mehrkian, K. 2015. The detrital record of late-Miocene to Pliocene surface uplift and exhumation of the Venezuelan Andes in the Maracaibo and Barinas foreland basins. *Basin Research*, pp. 1–26. 113, 116

Birindelli, J.L.O. 2014. Phylogenetic relationships of the South American Doradoidea (Ostariophysi: Siluriformes). *Neotropical Ichthyology*, 12:451–564. 97, 101

Bogan, S., Sidlauskas, B., Vari, R.P., and Agnolin, F.L. 2012. *Arrhinolemur scalabrinii* Ameghino, 1898, of the late Miocene - a taxonomic journey from the Mammalia to the Anostomidae (Ostariophysi: Characiformes). *Neotropical Ichthyology*, 10:555–560. 75

Boni, T.A., Padial, A.A., Prioli, S.M., Lucio, L.C., Maniglia, T.C., Bignotto, T.S., Panarari-Antunes, R.S., Prioli, R.A., and Prioli, A.J. 2011. Molecular differentiation of species of the genus *Zungaro* (Siluriformes, Pimelodidae) from the Amazon and Paraná-Paraguay River basins in Brazil. *Genetics and Molecular Research*, 10:2795–2805. 108

Caballero, V., Parra, M., Mora, A., and Mora Bohorquez, A.R. 2010. Levantamiento de la Cordillera Oriental de Colombia durante el Eoceno tardío-Oligoceno temprano: Proveniencia sedimentaria en el Sinclinal de Nuevo Mundo, Cuenca Valle Medio del Magdalena. *Geologia Colombiana*, 32:45–77. 74

Cadena, E. and Jaramillo, C.A. 2015a. Early to Middle Miocene Turtles from the Northernmost Tip of South America: Giant Testudinids, Chelids, and Podocnemidids from the Castilletes Formation, Colombia. *Ameghiniana*, 52:188–203. 74, 79

Cadena, E. and Jaramillo, C.A. 2015b. The first fossil skull of *Chelus* (Pleurodira: Chelidae, Matamata turtle) from the early Miocene of Colombia. *Palaeontologica Electronica*, 18.2:1–10. 74, 79

Calegari, B.B., Vari, R.P., and Reis, R.E. 2019. Phylogenetic systematics of the driftwood catfishes (Siluriformes: Auchenipteridae): a combined morphological and molecular analysis. *Zoological Journal of the Linnean Society*, 187:661–773. 97

Carolsfeld, J., Harvey, B., Ross, C., and Bae, A. 2003. *Migratory Fishes of South America. Biology, Fisheries, and Conservation Status*. International Development Research Centre, Ottawa. 74, 115

Carrillo, J.D., Amson, E., Jaramillo, C., Sánchez, R., Quiroz, L., Cuartas, C., Rincón, A.F., and Sánchez-Villagra, M.R. 2018. The Neogene Record of Northern South American Native Ungulates. *Smithsonian Contributions to Paleobiology*, pp. 1–67. 74, 79, 80, 83, 84, 85, 86

Carrillo, J.D., Forasiepi, A., Jaramillo, C., and Sánchez-Villagra, M.R. 2015. Neotropical mammal diversity and the great American biotic interchange: Spatial and temporal variation in South America's fossil record. *Frontiers in Genetics*, 5:1–11. 115

Carrillo-Briceño, J.D., Argyriou, T., Zapata, V., Kindlimann, R., and Jaramillo, C. 2016. A new early Miocene (Aquitanian) elasmobranchii assemblage from the la Guajira Peninsula, Colombia. *Ameghiniana*, 53:77–99. 79

Carrillo-Briceño, J.D., Luz, Z., Hendy, A., Kocsis, L., Aguilera, O., and Vennemann, T. 2019. Neogene Caribbean elasmobranchs: Diversity, paleoecology and paleoenvironmental significance of the Cocinetas Basin assemblage (Guajira Peninsula, Colombia). *Biogeosciences*, 16:33–56. 79, 80, 84, 85, 86

Castello, L. 2008. Lateral migration of Arapaima gigas in floodplains of the Amazon. *Ecology of Freshwater Fish*, 17:38–46. 74

Castresana, J. 2000. Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. *Molecular Biology and Evolution*, 17:540–552. 77

Cione, A.L. and Azpelicueta, M.d.l.M. 2013. The first fossil species of *Salminus*, a conspicuous South American freshwater predatory fish (Teleostei , Characiformes), found in the Miocene of Argentina. *Journal of Vertebrate Paleontology*, 33:1051–1060. 75

Cione, A.L., Azpelicueta, M.d.l.M., Bond, M., Carlini, A.A., Casciotta, J.R., Cozzuol, M.A., de la Fuente, M., Gasparini, Z., Goin, F.J., Noriega, J., Scillato-Yané, G.J., Soibelzon, L., Tonni, E.P., Verzi, D., and Vucetich, M.G. 2000. The Miocene vertebrates from Paraná, eastern Argentina Miocene vertebrates from Entre Ríos province, eastern Argentina. *Correlación Geológica*, 14:191–237. 75

Cione, A.L., Dahdul, W.M., Lundberg, J.G., and Machado-Allison, A. 2009. *Megapiranha paranensis*, a new genus and species of Serrasalmidae (Characiformes, Teleostei) from the upper Miocene of Argentina. *Journal of Vertebrate Paleontology*, 29:350–358. 75

Clavijo-Torres, J. and Barrera-Olmo, R. 2001. Planchas 44 y 52, Sincelejo y Sahagún. Escala 1:100.000. Memoria Explicativa. Technical report, Ingeominas. 81, 82

Dagosta, F.C.P. and de Pinna, M.C.C. 2019. The fishes of the Amazon: Distribution and biogeographical patterns, with a comprehensive list of species. *Bulletin of the American Museum of Natural History*, 431:1–163. 93, 101, 108

Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, 9:772–772. 77

de Porta, J. 1962. El terciario superior en los alrededores de Sincelejo. Technical report, Servicio Geológico Nacional. 80, 81

Diaz de Gamero, M.L. 1996. The changing course of the Orinoco River during the Neogene: A review. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 123:385–402. 74

Drummond, A.J., Suchard, M.A., Xie, D., and Rambaut, A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, 29:1969–1973. 77

Dueñas, H. and Duque-Caro, H. 1981. Geología del cuadrángulo F-8. *Boletin Geologico Ingeominas*, 24:1–35. 80, 81

Duque-Caro, H. 1966. Observaciones generales a la bioestratigrafía y geología regional en los Departamentos de Bolívar y Córdoba. 80, 81

Duque-Caro, H. 1967. Informe bioestratigráfico preliminar de los cuadrángulos D-8 y E-8. Technical report, Servicio Geológico Nacional. 80, 81

Duque-Caro, H. 1984. Structural style, diapirism, and accretionary episodes of the Sinú-San Jacinto terrane, southwestern Caribbean borderland. *Geological Society of America Memoirs*, 162:303–316. 80

Flórez, P., Zapata-Ramírez, P., and Klaus, J.S. 2018. Early Miocene shallow-water corals from La Guajira, Colombia: Part II, Mussidae–Siderastreidae and Milleporidae. *Journal of Paleontology*, pp. 1–21. 79

Flórez, P., Zapata-Ramírez, P., and Klaus, J.S. 2019. Early Miocene shallow-water corals from la Guajira, Colombia: Part I, Acroporidae-Montastraeidae. *Journal of Paleontology*, 93:1–24. 79

Forasiepi, A.M., Soibelzon, L.H., Suárez Gomez, C., Sánchez, R., Quiroz, L.I., Jaramillo, C.A., and Sánchez-Villagra, M.R. 2014. Carnivorans at the Great American Biotic Interchange: new discoveries from the northern neotropics. *Naturwissenschaften*, 101:965–74. 74, 79, 80, 86

Galvis, G., Mojica, J.I., and Camargo, M. 1997. *Peces del Catatumbo*. Asociación Cravo Norte, Santafé de Bogotá. xviii, 114

Goloboff, P.A., Farris, J.S., and Nixon, K.C. 2008. TNT, a free program for phylogenetic analysis. *Cladistics*, 24:774–786. 76

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology*, 59:307–321. 77

Hendy, A.J.W., Jones, D.S., Moreno, F., Zapata, V., and Jaramillo, C.A. 2015. Neogene molluscs, shallow marine paleoenvironments, and chronostratigraphy of the Guajira Peninsula, Colombia. *Swiss Journal of Palaeontology*, pp. 1–31. 79, 80

Kassem, T., Cáceres, C., and Cucalón, I. 1964. Geología del Cuadrángulo E-8 Sincelejo. Technical report, Servicio Geológico Nacional, Bogotá. 80

Katoh, K. and Standley, D.M. 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30:772–780. 77

Lundberg, J.G. 1997. Freshwater fishes and their paleobiotic implications. In R.F. Kay, R.H. Madden, R.L. Cifelli, and J.J. Flynn (eds.), *Vertebrate Paleontology in the Neotropics: The Miocene Fauna of La Venta, Colombia*, chapter 5, pp. 67–91. Smithsonian Institution Press. 74, 97, 113, 116

Lundberg, J.G. 2005. *Brachyplatystoma promagdalena*, new species, a fossil goliath catfish (Siluriformes: Pimelodidae) from the Miocene of Colombia, South America. *Neotropical Ichthyology*, 3:597–605. 103, 113, 116

Lundberg, J.G. and Aguilera, O.A. 2003. The late Miocene *Phractocephalus* catfish (Siluriformes: Pimelodidae) from Urumaco, Venezuela: additional specimens and reinterpretation as a distinct species. *Neotropical Ichthyology*, 1:97–109. 103

Lundberg, J.G. and Akama, A. 2005. *Brachyplatystoma capapretum*: a new species of Goliath Catfish from the Amazon basin, with a reclassification of allied catfishes (Siluriformes: Pimelodidae). *Copeia*, 2005:492–516. 103

Lundberg, J.G. and Chernoff, B. 1992. A Miocene Fossil of the Amazonian Fish *Arapaima* (Teleostei , Arapaimidae) from the Magdalena River Region of Colombia–Biogeographic and Evolutionary Implications. *Biotropica*, 24:2–14. 113

Lundberg, J.G., Machado-Allison, A., and Kay, R.F. 1986. Miocene characid fishes from Colombia; evolutionary stasis and extirpation. *Science*, 234:208–209. 113

Lundberg, J.G., Marshall, L.G., Guerrero, J., Horton, B., Malabarba, M.C.S.L., and Wesselingh, F. 1998. The stage for Neotropical Fish diversification: A history of Tropical South American rivers. In L.R. Malabarba, R.E. Reis, R.P. Vari, Z.M. Lucena, and C.A.S. Lucena (eds.), *Phylogeny and Classification of Neotropical Fishes*, pp. 13–48. Edipucrs, Porto Alegre. 74

Lundberg, J.G. and McDade, L.A. 1986. On the South American Catfish *Brachyrhamdia imitator* Myers (Siluriformes, Pimelodidae), with phylogenetic evidence for a large intrafamilial lineage. *Notulae Naturae*, 463. 75

Lundberg, J.G., Sabaj Pérez, M.H., Dahdul, W.M., and Aguilera, O.A. 2010. The Amazonian Neogene fish fauna. In C. Hoorn and F.P. Wesselingh (eds.), *Amazonia, Landscape and Species Evolution: A Look Into the Past*, pp. 281–301. Blackwell Publishing. 75, 103, 106, 110, 113, 115

Lundberg, J.G., Sullivan, J.P., and Hardman, M. 2011. Phylogenetics of the South American Catfish family Pimelodidae (Teleostei: Siluriformes) using nuclear and mitochondrial gene sequences. *Proceedings of the Academy of Natural Sciences of Philadelphia*, 161:153–189. 77, 90

Maddison, W.P. and Maddison, D.R. 2011. Mesquite: a modular system for evolutionary analysis. 76

Marshall, L.G., Hoffstetter, R., and Pascual, R. 1983. Mammals and stratigraphy: Geochronology of the Continental Mammal-bearing Tertiary of South America. *Paleovertebrata*, pp. 1–93. 82

Mattox, G.M.T. and Toledo-Piza, M. 2012. Phylogenetic study of the Characinae (Teleostei: Characiformes: Characidae). *Zoological Journal of the Linnean Society*, 165:809–915. 76

Moreno, F., Hendy, A.J.W., Quiroz, L., Hoyos, N., Jones, D.S., Zapata, V., Zapata, S., Ballen, G.A., Cadena, E., Cárdenas, A.L., Carrillo-Briceño, J.D., Carrillo, J.D., Delgado-Sierra, D., Escobar, J., Martínez, J.I., Martínez, C., Montes, C., Moreno, J., Pérez, N., Sánchez, R., Suárez, C., Vallejo-Pareja, M.C., and Jaramillo, C.A. 2015. Revised stratigraphy of Neogene strata in the Cocinetas Basin, La Guajira, Colombia. *Swiss Journal of Palaeontology*, 134:5–43. xvi, 78, 83, 85, 86, 88

Moreno-Bernal, J.W., Head, J., and Jaramillo, C.A. 2016. Fossil crocodilians from the High Guajira Peninsula of Colombia: Neogene faunal change in northernmost South America. *Journal of Vertebrate Paleontology*, 36:e1110586. 74, 79, 80, 86

Ochoa, D., Hoorn, C., Jaramillo, C.A., Bayona, G., Parra, M., and De la Parra, F. 2012. The final phase of tropical lowland conditions in the axial zone of the Eastern Cordillera of Colombia: Evidence from three palynological records. *Journal of South American Earth Sciences*, 39:157–169. 74

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., and Wagner, H. 2019. vegan: Community Ecology Package. 76

R Core Development Team. 2018. R: A language and environment for statistical computing. 76

Ramirez, J.L., Birindelli, J.L.O., and Jr, P.M.G. 2017. A new genus of Anostomidae (Ostariophysi: Characiformes): Diversity, phylogeny and biogeography based on cytogenetic, molecular and morphological data. *Molecular Phylogenetics and Evolution*, 107:308–323. 87

Renz, O. 1960. Geología de la parte sureste de la Peninsula de la Guajira (Republica de Colombia). *Memorias del Boletín Geológico*, 3:317–347. 78

Rincón, A.D., Solórzano, A., Macsotay, O., McDonald, H.G., and Núñez-Flores, M. 2016. A new Miocene vertebrate assemblage from the Río Yuca Formation (Venezuela) and the northernmost record of typical Miocene mammals of high latitude (Patagonian) affinities in South America. *Geobios*, 49:395–405. 106, 108

Rollins, J.F. 1965. Stratigraphy and structure of the Goajira Peninsula, northwestern Venezuela and northeastern Colombia. *University of Nebraska Studies New Series*, 30:1–103. 78

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., and Huelsenbeck, J.P. 2012. MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61:539–542. 77

Royo y Gomez, J. 1946. Los vertebrados del Terciario continental Colombiano. *Revista de la Academia Colombiana de Ciencias Exactas, Fisicas y Naturales*, 6:496–512. 81

Sabaj Pérez, M.H., Aguilera, O.A., and Lundberg, J.G. 2007. Fossil Catfishes of the Families Doradidae and Pimelodidae (Teleostei: Siluriformes) from the Miocene Urumaco Formation of Venezuela. *Proceedings of the Academy of Natural Sciences of Philadelphia*, 156:157–194. 106, 108

Saeid, E., Bakioglu, K.B., Kellogg, J., Leier, A., Martinez, J.A., and Guerrero, E. 2017. Garzón Massif basement tectonics: Structural control on evolution of petroleum systems

in upper Magdalena and Putumayo basins, Colombia. *Marine and Petroleum Geology*, 88:381–401. 116

Schaefer, S.A., Weitzman, S.H., and Britski, H.A. 1989. Review of the Neotropical Catfish genus Scoloplax (Pisces: Loricarioidea: Scoloplacidae) with comments on reductive characters in phylogenetic analysis. *Proceedings of the Academy of Natural Sciences of Philadelphia*, 141:181–211. 74

Sidlauskas, B.L. and Vari, R.P. 2008. Phylogenetic relationships within the South American fish family Anostomidae (Teleostei, Ostariophysi, Characiformes). *Zoological Journal of the Linnean Society*, 154:70–210. 87

Slobodian, V. and Pastana, M.N. 2018. Description of a new *Pimelodella* (Siluriformes: Heptapteridae) species with a discussion on the upper pectoral girdle homology of Siluriformes. *Journal of Fish Biology*, 93:901–916. 75

Stirton and R., A. 1947. A rodent and a peccary from the Cenozoic of Colombia, t. 7. *Serv. Geol. Nac., Compilacion Estud. Geol. Ofic. Colombia*, 7:317–324. 81

Stirton, R.A. 1953. Vertebrate paleontology and continental stratigraphy in Colombia. *Geological Society of America Bulletin*, 64:603–622. 81, 113

Suarez, C., Forasiepi, A.M., Goin, F.J., and Jaramillo, C.A. 2016. Insights into the Neotropics Prior to the Great American Biotic Interchange: New evidence of mammalian predators from the Miocene of Northern Colombia. *Journal of Vertebrate Paleontology*, 36:e1029581. 74, 79

Talavera, G. and Castresana, J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology*, 56:564–577. 77

Tejada-Lara, J.V., Salas-Gismondi, R., Pujos, F., Baby, P., Benammi, M., Brusset, S., de Franceschi, D., Espurt, N., Urbina, M., and Antoine, P.O. 2015. Life in proto-Amazonia: Middle Miocene Mammals from the Fitzcarrald Arch (Peruvian Amazonia). *Palaeontology*, 58:341–378. 75

Toledo-Piza, M. 2000. The Neotropical fish subfamily Cynodontinae (Teleostei: Ostariophysi: Characiformes): A phylogenetic study and a revision of *Cynodon* and *Rhaphiodon*. *American Museum Novitates*, 3286:1–88. 75, 76

Toledo-Piza, M., Menezes, N.A., and dos Santos, G.M. 1999. Revision of the neotropical fish genus *Hydrolycus* (Ostariophysi: Cynodontinae) with the description of two new species. *Ichthyological Exploration of Freshwaters*, 10:255–280. 93, 94

Villarroel, C. and Clavijo, J. 2005. Los mamíferos fósiles y las edades de las sedimentitas continentales del neógeno de la costa caribe colombiana. *Academia Colombiana de Ciencias Exactas, Fisicas y Naturales*, 29:345–356. 82

Werenfelds, A. 1926. Una sección estratigráfica a través del Terciario de Toluviejo, Colombia. *Ecole Geologique Helvetica*, 20:79–84. 80

Chapter 5

Statistical approaches in estimation of the time of separation between biogeographic areas

5.1 Abstract

Vicariance is one of the most important mechanisms and patterns in biogeography. Although the literature provides an large amount of studies on vicariance as both a patter and a process, comparatively less has been written on how to estimate the time component of such events. Divergence time estimation is a bayesian methodology for translating branch lengths into absolute time, with the potential to generate data on time of vicariant events. Although empirical studies often provide data on vicariant patterns along with their divergence time estimation as a way to estimate time for these events in a specific biological group, this constitutes the statistical analog of using a single data point in order to estimate a parameter as in the end, it is only one point in time. It is herein proposed that in presence of several, independent estimates of divergence for vicariant events, we can use statistical methods in order to provide adequate estimates of confidence regions for the timing of such events. Three different methods are herein proposed in order to estimate the date of area separation. The first methods constitutes an application of the analogy of stratigraphic intervals for area separation events in which the different divergence time estimation points behave as occurrences of the event of interest, and thus using their pattern of occurrence in order to estimate the confidence interval on endpoints parameters θ for such stratigraphic interval. The second method uses model fitting on cumulative density functions assuming that the x-intercept is the true point in time where areas became separated, and thus using the empirical cumulative density in order to estimate a confidence interval that parameter. The third method consists of resampling from the uncertainty regions of each divergence time estimation data point under the assumption that regardless of the original statistical distribution, we can design a resample procedure for estimating the probability density for

both the mean and the median. The advantages and issues for each method are discussed, in particular with respect to the assumption of measurement without error in the independent variable. These three methods are applied to divergence time estimations postulating separations between cis- and trans-Andean areas, one of the most prominent biogeographic events in the Neotropics. Differences in estimation are found and discussed with respect to the expected behavior of each method, and the relevance of the present estimations discussed with respect to fossil evidence suggesting a Pliocene age of the vicariant event between cis- and trans-Andean areas in northern South America. Given a combination of different estimation of confidence intervals, the vicariant event should have happened 2–5.8 Ma ago, a considerably younger age for this event that is estimated as 11–13 Ma in the literature. These methods are of potential application to any event in time and are not necessarily restricted to vicariant events, despite its primary application for these biogeographic events.
Keywords: Vicariance, Andes, Resampling, Cumulative density function, stratigraphic interval.

5.2 Introduction

The Andes is perhaps the most prominent geographic feature of South America along with the Amazon river. It consists of a mountain system with north-south trend running along the western flank of South America. The tectonic evolution of this system is the product of the interaction between the Nazca and South American tectonic plates, with contribution of the interaction with the Caribbean plate to the north (Gregory-Wodzicki, 2000). In addition to mountain building, the Andean uplift also affected the drainage evolution remodeling basins and is has long been pointed out as the driver of important river course shifts during the Cenozoic (Diaz de Gamero, 1996; Hoorn et al., 1995). Although the early literature on Andean uplift tended to see the process in a simplistic way in the sense that considered the uplift as general for the northern Andes and thus occurring in a specific point of geologic time, there is a growing body of literature suggesting that the tectonics of the northern Andes were very complex, spanning almost all of the Cenozoic, and proceeded by steps with successive events for specific massifs. Thus, while earlier papers suggested that the uplift of the Cordillera Oriental in Colombia occurred ca. 11-13 Ma (Diaz de Gamero, 1996; Guerrero, 1997), later contributions have demonstrated that some parts of the Cordillera suffered uplift as early as during the late Paleocene (e.g., the Santander massif, Bayona et al., 2013) while some should have suffered uplift during the Pliocene (e.g., the Garzón massif, Saeid et al., 2017). As a consequence, effective drainage separation resulting from such tectonic events should have occurred in these ages, either around 12 Ma as in the classical tectonic model for the Cordillera Oriental, or during the Pliocene (2–4 Ma) according to the modern models.

The consequences of Andean uplift for the evolution of the biota in the continent have been profound and well known, both for terrestrial (Antonelli et al., 2009; Richardson et al.,

2018; Smith et al., 2014) and aquatic organisms (Albert et al., 2006; Lundberg et al., 1998; Schaefer, 2011), with a general trend of recovery of sister-groups across the Andes, that is, between cis- and trans-Andean groups. Some studies have recovered a consistent pattern of vicariance either when looking at the phylogenetic relationships in specific groups (Řífčan et al., 2013) but also in larger analyses including phylogenies, occurrence data, and divergence time estimation (Hazzi et al., 2018).

Divergence time estimation is a multidisciplinary research area that involves the interplay among several fields such as systematics, molecular biology, paleontology, geochronology and statistics (Drummond and Bouckaert, 2015; Heath et al., 2014; Heled and Drummond, 2015; Rannala, 2016). The main goal of divergence time estimation is to use several sources of information in order to infer the absolute time at which a given speciation event took place, and therefore to generalize it to a whole phylogeny so that the time framework can be incorporated into evolutionary biology and biogeography. Such task is known as time calibration and can in principle be carried out using information from rates of mutation along branches of a phylogeny or by bracketing speciation events using fossil information and the time framework they are associated to.

Despite the wealth of information on all the topics described above (andean tectonics, biogeography, divergence time estimation), we still lack a quantitative approach at estimation of the time at which two areas suffered vicariance in the presence of information for specific biotic groups. The most common strategy is to plot the point estimates of divergence time along with their uncertainty and visually extract information on when such events should have happened (Hazzi et al., 2018). Another alternative has been to plot the mean of such data as a point estimate of general time of separation (Smith et al., 2014) as a way to summarize such general divergence time. In summary, we lack methods for inferring the age (and its uncertainty) at which the process responsible for all these vicariance events took place. The goal of this contribution is to propose three different methods for estimating the time of separation between two areas using information from the divergence times that occurred as a consequence. In order to explore the properties of such methods, the case of Andean uplift as a driver of vicariance between cis- and trans-Andean areas will be used; this does not mean that the methods are specific for this event in geologic time but instead that it can be applied to any event of vicariance in earth history for which we have information coming from divergence time estimation.

5.3 Divergence time estimation

Bayesian inference has been extensively used as a tool for incorporating several sources of calibration along with models of branching and sequence change in order to estimate divergence time for the nodes in a given phylogeny. The temporal component of such inferences has proven to be a key tool for studying several biotic phenomena such as diversification

rates (Lagomarsino et al., 2016; Rabosky et al., 2018) and biogeography (e.g., Landis, 2017; Tagliacollo et al., 2015) among its most prominent applications. An important property of bayesian divergence time estimation over alternative approaches maximum-likelihood counterparts (Britton et al., 2007) is that uncertainty is properly modeled through under a coherent probabilistic framework in the form of calibration priors (Barba-Montoya et al., 2017; Drummond and Stadler, 2016; Heath, 2012; Warnock et al., 2014).

A bayesian approach at divergence time estimation (Heath et al., 2014) uses prior information on time so that a phylogenetic tree can be calibrated, that is, its nodes located at a given scale that converts branch lengths into time:

$$f(\boldsymbol{r}, \boldsymbol{a}, \theta_r, \theta_a, \theta_s|D) = \frac{f(D|\boldsymbol{r}, \boldsymbol{a}, \theta_s)f(r|\theta_r)f(a|\theta_a)f(a|\theta_a)f(\theta_r)f(\theta_a)f(\theta_s)}{f(D)} \tag{5.1}$$

where D is the DNA sequence data, $f(r|\theta_r)f(\theta_r)$ is the molecular clock model translating branch length information into time, $f(\theta_s)$ is the model of molecular substitution, and $f(a|\theta_a)f(\theta_a)$ are the model of branching times, that is, the time at which each node of the tree τ diverges.

The result of this method is a posterior sample of topologies with posterior probabilities for nodes, and branch lengths representing point estimates of absolute time along with uncertainty in the form of a highest posterior density (HPD) (Drummond and Bouckaert, 2015).

5.4 The model

Let X be an arbitrary biogeographic area in geologic time bearing a biota B_X. At a given point in time physical earth processes divide the area in two daughter areas Y and Z in a process commonly known as vicariance

$$X \rightarrow Y, Z$$

The expectation is that the biota B_X also separates and reacts to this interruption of gene flow by speciation and thus creates a pattern of repeated sister-species-pairs between areas Y and Z, although it might also undergo extinction at a given rate, but for simplicity we will assume first that there is no extinction taking place, so that

$$B_X \rightarrow B_Y, B_Z$$

Note that the pattern preserved may also be subject to two obscuring processes: lack of speciation, and immigration from adjacent areas followed or not by speciation, thus creating non-Y, Z geographic patterns (Figure 5.1A).

Now imagine the ancestral area X previously consisting of continuous lowlands with a common drainage network, now undergoes vicariance, for instance through mountain uplift,

thus imposing a physical barrier to drainages that creates two different watersheds, and also, creates montane habitats. In sum, these two phenomena block gene flow, and promote speciation given that the lowland terrestrial biota is unable to get across mountain regions mostly due to physiological constrains. As the process also cuts the former drainage network creating two or more drainages that lack connectivity, the freshwater biota also looses gene flow, and respond speciating. This process is however not instantaneous in time, as both the geological processes will be in the scale of millions of years, and their response in the biota will also take some time after completion of area separation. This last phenomenon can be though of as a lag between actual area separation t_0 and the first speciation event t_1 that will depend on how fast a particular biological group will respond to blocking of gene flow. Then, a number of speciation events in time $t_{1:N}$ later than the true area separation t_0 will take place as a consequence of the vicariance event (Figure 5.1A). The vector of speciation events is implicitly assumed to be measured without error because the probability that either endpoint parameter is actually present within the observed interval is zero following the definition of the confidence intervals as expansions beyond the first and last observed occurrences in time (cf. Equations 5.2,5.3). Given this inferential issue, outlier removal procedures might be necessary in order to remove suspicious data that are contradicted by the overall distribution of data (e.g., a divergence estimate that is too old in geologic time, or based on misleading calibration information).

As we lack direct evidence of the event $X \rightarrow Y, Z$ we aim at using information preserved in the statistical distribution of events $t_{1:N}$ in order to estimate the time at which the vicariance event took place, t_0 in the notation above. Herein I propose two alternative ways of looking at the statistical structure of these phenomena in time: First, as a stratigraphic interval formed by the unobserved origination of the vicariance event and its end (maybe unnecessary), along with the collection of events in time that resulted from it (Figure 5.1B, section 5.5.1), and second, as the pure statistical distribution of these events in time (Figure 5.1C, section 5.5.2).

5.5 Approaches

5.5.1 Confidence intervals

Some methods are available for estimating the confidence intervals on the extreme points of a stratigraphic interval, that is, an estimation under a certain level of confidence for the point of appearance and disappearance of a given taxon in geologic time. These methods are relevant here since we can think of a biogeographic events in an analog way with a taxon in geologic time, that is, each biogeographic event has a birth and death in time in the same way as taxa appear and become extinct. This approach is so far inexistent in historical biogeography and seems to fit the main goal of determining the final point of connection between biogeographic areas. Marshall (2010) summarizes different approaches for estimat-

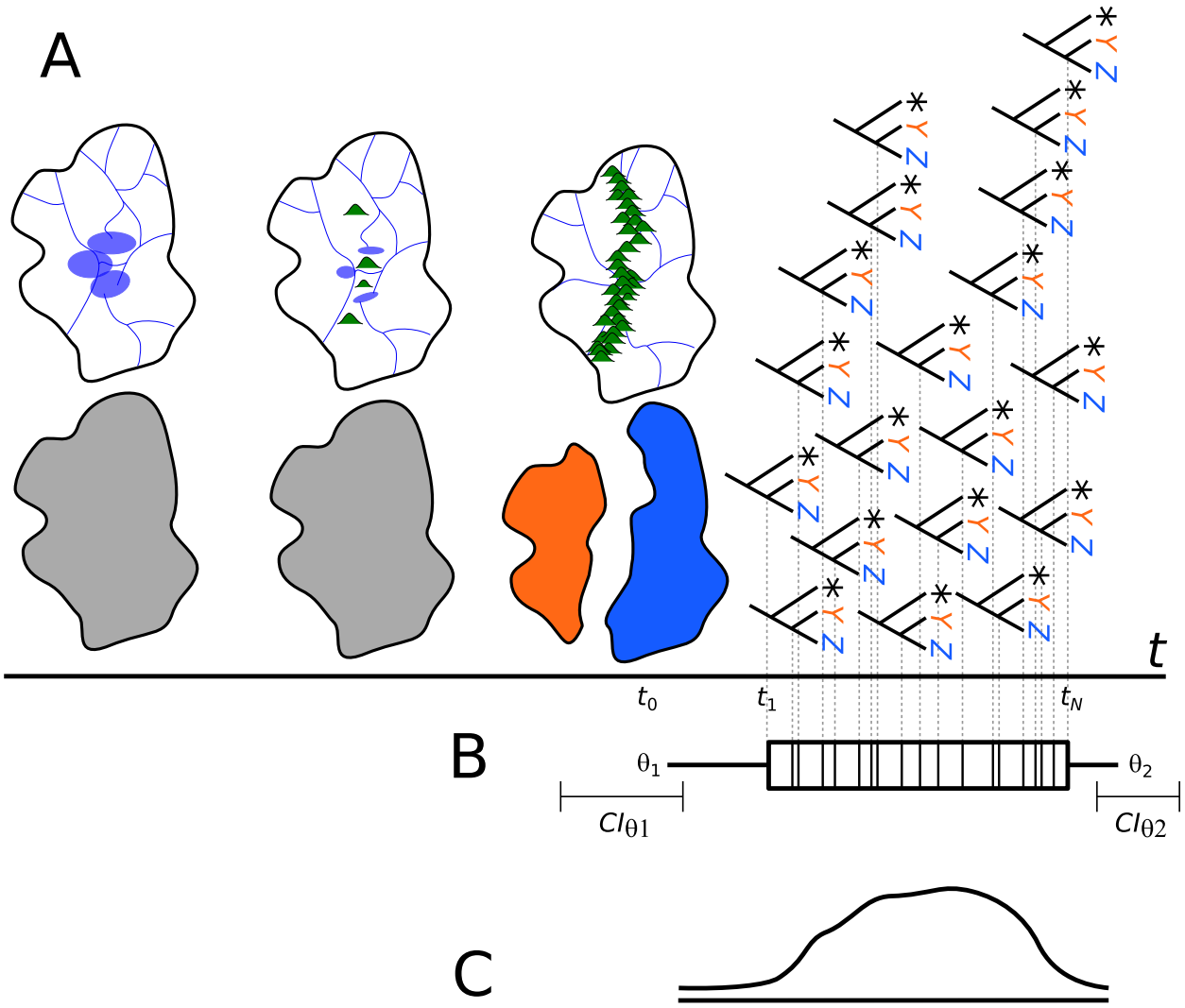


Figure 5.1: *The inferential model and two possible analytical strategies depicted. A) The separation of an arbitrary biogeographic area X along geologic time that produces daughter areas Y and Z; after a lag in time after vicariance, the biota responds speciating and the collection of sister pairs distributed in daughter areas inform about the original time of separation t_0. B) The specification of a stratigraphic interval (Section 5.5.1) and with endpoint parameters θ_1, θ_2, and all the occurrences in time $t_{1:N}$; with such information it is possible to construct confidence intervals at a nominal α value for said parameters, in particular θ_1. C) The empirical density of the divergence time events that allow to re-characterize it as a CDF and use regression for estimation of the x-intercept (Section 5.5.2).*

ing confidence intervals under certain sampling regimes, assumptions and confidence levels, ranging from classical to bayesian-based estimators. So far two of these have been coded and a third based on bayesian inference is currently under development. Summary information for each type of estimator is found in Table 5.1. Additionally, the mathematical details and speudocode describing the algorithms implemented for the first two methods are found in below and in Appendix E.5.

Table 5.1: *Methods for estimation of confidence intervals on the endpoints of stratigraphic intervals.*

Method	Preservation	Sampling	Reference
Classical CI	Random (uniform)	Continuous	Strauss and Sadler (1989)
Distribution-free	Non-random, unknown	Continuous	Marshall (1994)
Generalized CI	Non-random, known	Continuous	Marshall (2010)
Bayesian analytical solution	Uniform	Discrete	Marshall (2010)
Bayesian MCMC solution	Non-homogeneous Poisson	Discrete	Silvestro et al. (2014)

Classical CIs (Strauss and Sadler, 1989)

This method provide estimators for one- (θ_1) or two-parameter (θ_1, θ_2) cases are constructed adding and substracting a portion α of the magnitude of the observed stratigraphic interval $[y, z]$. Such constant depends on the confidence level wanted (e.g., 95%) and its calculation is different for each parameter case.

$$y - \alpha(z - y) < \theta_1 < y \tag{5.2}$$

$$z < \theta_2 < z + \alpha(z - y) \tag{5.3}$$

The constant α can be calculated either exactly (one-parameter case) or iteratively (two-parameter case) since the latter can not be solved for alpha analytically. In order to calculate it, the desired level of confidence (either p_1 or p_2)need to be provided.

$$\alpha = (1 - p_1)^{-1/(n-1)} - 1 \quad \text{For the } \theta_2 \text{ parameter case} \tag{5.4}$$

$$p_2 = 1 - 2(1 + \alpha)^{-(n-1)} + (1 + 2\alpha)^{-(n-1)} \quad \text{For the } \theta_1, \theta_2 \text{ parameter case} \tag{5.5}$$

These equations can be resolved with the following algorithms for both cases. In the case of the one-parameter case:

Algorithm 1: Pseudocode for the estimation of the one-parameter case

Data: A vector of dated occurrences $\{y, \ldots, z\}$ and the level of confidence p_1
Result: The latest occurrence z plus the upper limit estimate for θ_2
if *There are duplicate occurrences with the same age* **then**
  remove duplicates;
else
  Continue;
calculate $minAge = min\{y, \ldots, z\}$;
calculate $maxAge = max\{y, \ldots, z\}$;
calculate $R = maxAge - minAge$;
calculate $n = length\{y, \ldots, z\}$;
calculate α from Eq. (5.4);
return $maxAge + \alpha R$ or $minAge - \alpha R$;

For the two-parameter case an approximate solution based on values of α whose corresponding solution in (5.5) satisfies $|p_{2Calculated}(a_i) - p_{2Expected}| < 0.001$, that is, which value(s) of α will produce estimated values of p_2 most similar to the one expected. Despite being conceptually iterative, I used an approach of vectorized operations that calculates beforehand small values of alpha from 0 to 30 each 0.001 steps and then chooses the optimal values given the condition above:

Algorithm 2: Pseudocode for the estimation of the two-parameter case

Data: A vector of dated occurrences $\{y, \ldots, z\}$ and the level of confidence p_2
Result: A vector with the earliest occurrence minus the lower limit estimate for θ_1 and the latest occurrence z plus the upper limit estimate for θ_2
generate a vector of α values from 0 to 30 each 0.001 step;
for i *in* $\alpha = \{0, 0.001, \ldots, 29.999, 30\}$ **do**
  evaluate α_i in Eq. (5.5) and store these values in *alphae*;
filter values in *alphae* satisfying $|p_{2Calculated}(a_i) - p_{2Expected}| < 0.001$;
calculate $\alpha = mean(filteredVals)$;
if *There are duplicate occurrences with the same age* **then**
  remove duplicates;
else
  Continue;
calculate $minAge = min\{y, \ldots, z\}$;
calculate $maxAge = max\{y, \ldots, z\}$;
calculate $R = maxAge - minAge$;
calculate $n = length\{y, \ldots, z\}$;
return $\{minAge - \alpha R, maxAge + \alpha R\}$;

Distribution-free CIs (Marshall, 1994)

The method of distribution-free confidence intervals was developed without assuming any particular distribution of underlying gap sizes, instead, works calculating quantiles of gap size between occurrence points for an ordered vector of gap sizes. The major cost is that

such intervals are larger that classical CIs and on the other hand require much more data for constructing intervals of high confidence (e.g., 95%).

The method works constructing a confidence interval that has a level of confidence C (i.e., Cth percentile) for a γ confidence probability with the information from N gaps. As originally stated by Marshall (1994):

> The lower bound on the size of the confidence interval corresponding to a confidence level C is the $(x+1)th$ smallest gap of the N gaps in the stratigraphical range, where x is the largest integer that satisfies [Eq. (1) of Marshall (1994)]

Such lower bound is then the i_{th} gap in the ordered vector applied to $min(vector) - x_i$ where x_i is the magnitude of such i_{th} element in the vector. The formula is as follows:

$$\gamma > \sum_{x=0}^{x} \binom{N}{x} C^x (1-C)^{N-x} \quad \text{for } x \leq N \tag{5.6}$$

The upper limit can be constructed with the formula:

$$(1-\gamma) < \sum_{x=0}^{x} \binom{N}{x} C^x (1-C)^{N-x} \quad \text{for } x \leq N \tag{5.7}$$

The method makes use of the binomial distribution in order to find the quantiles that satisfy the condition of being smaller that γ and from these picking the largest (in the case of the lower bound) and the smallest (for the upper bound).

It is noteworthy that if only one x satisfies Eqs (5.6) or (5.7), no lower or upper confidence interval can be constructed. Also, if $x = N$ in (5.7), no upper confidence interval can be constructed since there is no $(x+1)_{th}$ gap in the vector.

The algorithm for solving Eq. (5.6) has been constructed in the following way:

Algorithm 3: Pseudocode for the estimation of the lower interval for the distribution-free confidence interval

Data: A vector of gap magnitudes between dated occurrence points $\{x_1 = n_2 - n_1, \ldots, x_N = n_N - n_{N-1}\}$ where n are occurrence points and x are their corresponding gaps, a given confidence probability $1 - 2\gamma$, and the level of confidence (or quantile) C

Result: The earliest occurrence minus the magnitude of the x_{th} gap in the ordered vector of occurrences satisfying Eq. (5.6)

initialize *rightside* vector;
for *i in* 0 *to* N **do**
    rightside = calculate the right side of the inequality;
    concatenate the result to the *rightside* vector;
initialize *sums* vector;
for *i in rightside* **do**
    sum *rightside* from the 1_{st} to the i_{th} element;
    concatenate this cumulative sum to *sums*;
count each element in sums that satisfy $test = sums < \gamma$;
identify x being the largest N satisfying Eq. (5.6) by summing the *tests* vector of 0's and 1's and assign it to the *output* vector;
if $output \geq 1$ **then**
    return $minAge - x_{output}$;
else
    break with message "Impossible to calculate lower bound";

Now the upper bound from Eq. (5.7) can be calculated in a similar way: The upper bound is found with a similar algorithm changing the magnitude to the left of the inequality and meeting the requirements in Eq. 2 of the original reference.

Algorithm 4: Pseudocode for the estimation of the upper interval for the distribution-free confidence interval

Data: A vector of gap magnitudes between dated occurrence points $\{x_1 = n_2 - n_1, \ldots, x_N = n_N - n_{N-1}\}$ where n are occurrence points and x are their corresponding gaps, a given confidence probability $1 - 2\gamma$, and the level of confidence (or quantile) C

Result: The latest occurrence plus the magnitude of the x_{th} gap in the ordered vector of occurrences satisfying Eq. (5.7)

initialize *rightside* vector;
for *i in* 0 *to* N **do**
    rightside = calculate the right side of the inequality;
    concatenate the result to the *rightside* vector;
initialize *sums* vector;
for *i in rightside* **do**
    sum *rightside* from the 1_{st} to the i_{th} element;
    concatenate this cumulative sum to *sums*;
count each element in sums that satisfy $test = sums > 1 - \gamma$;
identify x being the first element satisfying Eq. (5.6) indexing *tests* with the first element being TRUE and assign its index to the *output* vector;
if $output \in \{NA, N\}$ **then**
    break with message "Impossible to calculate upper bound";
else
    return $maxAge + x_{output}$;

Marshall (1994) presents a table with values of x satisfying either bound or both for a selected number of N, C and $1 - 2\gamma$. The algorithms already developed have been tested for several combinations of values and all they match the reported values in Marshall's Table 1.

5.5.2 Cumulative density functions

These approaches are based on the assumption that the cumulative the pattern of appearance of divergence time events from an unknown time t_0 onwards can inform on the value of such event, thus, a regression model with an additional parameter x-intercept may be thought of as an estimator for such point in time given the trend present in the linear model (Figure 5.2). Moreover, the estimation of uncertainty regions such as confidence intervals on the x-intercept can be even more informative than the point estimate itself; such regions can be constructed through fitting of empirical cumulative density function.

In probability theory, a function $f(x)$ that is a probability density function (PDF) has an associated cumulative density function (CDF) that describes the proportion of the distribution at a given value in x. For instance, given the exponential distribution

$$f(x) = \lambda e^{-\lambda x}$$

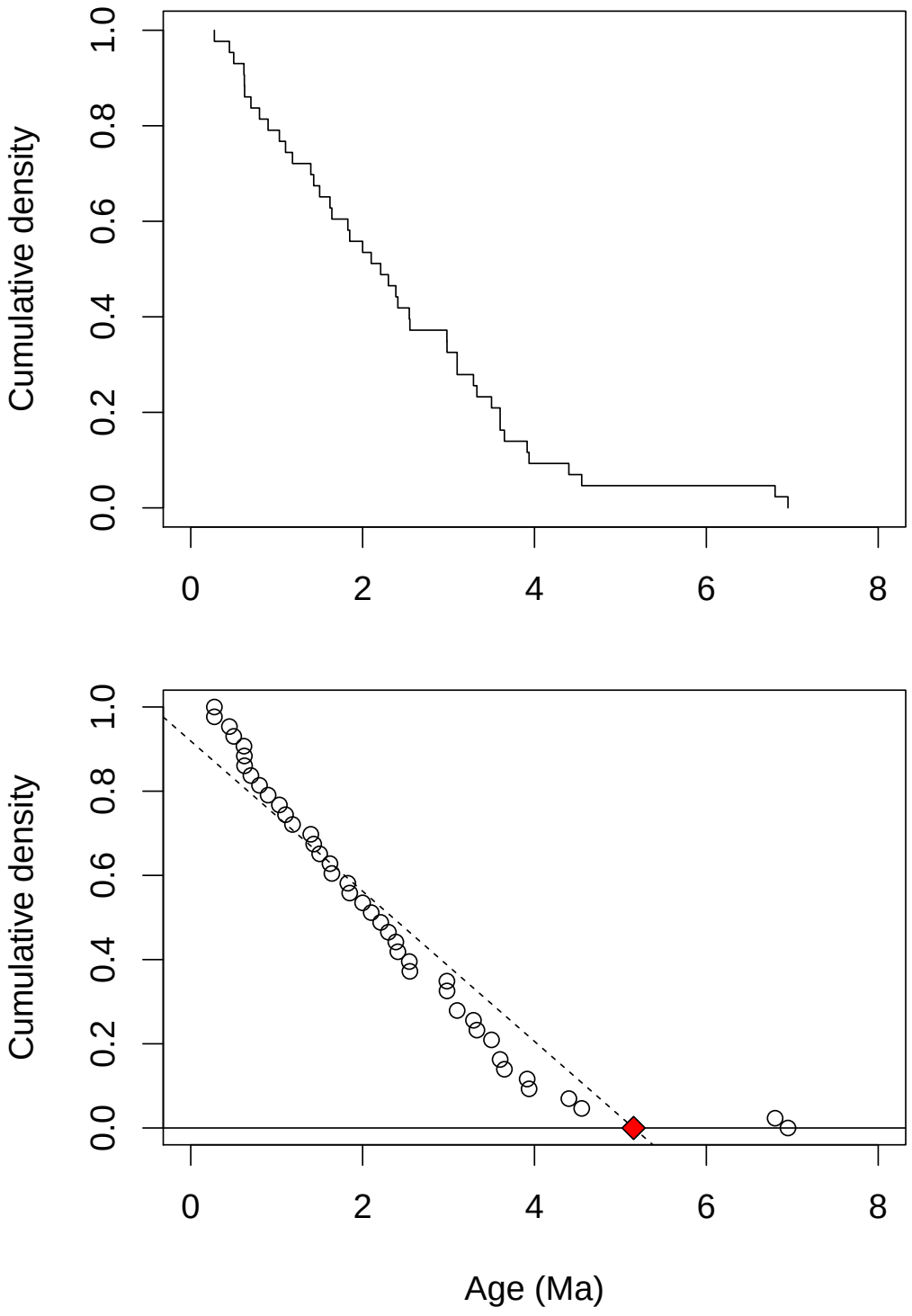


Figure 5.2: *Empirical cumulative distribution of divergence time events. Above, plot of the empirical density; below, model fit of the cumulative distribution. The interrupted line represents the regression model on the cumulative values, and the red diamond represents the x-intercept.*

the area under the curve ($P(X \leq x)$) up to an arbitrary value x is

$$P(X \leq x) = 1 - e^{-\lambda x}$$

Any CDF shows three properties:

- It is bound between 0 and 1
- It is strictly increasing, that is, its derivative is always positive regardless of where it is evaluated in x
- It can be linearized depending on the shape of the curve

Events in time can be described as a distribution regardless of its generator process. Thus, they will follow a PDF $f(t)$, with an associated CDF $P(X < x)$. Given these two facts we can fit an empirical CDF for which there are several methods available given the properties pf CDFs outlined above. However, the most practical method of curve fitting is the linear regression. The only case where a CDF would be truly linear would be in the case of a uniform distribution where

$$P(X \leq x) = \frac{x}{b-a} \text{ for } x \in [a, b]$$

which implies that the empirical data in time follow a uniform distribution. For any other case, a fitted CDF would be a non-linear, and thus a linearization should be applied to it.

There are are several methods available for linearizing data (e.g., applying the logarithm, the square and cubic roots, the arcsine, or the reciprocal to the raw data), often applied (even in excess) as a procedure for normalizing data that do not follow a normal distribution. In the case of CDF fitting, we need to guarantee that the data are linear regardless of the PDF in order to estimate parameters of the model and thus the procedure is appropriate. After model fitting, a back transformation would be necessary for representing the time value of interest in real space.

The linear model has the form

$$y = mx + b$$

where m is the slope and b is the y-intercept. However, we are not interested in any of these parameters but instead in a third one, the x-intercept, that is defined as the point in x where $y = 0$. This parameter can be defined by letting $y = 0$ in the equation and thus

$$x_{y=0} = -\frac{b}{m}$$

allows us to find the point at which the data began to be generated.

Several methods have been proposed for estimating the uncertainty on the x-intercept estimate, from the intuitive inverse regression X Y to bootstraping methods. According to

Seber and Lee (2013, pp. 146-147) this problem has long been discussed in the statistical literature without much consensus, although some specific alternatives are of interest due to their theoretical appeal, ease of implementation, or simplification of assumptions. Uncertainty on the estimate of the x-intercept can be obtained either through Taylor series simplification that gives

$$\left(\frac{\sigma_X}{X}\right)^2 = \left(\frac{\sigma_a}{a}\right)^2 + \left(\frac{\sigma_b}{b}\right)^2 + 2\left(\frac{\sigma_a}{a}\right)\left(\frac{\sigma_b}{b}\right)\rho_{ab}$$

for symmetric intervals. For asymmetric intervals, de Capitani and Pollastri (2012) proposed a method for calculating the cumulative density function and density function for the ratio of two normal, correlated random variables. Another approach by Draper and Smith (1998, pp.83-86) allows the calculation confidence intervals by projecting the values of the confidence region around the linear model on the x-axis by manipulation of the curves describing the confidence region, what gives

$$CI_{X_0} = \hat{X}_0 + \frac{(X_0 - \bar{X})g \pm (ts/b_1)\{[(\hat{X}_0 - \bar{X})^2 S_{XX}] + (1-g)/n\}^{1/2}}{1-g}$$

where $g = t^2 s^2/(b_1^2/S_{XX})$, b_1 is the slope, S_{XX} is the sum of squares of X, t is the t-statistic with n-2 degrees of freedom and at $1-\alpha/2$ significance level, s^2 is the standard deviation, $\hat{X}_0$ is the estimation of the x-intercept, and $\bar{X}$ is the mean of X. This method of estimation of the confidence interval has been implemented in the Appendix E.5.6.

Another approach is to use bootstrap resampling in order to estimate the variance and thus the confidence interval in asymmetric cases such as the present one, implemented in the Appendix E.5.6. Instead of using the estimates from a single fitted model on the coordinates of the empirical CDF, this method fits multiple models with random subsamples in order to construct a collection of estimated values for a parameter of interest, in this case, the x-intercept and this provide a confidence interval on the parameter.

5.5.3 Resampling

A more practical strategy than the mathematically-loaded approaches above is just to construct a distribution for parameters of interest such as the mean and the median. Let $f(t_i)$ be a PDF for the divergence time point t_i out in the vector of divergence time points $i \in 1:T$:

$$t_i \sim f(t_i)$$

such that each data instance has both a point estimate and an uncertainty region, both generated by the PDF. Two approaches are possible, either we can fit a given distribution, or sample from the empirical confidence region. Although the former is more mathematically sound, there is a universe of continuous distributions available in the statistical literature, with intricate relationships among them as some are special cases of others, and we can

arrive at different distributions through approximations of more complicated ones (Evans et al., 2001; Johnson et al., 1995; Krishnamoorthy, 2016). Thus, until a proper stochastic process can dictates the proper statistical distribution for divergence time data, we can use another approach that is distribution-agnostic and is to assume that the uncertainty on a given data point t_i follows a uniform distribution

$$f(t_i) \sim U(t_i; \alpha, \beta)$$

so that we can sample from such simple distribution without assuming beyond the most basic statement given by the posterior distribution of the bayesian analysis that generated each data point.

The following algorithm was constructed in order to estimate through resampling the mean and median of the distribution of divergence time data

Algorithm 5: Pseudocode for the estimation of the statistical distribution of mean and median of the divergence time data

Data: A vector of divergence times and a data frame with confidence regions, most often HPDs from bayesian analyses
Result: PDFs for both mean and median of the distribution of divergence time data
initialize N number of iterations, and placeholder vectors $meanConf$ and $medianConf$;
for i *in* 1 *to* N **do**
  df = choose a subsample 15 data point rows from the $confidences$ data frame;
  initialize the *estimates* vector matching number of rows of df;
  for j *in* 1 *to* $nrow(df)$ **do**
    $estimates_j \sim U(\alpha = df\$min_j, \beta = df\$max_j)$;
  $meanConf_i = mean(estimates_{j=1:nrow(df)})$;
  $medianConf_i = median(estimates)_{j=1:nrow(df)}$;

5.6 The separation of cis- and trans-Andean drainages

Divergence time estimation data were compiled from the literature that reported separation events between cis- and trans-Andean lineages, be it contemporary species or nodal divergences (Abe et al., 2014; Cheviron et al., 2005; Collins and Dubach, 2000; Cortés-Ortiz et al., 2003; D'Horta et al., 2013; Dick et al., 2003; Elias et al., 2009; Fernandes et al., 2014; Grau et al., 2005; Gutiérrez et al., 2014; Hardman and Lundberg, 2006; Hernández Torres, 2015; Machado et al., 2014; Miller et al., 2008; Patané et al., 2009; Patel et al., 2011; Picq et al., 2014; Ribas et al., 2005, 2007; Ruiz-García et al., 2015; Smith et al., 2014; Voss et al., 2013; Weir and Price, 2011; Říčan et al., 2013). A total of 54 data points were compiled along with their credible intervals, most frequently HPD intervals.

The distribution of divergence time data is very asymmetrical with a heavy right tail, product of several outliers older than 10 Ma (Figure 5.3). These outliers affect strongly those

methods who make a stronger assumption on the lack of uncertainty in the measurement of time values (e.g., methods based on stratigraphic intervals and less strongly, those based on cumulative densities). Given this analytical issue, these values must be excluded, what makes unnecessary the process of linearization through data transformation given that the data conform to a linear model after outlier remotion.

5.6.1 Estimation through CIs for stratigraphic intervals

The analogy of stratigraphic intervals for divergence time data produced results according with the theory of both methods implemented (Figure 5.4). The distribution-free method of Marshall (1994) produced a slightly wider confidence interval, although both were bound as expected by the oldest observed occurrence. This agrees with the implicit assumption that each time occurrences is measured without error, so that the probability that the true older parameter is actually younger than the oldest occurrence is 0. The procedure of outlier remotion is particularly justified here given the assumption mentioned above, given that this method is the most sensible to extreme values due to their effect on the specification of confidence intervals. As expected, the methods based on stratigraphic intervals produce older inferred intervals of about 7–9 Ma provided that the oldest occurrences has been measured with confidence.

5.6.2 Estimation through x-intercept in CDFs

These novel methods have produced results in line with those of the previous section, however, they allow estimated x-intercepts and consequently their confidence intervals that are not necessarily older than the oldest divergence time estimation event. Thus, these methods are robust to the assumption of measurement without error that is implicit in the methods based on stratigraphic intervals. However, they are still sensitive to the presence of strong outliers as preliminary analyses showed that the regressions on these data with outliers affected strongly the robust regressions.

Both methods produce similar results (Figure 5.5). Although the method based on robust regression produced a CI wider than the one based on ordinary least squares, the bootstrap CI estimate was as wide as the two former combined. There is no mathematical relationship between the magnitude of such confidence intervals, but the fact that the bootstraping based on robust estimates is of more general application, it seems more adequate for these situations. It is noteworthy that these inferences are based on the assumption that the events in time are measured without error, however, this assumption relaxes due to the effect of re sampling and uncertainty in the estimates. According to both estimators, the separation between cis- and trans-Andean areas took place between ca. 4.3–5.8 Ma.

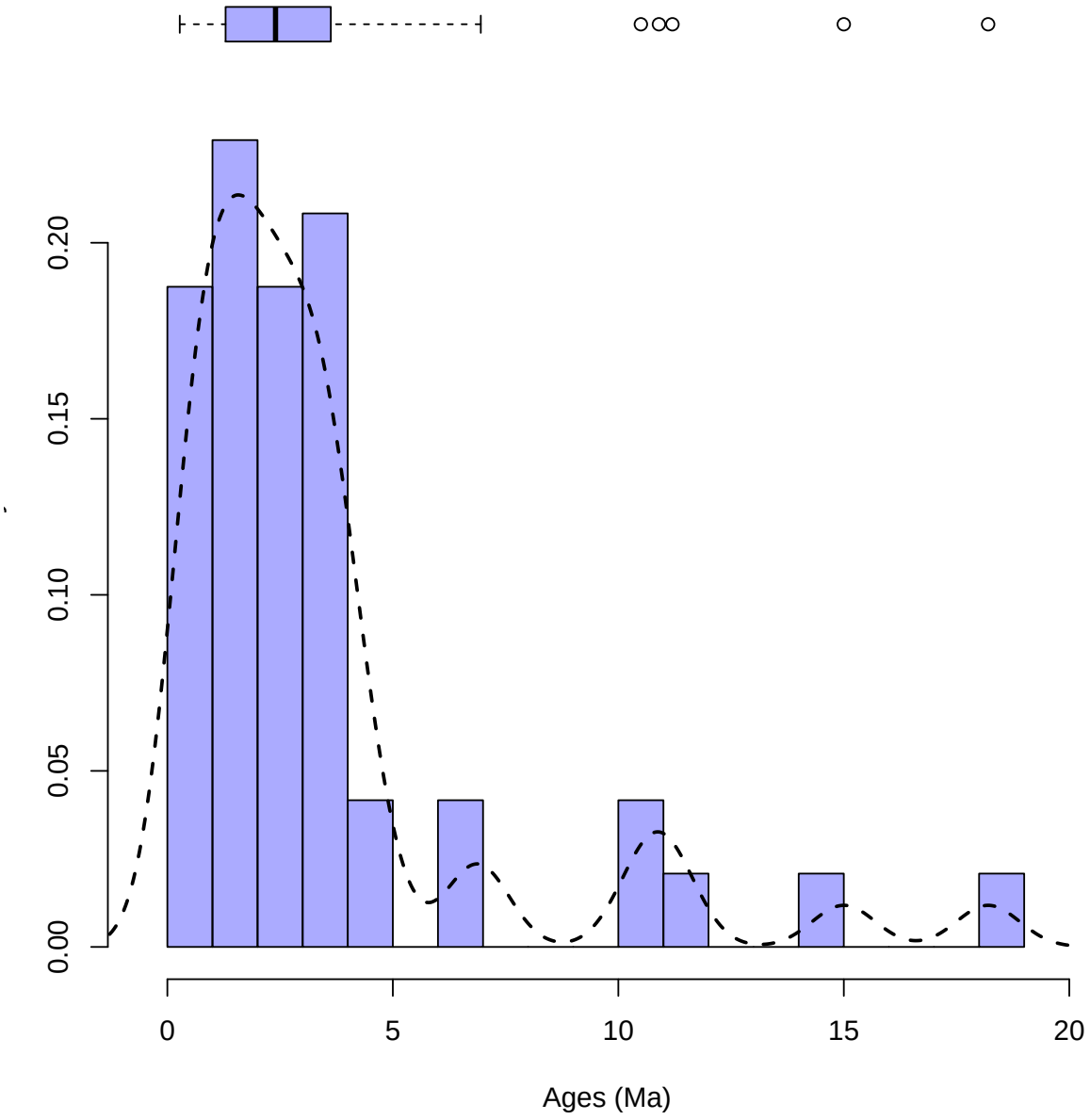


Figure 5.3: *Descriptive statistics of the divergence time data. the histogram illustrates that the multimodality in the density line is an artifact of discontinuity in the age data, while the boxplot on top of the figure shows the existence of multiple outliers to the right of the distribution. All three methods agree in showing the asymmetry in the distribution with a heavy right tail. Discontinuous line represents the empirical probability density.*

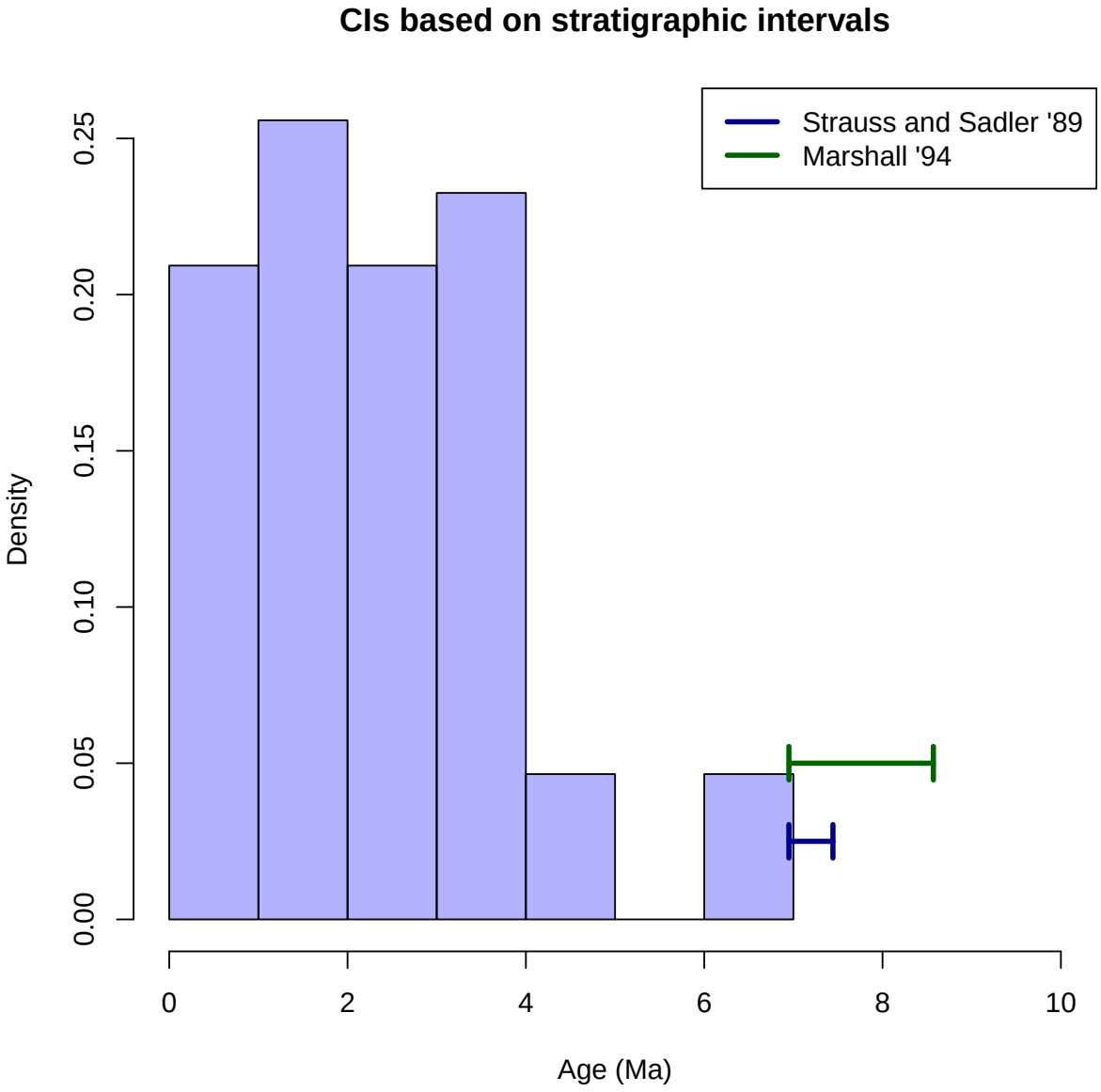


Figure 5.4: *Estimation of the time of separation between cis- and trans-Andean drainages based on the methods of stratigraphic intervals of Strauss and Sadler (1989) and Marshall (1994). The latter method produces a slightly larger confidence interval. Both confidence intervals are bounded by the oldest occurrence and are extensions beyond such points.*

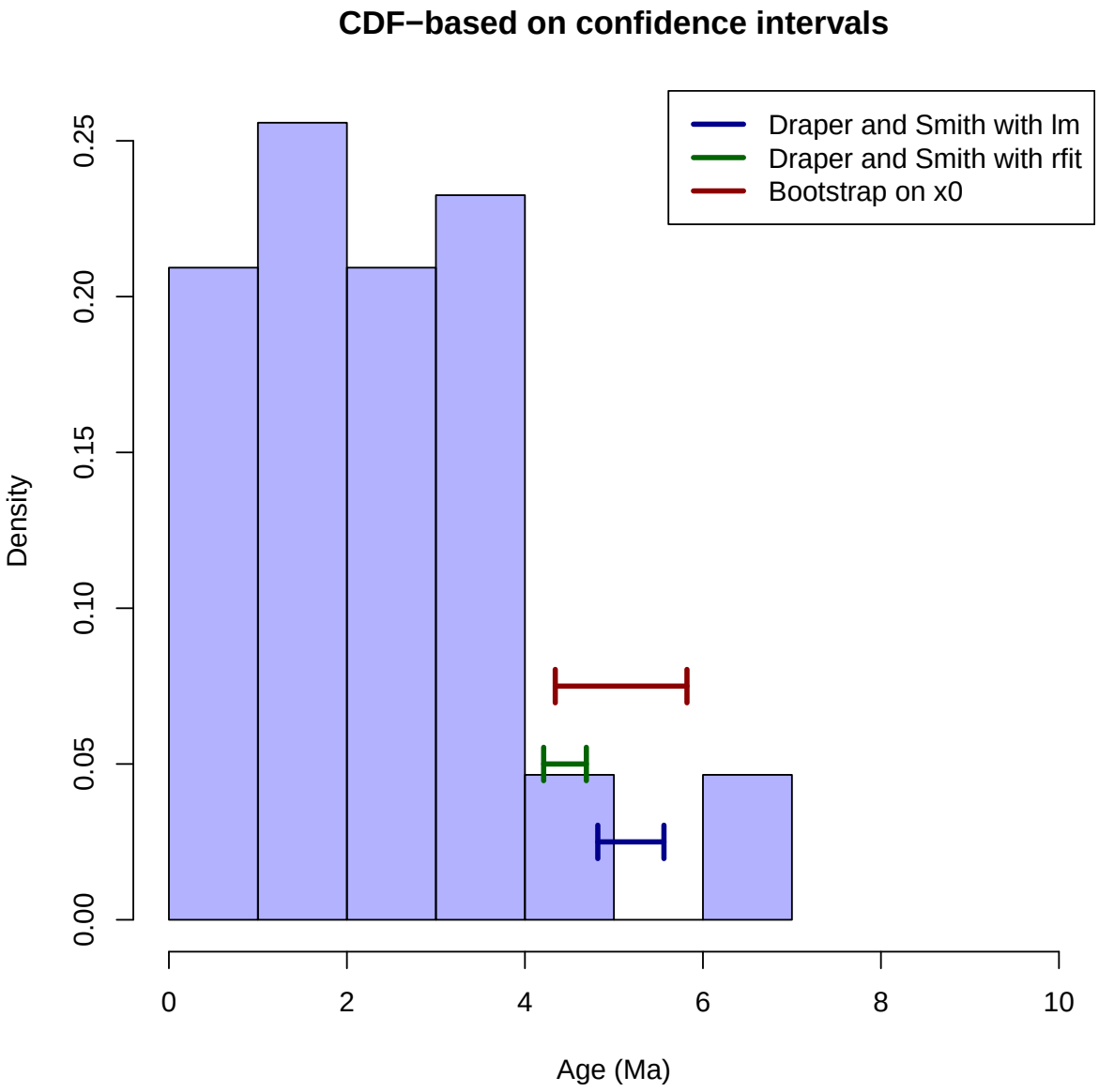


Figure 5.5: *Estimation of the time of separation between cis- and trans-Andean drainages based on the methods of CDF fitting using the estimator of Draper and Smith (1998) using two different approaches at model fitting, one using least squares traditional regression, and the other using robust regression (Kloke and McKean, 2014) and bootstraping techniques. Both confidence intervals are near but no bound by the oldest occurrence as in stratigraphic-interval-based approaches.*

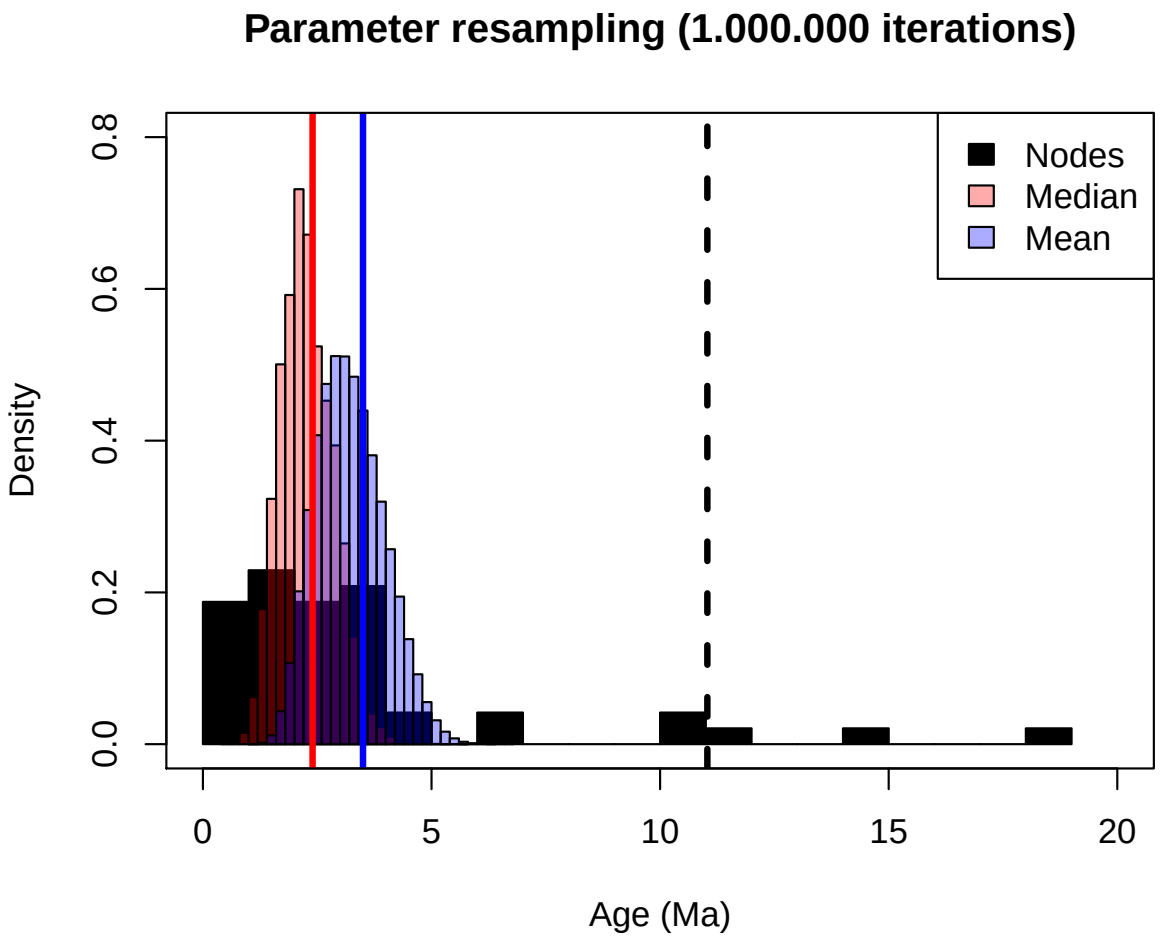


Figure 5.6: *Simulation of mean and median separation time between cis- and trans-Andean drainages based on the credible intervals in 54 divergence time estimation data. Bright red and blue lines represent point estimates calculated on nodal ages, not on simulated values, whereas the dark interrupted line represents the mean classical age of separation between the Magdalena and cis-Andean drainages (Diaz de Gamero, 1996; Flynn et al., 1997; Guerrero, 1997; Hoorn and Wesselingh, 2010).*

5.6.3 Estimation through resampling

Resampling allowed to construct distributions for the mean and the median of the divergence time data (Figure 5.6). As expected, these distributions were not directly affected by the position of the oldest and youngest divergence time data and thus provide robust results than the former methods. However, an implicit assumption is that each divergence time data point is a realization of a stochastic process for which the resampling is providing an estimate of the distribution for a given parameter. The interpretation of these estimates are that, given that each point in time is a realization of a stochastic process in time, their mean or median provide information for the vicariance event within the distribution and not outside of it as in previous approaches. This is an important analytical difference explaining why this method will always provide younger estimates than the alternative methods.

5.7 Discussion

Data quality becomes an issue for some of the methods herein proposed. Although it has long been recognized that calibration quality will affect the posterior estimates of divergence time estimation (Parham et al., 2012; Warnock et al., 2014), the empirical application of these methods is still poorly implemented by users, what generates a cascade of errors down the analysis that are often overlooked in the literature. It is rare to see any discussion on issues such as prior-posterior sensitivity, justification for choosing particular distributions and parametrizations, and the effect of the conjugate time prior for the results of divergence time estimation. The dataset herein considered includes literature from the past 20 years, and thus, spans the interval in which most of the theoretical developments and computational implementations of such methods were carried out. As a consequence, the quality of estimates is expected to vary considerably, thus being more or less sensitive to prior information such as the former model of Miocene northern Andean uplift that was used in the ichthyological literature until recently. These sources of information tend to explain the outliers in estimated separation times for fishes, while other sources use different calibration strategies such as mutation rates or other kinds of calibration (Hazzi et al., 2018).

The first two sets of methods make strong assumptions on the raw data and such inconvenience can be modeled to a more realistic scenario where each data point has associated a distribution of uncertainty. It is possible to use resampling in order to generate estimates of CI that take into account uncertainty in the x values, so that they are less contingent to the position of the oldest occurrence in time, that with high probability, can be an outlier. This approach is already implemented in the resampling method that explicitly models the distribution of mean and median of divergence time values, showing perhaps the more realistic scenario. With aid of further refinements, these three methods be used in different situations less contingent to assumption about the data.

Fossil evidence indicates drainage connections between cis- and trans-Andean areas during the middle Miocene (Ballen and Moreno-Bernal, 2019; Lundberg, 1997, Chapter 3 herein) and has long been used to corroborate geological paleodrainage models (Albert et al., 2006; Diaz de Gamero, 1996; Hoorn et al., 1995). This Miocene scenario corresponds to the null hypothesis depicted in Figure 1.3A, where drainage connection ceased during the middle Miocene, in which case the fossil freshwater fishes will document the latest phase of drainage connection between cis- and trans-Andean drainages. However, a growing body of fossil evidence suggests that such age is wrong as drainage connections are still suggested by fossil freshwater fishes of Pliocene age (Aguilera et al., 2013, Chapter 4 herein), what would reject the null hypothesis in favor of the alternative hypothesis of connection persistence until the Pliocene (2–4 Ma).

The different inferential methods herein proposed agree in variable ways with the alternative hypothesis of connection persistence until the Pliocene. Estimates of vicariance time provided by the methods based on CIs of stratigraphic intervals are those which make the

strongest assumption about the observed data, and consequently whose estimates of t_0 are the oldest. On the other hand, the intervals estimated with the methods based on fitting of CDFs are statistically indistinguishable from those based on resampling as they overlap with the HDI of the resampling distribution of mean and median (Figures 5.5,5.6). Estimates of the former group are shifted to slightly older dates by about 2 Ma. A possible consensus is to think of the separation time as the combination of possible dates by these two methods, what would produce an interval of 2–5.8 Ma. This is consistent with the fossil specimens studied in Chapter 4 as they suggest a persistent connection between ca. 2–4 Ma.

Such younger separation time has interesting consequences for the evolution of freshwater faunas, given that the high levels of endemism in trans-Andean freshwater fishes (Figure 1.2) should have been generated in a timespan of 2–4 Ma, a somewhat quick event that remodeled drastically the composition of freshwater communities in northern South America. It remains to be studied how Neotropical biotas could have responded to these drainage evolution events in fast ways, given that the existent literature always have considered the separation event as occurring 11–13 Ma.

Bibliography

Abe, K.T., Mariguela, T.C., Avelino, G.S., Foresti, F., and Oliveira, C. 2014. Systematic and historical biogeography of the Bryconidae (Ostariophysi: Characiformes) suggesting a new rearrangement of its genera and an old origin of Mesoamerican ichthyofauna. *BMC evolutionary biology*, 14:152. 142

Aguilera, O.A., Lundberg, J.G., Birindelli, J.L.O., Sabaj Pérez, M.H., Jaramillo, C.A., and Sánchez-Villagra, M.R. 2013. Palaeontological evidence for the last temporal occurrence of the ancient Western Amazonian river outflow into the Caribbean. *PLoS ONE*, 8:1–17. 148

Albert, J.S., Lovejoy, N.R., and Crampton, W.G.R. 2006. Miocene tectonism and the separation of cis- and trans-Andean river basins: Evidence from Neotropical fishes. *Journal of South American Earth Sciences*, 21:14–27. 130, 148

Antonelli, A., Nylander, J.A., Persson, C., and Sanmartín, I. 2009. Tracing the impact of the Andean uplift on Neotropical plant evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 106:9749–9754. 129

Ballen, G.A. and Moreno-Bernal, J.W. 2019. New records of the enigmatic Neotropical fossil fish *Acregoliath rancii* (Teleostei *incertae sedis*) from the middle Miocene Honda group of Colombia. *Ameghiniana*, 56:431–440. 148

Barba-Montoya, J., dos Reis, M., and Yang, Z. 2017. Comparison of different strategies for using fossil calibrations to generate the time prior in Bayesian molecular clock dating. *Molecular Phylogenetics and Evolution*, 114:386–400. 131

Bayona, G., Cardona, A., Jaramillo, C.A., Mora, A., Montes, C., Caballero, V., Mahecha, H., Lamus, F., Montenegro, O., Jimenez, G., Mesa, A., and Valencia, V. 2013. Onset of fault reactivation in the Eastern Cordillera of Colombia and proximal Llanos Basin; response to Caribbean - South American convergence in early Palaeogene time. *Geological Society, London, Special Publications*. 129

Britton, T., Lisa, C., Jacquet, D., Lundqvist, S., and Bremer, K. 2007. Estimating Divergence Times in Large Phylogenetic Trees Estimating Divergence Times in Large Phylogenetic Trees. *Systematic Biology*, 56:741–752. 131

Cheviron, Z.A., Hackett, S.J., and Capparella, A.P. 2005. Complex evolutionary history of a Neotropical lowland forest bird (*Lepidothrix coronata*) and its implications for historical hypotheses of the origin of Neotropical avian diversity. *Molecular Phylogenetics and Evolution*, 36:338–357. 142

Collins, A.C. and Dubach, J.M. 2000. Biogeographic and ecological forces responsible for speciation in Ateles. *International Journal of Primatology*, 21:421–444. 142

Cortés-Ortiz, L., Bermingham, E., Rico, C., Rodríguez-Luna, E., Sampaio, I., and Ruiz-García, M. 2003. Molecular systematics and biogeography of the Neotropical monkey genus, *Alouatta*. *Molecular Phylogenetics and Evolution*, 26:64–81. 142

de Capitani, L. and Pollastri, A. 2012. The R-code for computing the cdf and the df of the ratio of two correlated normal rv. Technical report, Università degli Studi di Milano-Bicocca. 141

D'Horta, F.M., Cuervo, A.M., Ribas, C.C., Brumfield, R.T., and Miyaki, C.Y. 2013. Phylogeny and comparative phylogeography of *Sclerurus* (Aves: Furnariidae) reveal constant and cryptic diversification in an old radiation of rain forest understorey specialists. *Journal of Biogeography*, 40:37–49. 142

Diaz de Gamero, M.L. 1996. The changing course of the Orinoco River during the Neogene: A review. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 123:385–402. xix, 129, 147, 148

Dick, C.W., Abdul-Salim, K., and Bermingham, E. 2003. Diversification of a Widespread Tropical Rain Forest Tree. *The American Naturalist*, 162:691–703. 142

Draper, N.R. and Smith, H. 1998. *Applied Regression Analysis*. Wiley Interscience, New York. xix, 141, 146

Drummond, A.J. and Bouckaert, R. 2015. *Bayesian Evolutionary Analysis with Beast2*. Cambridge University Press. 130, 131

Drummond, A.J. and Stadler, T. 2016. Bayesian phylogenetic estimation of fossil ages. *Philosophical Transactions of the Royal Society B*, 371:1–14. 131

Elias, M., Joron, M., Willmott, K., Silva-Brandão, K.L., Kaiser, V., Arias, C.F., Piñerez, L.M.G., Uribe, S., Brower, A.V.Z., Freitas, A.V.L., and Jiggins, C.D. 2009. Out of the Andes: patterns of diversification in clearwing butterflies. *Molecular Ecology*, 18:1716–1729. 142

Evans, M., Hastings, N., and Peacock, B. 2001. *Statistical Distributions, Third Edition*, volume 12. John Wiley & Sons. 142

Fernandes, A.M., Wink, M., Sardelli, C.H., and Aleixo, A. 2014. Multiple speciation across the Andes and throughout Amazonia: The case of the spot-backed antbird species complex (*Hylophylax naevius/Hylophylax naevioides*). *Journal of Biogeography*, 41:1094–1104. 142

Flynn, J.J., Guerrero, J., and Swisher III, C.C. 1997. Geochronology of the Honda Group. In R.F. Kay, R.H. Madden, R.L. Cifelli, and J.J. Flynn (eds.), *Vertebrate Paleontology in the Neotropics: the Miocene Fauna of La Venta, Colombia*, pp. 44–60. Smithsonian Institution Press. xix, 147

Grau, E.T., Pereira, S.L., Silveira, L.F., Höfling, E., and Wajntal, A. 2005. Molecular phylogenetics and biogeography of Neotropical piping guans (Aves: Galliformes): Pipile Bonaparte, 1856 is synonym of Aburria Reichenbach, 1853. *Molecular Phylogenetics and Evolution*, 35:637–645. 142

Gregory-Wodzicki, K.M. 2000. Uplift history of the Central and Northern Andes: A review. *Geological Society of America Bulletin*, 112:1091–1105. 129

Guerrero, J. 1997. Stratigraphy, sedimentary environments, and the Miocene uplift of the Colombian Andes. In R.F. Kay, R.H. Madden, R.L. Clifelli, and J.J. Flynn (eds.), *Vertebrate Paleontology in the Neotropics: The Miocene Fauna of La Venta, Colombia*, pp. 15–43. Smithsonian Institution Press. xix, 129, 147

Gutiérrez, E.E., Anderson, R.P., Voss, R.S., Ochoa-G., J., Aguilera, M., and Jansa, S.A. 2014. Phylogeography of *Marmosa robinsoni*: insights into the biogeography of dry forests in northern South America. *Journal of Mammalogy*, 95:1175–1188. 142

Hardman, M. and Lundberg, J.G. 2006. Molecular phylogeny and a chronology of diversification for "phractocephaline" catfishes (Siluriformes: Pimelodidae) based on mitochondrial DNA and nuclear recombination activating gene 2 sequences. *Molecular Phylogenetics and Evolution*, 40:410–418. 142

Hazzi, N.A., Moreno, J.S., Ortiz-Movliav, C., and Palacio, R.D. 2018. Biogeographic regions and events of isolation and diversification of the endemic biota of the tropical Andes. *Proceedings of the National Academy of Sciences of the United States of America*, 115:7985–7990. 130, 148

Heath, T.A. 2012. A hierarchical Bayesian model for calibrating estimates of species divergence times. *Systematic Biology*, 61:793–809. 131

Heath, T.A., Huelsenbeck, J.P., and Stadler, T. 2014. The fossilized Birth-Death Process: A coherent model of fossil calibration for divergence time estimation. *Proceedings of the National Academy of Sciences*, 111:E2957–E2966. 130, 131

Heled, J. and Drummond, A.J. 2015. Calibrated birth-death phylogenetic time-tree priors for Bayesian inference. *Systematic Biology*, 64:369–383. 130

Hernández Torres, C.L. 2015. *Species Delimitation, Phylogenetics, and Biogeography of the Catfish Genus Rhamdia Bleeker (Heptapteridae) of Central America and the Trans-Andean Region of Colombia*. Phd thesis, The University of Southern Mississippi. 142

Hoorn, C., Guerrero, J., Sarmiento, G.A., and Lorente, M.A. 1995. Andean tectonics as a cause for changing drainage patterns in Miocene northern South America. *Geology*, 23:237–240. 129, 148

Hoorn, C. and Wesselingh, F. 2010. *Amazonia: Landscape and Species Evolution*. Wiley-Blackwell. xix, 147

Johnson, N.J., Kotz, S., and Balakrishnan, N. 1995. *Continuous Univariate Distributions, Volume 1*, volume 1. John Wiley & Sons. 142

Kloke, J.D. and McKean, J.W. 2014. *Nonparametric Statistical Methods Using R*. 2nd edition. xix, 146

Krishnamoorthy, K. 2016. *Handbook of Statistical Distributions with Applications*. 142

Lagomarsino, L.P., Condamine, F.L., Antonelli, A., Mulch, A., and Davis, C.C. 2016. The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *New Phytologist*, 210:1430–1442. 131

Landis, M.J. 2017. Biogeographic dating of speciation times using paleogeographically informed processes. *Systematic Biology*, 66:128–144. 131

Lundberg, J.G. 1997. Freshwater fishes and their paleobiotic implications. In R.F. Kay, R.H. Madden, R.L. Cifelli, and J.J. Flynn (eds.), *Vertebrate Paleontology in the Neotropics: The Miocene Fauna of La Venta, Colombia*, chapter 5, pp. 67–91. Smithsonian Institution Press. 148

Lundberg, J.G., Marshall, L.G., Guerrero, J., Horton, B., Malabarba, M.C.S.L., and Wesselingh, F. 1998. The stage for Neotropical Fish diversification: A history of Tropical South American rivers. In L.R. Malabarba, R.E. Reis, R.P. Vari, Z.M. Lucena, and C.A.S. Lucena (eds.), *Phylogeny and Classification of Neotropical Fishes*, pp. 13–48. Edipucrs, Porto Alegre. 130

Machado, L.F., Leite, Y.L.R., Christoff, A.U., and Giugliano, L.G. 2014. Phylogeny and biogeography of tetralophodont rodents of the tribe Oryzomyini (Cricetidae: Sigmodontinae). *Zoologica Scripta*, 43:119–130. 142

Marshall, C.R. 1994. Confidence intervals on stratigraphic ranges: Partial relaxation of the assumption of randomly distributed fossil horizons. *Paleobiology*, 20:459–469. xix, 134, 135, 136, 138, 143, 145

Marshall, C.R. 2010. Using confidence intervals to quantify the uncertainty in the end-points of stratigraphic ranges. *Quantitative Methods in Paleobiology*, 16:291–316. 132, 134

Miller, M.J., Bermingham, E., Klicka, J., Escalante, P., do Amaral, F.S.R., Weir, J.T., and Winker, K. 2008. Out of Amazonia again and again: episodic crossing of the Andes promotes diversification in a lowland forest flycatcher. *Proceedings of the Royal Society B*, 275:1133–42. 142

Parham, J.F., Donoghue, P.C.J., Bell, C.J., Calway, T.D., Head, J.J., Holroyd, P.A., Inoue, J.G., Irmis, R.B., Joyce, W.G., Ksepka, D.T., Patané, J.S.L., Smith, N.D., Tarver, J.E., van Tuinen, M., Yang, Z., Angielczyk, K.D., Greenwood, J.M., Hipsley, C.A., Jacobs, L., Makovicky, P.J., Müller, J., Smith, K.T., Theodor, J.M., Warnock, R.C.M., and Benton, M.J. 2012. Best practices for justifying fossil calibrations. *Systematic biology*, 61:346–59. 148

Patané, J.S.L., Weckstein, J.D., Aleixo, A., and Bates, J.M. 2009. Evolutionary history of *Ramphastos* Toucans: Molecular phylogenetics, temporal diversification, and biogeography. *Molecular Phylogenetics and Evolution*, 53:923–934. 142

Patel, S., Weckstein, J.D., Patané, J.S.L., Bates, J.M., and Aleixo, A. 2011. Temporal and spatial diversification of *Pteroglossus* Araçaris (Aves: Ramphastidae) in the neotropics: Constant rate of diversification does not support an increase in radiation during the Pleistocene. *Molecular Phylogenetics and Evolution*, 58:105–115. 142

Picq, S., Alda, F., Krahe, R., and Bermingham, E. 2014. Miocene and Pliocene colonization of the Central American Isthmus by the weakly electric fish *Brachyhypopomus occidentalis* (Hypopomidae, Gymnotiformes). *Journal of Biogeography*, 41:1520–1532. 142

Rabosky, D.L., Chang, J., Title, P.O., Cowman, P.F., Sallan, L., Friedman, M., Kaschner, K., Garilao, C., Near, T.J., Coll, M., and Alfaro, M.E. 2018. An inverse latitudinal gradient in speciation rate for marine fishes. *Nature*, 559:392–395. 131

Rannala, B. 2016. Conceptual issues in Bayesian divergence time estimation. *Philosophical Transactions of the Royal Society of London B*, 371:20150134. 130

Ribas, C.C., Gaban-Lima, R., Miyaki, C.Y., and Cracraft, J. 2005. Historical biogeography and diversification within the Neotropical parrot genus *Pionopsitta* (Aves: Psittacidae). *Journal of Biogeography*, 32:1409–1427. 142

Ribas, C.C., Moyle, R.G., Miyaki, C.Y., and Cracraft, J. 2007. The assembly of montane biotas: linking Andean tectonics and climatic oscillations to independent regimes of diversification in *Pionus* parrots. *Proceedings. Biological sciences / The Royal Society*, 274:2399–2408. 142

Richardson, J.E., Madriñán, S., Gómez-Gutiérrez, M.C., Valderrama, E., Luna, J., Banda-R., K., Serrano, J., Torres, M.F., Jara, O.A., Aldana, A.M., Cortés-B., R., Sánchez, D., and Montes, C. 2018. Using dated molecular phylogenies to help reconstruct geological, climatic, and biological history: Examples from Colombia. *Geological Journal*, pp. 1–9. 129

Ruiz-García, M., Vásquez, C., Sandoval, S., Kaston, F., Luengas-Villamil, K., and Shostell, J.M. 2015. Phylogeography and spatial structure of the lowland tapir (*Tapirus terrestris* , Perissodactyla: Tapiridae) in South America. *Mitochondrial DNA*, 1736:1–9. 142

Saeid, E., Bakioglu, K.B., Kellogg, J., Leier, A., Martinez, J.A., and Guerrero, E. 2017. Garzón Massif basement tectonics: Structural control on evolution of petroleum systems in upper Magdalena and Putumayo basins, Colombia. *Marine and Petroleum Geology*, 88:381–401. 129

Schaefer, S.A. 2011. The Andes. Riding the Tectonic Uplift. In J.S. Albert and R.E. Reis (eds.), *Historical Biogeography of Neotropical Freshwater Fishes*, pp. 259–278. University of California Press. 130

Seber, G.A.F. and Lee, A.J. 2013. *Linear Regression Analysis*. 2nd editio edition. Wiley Interscience, Hoboken, NJ. 141

Silvestro, D., Schnitzler, J., Liow, L.H., Antonelli, A., and Salamin, N. 2014. Bayesian estimation of speciation and extinction from incomplete fossil occurrence data. *Systematic Biology*, 63:349–367. 134

Smith, B.T., Mccormack, J.E., Cuervo, A.M., Hickerson, M.J., Aleixo, A., Cadena, C.D., Pérez-Emán, J., Burney, C.W., Xie, X., Harvey, M.G., Faircloth, B.C., Glenn, T.C., Derryberry, E.P., Prejean, J., Fields, S., and Brumfield, R.T. 2014. The drivers of tropical speciation. *Nature*, pp. 1–8. 130, 142

Strauss, D. and Sadler, P.M. 1989. Classical confidence intervals and Bayesian probability estimates for ends of local taxon ranges. *Mathematical Geology*, 21:411–427. xix, 134, 145

Tagliacollo, V.A., Roxo, F.F., Duke-Sylvester, S.M., Oliveira, C., and Albert, J.S. 2015. Biogeographical signature of river capture events in Amazonian lowlands. *Journal of Biogeography*, 42:2349–2362. 131

Voss, R.S., Hubbard, C., and Jansa, S.A. 2013. Phylogenetic Relationships of New World Porcupines (Rodentia, Erethizontidae): Implications for Taxonomy, Morphological, and Biogeography. *American Museum Novitates*, pp. 1–36. 142

Warnock, R.C.M., Parham, J.F., Joyce, W.G., Lyson, T.R., and Donoghue, P.C.J. 2014. Calibration uncertainty in molecular dating analyses: there is no substitute for the prior evalu-

ation of time priors. *Proceedings of the Royal Society B: Biological Sciences*, 282:20141013–20141013. 131, 148

Weir, J.T. and Price, M. 2011. Andean uplift promotes lowland speciation through vicariance and dispersal in *Dendrocincla* woodcreepers. *Molecular Ecology*, 20:4550–4563. 142

Říčan, O., Piálek, L., Zardoya, R., Doadrio, I., and Zrzavý, J. 2013. Biogeography of the Mesoamerican Cichlidae (Teleostei: Heroini): colonization through the GAARlandia land bridge and early diversification. *Journal of Biogeography*, 40:579–593. 130, 142

Appendix A

Terminology

A.1 References assessed for terminology

Table A.1: *Terms used for description of anterior and posterior ornaments in dorsal and pectoral spines.*

Name	Spine	Reference
Asserraciones	Dorsal	Royero (1999)
Corrugaciones	Dorsal	Royero (1999)
Crenulated	Dorsal	Eigenmann and Eigenmann (1890)
Crête anterieur médian	Dorsal	Gayet and van Neer (1990)
Dentations	Dorsal	Lucena et al. (1992); Lundberg (1975); Parisi et al. (2006)
Dentelures	Dorsal	Cuvier and Valenciennes (1840)
Denticulations	Dorsal	Vigliotta (2008)
Dentículos	Dorsal	Royero (1999)
Erect	Dorsal	Parisi et al. (2006)
Granulate	Dorsal	Thomson and Page (2006)
Granulée	Dorsal	Cuvier and Valenciennes (1840)
Granules	Dorsal	Eigenmann and Eigenmann (1890)
Granuloso	Dorsal	Royero (1999)
Grooves	Dorsal	Eigenmann (1925)
Keels	Dorsal	Eigenmann (1925)
Longitudinal keel	Dorsal	Greenwood (1959)
Longitudinal ridge	Dorsal	Divay and Murray (2015)
Median crest	Dorsal	Otero et al. (2007); Pinton and Otero (2010)
Pits between striations	Dorsal	Divay and Murray (2015)
Recurved notches	Dorsal	Eigenmann and Eigenmann (1890)

Continuing on the next page

Table A.1 – *Continuing from the previous page*

Name	**Spine**	**Reference**
Recurved teeth	Dorsal	Eigenmann and Eigenmann (1890)
Ridges	Dorsal	Eigenmann (1925); Lundberg (1975)
Serrae	Dorsal	Parisi et al. (2006); Thomson and Page (2006); Vigliotta (2008); Wright (2009)
Serrate	Dorsal	Cuvier and Valenciennes (1840); Eigenmann (1925); Eigenmann and Eigenmann (1890)
Serrations	Dorsal	Ferraris and Mago-Leccia (1989); Greenwood (1959); Reed (1924); Sabaj Pérez and Birindelli (2008); Trapani (2008); Watanabe and Uyeno (1999)
Sierras romas y poco desarrolladas	Dorsal	Royero (1999)
Sinuous striations	Dorsal	Divay and Murray (2015)
Straight teeth	Dorsal	Eigenmann and Eigenmann (1890)
Striae	Dorsal	Cione et al. (2005)
Striate	Dorsal	Eigenmann (1925); Eigenmann and Eigenmann (1890); Lundberg (1975)
Teeth	Dorsal	Mees (1974)
Thorns	Dorsal	Eigenmann (1925)
Tubercles	Dorsal	Lundberg (1975); Pinton and Otero (2010)
Tubercules	Dorsal	Gayet and van Neer (1990)
Barbs	Pectoral	Paruch (1986)
Bifid dentations	Pectoral	Bennett (1979)
Bifid serrations	Pectoral	Ohara et al. (2016); Tencatt et al. (2016); Wosiacki et al. (2014)
Bristly	Pectoral	Eigenmann and Eigenmann (1890)
Bumps	Pectoral	Smith (1987)
Conical dentations	Pectoral	Lundberg (1975)
Crenelée	Pectoral	Cuvier and Valenciennes (1840)
Dens	Pectoral	Cuvier and Valenciennes (1840)
Dentate	Pectoral	Eigenmann and Eigenmann (1890); Hubbs and Hibbard (1951); Jordan (1880)

Continuing on the next page

Table A.1 – *Continuing from the previous page*

Name	Spine	Reference
Dentations	Pectoral	Bennett (1979); Bornbusch (1991); Diogo (2007); Hubbs and Hibbard (1951); Lucena et al. (1992); Lundberg (1975); Lundberg et al. (2004); Lundberg and McDade (1986); Paloumpis (1963); Parisi et al. (2006); Smith (1987); Vanscoy et al. (2015)
Dentato	Pectoral	Linnaeus (1758)
Dentelures	Pectoral	Cuvier and Valenciennes (1840)
Denticles	Pectoral	Pinton et al. (2006)
Denticulations	Pectoral	Vigliotta (2008)
Denticules	Pectoral	Gayet and van Neer (1990)
Dientes	Pectoral	Bisbal and Gomez (1986)
Épines	Pectoral	Gayet and van Neer (1990)
Erect	Pectoral	Hubbs and Hibbard (1951); Lundberg (1975)
Extrorse	Pectoral	Eigenmann and Eigenmann (1890)
Granular	Pectoral	Eigenmann and Eigenmann (1888, 1890)
Granulate	Pectoral	Thomson and Page (2006)
Granulée	Pectoral	Cuvier and Valenciennes (1840)
Hispid	Pectoral	Eigenmann and Eigenmann (1890)
Hooked dentations	Pectoral	Vanscoy et al. (2015)
Hooks	Pectoral	Eigenmann (1925); Eigenmann and Eigenmann (1888, 1890); Mees (1974); Parisi et al. (2006); Shibatta and Benine (2005)
Lamellae	Pectoral	Eigenmann and Eigenmann (1890)
Laminar serrations	Pectoral	Tencatt and Ohara (2016)
Multicuspid dentations	Pectoral	Vanscoy et al. (2015)
Multifid dentations	Pectoral	Lundberg (1975); Smith (1987)
Notches	Pectoral	Eigenmann and Eigenmann (1890); Paloumpis (1963); Paruch (1986)
Pectinate	Pectoral	Eigenmann and Eigenmann (1890)
Protuberancia	Pectoral	Bisbal and Gomez (1986)
Retrodentatus	Pectoral	Linnaeus (1758)
Roughened	Pectoral	Eigenmann and Eigenmann (1888)

Continuing on the next page

Table A.1 – *Continuing from the previous page*

Name	**Spine**	**Reference**
Roughish	Pectoral	Jordan (1880)
Roughness	Pectoral	Page et al. (2007)
Rugose	Pectoral	Hubbs and Hibbard (1951)
Serrae	Pectoral	Bornbusch (1991); Bosher et al. (2006); Diogo et al. (2003); Eigenmann and Eigenmann (1890); Evermann and Kendall (1898); Geerinckx et al. (2004); Jordan (1880); Lundberg (1975); Lundberg and McDade (1986); Mees (1974); Page et al. (2007); Parisi et al. (2006); Smith (1987); Vanscoy et al. (2015); Vigliotta (2008); Wright (2009)
Serrata	Pectoral	Linnaeus (1758)
Serrated	Pectoral	Chen and Lundberg (1995); Eigenmann (1925); Eigenmann and Eigenmann (1888, 1890); Ferraris and Mago-Leccia (1989); Geerinckx et al. (2004); Jordan (1880); Mees (1974); Parmentier et al. (2010); Pinton and Otero (2010); Teugels and Adriaens (2003); Ünlü et al. (2012); van Neer (1992); Wosiacki et al. (2014)
Serrations	Pectoral	Cione et al. (2005); Divay and Murray (2015); Egge and Simons (2011); Eigenmann (1925); Kaatz et al. (2010); Ng (2003); Ohara et al. (2016); Page et al. (2007); Paloumpis (1963); Paruch (1986); Pinton et al. (2011); Reed (1924); Royero (1999); Sabaj Pérez and Birindelli (2008); Shibatta and Benine (2005); Smith (1987); Tencatt et al. (2016); Thomson and Page (2006); Trapani (2008); Watanabe and Uyeno (1999); Wosiacki et al. (2014)
Serratum	Pectoral	Arctaedius (1738)

Continuing on the next page

Table A.1 – *Continuing from the previous page*

Name	Spine	Reference
Setiferous	Pectoral	Eigenmann and Eigenmann (1890)
Spines	Pectoral	Boulenger (1900); Eigenmann and Eigenmann (1890); Mees (1974)
Spinosus	Pectoral	Linnaeus (1758)
Spinous	Pectoral	Eigenmann and Eigenmann (1888, 1890)
Spiny	Pectoral	Eigenmann and Eigenmann (1890)
Striate	Pectoral	Eigenmann and Eigenmann (1890); Hubbs and Hibbard (1951); Parisi et al. (2006)
Striations	Pectoral	Divay and Murray (2015)
Striée	Pectoral	Cuvier and Valenciennes (1840)
Stries	Pectoral	Gayet and van Neer (1990)
Teeth	Pectoral	Eigenmann and Eigenmann (1888, 1890); Mees (1974); Paloumpis (1963); Shibatta and Benine (2005)
Thorns	Pectoral	Myers (1927)
Tubercles	Pectoral	Otero et al. (2009); Pinton et al. (2006); Pinton and Otero (2010)
Tubercules	Pectoral	Divay and Murray (2015); Eigenmann and Eigenmann (1890); Gayet and van Neer (1990)
Vermiculée	Pectoral	Cuvier and Valenciennes (1840)

A.2 Cost calculations and optimal terminology

Table A.2: *Results of the cost calculations highlighting the optimal terms. Whenever two or more terms showed ties, choices were based on other arguments.*

Structure	Terms	Frequency	Total	indCost
Pdp	2	1	23	22
Pdp	arched crest	1	23	22
Pdp	Articular plateau	3	23	20
Pdp	Cleithral process	1	23	22
Pdp	Curved flange	1	23	22

Continuing on the next page

Table A.2 – *Continuing from the previous page*

Structure	Terms	Frequency	Total	indCost
Pdp	delta	1	23	22
Pdp	Dorsal articulating process	2	23	21
Pdp	Dorsal articulating surface	1	23	22
Pdp	Dorsal condyle	2	23	21
Pdp	**Dorsal process**	**5**	**23**	**18**
Pdp	Dorsal process of spine	2	23	21
Pdp	Expanded flange	1	23	22
Pdp	Helicoidal semidisk	1	23	22
Pdp	Surface cleithrale	1	23	22
Dlas	Articular facet for posterior nuchal plate	1	4	3
Dlas	**Lateral articular surface**	**2**	**4**	**2**
Dlas	Surface articulaire de l'aile lateral	1	4	3
Pmp	Anterior process of dorsal condyle	1	12	11
Pmp	Axial process	2	12	10
Pmp	**Median process of dorsal condyle**	**6**	**12**	**6**
Pmp	Processus axial	1	12	11
Pmp	Proximal tubercle	1	12	11
Pmp	Rotator process	1	12	11
Pafvp	**Articular facet of ventral process**	**NA**	**NA**	**Herein proposed**
Base	**Base**	**4**	**9**	**5**
Base	Cabeza	1	9	8
Base	Head	3	9	6
Base	Tête articulaire	1	9	8
Ppp	Dorsomedian process	1	3	2
Ppp	**Posterior process**[1]	**1**	**3**	**2**
Ppp	Proximal process	1	3	2
Pafpp	**Articular facet of posterior process**	**NA**	**NA**	**Herein proposed**

Continuing on the next page

[1]This term was chosen in coordination to 'Anterior process' being picked for term K where it was one of the two optimal terms. See footnote 3.

Table A.2 – *Continuing from the previous page*

Structure	Terms	Frequency	Total	indCost
Pafs	**Articular facet for scapulocoracoid**	**2**	**6**	**4**
Pafs	Articular groove	1	6	5
Pafs	Articular notch (fossa)	1	6	5
Pafs	Central articulating surface	1	6	5
Pafs	Cotyle of spine	1	6	5
Pafdp	**Articular facet of dorsal process**	**NA**	**NA**	**Herein proposed**
E	Central articulating surface	1	2	1
Ppk	**Posterior keel**	**NA**	**NA**	**Herein proposed**
E	Posterior fossa	1	2	1
Dlf	**Lateral fossae**	**NA**	**NA**	**Herein proposed**
F	Furrow	1	5	4
F	**Posterior groove**	**3**	**5**	**2**
F	Sulcus	1	5	4
FF	**Left hemitrichium**	**NA**	**NA**	**Herein proposed**
G	**Basal recess**	**3**	**13**	**10**
G	Foramen	1	13	12
G	Fosse interne	1	13	12
G	Hendidura basal	1	13	12
G	Inner fossa	2	13	11
G	Inner hole	1	13	12
G	Nutritive canal	1	13	12
G	Posterior basal recess	1	13	12
G	Proximal foramen	2	13	11
GG	**Right hemitrichium**	**NA**	**NA**	**Herein proposed**
H	**Basal process**	**1**	**1**	**0**

Continuing on the next page

Table A.2 – *Continuing from the previous page*

Structure	Terms	Frequency	Total	indCost
HH	**Attachment surface for the *ligamentum interspinalis*[2]**	**NA**	**NA**	**Herein proposed**
I	Distal lobe	1	2	1
I	**Medial flange for insertion of superficialis abductor muscle[3]**	**1**	**2**	**1**
II	**Posterior fossae**	**NA**	**NA**	**Herein proposed**
J	Anterior basal recess	1	4	3
J	**Anterior fossa[4]**	**1**	**4**	**3**
J	Insertion for arrector dorsalis dorsal division muscle	1	4	3
J	Insertion surface for ventral arrector part z muscle	1	4	3
K	Anterior articular process	1	13	12
K	Anterior condyle	1	13	12
K	**Anterior process[5]**	**3**	**13**	**10**
K	Anteroventral process	1	13	12
K	Dorsolateral process	3	13	10
K	Dorsomedial process	1	13	12
K	Processus dorso-latéral	1	13	12
K	Proximoventral process	1	13	12
K	Ventral process	1	13	12
L	Distoventral process	1	12	11
L	Posteroventral process	1	12	11
L	Processus ventro-latéral	1	12	11
L	Ventral articular process	1	12	11
L	Ventral condyle	1	12	11
L	**Ventral process**	**4**	**12**	**8**

Continuing on the next page

[2]New name for the bony trait, new name for the ligament too

[3]This term needs modification = flange of ventral process (as herein proposed in order to avoid redundant definitions)

[4]This term is associated to 'Posterior fossa' in E, and therefore takes precedence over the others

[5]This term was chosen in coordination to 'Posterior process' being picked for term C where it was one of the two optimal terms. See footnote 1.

Table A.2 – *Continuing from the previous page*

Structure	Terms	Frequency	Total	indCost
L	Ventrolateral process	2	12	10
L	Ventromedial process	1	12	11
M	**Anteroventral emargination**[6]	**2**	**3**	**1**
M	Lateral cotyle	1	3	2
N	Depression for the arrector ventralis muscle	1	2	1
N	**Ventral fossa**[7]	**1**	**2**	**1**
O	Mur inféro-interne	1	2	1
O	**Rama ventral**[8]	**1**	**2**	**1**
P	Mur supéro-interne	1	2	1
P	**Rama dorsal**[9]	**1**	**2**	**1**
Q	**Proximal crest**[10]	**1**	**1**	**0**
R	**Insertion for arrector ventralis**[11]	**1**	**1**	**0**
S	1	1	3	2
S	beta	1	3	2
S	**Proximal articulating surface**[12]	**1**	**3**	**2**
Shaft	Cuerpo	1	8	7
Shaft	**Shaft**	**7**	**8**	**1**
T	Aile latéral	1	6	5
T	Basal condyle	1	6	5
T	Condilo lateral	1	6	5
T	**Lateral condyle**	**1**[13]	**6**	**5**
T	Lateral wing	2	6	4

Continuing on the next page

[6]This shall be called Proximal emargination as herein proposed because anteroventral is a misleading positional term despite this term has the lowest cost.

[7]Despite having the same cost as the first term, this one was chosen for coordination with the anterior and posterior fossae

[8]There is a better term for this structure and is ventral hemitrichium, that unfortunately no reference use, there is no need for further terms.

[9]There is a better term for this structure and is ventral hemitrichium, that unfortunately no reference use, there is no need for further terms.

[10]This structure needs to be renamed 'Flange of anterior process' because of its topographical location and because there is also a flange on the ventral process.

[11]This structure is the same as N and therefore should be termed 'Ventral fossa'

[12]This should be called articular facet of anterior process as herein proposed

[13]This term should be chosen since it is the same as the spanish 'condilo lateral' and therefore it would be the lower in cost. If anything else fails as argumentation, I chose it as reviewer because I felt it best described the structure. Period.

Table A.2 – *Continuing from the previous page*

Structure	Terms	Frequency	Total	indCost
U	**Anterior articular surface**[14]	**1**	**7**	**6**
U	Articular median process	1	7	6
U	Condilo medial	1	7	6
U	Median articular process	1	7	6
U	Processus articulaire médian	1	7	6
U	Processus médian d'articulation de l'épine	1	7	6
U	rd2 (radialis distalis)	1	7	6
V	**Foseta para inserción de m. erector spina**[15]	**1**	**1**	**0**
W	Articular foramen	1	7	6
W	Basal foramen	1	7	6
W	**Foramen**[16]	**3**	**7**	**4**
W	Foramen de la base de la espina	1	7	6
W	Foramen médian	1	7	6
X	**Inflexion point of the median crest**[17]	**1**	**1**	**0**
Y	Canal central	1	4	3
Y	Deep median groove	1	4	3
Y	Median furrow	1	4	3
Y	**Sulcus**[18]	**1**	**4**	**3**
Z	**Posterior process**	**1**	**2**	**1**
Z	Processus posterieur de blocage	1	2	1

[14]The proper term here should be 'Anterior articular surface' because it is not a condile with a distinct peduncle and it is sometimes absent or restricted to a rugose surface

[15]To be renamed 'Anterior fossae' as it is a bilateral structure. Not to be confounded with the lateral fossae that are located beside the anterior longitudinal ridge.

[16]Maybe 'Basal foramen' could be a better description

[17]To be renamed 'Inflexion point of the anterior longitudinal ridge' or something else.

[18]In fact, this could be modified as 'Posterior sulcus'.

Bibliography

Arctaedius, P. 1738. *Ichthyologia sive opera omnia de piscibus, scilicet: Bibliotheca ichthyologica. Philosophia ichthyologica. Genera piscium. Synonymia specierum. Descriptiones specierum. Omnia in hoc genere perfectiora, quam antea ulla.* 160

Bennett, D.K. 1979. Late Cenozoic fish faunas from Nebraska. *Transactions of the Kansas Academy of Science*, 82:146–177. 158, 159

Bisbal, G.A. and Gomez, S.E. 1986. Morfolofgía comparada de la espina pectoral de algunos Siluriformes Bonaerenses (Argentina). *Physis*, 44:81–93. 159

Bornbusch, A.H. 1991. Redescription and reclassification of the Silurid Catfish *Apodoglanis furnessi* Fowler (Siluriformes: Siluridae), with diagnoses of three intrafamilial Silurid subgroups. *Copeia*, 1991:1070–1084. 159, 160

Bosher, B.B.T., Newton, S.H.S., and Fine, M.L.M. 2006. The spines of the Channel Catfish, *Ictalurus punctatus*, as an anti-predator adaptation: an experimental study. *Ethology*, 112:188–195. 160

Boulenger, G.A. 1900. On some little-known African Silurid fishes of the subfamily Doradinae. *Annals and Magazine of Natural History*, 6:520–529. 161

Chen, X. and Lundberg, J.G. 1995. *Xiurenbagrus*, a new genus of Amblycipitid Catfishes (Teleostei: Siluriformes), and phylogenetic relationships among the genera of Amblycipitidae. *Copeia*, 1995:780–800. 160

Cione, A.L., Azpelicueta, M.d.l.M., Casciotta, J.R., and Dozo, M.T. 2005. Tropical freshwater teleosts from Miocene beds of eastern Patagonia, southern Argentina. *Geobios*, 38:29–42. 158, 160

Cuvier, G. and Valenciennes, A. 1840. *Histoire Naturelle des Poissons. Volume 15, Book 17.* Strasbourg. 157, 158, 159, 161

Diogo, R. 2007. Osteology and myology of the cephalic region and pectoral girdle of *Heptapterus mustelinus*, comparisons with other heptapterins, and discussion on the synapomorphies and phylogenetic relationships of the Heptapterinae and the Pimelodidae. *International Journal of Morphology*, 25:735–748. 159

Diogo, R., Chardon, M., and Vandewalle, P. 2003. Osteology and myology of the cephalic region and pectoral girdle of *Erethistes pusillus*, comparison with other erethistids, and comments on the synapomorphies and phylogenetic relationships of the Erethistidae (Teleostei: Siluriformes). *Journal of Fish Biology*, 63:1160–1175. 160

Divay, J.D. and Murray, A.M. 2015. The late Eocene-early Oligocene ichthyofauna from the Eastend Area of the Cypress Hills Formation, Saskatchewan, Canada. *Journal of Vertebrate Paleontology*, 35:e956877. 157, 158, 160, 161

Egge, J.J.D. and Simons, A.M. 2011. Evolution of venom delivery structures in Madtom Catfishes (Siluriformes: Ictaluridae). *Biological Journal of the Linnean Society*, 102:115–129. 160

Eigenmann, C.H. 1925. A review of the Doradidae, a family of South American Nematognathi, or Catfishes. *Transactions of the American Philosophical Society, New Series*, 22:279–365. 157, 158, 159, 160

Eigenmann, C.H. and Eigenmann, R.S. 1888. Preliminary notes on South American Nematognathi. I. *Proceedings of the California Academy of Sciences*, 2:119–172. 159, 160, 161

Eigenmann, C.H. and Eigenmann, R.S. 1890. South American Nematognathi or Cat-Fishes. *Occassional Papers of the California Academy of Sciences*, 1:1–508. 157, 158, 159, 160, 161

Evermann, B.W. and Kendall, W.C. 1898. Descriptions of new or little-known genera and species of Fishes from the United States. *Bulletin of the United States Fish Commission*, 17:125–133. 160

Ferraris, C.J. and Mago-Leccia, F. 1989. A new genus and species of Pimelodid Catfish from the Río Negro and Río Orinoco drainages of Venezuela. *Copeia*, 1989:166–171. 158, 160

Gayet, M. and van Neer, W. 1990. Caractères diagnostiques des épines de quelques silures africains. *Journal of African Zoology*, 104:241–252. 157, 158, 159, 161

Geerinckx, T., Adriaens, D., Teugels, G.G., and Verraes, W. 2004. A systematic revision of the African catfish genus *Parauchenoglanis* (Siluriformes: Claroteidae). *Journal of Natural History*, 38:775–803. 160

Greenwood, P.H. 1959. Quaternary fish-fossils. *Parc National Albert Mission J. de Heinzelin de Braucourt*, 4:5–80. 157, 158

Hubbs, C.H. and Hibbard, C.W. 1951. *Ictalurus lambda*, a new Catfish, based on a pectoral spine from the lower Pliocene of Kansas. *Copeia*, 1951:8–14. 158, 159, 160, 161

Jordan, D.S. 1880. *Manual of the Vertebrates of the Northern United States, Including the District East of the Mississippi River, and North of North Carolina and Tennessee, Exclusive of Marine Species*. 3rd edition. Jansen, McClurg & Company. 158, 160

Kaatz, I.M., Stewart, D.J., Rice, A.N., and Lobel, P.S. 2010. Differences in pectoral fin spine morphology between vocal and silent clades of catfishes (Order Siluriformes): Ecomorphological implications. *Current Zoology*, 56:73–89. 160

Linnaeus, C. 1758. *Systema naturae per regna tria naturae: Secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis*. 10th edition. 159, 160, 161

Lucena, C.A.S., Malabarba, L.R., and Reis, R.E. 1992. Resurrection of the Neotropical Pimelodid Catfish *Parapimelodus nigribarbis* (Boulenger), with a phylogenetic diagnosis of the genus *Parapimelodus* (Teleostei: Siluriformes). *Copeia*, 1992:138–146. 157, 159

Lundberg, J.G. 1975. The fossil Catfishes of North America. *University of Michigan, Papers on Paleontology*, 11:1–51. 157, 158, 159, 160

Lundberg, J.G., Berra, T.M., and Friel, J.P. 2004. First description of small juveniles of the primitive catfish *Diplomystes* (Siluriformes: Diplomystidae). *Ichthyological Exploration of Freshwaters*, 15:71–82. 159

Lundberg, J.G. and McDade, L.A. 1986. On the South American Catfish *Brachyrhamdia imitator* Myers (Siluriformes, Pimelodidae), with phylogenetic evidence for a large intrafamilial lineage. *Notulae Naturae*, 463. 159, 160

Mees, G.F. 1974. The Auchenipteridae and Pimelodidae of Suriname (Pisces, Nematognathi). *Zoologische Verhandelingen*, 132:1–246. 158, 159, 160, 161

Myers, G.S. 1927. Descriptions of new South American Fresh-water Fishes collected by Dr. Carl Ternetz. *Bulletin of the Museum of Comparative Zoology at Harvard College*, 68:107–135. 161

Ng, H.H. 2003. A revision of the South Asian sisorid catfish genus *Sisor* (Teleostei: Siluriformes). *Journal of Natural History*, 37:2871–2883. 160

Ohara, W.M., Tencatt, L.F.C., and Britto, M.R. 2016. Wrapped in flames: *Corydoras hephaestus*, a new remarkably colored species from the Rio Madeira basin (Teleostei: Callichthyidae). *Zootaxa*, 4170:539–552. 158, 160

Otero, O., Likius, A., Vignaud, P., and Brunet, M. 2007. A new claroteid Catfish (Siluriformes) from the upper Miocene of Toros-Menalla, Chad: *Auchenoglanis soye*, sp.nov. *Journal of Vertebrate Paleontology*, 27:285–294. 157

Otero, O., Pinton, A., Mackaye, H.T., Likius, A., Vignaud, P., and Brunet, M. 2009. First description of a Pliocene ichthyofauna from Central Africa (site KL2, Kolle area, Eastern Djurab, Chad): What do we learn? *Journal of African Earth Sciences*, 54:62–74. 161

Page, L.M., Hadiaty, R.K., López, J.A., Rachmatika, I., and Robins, R.H. 2007. Two new species of the *Akysis variegatus* species sroup (Siluriformes: Akysidae) from Southern Sumatra and a Redescription of *Akysis variegatus* Bleeker, 1846. *Copeia*, 2007:292–303. 160

Paloumpis, A.A. 1963. A key to the Illinois species of *Ictalurus* (Class Pisces) based on pectoral spines. *Transactions of te Illinois Academy of Science*, 56:129–133. 159, 160, 161

Parisi, B.M., Lundberg, J.G., and Donascimiento, C. 2006. *Propimelodus caesius* a new species of long-finned pimelodid catfish (Teleostei: Siluriformes) from the Amazon Basin, South America. *Proceedings of the Academy of Natural Sciences of Philadelphia*, 155:67–78. 157, 158, 159, 160, 161

Parmentier, E., Fabri, G., Kaatz, I., Decloux, N., Planes, S., and Vandewalle, P. 2010. Functional study of the pectoral spine stridulation mechanism in different mochokid catfishes. *Journal of Experimental Biology*, 213:1107–1114. 160

Paruch, W. 1986. Identification of Wisconsin Catfishes (Ictaluridae): A kew based on pectoral fin spines. *Wisconsin Academy of Sciences, Arts and Letters*, 74:58–62. 158, 159, 160

Pinton, A., Fara, E., and Otero, O. 2006. Spine anatomy reveals the diversity of catfish through time: a case study of *Synodontis* (Siluriformes). *Naturwissenschaften*, 93:22–26. 159, 161

Pinton, A. and Otero, O. 2010. The bony anatomy of Chadian *Synodontis* (Osteichthyes, Teleostei, Siluriformes, Mochokidae): Interspecific variations and specific characters. *Zoosystema*, 32:173–231. 157, 158, 160, 161

Pinton, A., Otero, O., Likius, A., Mackaye, H.T., Vignaud, P., and Brunet, M. 2011. Giants in a minute Catfish genus: First description of fossil *Mochokus* (Siluriformes, Mochokidae) in the late Miocene of Chad, Including *M. gigas*, sp. nov. *Journal of Vertebrate Paleontology*, 31:22–31. 160

Reed, H.D. 1924. The morphology and growth of the spines of Siluroid fishes. *Journal of Morphology*, 38:431–451. 158, 160

Royero, R. 1999. *Studies on the Systematics of the Catfish Family Auchenipteridae (Teleostei: Siluriformes)*. Ph.D. thesis, University of Bristol. 157, 158, 160

Sabaj Pérez, M.H. and Birindelli, J.L.O. 2008. Taxonomic revision of extant *Doras* Lacepède, 1803 (Siluriformes: Doradidae) with descriptions of three new species. *Proceedings of the Academy of Natural Sciences of Philadelphia*, 157:189–233. 158, 160

Shibatta, O.A. and Benine, R.C. 2005. A new species of *Microglanis* (Siluriformes: Pseudopimelodidae) from upper rio Paraná basin, Brazil. *Neotropical Ichthyology*, 3:579–585. 159, 160, 161

Smith, M.L. 1987. Osteology and systematics of Fossil Catfishes (Genus *Ictalurus*) of Central Mexico. *Journal of Paleontology*, 61:380–387. 158, 159, 160

Tencatt, L.F.C., Britto, M.R., and Pavanelli, C.S. 2016. Revisionary study of the armored catfish *Corydoras paleatus* (Jenyns, 1842) (Siluriformes: Callichthyidae) over 180 years after its discovery by Darwin , with description of a new species. *Neotropical Ichthyology*, 14:e150089. 158, 160

Tencatt, L.F.C. and Ohara, W.M. 2016. Two new species of *Corydoras* Lacépède, 1803 (Siluriformes: Callichthyidae) from the rio Madeira basin, Brazil. *Neotropical Ichthyology*, 14:e150063. 159

Teugels, G.G. and Adriaens, D. 2003. Taxonomy and phylogeny of Clariidae-An overview. In G. Arratia, B.G. Kapoor, M. Chardon, and R. Diogo (eds.), *Catfishes*, chapter 16, pp. 465–487. Friderich Pfeill Verlag. 160

Thomson, A.W. and Page, L.M. 2006. *Genera of the Asian catfish families Sisoridae and Erethistidae (Teleostei: Siluriformes)*. 1345. 157, 158, 159, 160

Trapani, J. 2008. Quaternary fossil fish from the Kibish Formation, Omo Valley, Ethiopia. *Journal of Human Evolution*, 55:521–530. 158, 160

Ünlü, E., Deger, D., and Çiçek, T. 2012. Comparision of morphological and anatomical characters in two catfish species, *Silurus tristegus* Heckel, 1843 and *Silurus glanis* L., 1758 (Siluridae, Siluriformes). *North-Western Journal of Zoology*, 8:119–124. 160

van Neer, W. 1992. New late Tertiary fish fossils from the Sinda region, Eastern Zaire. *African Study Monographs. Supplementary Issue*, 17:27–47. 160

Vanscoy, T., Lundberg, J.G., and Luckenbill, K.R. 2015. Bony ornamentation of the catfish pectoral-fin spine: comparative and developmental anatomy, with an example of fin-spine diversity using the Tribe Brachyplatystomini (Siluriformes, Pimelodidae). *Proceedings of the Academy of Natural Sciences of Philadelphia*, 164:177–212. 159, 160

Vigliotta, T.R. 2008. A phylogenetic study of the African catfish family Mochokidae (Osteichthyes, Ostariophysi, Siluriformes), with a key to genera. *Proceedings of the Academy of Natural Sciences of Philadelphia*, 157:73–136. 157, 158, 159, 160

Watanabe, K. and Uyeno, T. 1999. Fossil bagrid catfishes from Japan and their zoogeography, with description of a new species, *Pseudobagrus ikiensis*. *Ichthyological Research*, 46:397–412. 158, 160

Wosiacki, W.B., Pereira, T.d.G., and Reis, R.E. 2014. Description of a new species of *Aspidoras* (Siluriformes, Callichthyidae) from the Serra dos Carajás, Lower Tocantins River Basin, Brazil. *Copeia*, 2014:309–316. 158, 160

Wright, J.J. 2009. Diversity, phylogenetic distribution, and origins of venomous catfishes. *BMC Evolutionary Biology*, 9:282–293. 158, 160

Appendix B

Material examined

ACTINOPTERYGII
CHARACIFORMES

ACESTRORHYNCHIDAE

Acestrorhynchus falcirostris: MZUSP 92991 (1 out of 2, DS). *Acestrorhynchus heterolepis*: MZUSP 82896 (1, DS). *Acestrorhynchus lacustris*: MZUSP 83376 (1, DS). *Gilbertolus alatus*: ICNMHN 16753 (16, 99.9–120.0 mm SL), BMNH 1924.3.3.46–48 (1, CS). *Gilbertolus atratoensis*: MZUSP 10663 (1, CS). *Roestes molossus*: INPA 11068 (1, CS), MZUSP 52063 (1, CS). *Roestes ogilviei*: ICNMHN 144699 (1, 120.0 mm SL).

CHARACIDAE

Cynopotamus atratoensis: USNM 310494 (2, 142.1–144.3 mm SL).

CYNODONTIDAE

Cynodon gibbus: MZUSP 15641 (1, CS), USNM 222848 (1, 205.0 mm SL), USNM 257564 (1, 197.0 mm SL), USNM 305368 (2, 167.0–240.0 mm SL), USNM 403923 (1, 250 mm SL). *Cynodon meionactis*: MZUSP 99683 (1, 151.9 mm SL). *Cynodon* sp.: MZUSP 91687 (1, DS). *Cynodon septenarius*: MZUSP 92544 (3, 189.2–195.5 mm SL). *Hydrolycus armatus*: MZUSP 94477 (1, DS), 89507 (2, DS), 95879 (1, DS), USNM 403920 (1, 245.0 mm SL), USNM 403921 (1, 235.0 mm SL). *Hydrolycus scomberoides*: ANSP 159099 (1, DS), MZUSP 6971 (1, CS), 26177 (1, CS), 32093 (1, CS), 32616 (1 out of 8, 175.3 mm SL, dentary as DS preparation), USNM 403923 (1, 250.0 mm SL). *Hydrolycus tatauaia*: MZUSP 94086 (3 out of 4, DS), 32637 (1, CS), USNM 402147 (1, 200.0 mm SL). *"Hydrolycus" wallacei*: MZUSP 32634 (1, 269.0 mm SL, dentary as DS preparation), 32638 (1, CS). *Rhaphiodon vulpinus*: MZUSP 117050 (2, DS), 92008 (1, DS), USNM 52549 (4, 250.0–300.0 mm SL), USNM 126662 (5, 128.8–184.0 mm SL), USNM 222863 (1, 233.0 mm SL), USNM 310941 (2, 198.0–265.0 mm SL).

SERRASALMIDAE

Acnodon sp.: MZUSP 20361. *Catoprion mento* MZUSP 20246 (10) *Metynnis fasciatus*: MZUSP 20555. *Mylesinus schomburgkii*: MZUSP 101455. *Myleus setiger*: MZUSP 15673. *Myleus* sp.: MZUSP 91685. *Myloplus rhomboidalis*: MZUSP 102136. *Myloplus schomburgkii*:

MZUSP 97639. *Myloplus torquatus*: MZUSP 95388. *Mylossoma acanthogaster*: IAvH-P 15900. *Mylossoma duriventre*: MZUSP 89523. *Mylossoma paraguayensis* MZUSP 895232. *Piaractus brachypomus*: MZUSP 20376. *Piaractus mesopotamicus*: MZUSP 89508. *Pygocentrus nattereri*: MZUSP 89511. *Pygopristis denticulata*: MZUSP 57580. *Tometes kranponhah*: MZUSP 110948. *Utiaritichthys sennaebragai*: MZUSP 93692

SILURIFORMES

ARIIDAE

Sciades proops: MZUSP 52842.

ASTROBLEPIDAE

Astroblepus sp.: MZUSP 22326.

ASPREDINIDAE

Bunocephalus coracoideus: MZUSP 103254.

AUCHENIPTERIDAE

Ageneiosus pardalis: ICNMHN 10014. *Liosomadoras morrowi*: ICNMHN 14026. *Liosomadoras oncinus*: MZUSP 105828. *Tetranematichthys quadrifilis*: ICNMHN 12385. *Tocantinsia piresi*: MZUSP 100031, 98301, 103244, 103415. *Trachelyopterichthys anduzei*: ICNMHN 16904. *Trachelyopterichthys taeniatus*: MZUSP 6642. *Trachelyopterus striatulus*: MZUSP 86012. *Trachelyopterus insignis*: ICNMHN uncat. *Trachycorystes trachycorystes*: ICNMHN 2761, MZUSP 104547, 111057. *Trachycorystes porosus*: MZUSP 101876. *Trachycorystes* sp.: MZUSP 91659.

BAGRIDAE

Mystus gulio: MZUSP 63581.

CALLICHTHYIDAE

Aspidoras albater: MZUSP 28599, 28, 14.59–39.76; 50157, 1, cs. *Aspidoras microgalaeus*: MZUSP 86842, 1, cs. *Callichthys callichthys*: MZUSP 43599, 1, 121.12; MZUSP 93043, 8, 47.68–72.60; MZUSP 109087, 2, 116.71–145.12; MZUSP 84199, 2, cs. *Callichthys serralabiatum*: MZUSP 93168, 1, 104.65. *Corydoras araguaiaensis*: MZUSP 86269, 1, cs. *Corydoras ehrhardti*: MZUSP 81572. *Corydoras* cf. *guianensis*: MZUSP 107151, 2, cs. *Corydoras splendens*: MZUSP 89377, 5, 47.16–57.47. *Dianema* sp.: MZUSP 30862, 2, cs. *Hoplosternum littorale*: MZUSP 85987, 1, 84.91; MZUSP 94658, 1, 95.46; MZUSP 117107, 1, cs. *Lepthoplosternum pectorale*: MZUSP 83608, 2, cs; MZUSP 112403, 1, 57.13. *Megalechis thorocata*: MZUSP 25451, 2, cs. *Scleromystax barbatus*: MZUSP 37723, 2, cs.

CETOPSIDAE

Cetopsis gobioides: MZUSP 99279.

CLARIIDAE

Clariallabes longicauda: BMNH 1937.12.12:22. *Clariallabes melas*: MRAC 51599-611. *Dinotopterus cunningtoni*: UMMZ 199855. *Clarias* sp.: MZUSP 91656.

CLAROTEIDAE

Chrysichthys auratus: BMNH 1907.13.2:2082-7. *Phyllonemus filinemus*: MRAC 90257.

DIPLOMYSTIDAE

Diplomystes camposensis: MZUSP 88533. *Diplomystes viedmensis*: ANSP 192904.

CRANOGLANIDIDAE

Cranoglanis bouderius: ANSP 164978 (CT available at http://catfishbone.acnatsci.org/Cranoglanididae/Cranoglanis/bouderius/catscan.html).

DORADIDAE

Acanthodoras cataphractus: MZUSP 31098 (1). *Acanthodoras spinosissimus*: ICNMHN 8266, MZUSP 112857, MZUSP 29660 (1). *Agamyxis albomaculatus*: ICNMHN 1319, MZUSP 88607. *Agamyxis pectinifrons*: MZUSP 51777 (1). *Amblydoras affinis*: ICNMHN 8287, MZUSP 31699 (1). *Amblydoras bolivarensis*: MZUSP 88610. *Anadoras grypus*: ICNMHN 8274 MZUSP 86818, MZUSP 36119 (7). *Anadoras weddellii*: MZUSP 31701 (1). *Anduzedoras oxyrhynchus*: MZUSP 29025 (2). *Astrodoras asterifrons*: MZUSP 29049, MZUSP 111428 (2). *Centrodoras brachiatus*: MZUSP 103891 (2). *Centrodoras hasemani*: MZUSP 91675 (1). *Centrochir crocodili*: ICNMHN 5698. *Doras carinatus*: MZUSP 108947 (1). *Doras higuchii*: MZUSP 101627 (1). *Doras zuanoni*: MZUSP 110808 (1). *Franciscodoras marmoratus*: MZUSP 2201 (7). *Hassar affinis*: MZUSP 74809 (6). *Hassar gabiru*: MZUSP 111342 (10). *Hassar orestis*: MZUSP 22057 (1), MZUSP 82256 (1), MZUSP 21915 (1). *Hassar wilderi*: MZUSP 62998 (5). *Hemidoras morei*: MZUSP MZUSP 27844 (1). *Hemidoras morrisi*: MZUSP 82890 (4).x *Hemidoras stenopeltis*: ICNMHN 5936, MZUSP 42772 (4), MZUSP 7541 (4). *Hemidoras stuebelii*: ICNMHN 17104. *Hypodoras forficulatus*: ICNMHN 10143. *Kalyptodoras bahiensis*: MZUSP 100737 (4). *Leptodoras acipenserinus*: MZUSP 49774 (1). *Leptodoras cataniai*: MZUSP 82978 (2). *Leptodoras hasemani*: MZUSP 111855 (3). *Leptodoras juruensis*: MZUSP 82985 (1), MZUSP 82880 (2). *Leptodoras nelsoni*: ICNMHN 413. *Leptodoras oyakawai*: MZUSP 98216 (3). *Leptodoras praelongus*: MZUSP 31694 (1). *Leptodoras* sp.: ICNMHN 8422. *Lithodoras dorsalis*: MZUSP 46587 (1), MZUSP 74677 (4). *Megalodoras uranoscopus*: ICNMHN 6996, MZUSP 46007 (1), MZUSP 55838 (5). *Merodoras nheco*: MZUSP 109766 (1). *Nemadoras elongatus*: ICNMHN 8399, MZUSP 83211 (2). *Nemadoras hemipeltis*: MZUSP 83262 (1). *Nemadoras humeralis*: MZUSP 103873 (1). *Orinocodoras eigenmanni*: ICNMHN 8441, MZUSP 86807 (1). *Ossancora eigenmanni*: MZUSP 95024 (66), MZUSP 31709 (1). *Ossancora fimbriata*: MZUSP 56703 (39), MZUSP 52534 (3). *Ossancora punctata*: ICNMHN 16705, MZUSP 82894 (12). *Oxydoras niger*: ICNMHN 8445, MZUSP 56870 (1), MZUSP 86819 (3). *Oxydoras sifontesi*: MZUSP 76540 (1). *Physopixys lyra*: MZUSP 62709, MZUSP 121168 (6). *Physopyxis ananas*: MZUSP 117504 (9). *Platydoras armatulus*: ICNMHN 10025, MZUSP 90425 (1), MZUSP 116542 (5). *Platydoras costatus*: MZUSP 37008 (1). *Pterodoras granulosus*: ICNMHN 8300, MZUSP 91655, MZUSP 36100 (2). *Pterodoras rivasi*: MZUSP 88613 (1). *Rhinodoras armbrusteri*: MZUSP 125377 (1). *Rhinodoras dorbignyi*: MZUSP 106761 (3). *Rhinodoras thomersoni*: ICNMHN 2172. *Rhynchodoras woodsi*: MZUSP 86816 (1). *Scorpiodoras heckelii*: ICNMHN 12962, MZUSP 112942, MZUSP 85494 (1), MZUSP 36251 (1). *Tenellus leporhinus*: MZUSP 97520 (1). *Tenellus trimaculatus*: MZUSP 92344 (1). *Tenellus ternetzi*: MZUSP 55760 (3). *Trachydoras nattereri*: MZUSP 92372 (2). *Trachydoras paraguayensis*:

MZUSP 62700 (2). *Trachydoras steindachneri*: MZUSP 49526 (5). *Wertheimeria maculata*: MZUSP 106759 (1).

HEPTAPTERIDAE

Goeldiella eques: MZUSP 33190. *Pimelodella chagresi*: ICNMHN 1623. *Rhamdia quelen*: MZUSP 102718.

HETEROPNEUSTIDAE

Heteropneustes fossilis: USNM 273737.

LORICARIIDAE

Ancistrus centrolepis: IAvH-P 10473, ICNMHN 104, 189, 1632, 3153. *Ancistrus martini*: ICNMHN 1206, 17647, 17648, 17653. *Ancistrus triradiatus*: ICNMHN 17649, 17654. *Baryancistrus niveatus*: MNRJ 19344. *Chaetostoma alternifasciatum*: ANSP 71711 (holotype, photograph, xray). *Chaetostoma anale*: ANSP 70525 (holotype), ICNMHN 13397, 17634. *Chaetostoma anomalum*: USNM 133135 (syntype, photograph, xray). *Chaetostoma breve*: BMNH 1898.11.4.33-36 (syntypes, photograph). *Chaetostoma brevilabiatum*: ICNMHN 6134 (holotype). *Chaetostoma carrioni*: BMNH 1933.5.29.1 (holotype, photograph). *Chaetostoma dorsale*: ICNMHN 1183, 3372, 3535, 7997, 8011, 8013, 8027, 8031, 17499, 17646, MLS 588, 747, 604. *Chaetostoma jegui*: INPA Uncatalogued (photograph). *Chaetostoma loborhynchos*: MUSM 20291 (photograph), CZUT-IC 5551, 5552. *Chaetostoma microps*: BMNH 1860.6.16.137-143 (syntypes, photograph). *Chaetostoma milesi*: ANSP 69330 (holotype), ICNMHN 10420, 15528, 16123, 16268, 16291, 16923, MLS 562. *Chaetostoma pearsei*: ICNMHN 10361. *Chaetostoma platyrhynchus*: ICNMHN 5488, 5492, 7971, 9417, 17624, 17625, 17626, 17628, 17629, 17630, 17631. *Chaetostoma sovichthys*: ICNMHN 2381, 16221, 16223, MLS 568, 590, 600, USNM 121053 (holotype, photograph, xray). *Chaetostoma* sp. "Perú": AUM 45597, 45634. *Chaetostoma tachiraense*: MLS 797, 799, 805, USNM 121052 (holotype, photograph, xray). *Chaetostoma vagum*: ANSP 70521 (holotype, photograph). *Cordylancistrus daguae*: ICNMHN 3515, 17643, 17644, 17645. *Cordylancistrus perijae*: ANSP 168917 (paratype), ICNMHN 17502 (ex CAR 270, ex MBUCV-V 21745, paratype). *Cordylancistrus platycephalus*: BMNH 1898.11.4:42 (holotype, photograph). *Cordylancistrus tayrona*: ICNMHN 17503 (ex CAR 370), MLS 541. *Cordylancistrus* sp2. "Pacífico": BMNH 1908.5.29.70-79. *Cordylancistrus* sp3. "Magdalena": CP-UCO 1060, 1062. *Cordylancistrus* sp4.: FMNH 76213 (cs). *Cordylancistrus torbesensis*: USNM 121001 (holotype, xray). *Dekeyseria niveata*: ANSP 185259. *Dekeyseria pulcher*: ANSP 185298. *Dekeyseria scaphirhyncha*: ICNMHN 12787, 12788. *Dolichancistrus atratoensis*: CIUA 768, 769, 771, 772, IAvH-P 6630, ICNMHN 51 (holotype), 46 (paratypes), 74 (paratype), 3460. *Dolichancistrus carnegiei*: ICNMHN 591, 1822, 3235, 3571, 5445, 16016, 16017, 16018, 17498, 17500, 17501, MLS 522, 542, 543, 550. *Dolichancistrus cobrensis*: AUM 30377, 46306, MCNG 541, USNM 121036 (holotype), 121037 (paratypes). *Dolichancistrus fuesslii*: IAvH-P 7931, 11381, 3939, 3940, 9230, 9605, ICNMHN 2638, 2817, 2839, 3212, 3641, 14582, 16811, NMW 48026. *Hemiancistrus sabaji*: ANSP 185153. *Hemiancistrus guahiborum*: ICNMHN 5323, 11915. *Hemiancistrus punctulatus*: ANSP 170168. *Hopliancistrus tricornis*: AUM 39853. *Hypan-*

cistrus contradens: ICNMHN 11917, 11918. *Hypancistrus debilittera*: ICNMHN 10691. *Hypostomus* cf. *wachereri*: MZUSP 87480. *Lasiancistrus caucanus*: ICNMHN 8763. *Lasiancistrus guacharote*: ICNMHN 16916. *Lasiancistrus triactis*: ZMA 120774. *Leporacanthicus galaxias*: AUM 42144. *Leptoancistrus canensis*: USNM 78300 (paratypes, xray). *Leptoancistrus* cf. *cordobensis*: CIUA 774, 775, 776, 777, 778, 779, 780, 781. *Lithoxus jantjae*: ANSP 182809 (paratypes). *Lithoxus lithoides*: ANSP 39121 (paratype). *Megalancistus aculeatus*: USNM 52594. *Neblinichthys pilosus*: ANSP 157587 (paratypes). *Neblinichthys roraima*: ANSP 174914 (paratypes). *Panaque cochliodon*: ICNMHN 369. *Panaque maccus*: ICNMHN 15728. *Peckoltia bachi*: ICNMHN 13952. *Peckoltia brevis*: ICNMHN 7952. *Peckoltia vittata*: ICNMHN 7954, 12792. *Pseudacanthicus leopardus*: AUM 35550, USNM 197105. *Pseudacanthicus spinosus*: USNM 52594. *Pseudancistrus sidereus*: ANSP 185297. *Pseudolithoxus dumus*: ANSP 185255. *Spectracanthicus punctatissimus*: MNHN 1999-0021.

MALAPTERURIDAE

Malapterurus beninensis: MZUSP 84464, USNM 303492.

MOCHOCKIDAE

Chiloglanis swierstrai: MZUSP 65801. *Synodontis schall*: MZUSP 84468.

NEMATOGENYIDAE

Nematogenys inermis: MZUSP 75256.

PANGASIIDAE

Pangasius macronema: MZUSP 62606. *Pangasius pangasius*: UMMZ 208434.

PIMELODIDAE

Bergiaria westermanni: MZUSP 85627. *Brachyplatystoma capapretum*: MZUSP 53262, ANSP 178524, 179733, 178524. *Brachyplatystoma juruense*: ANSP 178514-1,3,4,6, DUF 1071. *Brachyplatystoma platynemum*: DUF 993, 1076, ANSP uncat, 187321. *Brachyplatystoma rousseauxii*: ANSP 179794, 179793, 179233, DUF 1051, 981, 1078. *Brachyplatystoma vaillanti*: ANSP 179799, 178525, 179474, DUF 994, 1164. *Brachyplaystoma filamentosum*: DUF 1079, 1080, ANSP 179776. *Calophysus macropterus*: DUF 1199, 1049, 403, ANSP 199813, 178164, 178260, ICNMHN uncat. *Duopalatinus emarginatus*: MZUSP 85622. *Exallodontus aguanai*: ANSP 18947, uncat. *Hemisorubim platyrhynchos*: ANSP uncat, 179234, ICNMHN 7909, MZUSP 7009. *Hypophthalmus* sp. "curved": ANSP uncat. *Hypophthalmus* sp. "straight": ANSP 180993, 178512, 180993, 187103. *Iheringichthys labrosus*: ANSP 180505, MZUSP 78459, 25102. *Leiarius marmoratus*: MZUSP 108333. *Leiarius perruno*: ANSP uncat-I, uncat-II, ICNMHN 2158. *Leiarius pictus*: MZUSP 82577. *Leiarius* sp.: ANSP 178526, 178527, uncat without skull, DUF 1054, 1056, 1036, 1055, 1037. *Luciopimelodus pati*: ANSP 178798, MZUSP 78464, 78457. *Megalonema platanum*: MZUSP 78465. *Megalonema platycephalum*: ANSP 179249, 178515. *Megalonema* sp.: MZUSP 92604. *Parapimelodus nigribarbis*: MZUSP 78451. *Parapimelodus valenciennis*: MZUSP 78466, ANSP 178800. *Phractocephalus hemioliopterus*: ANSP 179559, 179553, 179554, ICN uncat. *Pimelodina flavipinnis*: ANSP uncat, 178513, 178516. *Pimelodus argenteus*: ANSP 181017, uncat. *Pimelodus fur*: MZUSP 22566. *Pimelodus grosskopfii*: ICNMHN 6867. *Pimelodus maculatus*: MZUSP

85486, 110379. *Pimelodus microstoma*: MZUSP 22696, 22712 (paratype of *Pimelodus heraldoi*). *Pimelodus mysteriosus*: ANSP 180506, MZUSP 90595. *Pimelodus ortmanni*: MZUSP 50053. *"Pimelodus" ornatus*: ANSP 178452, 180985, MZUSP 34480, 109128. *"Pimelodus"* sp.: MZUSP 58328. *Pinirampus argentina*: ANSP 181016. *Pinirampus pirinampu*: ANSP uncat, 178530. *Platynematichthys notatus*: ANSP uncat, 178528. *Platysilurus malarmo*: ANSP uncat, IAvH-P 11815, IAvH-P 11819, IAvH-P 15894. *Platysilurus mucosus*: DUF 986, uncat, ANSP 178508, ANSP 178509, IAvH-P 17638, IAvH-P 10850, IAvH-P 10856, IAvH-P 6006. *Platystomatichthys sturio*: ANSP uncat, SU 22463, ICNMHN 15201. *Platystomatichthys* sp. "Xingu": MZUSP 58364. *Propimelodus* sp.: ANSP 180939. *Pseudoplatystoma corruscans*: ANS 188913, MZUSP 78477. *Pseudoplatystoma fasciatum*: ANSP 177346. *Pseudoplatystoma magdaleniatum*: ICNMHN 6860. *Pseudoplatystoma metense*: ANSP 149541. *Pseudoplatystoma reticulatum*: ANSP 188912. *Pseudoplatystoma* sp.: DUF 1125, ICNMHN uncat. *Pseudoplatystoma tigrinum*: DUF 921, uncat, ANSP 187010. *Sorubim cuspicaudus*: MBUCV uncat, DUF 932. *Sorubim elongatus*: ICNMHN 15039. *Sorubim lima*: ANSP 178507. *Sorubim trigonocephalus*: ANSP 188824. *Sorubimichthys planiceps*: 179235, 17850. *Steindachneridion scripta*: MZUSP 78463. *Zungaro zungaro*: ANSP uncat, DUF 982, ICN uncat, MPUJ 13213, MZUSP 96290, 108335, 94859, 96188, 151727.

PLOTOSIDAE

Plotosus lineatus: MZUSP 22179.

PSEUDOPIMELODIDAE

Batrochoglanis villosus: MZUSP 7356, 23864. *Cephalosilurus apurensis*: DU-F 924, 980, 1102, 1040, ICNMHN uncat. *Cephalosilurus fowleri*: MZUSP 73756. *Pseudopimelodus mangurus*: MZUSP 40684.*Pseudopimelodus raninus*: ICNMHN 6865. *Pseudopimelodus* sp.: MZUSP 78458. *Pseudopimelodus bufonius*: ANSP uncat, MBUCV 2001-12.6.

SCHILBEIDAE

Schilbe intermedius: MZUSP 62627. *Schilbe* sp.: MZUSP 65820.

SILURIDAE

Ompok pabo: MZUSP 63582. *Wallago leerii*: MZUSP 63357.

SISORIDAE

Bagarius cf. *yarrelli*: MZUSP 63119. *Erethistes hara*: MZUSP 48644. *Gagata cenia*: MZUSP 42052. *Pseudecheneis sulcata*: MZUSP 50844.

TRICHOMYCTERIDAE

Trichomycterus iheringi: MZUSP 120974.

DIPNOI

LEPIDOSIRENIDAE

Lepidosiren paradoxa: MZUSP 35634, 41101, 50036, 35633.

Appendix C

Phylogenetic analysis of the Pimelodidae

C.1 Introduction

The present analysis puts the fossil occurrences of *Phractocephalus* from the Sincelejo Fm. in a phylogenetic context. We included the morphological characters from Aguilera et al. (2008) and Azpelicueta and Cione (2016) along with our own characters relevant to the systematics of the genus, that includes a single extant, and three extinct species. We carried out a combined bayesian phylogenetic analysis placing the fossil taxa along with recent taxa, using both morphological and molecular data using bayesian phylogenetic inference.

C.2 Morphological characters

Previous studies have used morphological characters for inferring the interrelationships among species of *Phractocephalus*. We extended the original matrix of Aguilera et al. (2008) including the recently-described *Phractocephalus ivy* (Azpelicueta and Cione, 2016) and the specimens from the Sincelejo formation in order to assess their phylogenetic placement (Table C.1). Characters were coded for *Phractocephalus acreornatus*, *P. ivy*, and *P. nassi* following Aguilera et al. (2008), Azpelicueta and Cione (2016), and Lundberg and Aguilera (2003) respectively. Extant species were coded examining osteological specimens (material examiend in Section B). We included two additional new characters to this matrix.

C.2.1 Characters

Characters 1 to 6 were taken from Aguilera et al. (2008) with the inclusion of a new one (character 7) herein proposed. An additional character was added from recent examination of cranial osteology in the Pimelodidae (character 8). Character definitions and states are described below.

Character 1: Skull ornamentation near midline. **States:** Weakly or strongly ridged and grooved (0). Mostly reticulating ridges and pits (1). **Comments:** *Phractocephalus ivy* was

coded from figure 3 in Azpelicueta and Cione (2016). Impossible to code in our specimens of *Phractocephalus* Corozal.

Character 2: Mesethmoid width. **States:** Narrow (0). Very broad (1). **Comments:** *Phractocephalus ivy* was coded as missing (?) given that the character is not preserved in the available specimens of the original description. Impossible to code in our specimens of *Phractocephalus* Corozal.

Character 3: Anterior and lateral expansions of lateral ethmoid. **States:** Expansions absent (0). Expansions present but weak (1). Expansions present and strong (2). **Comments:** Said to be the same condition in *P. ivy* as in *Phractocephalus hemioliopterus* (Azpelicueta and Cione, 2016, p. 255). Impossible to code in our specimens of *Phractocephalus* Corozal.

Character 4: Length of supraoccipital process as compared to the length of Weberian apparatus. **States:** Supraoccipital process shorter than Weberian process (0). Supraoccipital process longer than Weberian process (1). **Comments:** Coded as missing (?) for *P. ivy* given that the character is not preserved in the available specimens of the original description. This character was coded as state 2 for *P. acreornatus* and *P. nassi* (Aguilera et al., 2008, table 2); however, the character definition only includes two states, 0 and 1. This coding error was corrected in the present matrix where the species in question should have state 1 instead. Impossible to code in our specimens of *Phractocephalus* Corozal.

Character 5: Degree of development of ornamentation on opercle. **States:** Localized and poorly-developed (0). Covering most of the opercle but moderately-developed (1). Covering most of the opercle and strongly-developed (2). **Comments:** Said to be the same condition in *P. ivy* as in *P. nassi* and *P. acreornatus* (Azpelicueta and Cione, 2016, p. 255). Impossible to code in our specimens of *Phractocephalus* Corozal.

Character 6: Ornamentation on the pectoral-spine shaft. **States:** Consisting of semi-parallel ridges or shaft surface smooth (0). Consisting of bony reticulations and some pits inbetween (1). Consisting of strong reticulations that form extensive cover of pits on the spine shaft (2). **Comments:** Redefined in the present work and coded from the figure 9 in Azpelicueta and Cione (2016) for *P. ivy*.

Character 7: Bifid spinules on proximal section of posterior pectoral-spine shaft. **States:** Absent (0). Present (1). **Comments:** New character, coded from figure 9 of (Azpelicueta and Cione, 2016) for *P. ivy*, from figure 5 and explicitly from text (p.104) in Lundberg and Aguilera (2003) for *P. nassi*, and from figure 10 in Aguilera et al. (2008) for *P. acreornatus*.

Character 8: Contribution of the lateral ethmoid to the margin of the orbit. **States:** Less than half of the dorsal margin of the orbit (0). More than half of the dorsal margin of the orbit (1). **Comments:** New character, coded from specimens examined in extant species, and from figure 7 of Azpelicueta and Cione (2016) for *P. ivy*, from figure 3 of Lundberg and Aguilera (2003) for *P. nassi*, and from figure 2 of Aguilera et al. (2008) for *P. acreornatus*. Impossible to code in our specimens of *Phractocephalus* Corozal.

C.2.2 Morphological matrix (modified from Aguilera et al. (2008))

The modified matrix is herein presented with codes indicating the specific source of coding.

Citation codes for the table:

- AC: Azpelicueta and Cione (2016)
- AE: Aguilera et al. (2008)
- LA: Lundberg and Aguilera (2003)

Table C.1: *Morphological characters used in the phylogenetic analysis. Missing data (?) are indicated whenever the condition could not be unambiguously coded. The codes by each character state indicate the figure f or page p where the state can be evidenced in the relevant references.*

Taxon and characters	Ch1	Ch2	Ch3	Ch4	Ch5	Ch6	Ch7	Ch8
Outgroup (composite)	0	0	0	0	0	0	0	0
P. acreornatus	0	1	2	1	2	2	1 AEf10	1 AEf2
P. ivy	0 ACf3	?	0 ACp255	?	2 ACp255	1 ACf9	1 ACf9	1 ACf7
P. nassi	0	1	1	1	2	1	1 LAf5p104	1 LAf3
P. hemioliopterus	1	0	0	1	1	2	1	1
Phractocephalus Corozal	?	?	?	?	?	1	1	?

Please note that while Aguilera et al. (2008) treated the outgroup as a truly composite entry in the matrix for inference, our current instance is composite only for brevity as it is invariant in all the characters presented herein, it will be coded accordingly when concatenated to the molecular partitions.

C.3 Phylogenetic systematics

We analyzed a combined molecular and morphological dataset with 84 terminals representing 24 genera of the Pimelodidae and two genera of the Pseudopimelodidae and the Heptapteridae respectively, while the fossil taxon set was composed of the three known extinct species of *Phractocephalus* as well as the specimens herein studied, for a total of four fossil terminals. DNA sequences for the molecular partitions were taken from Lundberg et al. (2011), who included the single-copy nuclear recombination activating genes (*rag1* and *rag2*), a mitochondrial region comprising the *12S* rRNA, tRNA-val and *16S* rRNA genes, and the cytochrome-b *cytb* gene (including Threonine tRNA and partial Proline tRNA regions), for a total of four molecular partitions. All available sequences were retrieved from GenBank and aligned with `MAFFT` v.7.271 (Katoh and Standley, 2013) using the `G-INS-i` algorithm with 1000 iterations except for *rag1* where we used the `L-INS-i` algorithm due to the unequal size of sequences in this partition. Minor manual adjustments were carried out for the *rag1* and *rag2* alignments. Additionally, we used `Gblocks` v.0.91 (Castresana, 2000; Talavera and Castresana, 2007) for reproducible exclusion of ambiguous and hypervariable

positions of the *12S* alignment with the following settings: `-t=d -b=a -d=y`; a total of 84 (4%) positions were removed from the original alignment. Accession numbers are available in Section C.2.2; missing data were coded as "`?`". A total dataset with 84 terminals and 7578 characters was assembled for analysis.

Phylogenetic analysis was performed using Bayesian Inference as implemented in `MrBayes` v.3.2.6 (Ronquist et al., 2012). Substitution models for each partition were selected using `JModeltest` 2.1.10 v.20160303 (Darriba et al., 2012) and `PhyML` 3.0 v.20131022 (Guindon et al., 2010) using the Bayesian Information Criterion (BIC) following Darriba et al. (2012). The best-fit substitution models were GTR+I+Γ (*rag1*); K80+I+Γ (*rag2*); GTR+I+Γ (*12S*), and HKY+I+Γ (*cytb*). Parameters other than topology were unlinked across partitions. Two runs with eight independent Markov chains were run in parallel for 2.000.000 generations and sampling every 2.000 generations. Convergence of runs was determined based on the average standard deviation of the split frequencies (ASDSF) < 0.01, and the effective sample size (ESS), calculated using `Tracer` v1.6.0 (Drummond et al., 2012), that was > 300 for all parameters. Additionally, the potential scale reduction factor (PSRF) approached 1.0, suggesting convergence in the estimation of the posterior probabilities of nodes and branch length parameters. The posterior density graphs of the two independent runs were also examined in `Tracer`, and not visual differences were found between them. The 25% of trees were discarded as burn-in and a 50% majority-rule tree and posterior probabilities (PP) for node support were calculated using the remaining trees. The dataset (in nexus format) and scripts (in `R`, `bash`, and `MrBayes`) for reproducing the analysis are available in Appendix E.2 as well as on GitLab https://gitlab.com/gaballench/phractocephalus.

C.3.1 Dataset in **nexus** format

This preamble is important (special parts in letters, e.g., A, B, etc.)

```
#NEXUS
BEGIN DATA;
[number of taxa in ntax and number of characters (total) in nchar]
dimensions ntax=84 nchar=7578;
[Standard are the morpho chars and DNA the molec chars, they must start in
7571 and end in 7578 position (morpho) and start in 1 and end in 7570 (molec)]
['interleave' indicates that the matrix can be specified in blocks in order
to avoid pasting all characters in a single string]
format datatype=mixed(DNA: 1-7570, Standard: 7571-7578) gap=- missing=? interleave=yes;
```

When specifying the partitions by position (the `charset` command) use the following commands. This will count the partitions (Nparts = morphology + molecular partitions) and define set them to be combined with the partition name used below in `charset`. unlink parameters is important in order to allow them to be governed by their own parameters during likelihood calculations:

```
CHARSET morphology = A-B;
...
```

```
partition combined = Nparts :  morphology, molecular_partitions;
set partition = combined;
unlink shape = ( all ) pinvar = ( all ) statefreq = ( all ) revmat = ( all );
```

Also, it is important to allow al partitions to have their own rates, so the next line is important:

```
[allow separate gamma parameters for each partition]
prset applyto=(all) ratepr=variable;
```

C.3.2 Morphological substitution model

There are two substitution models used in MrBayes 3: Lewis 2001 and parsimony. The default is Lewis (2001) that is analog to JC except for a variable number of states, up to 10. On the other hand, the "parsimony" model is much parameter-rich than other models and therefore is not parsimonious in statistical terms given that each character will require its own parameter set, and therefore the whole model will have $r * n$ parameters where r is the number of parameter for each character, and n is the number of characters. Therefore, the best agnostic model would be Lewis (2001) where the whole morphological partition is governed by the same set of parameters.

One important aspect of determining the morphological partition and specifying its substitution model is to define the sampling bias, that is, how did we sample the morphological characters. If we sampled continuously so that theoretically every morphological character had the same probability of being sampled, the parameter `coding` should take the value `all`. On the other hand, if we sampled so that characters that show only 0 or 1 in all taxa can not be observed, the value should be `variable`. Finally, if we sampled only parsimony-informative characters, the value should be set to `informative`. This is really important in model-based approaches where the frequencies of states are expected to inform about the real frequencies and therefore guide the model updating during MCMC. This is called ascertainment bias when morphology is sampled in the classical way. There seems to be no much to do besides specifying either `variable` or `informative`. The important difference here between the latter two values is that variable characters need not to be informative under parsimony. This is specified under the `lset` command:

```
[morphology substitution model]
lset applyto=(1) coding=variable;
```

There seems to be no need to specify the substitution model as the default (`Lewis 2001`) is enough for this partition.

C.3.3 Taxon sets

The following taxa are considered in the analysis:

Brachyplatystoma_capapretum
Brachyplatystoma_filamentosum_179214
Brachyplatystoma_filamentosum_179446
Brachyplatystoma_filamentosum_187105
Brachyplatystoma_juruense_179217
Brachyplatystoma_juruense_179454
Brachyplatystoma_platynemum
Brachyplatystoma_rousseauxii_179233
Brachyplatystoma_vaillantii_178525
Brachyplatystoma_vaillantii_187107
Calophysus_macropterus
Cheirocerus_abuelo
Cheirocerus_eques
Duopalatinus_peruanus
Exallodontus_aguanai_179220
Exallodontus_aguanai_MBUCV_VCD_8
Goeldiella_eques
Hemisorubim_platyrhynchos
Hypophthalmus_cf_edentatus
Hypophthalmus_cf_marginatus
Hypophthalmus_fimbriatus
Iheringichthys_labrosus
Leiarius_marmoratus
Leiarius_pictus
Luciopimelodus_pati
Megalonema_amaxanthum
Megalonema_platanum
Megalonema_platycephalum_178450
Parapimelodus_nigribarbis
Parapimelodus_valenciennis
Perrunichthys_perruno
Phractocephalus_hemioliopterus
Pimelodidae_sp.1
Pimelodidae_sp.2
Pimelodidae_sp.3
Pimelodidae_sp.4
Pimelodina_flavipinnis
Pimelodus_albicans
Pimelodus_argenteus
Pimelodus_blochii_179455
Pimelodus_blochii_179456
Pimelodus_blochii_179457
Pimelodus_blochii_49101
Pimelodus_blochii_54570
Pimelodus_cf_blochii_179186
Pimelodus_cf_altissimus_179190
Pimelodus_cf_altissimus_179226
Pimelodus_coprophagus
Pimelodus_maculatus_181109
Pimelodus_maculatus_78452
Pimelodus_ornatus_178452
Pimelodus_ornatus_22394
Pimelodus_ornatus_49102
Pimelodus_ornatus_5131
Pimelodus_pictus_178168
Pimelodus_pictus_179448
Pinirampus_pirinampu_178265
Pinirampus_pirinampu_180992
Pinirampus_pirinampu_181016
Platynematichthys_notatus

```
Platysilurus_malarmo
Platysilurus_mucosus_179215
Platysilurus_mucosus_22654
Platystomatichthys_sturio
Pseudopimelodus_mangurus
Pseudoplatystoma_corruscans
Pseudoplatystoma_fasciatum
Pseudoplatystoma_magdaleniatum
Pseudoplatystoma_tigrinum
Sorubim_cuspicaudus
Sorubim_elongatus
Sorubim_lima
Sorubim_maniradii
Sorubimichthys_planiceps
Steindachneridion_scriptum
Zungaro_zungaro
Brachyplatystoma_rousseauxii_187109
Brachyplatystoma_tigrinum
Megalonema_platycephalum_179458
Leiarius_longibarbis
Steindachneridion_scripta
Hypophthalmus_edentatus
```

Some taxa were unavailable for analysis due to the lack of osteological specimens for examination. These are *Brachyplatystoma tigrinum*, *Cheirocerus abuelo*, *Cheirocerus eques*, *Hypophthalmus fimbriatus*, *Leiarius longibarbis*, *Megalonema amaxanthum*, *Pimelodus albicans*, *Pimelodus* cf. *altissimus*, *Pimelodus coprophagus*, *Pimelodus pictus*, and *Sorubim maniradii*. However, there were molecular data for these and consequently they were included in the analysis applying missing data to the morphological partition in their respective positions.

Note: This list was produced by removing whatever between blank spaces and end of line with the regexp `\s-\(.*\)$` with the keystroke `C-M-%` (or `Ctrl+Alt+Shift+5`) for `replace-regexp` in Emacs.

C.3.4 Molecular data

Fossil terminals with missing data for molecular partitions were filled with `(make-string N ??)` where N is the number of missing data in a row, and `??` represents "?" in the Emacs function `make-string`. Given that *Phractocephalus hemioliopterus* is already represented in the molecular alignment, only the three extinct species and the Corozal specimen needed to be placed after all other taxa in the molecular matrix:

```
...
Steindachneridion_scripta           ---------------------------------
Hypophthalmus_edentatus             ---------------------------------
Phractocephalus_acreornatus         ????????????????????????????????
Phractocephalus_ivy                 ????????????????????????????????
Phractocephalus_nassi               ????????????????????????????????
Phractocephalus_Corozal             ????????????????????????????????
...
```

C.4 Analysis implementation

See Appendix E.2

C.4.1 Analysis results

The following is an excerpt of relevant sections of the `.log` file:

```
...
Average standard deviation of split frequencies: 0.004698

Analysis completed in 4 hours 17 mins 20 seconds
Analysis used 15433.77 seconds of CPU time on processor 0
Likelihood of best state for "cold" chain of run 1 was -61192.14
Likelihood of best state for "cold" chain of run 2 was -61196.45
...
Summary statistics for partitions with frequency >= 0.10 in at least one run:
    Average standard deviation of split frequencies = 0.004698
    Maximum standard deviation of split frequencies = 0.032954
    Average PSRF for parameter values (excluding NA and >10.0) = 1.000
    Maximum PSRF for parameter values = 1.008
...
# 98% HPD interval remomved from this excerpt for brevity...
Parameter       Mean      Variance     Median    min ESS*  avg ESS    PSRF+
-------------------------------------------------------------------------
TL{all}       2.400185    0.009124    2.395997    626.76    688.88    1.002
kappa{2}      3.989707    0.105469    3.975593    458.80    526.19    1.000
kappa{4}     12.798448    0.379460   12.784600    621.71    631.22    0.999
r(A<->C){1}   0.089537    0.000053    0.089183    671.20    711.10    0.999
r(A<->G){1}   0.280846    0.000213    0.280674    648.14    670.27    1.000
r(A<->T){1}   0.085786    0.000054    0.085905    554.95    652.97    1.001
r(C<->G){1}   0.065785    0.000042    0.065631    751.00    751.00    1.000
r(C<->T){1}   0.419569    0.000282    0.419714    628.79    683.99    1.001
r(G<->T){1}   0.058478    0.000038    0.058417    751.00    751.00    0.999
r(A<->C){3}   0.088174    0.000027    0.088035    603.87    677.44    1.000
r(A<->G){3}   0.222331    0.000183    0.222310    336.91    422.57    1.000
r(A<->T){3}   0.064775    0.000023    0.064663    682.01    695.35    1.000
r(C<->G){3}   0.005524    0.000004    0.005292    621.49    686.25    1.004
r(C<->T){3}   0.603542    0.000234    0.603397    484.89    527.91    1.001
r(G<->T){3}   0.015654    0.000009    0.015397    633.68    692.34    0.999
pi(A){1}      0.290219    0.000065    0.290229    663.63    683.70    1.001
pi(C){1}      0.221970    0.000050    0.221863    565.40    658.20    1.000
pi(G){1}      0.238408    0.000055    0.238417    704.18    727.59    0.999
pi(T){1}      0.249403    0.000057    0.249463    609.71    680.35    0.999
pi(A){2}      0.268564    0.000126    0.268310    751.00    751.00    0.999
pi(C){2}      0.233198    0.000124    0.233258    751.00    751.00    0.999
pi(G){2}      0.233218    0.000122    0.233088    649.40    700.20    0.999
pi(T){2}      0.265019    0.000118    0.265256    710.85    730.93    1.000
pi(A){3}      0.363814    0.000068    0.363738    695.39    723.20    1.002
pi(C){3}      0.237466    0.000046    0.237286    549.20    605.30    1.000
pi(G){3}      0.185158    0.000053    0.185276    579.42    611.87    1.001
pi(T){3}      0.213562    0.000042    0.213632    570.11    623.77    0.999
pi(A){4}      0.335488    0.000091    0.335634    751.00    751.00    0.999
pi(C){4}      0.367595    0.000072    0.367705    669.46    710.23    0.999
pi(G){4}      0.083074    0.000012    0.082991    751.00    751.00    1.000
pi(T){4}      0.213843    0.000036    0.213932    682.69    716.85    0.999
alpha{1}      0.929507    0.021963    0.916908    651.90    701.45    1.001
alpha{2}      0.845560    0.039927    0.824301    665.83    708.41    1.000
```

```
alpha{3}      0.628006    0.004524    0.624135    643.79    693.45    1.001
alpha{4}      0.858176    0.003148    0.856519    751.00    751.00    0.999
pinvar{1}     0.393908    0.001537    0.396444    611.27    661.28    1.000
pinvar{2}     0.296148    0.004579    0.302666    574.05    628.66    1.003
pinvar{3}     0.558731    0.000429    0.559689    648.29    653.72    1.000
pinvar{4}     0.479202    0.000267    0.479482    751.00    751.00    1.000
m{1}          0.344125    0.000267    0.343594    751.00    751.00    1.001
m{2}          0.437752    0.000812    0.436977    628.68    656.28    1.000
m{3}          0.885944    0.001246    0.886473    581.53    593.41    1.000
m{4}          3.066338    0.009307    3.066057    569.30    660.15    1.000
m{5}          0.545579    0.063024    0.516275    616.93    657.73    0.999
----------------------------------------------------------------------------
* Convergence diagnostic (ESS = Estimated Sample Size); min and avg values
  correspond to minimal and average ESS among runs.
  ESS value below 100 may indicate that the parameter is undersampled.
+ Convergence diagnostic (PSRF = Potential Scale Reduction Factor; Gelman
  and Rubin, 1992) should approach 1.0 as runs converge.
```

C.5 Material examined

See Appendix B

C.6 Accession numbers

Catalog Number	**Species**	***cytb***	***rag1***	***rag2***	***12S***
ANSP 179218	*Brachyplatystoma capapretum*	JF898520	JF898598	JF898745	JF898674
ANSP 179214	*Brachyplatystoma filamentosum*	JF898521	JF898599	JF898746	JF898675
ANSP 179446	*Brachyplatystoma filamentosum*	JF898522	JF898600	JF898747	JF898676
ANSP 187105	*Brachyplatystoma filamentosum*	JF898523	JF898601	JF898748	JF898677
ANSP 179217	*Brachyplatystoma juruense*	JF898516	JF898594	JF898741	JF898670
ANSP 179454	*Brachyplatystoma juruense*	JF898517	JF898595	JF898742	JF898671
ANSP 179230	*Brachyplatystoma platynemum*	JF898515	JF898593	JF898740	JF898669
ANSP 179233	*Brachyplatystoma rousseauxii*	JF898518	JF898596	JF898743	JF898672
ANSP 187109	*Brachyplatystoma rousseauxii*	JF898519	-	JF898744	JF898673
ANSP 179236	*Brachyplatystoma tigrinum*	JF898514	-	JF898739	JF898668
ANSP 178525	*Brachyplatystoma vaillantii*	JF898512	JF898590	JF898737	JF898666
ANSP 187107	*Brachyplatystoma vaillantii*	JF898513	JF898591	JF898738	JF898667
ANSP 179229	*Calophysus macropterus*	JF898528	JF898607	JF898753	JF898682
MBUCV 2001.12.6.L03	*Cheirocerus abuelo*	JF898533	JF898612	JF898758	JF898687
INHS 52717	*Cheirocerus eques*	JF898534	JF898613	JF898759	JF898688
ANSP 179221	*Duopalatinus peruanus*	JF898539	JF898618	JF898763	JF898693
ANSP 179220	*Exallodontus aguanai*	JF898542	JF898621	JF898765	JF898696
MBUCV-VCD#8	*Exallodontus aguanai*	JF898543	JF898622	JF898766	JF898697

Catalog Number	**Species**	***cytb***	***rag1***	***rag2***	***12S***
INHS 49299	*Goeldiella eques*	JF898565	JF898644	DQ492368	JF898719
ANSP 179234	*Hemisorubim platyrhynchos*	-	JF898588	JF898735	JF898664
INHS 52182 (cytb, rag1, 12S) / INHS 5218 (rag2)	*Hypophthalmus* cf. *edentatus*	JF898505	JF898583	DQ492362	JF898661
ANSP 179219	*Hypophthalmus* cf. *marginatus*	JF898504	JF898582	JF898730	JF898660
ANSP 179185	*Hypophthalmus fimbriatus*	-	JF898581	JF898729	JF898659
ANSP 180505	*Iheringichthys labrosus*	JF898550	JF898629	JF898773	JF898704
ANSP 178109	*Leiarius longibarbis*	DQ486766	-	DQ486784	-
ANSP 178109	*Leiarius marmoratus*	-	JF898568	-	JF898647
INHS 54805	*Leiarius perruno*	DQ486767	JF898570	DQ486785	-
ANSP 178108	*Leiarius pictus*	DQ486764	JF898569	DQ486782	JF898648
MZUSP 78457	*Luciopimelodus pati*	JF898529	JF898608	JF898754	JF898683
ANSP 179213	*Megalonema amaxanthum*	JF898535	JF898614	-	JF898689
MZUSP 78465	*Megalonema platanum*	JF898536	JF898615	JF898760	JF898690
ANSP 179458	*Megalonema platycephalum*	JF898537	-	JF898761	JF898691
ANSP 178450	*Megalonema platycephalum*	JF898538	JF898617	JF898762	JF898692
MZUSP 78451	*Parapimelodus nigribarbis*	JF898551	JF898630	JF898774	JF898705
MZUSP 78466	*Parapimelodus valenciennis*	JF898552	JF898631	-	JF898706
ANSP 179554 (cytb, rag1, 12S) / ANSP 179452 (rag2)	*Phractocephalus hemioliopterus*	JF898492	JF898567	DQ486781	JF898646
ANSP 179223	Pimelodidae sp. 1	-	JF898623	JF898767	JF898698
ANSP 179231	Pimelodidae sp. 2	JF898545	JF898624	JF898768	JF898699
ANSP 185247	Pimelodidae sp. 3	JF898546	JF898625	JF898769	JF898700
ANSP 179187	Pimelodidae sp. 4	JF898547	JF898626	JF898770	JF898701
ANSP 179225	*Pimelodina flavipinnis*	JF898527	JF898606	JF898752	JF898681
ANSP 178802	*Pimelodus albicans*	JF898553	JF898632	JF898776	JF898707
ANSP 181017	*Pimelodus argenteus*	JF898557	JF898636	JF898780	JF898711
INHS 49101	*Pimelodus blochii*	JF898559	JF898638	JF898782	JF898713
ANSP 179455	*Pimelodus blochii*	JF898560	JF898639	JF898783	JF898714
INHS 54570	*Pimelodus blochii*	JF898561	JF898640	JF898784	JF898715
ANSP 179456	*Pimelodus blochii*	JF898562	JF898641	JF898785	JF898716
ANSP 179457	*Pimelodus blochii*	JF898563	JF898642	JF898786	JF898717
ANSP 179190	*Pimelodus* cf. *altissimus*	JF898540	JF898619	-	JF898694
ANSP 179226	*Pimelodus* cf. *altissimus*	JF898541	JF898620	JF898764	JF898695
ANSP 179186	*Pimelodus* cf. *blochii*	JF898558	JF898637	JF898781	JF898712
MBUCV 2001.12.6.L10	*Pimelodus coprophagus*	JF898556	JF898635	JF898779	JF898710
ANSP 181109	*Pimelodus maculatus*	JF898554	JF898633	JF898777	JF898708
MZUSP 78452	*Pimelodus maculatus*	JF898555	JF898634	JF898778	JF898709
INHS 49102	*Pimelodus ornatus*	JF898524	JF898602	DQ492363	JF898678
ANSP 178452	*Pimelodus ornatus*	JF898525	JF898603	JF898749	JF898679
AUM 22394	*Pimelodus ornatus*	JF898526	JF898604	JF898750	-
MLP 5131	*Pimelodus ornatus*	-	JF898605	JF898751	JF898680
ANSP 178168	*Pimelodus pictus*	JF898548	JF898627	JF898771	JF898702
ANSP 179448	*Pimelodus pictus*	JF898549	JF898628	JF898772	JF898703
ANSP 178265	*Pinirampus pirinampu*	JF898530	JF898609	JF898755	JF898684
ANSP 180992	*Pinirampus pirinampu*	JF898531	JF898610	JF898756	JF898685
ANSP 181016	*Pinirampus pirinampu*	JF898532	JF898611	JF898757	JF898686
ANSP 179227	*Platynematichthys notatus*	JF898511	JF898589	JF898736	JF898665
MBUCV 2001.12.6.L17	*Platysilurus malarmo*	JF898507	JF898585	JF898732	JF898662
ANSP 179215	*Platysilurus mucosus*	JF898508	JF898586	JF898733	-
AUM 22654	*Platysilurus mucosus*	JF898509	JF898587	JF898734	JF898663
ANSP 179216	*Platystomatichthys sturio*	JF898506	JF898584	JF898731	-

Catalog Number	**Species**	***cytb***	***rag1***	***rag2***	***12S***
MZUSP 78458	*Pseudopimelodus mangurus*	JF898564	JF898643	DQ492360	JF898718
MZUSP uncat. skel	*Pseudoplatystoma corruscans*	JF898500	JF898578	JF898727	JF898657
ANSP 179449	*Pseudoplatystoma fasciatum*	JF898501	JF898579	-	-
ANSP 188882	*Pseudoplatystoma magdaleniatum*	JF898499	JF898577	JF898726	JF898656
ANSP 179450	*Pseudoplatystoma tigrinum*	JF898502	JF898580	JF898728	JF898658
AUM 22581	*Sorubim cuspicaudus*	JF898494	JF898572	JF898721	JF898651
ANSP 178315	*Sorubim elongatus*	JF898497	JF898575	JF898724	JF898654
ANSP 178507	*Sorubim lima*	JF898496	JF898574	JF898723	JF898653
ANSP 178366	*Sorubim maniradii*	JF898495	JF898573	JF898722	JF898652
INHS 54701	*Sorubimichthys planiceps*	JF898498	JF898576	JF898725	JF898655
MZUSP 78463	*Steindachneridion scriptum*	DQ486765	JF898566	DQ486783	JF898645
INHS 43365	*Zungaro zungaro*	JF898493	JF898571	JF898720	JF898650

Bibliography

Aguilera, O.A., Bocquentin, J., Lundberg, J.G., and Maciente, A. 2008. A new cajaro catfish (Siluriformes: Pimelodidae: *Phractocephalus*) from the Late Miocene of southwestern Amazonia and its relationship to *Phractocephalus nassi* of the Urumaco Formation. *Palaeontologische Zeitschrift*, 82:231–245. x, 179, 180, 181

Azpelicueta, M.d.l.M. and Cione, A.L. 2016. A southern species of the tropical catfish genus *Phractocephalus* (Teleostei: Siluriformes) in the Miocene of South America. *Journal of South American Earth Sciences*, 67:221–230. 179, 180, 181

Castresana, J. 2000. Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. *Molecular Biology and Evolution*, 17:540–552. 181

Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, 9:772–772. 182

Drummond, A.J., Suchard, M.A., Xie, D., and Rambaut, A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, 29:1969–1973. 182

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology*, 59:307–321. 182

Katoh, K. and Standley, D.M. 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30:772–780. 181

Lewis, P.O. 2001. A Likelihood Approach to Estimating Phylogeny from Discrete Morphological Character Data. *Systematic Biology*, 50:913–925. 183

Lundberg, J.G. and Aguilera, O.A. 2003. The late Miocene *Phractocephalus* catfish (Siluriformes: Pimelodidae) from Urumaco, Venezuela: additional specimens and reinterpretation as a distinct species. *Neotropical Ichthyology*, 1:97–109. 179, 180, 181

Lundberg, J.G., Sullivan, J.P., and Hardman, M. 2011. Phylogenetics of the South American Catfish family Pimelodidae (Teleostei: Siluriformes) using nuclear and mitochondrial gene

sequences. *Proceedings of the Academy of Natural Sciences of Philadelphia*, 161:153–189. 181

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., and Huelsenbeck, J.P. 2012. MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61:539–542. 182

Talavera, G. and Castresana, J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology*, 56:564–577. 181

Appendix D

Phylogenetic analysis of the Cynodontidae

D.1 Materials and Methods

A parsimony reanalysis of the matrix published in Toledo-Piza (2000) was carried out using TNT v.1.5 (Goloboff et al., 2008) including 11 new characters that are proposed here, that add up to a total of 83 characters including previous 72 characters available in Toledo-Piza (2000). That author excluded some of the 72 characters because they were recovered as uninformative to establish relationships among cynodontids. In contrast, I carried out our phylogenetic analyses including all 83 characters. The roestines were excluded following Mattox and Toledo-Piza (2012) that classified the subfamily as a tribe of the Heterocharacinae of the Characidae. A total of seven species (*Cynodon gibbus*, *C. septenarius*, *Hydrolycus armatus*, *H. scomberoides*, *H. tatauaia*, *H. wallacei*, and *Rhaphiodon vulpinus*) were included in the ingroup, while *Acestrorhynchus* and Toledo-Piza's composite outgroup (i.e., including *Acanthocharax*, *Acestrorhynchus*, *Agoniates*, *Boulengerella*, *Brycon*, *Carnegiella*, *Charax*, *Ctenolucius*, *Erythrinus*, *Galeocharax*, *Gasteropelecus*, *Gnathocharax*, *Hepsetus*, *Heterocharax*, *Hoplerythrinus*, *Hoplias*, *Hoplocharax*, *Hydrocynus*, *Lebiasina*, *Lonchogenys*, *Oligosarcus*, *Pyrrhulina*, *Roeboexodon*, *Roeboides*, and *Xenocharax*) were used as the outgroup. Nodal support values (i.e., bootstrap and jackknife values) were calculated using 10000 replications. Consistency and retention indexes were calculated using the script STATSALL.run v.1.3. TNT is available at http://www.lillo.org.ar/phylogeny/tnt/. I carried out an exact search with the characters unordered and rooting at Toledo-Piza's outgroup composite taxon. Characters of interest were mapped using Mesquite v.3.03, available at http://mesquiteproject.wikispaces.com/ (Maddison and Maddison, 2011). Additional tree formatting was carried out in R v.3.4 (R Core Development Team, 2018), available at http://www.r-project.org/. Code and data are available in the Appendix E.4.

D.2 Results

D.2.1 Character definitions

Character definitions and character matrix. New characters are numbered from 73 to 83, characters 1 to 72 are from Toledo-Piza (2000).

Character 73—Dentary, symphyseal teeth, number: 0=two; 1=one; 2=six to seven

Character 74—Dentary, leading canine, insertion plane: 0=labio-lingual; 1=commisuro-symphyseal; 2=straight, perpendicular to dentary horizontal plane.

Character 75—Dentary, leading canine, texture of lingual surface: 0=smoothly curve; 1=with a sharp and apically distinct cutting edge.

Character 76—Dentary, area posterior to tooth row, accessory teeth postertior to main tooth row: "-"=inapplicable; 0=absent; 1=extensive patch; 2=1-4 teeth in a restricted patch.

Character 77—Dentary-anguloarticular, coronoid process, presence: 0=absent; 1=present.

Character 78—Dentary, lateral sulcus on anterior portion below leading canine, presence: 0=absent; 1=present.

Character 79—Dentary, lateral depresion between leading canine and next hypertrophied canine, presence: 0=absent; 1=present.

Character 80—Dentary, leading canine, base of enameloid, labial surface, texture: 0=smooth, 1=present in the form of parallel/oblique ridges.

Character 81—Dentary, strong canines between leading canine and next hypertrophied canine number: "-"=inapplicable as dentary teeth are of comparable length; 0=1 canine; 1=2 canines.

Character 82—Dentary, anterior margin in lateral view, outline; 0=oblique, 1=straight, 2=round.

Character 83—Dentary, canine between leading canine and next hypertrophied canine, relative position: 0=midway between canines, 1=contiguous to posterior canine.

D.2.2 List of apomorphies

List of synapomorphies for the most-parsimonious tree. Synapomorphies are listed as transformation series in the form 'Character: plesiomorphic state → apomorphic state'. Homoplastic characters are indicated by an asterisk after the transformation series. Node numbers follow their designation in Figure 4.7. *Hydrolycus scomberoides* sensu stricto represents the clade containing the extinct and extant specimens; *Hydrolycus sensu stricto* represents the composition herein proposed with the exclusion of *"H." wallacei*.

Acestrorhynchus:

8: 0 → 1 *

Cynodontidae (node 2):

2: 0 → 1

5: 0 → 1

13: 0 → 1
14: 0 → 1
19: 0 → 1
23: 0 → 1
25: 0 → 1
27: 0 → 1
30: 0 → 1
31: 0 → 1
33: 0 → 1
36: 0 → 1
37: 0 → 1
40: 0 → 1
43: 0 → 1
44: 0 → 1
45: 0 → 1
50: 0 → 1
51: 0 → 1
53: 0 → 1
54: 0 → 1
55: 0 → 1
56: 0 → 1
57: 0 → 1
58: 0 → 1
60: 0 → 1
61: 0 → 1
62: 0 → 1
66: 0 → 1

Cynodon (node 3):

3: 0 → 2
9: 0 → 1
16: 0 → 1
18: 0 → 2
24: 0 → 1
32: 0 → 1
42: 0 → 1
46: 0 → 2 *
49: 0 → 1 *
59: 0 → 1 *
64: 0 → 1
77: 0 → 1

Cynodon gibbus:
No autapomorphies
Cynodon septenarius:
No autapomorphies
Node 4:
11: 0 → 1
15: 0 → 1
26: 0 → 1
28: 0 → 1
"Hydrolycus" wallacei:
4: 0 → 1
6: 0 → 1
81: 1 → 0 *
Node 5:
39: 0 → 1
46: 0 → 1
75: 0 → 1
82: 0 → 1
Hydrolycus s.s. (node 6):
4: 0 → 2
10: 0 → 1
35: 0 → 1
70: 2 → 1
73: 0 → 1
79: 0 → 1
Node 8:
1: 0 → 1
20: 2 → 1
80: 0 → 1
81: 1 → 0 *
82: 1 → 2 *
Hydrolycus armatus:
No autapomorphies
Hydrolycus scomberoides (node 7):
83: 0 → 1
Hydrolycus tatauaia:
No autapomorphies
Rhaphiodon:
3: 0 → 1
18: 0 → 2 *

45: 1 → 3
47: 0 → 1
48: 0 → 1
49: 0 → 1 *
51: 1 → 2
52: 0 → 1
53: 1 → 0
59: 0 → 1 *
62: 1 → 2
65: 1 → 2
67: 0 → 1
69: 0 → 1
71: 2 → 3
78: 0 → 1

D.3 Phylogenetic analysis

Eleven characters were identified in the present study both in fossil and extant specimens, all of them from the morphology of the dentary and leading canines. Despite the incompleteness of the fossil specimens, they were included in the cladistic analysis in order to assess its position among cynodontids.

The phylogenetic analysis recovered a single most parsimonious tree with length 117, CI=0.872 and RI=0.862 (Figure 4.7). A list of the synapomorphies for the different clades is found in Appendix 3. Our topology differs from that of Toledo-Piza (2000) in the position of *"Hydrolycus" wallacei*. I found it to be sister to *Hydrolycus* + *Rhaphiodon* (supported by characters 11, 15, 26, and 28). Furthermore, *Rhaphiodon* was found to be sister to *Hydrolycus* s.s. (characters 39, 46, 75, and 82) instead of *Cynodon* as proposed by Toledo-Piza (2000). The fossil specimen was found to be sister to *Hydrolycus scomberoides* on the basis of one character (character 83). Both, the genus *Hydrolycus* exclusive of *"H." wallacei* and the genus *Cynodon* were found to be monophyletic (Figure 4.7).

D.4 Discussion

The inclusion of the new characters proposed in this study recover a different topology than that of Toledo-Piza (2000). While that author found *"Hydrolycus" wallacei* to be the most basal taxon of the genus *Hydrolycus*, our topology indicates that *"H." wallacei* is sister to a clade composed by *Hydrolycus* + *Raphiodon*. Toledo-Piza also found *Rhaphiodon vulpinus* to be sister to *Cynodon* while I found that it is nested within the genus *Hydrolycus* (Figure 4.7). Toledo-Piza (2000) found *Hydrolycus* (including *H. wallacei*) to be mono-

phyletic based on four synapomorphies (her characters 4, 26, 28, and 54, only the first was non-homoplastic). In contrast I recovered *Hydrolycus* (exclusive of *H. wallacei*) as sister to *Rhaphiodon* based on four synapomorphies (our characters 39, 46, 75, and 82), all of which were non-homoplastic. Two of these synapomorphies were already proposed by Toledo-Piza, while the other two come from the characters herein proposed.

The characters recovered as informative for the position of the fossil among cynodontids are 73, 74, 75, 79 and 83, whose distribution in the most parsimonious cladogram indicate that the derived conditions are synapomorphic for Hydrolycus in characters 73, 74 and 79, synapomorphic for *Hydrolycus* + *Rhaphiodon* in character 75, and autapomorphic for *H. scomberoides* in character 83. Therefore the position of the fossil is well defined within the genus *Hydrolycus* and furthermore indicates it as member of the extant species *H. scomberoides* based on the presence of an autapomorphy and on character states that despite not being autapomorphic, are the same in the examined specimens of the extant species and the fossil.

In our topology *Hydrolycus* is recovered as paraphyletic. A solution could be propose a new genus for *"H." wallacei*, as *Hydrolycus* and *Rhaphiodon* are widely accepted genera, clearly diagnosable on the basis of external morphology. The alternative solution of synonymizing *Hydrolycus* and *Rhaphiodon*, would be far more problematic as it would produce many nomenclature changes. Although *"Hydrolycus" wallacei* is clearly diagnosable from other cynodontids, I prefer not to propose a new genus because there is a the low support for the node *Hydrolycus* + *Rhaphiodon* and *"H." wallacei* (bootstrap = 16). Additional data, including molecular markers are needed to test the topology I propose herein.

Bibliography

Goloboff, P.A., Farris, J.S., and Nixon, K.C. 2008. TNT, a free program for phylogenetic analysis. *Cladistics*, 24:774–786. 192

Maddison, W.P. and Maddison, D.R. 2011. Mesquite: a modular system for evolutionary analysis. 192

Mattox, G.M.T. and Toledo-Piza, M. 2012. Phylogenetic study of the Characinae (Teleostei: Characiformes: Characidae). *Zoological Journal of the Linnean Society*, 165:809–915. 192

R Core Development Team. 2018. R: A language and environment for statistical computing. 192

Toledo-Piza, M. 2000. The Neotropical fish subfamily Cynodontinae (Teleostei: Ostariophysi: Characiformes): A phylogenetic study and a revision of *Cynodon* and *Rhaphiodon*. *American Museum Novitates*, 3286:1–88. 192, 193, 196, 197

Appendix E

Computer code

E.1 Terminology and anatomy of Siluriform spines

E.1.1 Cost function for choosing optimal terminologies

```
### Load packages
library(magrittr)
library(dplyr)

### Read the dataset
dataset <- read.delim(file = "nonShaft.tab", stringsAsFactors = FALSE)

### Summarize data in tables
tables <- apply(X = dataset, MARGIN = 2, FUN = table)

### Construct the different combinations corresponding to different terminology systems.
### tic-toc indicates ca. 0.05 secs running
termSystems <- vector(length = length(tables$Spine) * length(tables$Structure) * length(tables$Term))

tic <- Sys.time()
counter <- 1
for (i in names(tables$Spine)) {
    for (j in names(tables$Structure)) {
        for (k in names(tables$Term)) {
            termSystems[counter] <- paste(i, j, k, sep = ", ")
            counter <- counter + 1
        }
    }
}
toc <- Sys.time()
toc - tic

### Calculate cost for each alternative term
freqs <- vector()
totals <- vector()
structs <- vector()
termins <- vector()

counter <- 1
for (i in names(tables$Structure)) {
    terms <- table(dataset[dataset$Structure == i, "Term"])
    for (j in seq_along(names(terms))) {
        freqs <- c(freqs, unlist(terms)[j] %>% unname)
```

```
        totals <- c(totals, sum(unlist(terms)))
        structs <- c(structs, i)
        termins <- c(termins, names(terms)[j])
        cat("Structure ", i, ", term ", names(terms)[j], ", frequency = ",
            freqs[counter], ", against ", totals[counter], "\n", sep = "")
        counter <- counter + 1
    }
}

output <- data.frame(Structure = structs, Terms = termins, Frequency = freqs,
                     Total = totals, stringsAsFactors = FALSE)

output <- mutate(output, indCost = Total - Frequency)

### Determine argmin{C(t)}
argmins <- aggregate(x = output$indCost, by = list(output$Structure, output$Terms), FUN = min)

aggregate(indCost ~ Structure, data = output, min)

### identify optimal terms. This still needs to be done visually because there could
### be more than one optimal term
output[order(output$Structure, output$indCost),] %>% View

### Write this table for decision making
write.table(x = output, file = "optimalTerms.tab", quote = FALSE, sep = "\t", row.names = FALSE)
```

E.2 Phylogenetic context of the *Phractocephalus* occurrences form the Sincelejo Formation

E.2.1 `jmodeltest_script.sh`

```
#!/usr/bin/bash

for i in $(ls .fasta)
do
    echo "Running jModelTest2 on $i\n"
    java -jar /bioprograms/jmodeltest2/dist/jModelTest.jar -BIC -d $i -f -i -g
                                                           -s 11 -t ML -tr 22
                                                           -o $i`date +"%H.%M_%d_%m_%Y"`.out'
done
```

E.2.2 `phylogeny_mrbayes.mb`

This script runs MrBayes and specifies the bayesian phylogenetic model as well as the heuristic Markov Chain Monte Carlo process. This script was jointly developed with Sandra Reinales.

```
begin mrbayes;
[Script documentation carried out using comments]

[log the analysis]
log start filename = phylogeny_mrbayes.log;
[read the matrix all_markers_ed_nom.nex]
execute all_markers_ed_nom.nex;
```

```
[close analysis at end]
set autoclose = yes;
[set Goeldiella_eques as outgroup]
outgroup Goeldiella_eques;
[This command shows the status of all the taxa, according to the documentation]
taxastat;

[definition of individual partitions per marker]
charset rag1 = 1-2676;
charset rag2 = 2677-3607;
charset 12S = 3608-6316;
charset cytb = 6317-7570;
charset morphology = 7571-7578;

[definition of combined dataset]
partition combined = 5: rag1, rag2, 12S, cytb, morphology;

[specification of substitution models and morphological model]
set partition = combined;
lset applyto = (1) nst = 6 rates = invgamma; [GTR+I+G]
lset applyto = (2) nst = 2 rates = invgamma; [K80+I+G = K2P+I+G]
lset applyto = (3) nst = 6 rates = invgamma; [GTR+I+G]
lset applyto = (4) nst = 2 rates = invgamma; [HKY+I+G]
lset applyto = (5) coding = variable; [Lewis2001 model]

[unlink parameters across partitions]
unlink shape = (all) pinvar = (all) statefreq = (all) revmat = (all) tratio = (all);

[allow separate gamma parameters for each partition]
prset applyto=(all) ratepr=variable;

[show the model just specified for each partition]
showmodel;

[set up the MCMC, with this setting the analysis will need not less than 16 threads]
mcmcp nruns = 2 ngen = 4000000 nchains = 8 samplefreq = 4000 printfreq = 100;
[run the MCMC]
mcmc;

[summarize the posterior trees]
sumt nruns = 2 relburnin = yes burninfrac = 0.25;
plot;

[summarize parameter posteriors]
sump;

log stop;
end;
```

E.2.3 `beginExec.R` and `finishExec.R`

These two scripts are executed bracketing the main MrBayes script and send an email to a given recipient address (`"recipient_email"`) from a given sender (`"sender_email"`). Please note that the password for the sender account must be in the environment (the `.Renviron` file) and saved as `PASSWD` so that the script securely has access to the password.

```
# beginExec.R
```

```
library(mailR)
sender <- "sender_email@mail.com"
recipient <- "recipient_email@mail.com"
notification <- send.mail(from = sender,
to = recipient,
subject = "Started execution",
Sys.Date(), "{}",
body = "MrBayes execution on 16 processors for the Phractocephalus analysis",
encoding = "utf-8",
smtp = list(host.name = "smtp.gmail.com", port = 465, user.name = "sender_email@mail.com",
            passwd = Sys.getenv("PASSWD"), ssl = TRUE),
authenticate = TRUE,
attach.files = NULL,
send = TRUE,
html = TRUE,
inline = TRUE, debug = FALSE
)
# finishExec.R
library(mailR)
sender <- "sender_email@mail.com"
recipient <- "recipient_email@mail.com"
notification <- send.mail(from = sender,
to = recipient,
subject = "Finished execution",
Sys.Date(), "{}",
body = "MrBayes execution on 16 processors for the Phractocephalus analysis",
encoding = "utf-8",
smtp = list(host.name = "smtp.gmail.com", port = 465, user.name = "sender_email@mail.com",
            passwd = Sys.getenv("PASSWD"), ssl = TRUE),
authenticate = TRUE,
attach.files = NULL,
send = TRUE,
html = TRUE,
inline = TRUE, debug = FALSE
)
```

E.2.4 `run_mrbayes_analysis.sh`

```
#!/bin/bash
# dependencies: mpirun (openmpi), mrbayes compiled for mpi support AND beagle
# hardware: a machine with at least 16 processors. Modify the .mb script in
# order to use less that that amount
Rscript beginExec.R # R script for sending an email when the analysis starts
time mpirun -n 16 mb *.mb # it is assumed that mb has mpi-capabilities and uses beagle
Rscript finishExec.R # R script for sending an email when the analysis finishes
```

E.3 Middle miocene fossil fishes of the Guajira Peninsula

E.3.1 Analysis of faunal similarity

`neogeneFishOccs.tab`

In order to reconstruct the tabular structure of the dataset, search and replace four spaces and newline (`M-S-%` search `Î` and replace with nothing in Emacs). Field separators (semicolons `;`) should be replaced to tabulations with the same approach.

Occurrences compiled from the following sources: Aguilera et al. (2013a,b); Antoine et al. (2016); Azpelicueta and Cione (2016); Ballen and Moreno-Bernal (2019); Bogan et al. (2012); Cione and Azpelicueta (2013); Cione et al. (2000, 2009); Lundberg et al. (2010); Tejada-Lara et al. (2015).

```
Family;Taxon;Environment;Castillo;La Venta;Makaraipao;Castilletes marine;
    Loyola Mangan;Rio Acre;Solimoes Pebas;Urumaco;Utuquina;Pirabas;Cantaure;
    Ituzaingo;Rio Yuca;Fitzcarrald;Contamana;Reference
Acregoliathidae;Acregoliath;Freshwater;0;1;0;0;0;1;0;0;0;0;0;0;0;1;0;
    \citet{Ballen2019a,Lundberg2010b,Tejada-Lara2015}
Arapaimidae;Arapaima;Freshwater;0;1;0;0;0;1;0;0;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Anostomidae;Leporinus;Freshwater;0;1;0;0;1;0;1;0;0;0;0;1;0;0;1;
    \citet{Antoine2016,Bogan2012,Lundberg2010b}
Characidae;Salminus;Freshwater;0;0;0;0;0;0;0;0;0;0;0;1;0;0;0;\citet{Cione2013}
Serrasalmidae;Colossoma;Freshwater;1;1;0;0;0;1;0;1;0;0;0;1;0;0;0;
    \citet{Cione2000,Lundberg2010b}
Serrasalmidae;Megapiranha;Freshwater;0;0;0;0;0;0;0;0;0;0;0;1;0;0;0;\citet{Cione2009}
Serrasalmidae;Mylossoma;Freshwater;1;0;1;0;0;0;0;0;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Serrasalmidae;Piaractus;Freshwater;0;0;1;0;0;0;0;1;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Cynodontidae;Hydrolycus;Freshwater;0;1;0;0;0;0;0;0;0;0;0;1;0;1;1;
    \citet{Antoine2016,Cione2000,Lundberg2010b,Tejada-Lara2015}
Erythrinidae;Paleohoplias;Freshwater;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Erythrinidae;Hoplias;Freshwater;0;1;0;0;1;0;1;0;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Callichthyidae;Hoplosternum;Freshwater;0;1;0;0;0;1;0;0;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Loricariidae;Acanthicus;Freshwater;0;1;0;0;0;0;0;1;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Doradidae;Doraops;Freshwater;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Doradidae;Doras;Freshwater;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Doradidae;Oxydoras;Freshwater;0;0;0;0;0;0;0;0;1;0;0;0;0;0;0;\citet{Lundberg2010b}
Doradidae;Rhinodoras;Freshwater;0;0;0;0;0;0;0;0;1;0;0;0;0;0;0;\citet{Lundberg2010b}
Pimelodidae;Brachyplatystoma;Freshwater;0;1;0;0;0;0;0;1;0;0;0;0;0;0;0;
    \citet{Aguilera2013d,Lundberg2010b}
Pimelodidae;Phractocephalus;Freshwater;0;1;1;0;0;1;0;1;0;0;0;1;1;1;1;
    \citet{Antoine2016,Azpelicueta2016,Lundberg2010b,Tejada-Lara2015}
Pimelodidae;Platysilurus;Freshwater;0;0;0;0;0;0;0;1;0;0;0;0;1;0;0;\citet{Lundberg2010b}
Pimelodidae;Zungaro;Freshwater;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Ariidae;Amphiarius;Marine;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Ariidae;Aspistor;Both;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Ariidae;Bagre;Marine;1;0;0;1;0;0;0;1;0;0;1;0;0;0;0;\citet{Aguilera2013a,Lundberg2010b}
Ariidae;Cathorops;Marine;0;0;0;0;0;0;0;0;0;1;0;0;0;0;0;\citet{Aguilera2013a}
Ariidae;Cantarius;Marine;0;0;0;1;0;0;0;0;0;0;1;0;0;0;0;\citet{Aguilera2013a}
Ariidae;Notarius;Marine;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Ariidae;Sciades;Both;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Sciaenidae;Ctenosciaena;Marine;1;0;0;0;0;0;0;0;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Sciaenidae;Cynoscion;Both;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Sciaenidae;Equetus;Marine;1;0;0;0;0;0;0;1;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Sciaenidae;Larimus;Marine;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Sciaenidae;Micropogonias;Marine;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Sciaenidae;Nebris;Marine;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Sciaenidae;Ophioscion;Marine;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Sciaenidae;Pachypops;Freshwater;0;0;0;0;0;0;1;0;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Sciaenidae;Paralonchurus;Marine;1;0;0;0;0;0;0;0;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Sciaenidae;Plagioscion;Freshwater;1;0;0;0;0;0;1;1;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Sciaenidae;Protosciaena;Marine;1;0;0;0;0;0;0;0;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Sciaenidae;Xenotolithus;NA;0;0;0;0;0;0;1;0;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Serranidae;Epinephelus;Marine;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Sphyraenidae;Sphyraena;Marine;1;0;0;0;0;0;0;0;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Scombridae;Acanthocybium;Marine;1;0;0;0;0;0;0;0;0;0;0;0;0;0;0;\citet{Lundberg2010b}
Lepidosirenidae;Lepidosiren;Freshwater;0;1;1;0;0;1;0;0;0;0;0;0;0;1;1;
    \citet{Antoine2016,Lundberg2010b,Tejada-Lara2015}
```

similarity.R

```
library(vegan)

### faunal similarity analyses for the Marakaipao fish fauna
dataset <- read.delim(file = "neogeneFishOccs.tab", stringsAsFactors = FALSE)

#################### only freswater taxa
dataset <- dataset[which(dataset$Environment == "Freshwater" | dataset$Environment == "Both"), ]

### Solimões-Pebas does not seem to be a fauna but a collection of different faunas across
### the Amazon, remove it
dataset <- dataset[, -grep(pattern = "Pebas", x = colnames(dataset))]

comMatrix <- dataset[, -c(1, 3, 18)]
rownames(comMatrix) <- comMatrix$Taxon
comMatrix <- comMatrix[-1]
comMatrix <- t(comMatrix)

### number of occurrences per fauna
sort(x = apply(X = comMatrix, MARGIN = 2, FUN = sum), decreasing = TRUE)

### include only those faunas with at least the number of occurrences
### of Makaraipao
selectFaunas <- names(which(apply(X = comMatrix, MARGIN = 1, FUN = sum) >=
                            apply(X = comMatrix, MARGIN = 1, FUN = sum)["Makaraipao"]))
comMatrix <- comMatrix[selectFaunas, ]

### remove zero-sum species after faunal selection
selectSpp <- names(which(apply(X = comMatrix, MARGIN = 2, FUN = sum) > 0))
comMatrix <- comMatrix[, selectSpp]

### calculate the distance matrix using Bray-Curtis' method
distMatrixBray <- vegan::vegdist(comMatrix, method = "bray", binary = TRUE)
# rename labels in order to replace dots with spaces
attr(distMatrixBray, "Labels") <- gsub(pattern = "\\.", replacement = " ",
                                       x = attr(distMatrixBray, "Labels"))
# rename trans-Andean labels in order to place a leading asterisk
attr(distMatrixBray, "Labels") <- gsub(pattern = "Urumaco", replacement = "( T ) Urumaco",
                                       x = attr(distMatrixBray, "Labels"), fixed = TRUE)
attr(distMatrixBray, "Labels") <- gsub(pattern = "La Venta", replacement = "( T ) La Venta",
                                       x = attr(distMatrixBray, "Labels"), fixed = TRUE)
attr(distMatrixBray, "Labels") <- gsub(pattern = "Makaraipao", replacement = "( T ) Makaraipao",
                                       x = attr(distMatrixBray, "Labels"), fixed = TRUE)
### plot the dendrogram
pdf(file = "faunalSim.pdf")
plot(hclust(as.dist(distMatrixBray), method = "average"), main = "Faunal similarity",
     sub = NA, xlab = NA, ylab = "Bray-Curtis coefficient")
dev.off()
```

E.4 Phylogenetic analysis of the Cynodontidae

E.4.1 Readme

Dependencies

Please note that in order to run these scripts and reproduce the results you will need:

- TNT v.1.5 or higher for phylogenetic analysis, support statistics calculation, and tree production. Freely available at http://www.lillo.org.ar/phylogeny/tnt/. Please keep in mind that TNT starts numbering from 0, see below.
- R v.3.3.1 or higher for tree format conversion. Freely available at https://www.r-project.org/.
- (Optional) FigTree v.1.4.2 or higher for tree edition
- (Optional) Inkscape 0.91 or higher for fine graphical edition of the trees.

Fast run

On a UNIX system just run the `automaticPipeline` script. This can be achieved by running the command `bash automaticPipeline` or `bash ./automaticPipeline`. This script will return no visual output to the screen by itself but will show some internal TNT messages during bootstraping. These will disappear from the screen after running. The tree file format conversion is completely silent. After running you will have the output files described below. This alternative is designed for fast runs without going through each individual file.

Further details and comments are found in each script for those interested.

Input files

These files must be run following these instructions. These are specific for command-line TNT:

- Navigate to the path where the files are located,`cd PATH` for Unix users.
- Run TNT on the background with the command`tnt bground p complete.run`. Please note that the script was designed as a stand-alone tool so once called from TNT it will exit once the analysis has finished. See the file`complete.run` file for further comments and detailed explanation of each step in the phylogenetic analysis and construction of the output tree in .tre format.
- If there are already output files these will be rewritten by TNT with the respective warning being recorded in the`complete.out` output file. Otherwise such file should not contain any warning.

`complete.tnt`

This file contains the data matrix in tnt format. The penultimate line of the file contains the instruction`collapse [;` in order to collapse branches supported by no apomorphies, contrary to the default in TNT that always presents a dichotomous result.

`complete.run`

Script for carrying out the analysis, it contains the following steps as already documented inside the script:

- Tell TNT to use 1 Gb of RAM and store a large number of tree in memory.
- Save all output to`complete.out`
- Read the matrix`complete.tnt`
- Initialize the output tree file`complete.tre`
- Set outgroup to`Outgroup`
- Use implicit enumeration for exact search
- Plot the most-parsimonious tree (in this case, there was only one MPT) along with node numbers
- Initialize tree annotation
- Calculate CI and RI using the`STATSALL.RUN` script
- Calculate bootstrap statistics
- Calculate jackknife statistics
- Save annotations to tree file in parenthetical notation
- Plot MPT to log file
- Close tree file
- List synapomorphies for each clade. Please note that TNT starts numbering from 0, so you will need to edit the synapomorphies list by adding 1 to all node AND character numbers (e.g., Node 0 is actually node 1, and character 76 is actually character 77).
- Close log file
- Close TNT

`STATSALL.RUN`

This script was developed by Peterson Lopes (Universidade de São Paulo) and is available in a number of sources over the internet. For reproducibility purporses I am linking to a thread in the TNT user group on google where such file can be found. Unfortunately there seems to be no official source for this script since the TNT wiki website became unavailable a couple years ago. The script can be found here https://groups.google.com/d/msg/tnt-tree-analysis-using-new-technology/qPdCzlk_at8/YusQvlXCahwJ. I claim NO AUTHORSHIP for this script. Please contact Peterson Lopes directly for further information.

Output files

The script `complete.run` generates the following files as already noted. Tree edition was done in FigTree (http://tree.bio.ed.ac.uk/software/figtree/) and fine-edited with Inkscape.

`complete.out`

This file contains every output generated during analysis including graphic output such as trees and statistics. It will also contain any warning generated during analysis. Further information aside from graphic cladograms can be found here.

`complete.tre`

This file contains the output tree file to be converted to newick format in R (see below). Without any reformatting, this file can only be opened by TNT (and maybe Mesquite, did not test as it was unnecessary; feedback on this is greatly acknowledged).

Scripts for format conversion

An additional step is required before the `complete.tre` file can be edited in FigTree, given that TNT's output tree format differs from a newick standard format. An R script was designed for carrying out format conversion automatically.

In order to convert the file you need to load the function file with `source("treToNewick.R")` in the R console. Afterwards you are ready for converting between to newick format with the `treToNewick` function. Finally, you can open the .newick file in FigTree. Please keep in mind that the tree file has support values as annotations, so when opening FigTree you will be asked to name the labels (the defaul name labels is enough) so that you can activate them in the 'labels' section of FigTree.

`treToNewick.R`

This script contains the function for format conversion from TNT .tre to standard .newick. The function has three arguments:

- `file`: Character. The input file containing the .tre file as exported by TNT. Name this file including the extension (e.g., "example.tre").
- `output`: Character. The name for the output file. Use the .newick extension for this file (e.g., "output.newick").
- `subsetting`: Logical. Optional argument for fast file conversion. This argument subsets the tree in order to remove the TNT header and final lines. It should speed a bit the conversion for very large tree files. Defaults to TRUE.

Final steps

After obtaining a .newick file you will be ready to edit the tree in FigTree in order to add annotations, change fonts, colors, and some other basic edition tasks. For a fully edited

and good-looking tree it is suggested to save the tree in the vectorial svg format and then to further edit it with Inkscape. This way the cladogram published in the paper can be obtained.

E.4.2 `complete.run`

```
macro =;

/* Set max RAM to 1 Gb*/
mxram 1000;

/* Collapse all unsupported branches*/
collapse [;

/*
Save all the output to the file 'complete.out'
It can be opened in any text editor
 */
log complete.out;

/*
Read the matrix 'complete.tnt'
*/
proc complete.tnt;

/*
Number of trees to be held in memory
*/
hold 1000000;

/*
Set random seed
*/
rseed 0;

/*
Open tree file in parenthetical notation
*/
tsave * complete.tre;

/*
Set the composite taxon 'Outgroup' as the root
*/
outgroup Outgroup;

/*
Given the number of taxa, carry out an exact search
with implicit enumeration
*/
ienum;

/*
Taxon names ON
*/
taxname =;

/*
Plot MPT with node numbers
```

```
*/
naked -;
tplot /;
naked =;

/*
Tree tags ON
*/
ttags =;

/*
Overall CI and RI
*/
run STATSALL.RUN 1;

/*
Bootstrap and Jackknife branch support values
*/
resample boot rep 10000;
resample jak rep 10000;

/*
Save tags to tree file in parenthetical notation
*/
save *;

/*
Plot consensus tree to log file
*/
tplot /;

/*
Close tree files
*/
tsave /;

/*
List synapomorphies common to all trees
*/
apo -;

/*
Close log file
*/
log /;

/*
Close TNT
*/
quit;
```

E.4.3 `treToNewick.R`

```
# Script for carrying out format conversion from .tre to .newick
# There are two approaches, replacements with grep-like native R
# functions or using sed from a system call.
# Maybe it would be interesting to explore both approachs
# There is even an easier approach for the first search and is using subsetting
# However subsetting seems risky as perhaps my .tre files are formatted in a different way as
```

```
# those from other analyses.

treToNewick <- function(file, output, subsetting = TRUE){
    tree <- readLines(file)
    if(subsetting) {
        tree <- tree[-c(1,length(tree))]
    } else {
        tree <- gsub(pattern = "tread.*", replacement = "", x = tree, ignore.case = TRUE)
        tree <- gsub(pattern = "proc-;.*", replacement = "", x = tree, ignore.case = TRUE)
        tree <- gsub(pattern = "proc/.*", replacement = "", x = tree, ignore.case = TRUE)
    }

    tree <- gsub(pattern = "*", replacement = ";", x = tree, fixed = TRUE)
    tree <- gsub(pattern = " ", replacement = ",", x = tree, fixed = TRUE)
    tree <- gsub(pattern = ")(", replacement = "),(", x = tree, fixed = TRUE)
    tree <- gsub(pattern = ",)", replacement = ")", x = tree, fixed = TRUE)
    tree <- gsub(pattern = "_", replacement = " ", x = tree, fixed = TRUE)
    tree <- gsub(pattern = ",=", replacement = ":", x = tree, fixed = TRUE)
    tree <- gsub(pattern = "=", replacement = ":", x = tree, fixed = TRUE)
    tree <- gsub(pattern = " =", replacement = ":", x = tree, fixed = TRUE)
    tree <- gsub(pattern = "=", replacement = ":", x = tree, fixed = TRUE)
    tree <- gsub(pattern = ":/", replacement = "", x = tree, fixed = TRUE)
    tree <- gsub(pattern = ":", replacement = " ", x = tree, fixed = TRUE)
    if(length(which(tree == "")) == 0) {
        writeLines(text = tree, con = output)
    } else {
        tree <- tree[-which(tree == "")]
        writeLines(text = tree, con = output)
    }

}
```

E.4.4 `complete.tnt`

In order to reconstruct the structure of the dataset, cut the character string (e.g., the line with `???? ... 111` just below `Hydrolycus_sp._Fossil`) and paste it after the previous line that contains the taxon name (e.g., `Hydrolycus_sp._Fossil` in this case). The line should then contain the structure: `taxon_name` and `character_string` separated by a space (e.g., `Hydrolycus_armatus 110210 ... 020`. The matrix includes the characters originally proposed by Toledo-Piza (2000) and the new ones herein proposed (See Appendix D).

```
xread
83 10
Hydrolycus_sp._Fossil
??????????????????????????????????????????????????????????????????????????111??010111
Hydrolycus_armatus
110210000111111000110010111111101111111111011110001101111110111001100012111110011020
Hydrolycus_tatauaia
110210000111111000110010111111101111111111011110001101111110111001100012111110011020
Hydrolycus_wallacei
010111?000111110001201101111111011011101101110000110111111011100110002210002000000 0
Hydrolycus_scomberoides
010210010111111011121410111111101111111111011220001101111110111101101012111100010111
Rhaphiodon_vulpinus
011010000011111?012121310111111101101111110113111112110111111120021101231021201001 10
```

```
Cynodon_gibbus
0120101010001101121212111010111111011101211112001110111111111110111000210200-1000100
Cynodon_septenarius
012010101000110112121211101011111101110121111200111011111111111011100???200-1000100
Acestrorhynchus
00000??1000?0000000??1000000?000010001000000000000000000000000000?0000???-0000000-20
Outgroup
00000??0000?000000000000000000000000000?000000000000000000000000000000???-2000000---
;
collapse [;
proc /;
```

E.4.5 `automaticPipeline.sh`

```
#!/bin/bash

# run the TNT script 'complete.run' on the background. This will carry out the phylogenetic
# analysis and draw the annotated cladogram. Please refer to the file 'complete.run' for
# further details and comments.
tnt bground p complete.run

# Load the 'treToNewick' function in R and convert the 'complete.tre' output file to the
# newick format ('complete.newick'). Afterwards just open it in figtree or your preferred
# program for tree edition.
Rscript -e 'source("treToNewick.R")' -e 'treToNewick("complete.tre",
                                                     "complete.newick",
                                                     subsetting = FALSE)'
```

E.4.6 Analysis of faunal similarity for the Ware fauna

`neogeneFishOccs_withWare.tab`

In order to reconstruct the tabular structure of the dataset, search and replace four spaces and newline (`M-S-%` search `Î ` and replace with nothing in Emacs). Field separators (semicolons `;`) should be replaced to tabulations with the same approach.

Occurrences compiled from the following sources: Aguilera et al. (2013a,b); Antoine et al. (2016); Azpelicueta and Cione (2016); Ballen and Moreno-Bernal (2019); Bogan et al. (2012); Cione and Azpelicueta (2013); Cione et al. (2000, 2009); Lundberg et al. (2010); Tejada-Lara et al. (2015).

```
Family;Taxon;Environment;Castillo;La Venta;Makaraipao;Castilletes marine;
    Loyola Mangan;Rio Acre;Solimoes Pebas;Urumaco;Utuquina;Pirabas;Cantaure;
    Ituzaingo;Rio Yuca;Fitzcarrald;Contamana;Sincelejo;Ware;Reference
Acregoliathidae;Acregoliath;Freshwater;0;1;0;0;0;1;0;0;0;0;0;0;0;1;0;0;0;
    \citet{Ballen2019a,Lundberg2010b,Tejada-Lara2015}
Arapaimidae;Arapaima;Freshwater;0;1;0;0;0;1;0;0;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Anostomidae;Leporinus;Freshwater;0;1;0;0;1;0;1;0;0;0;0;1;0;0;1;1;0;
    \citet{Antoine2016,Bogan2012,Lundberg2010b}
Characidae;Salminus;Freshwater;0;0;0;0;0;0;0;0;0;0;0;1;0;0;0;0;0;
    \citet{Cione2013}
Serrasalmidae;Colossoma;Freshwater;1;1;0;0;0;1;0;1;0;0;0;1;0;0;0;0;0;
    \citet{Cione2000,Lundberg2010b}
Serrasalmidae;Megapiranha;Freshwater;0;0;0;0;0;0;0;0;0;0;0;1;0;0;0;0;0;
    \citet{Cione2009}
```

```
Serrasalmidae;Mylossoma;Freshwater;1;0;1;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Serrasalmidae;Piaractus;Freshwater;0;0;1;0;0;0;0;1;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Serrasalmidae;Pygtocentrus;Freshwater;0;1;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
    This study
Serrasalmidae;Serrasalmus;Freshwater;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;1;
    This study
Cynodontidae;Rhaphiodon;Freshwater;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;1;
    This study
Cynodontidae;Hydrolycus;Freshwater;0;1;0;0;0;0;0;0;0;0;0;1;0;1;1;0;1;
    \citet{Antoine2016,Cione2000,Lundberg2010b,Tejada-Lara2015}, this study
Erythrinidae;Paleohoplias;Freshwater;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Erythrinidae;Hoplias;Freshwater;0;1;0;0;1;0;1;0;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Auchenipteridae;Trachelyopterichthys;Freshwater;0;0;0;0;0;0;0;0;0;0;0;0;0;
    0;0;0;1;This study
Callichthyidae;Hoplosternum;Freshwater;0;1;0;0;0;1;0;0;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Loricariidae;Acanthicus;Freshwater;0;1;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Doradidae;Doraops;Freshwater;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Doradidae;Doras;Freshwater;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Doradidae;Hemidoras;Freshwater;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;1;
    This study
Doradidae;Oxydoras;Freshwater;0;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;1;
    \citet{Lundberg2010b}, this study
Doradidae;Rhinodoras;Freshwater;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Pimelodidae;Brachyplatystoma;Freshwater;0;1;0;0;0;0;0;1;0;0;0;0;0;0;0;
    0;0;1;\citet{Aguilera2013d,Lundberg2010b}, this study
Pimelodidae;Phractocephalus;Freshwater;0;1;1;0;0;1;0;1;0;0;0;1;1;1;1;
    1;0;\citet{Antoine2016,Azpelicueta2016,Lundberg2010b,Tejada-Lara2015}
Pimelodidae;Platysilurus;Freshwater;0;0;0;0;0;0;0;1;0;0;0;0;1;0;0;0;1;
    \citet{Lundberg2010b}, this study
Pimelodidae;Zungaro;Freshwater;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;0;1;
    \citet{Lundberg2010b}, this study
Ariidae;Amphiarius;Marine;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Ariidae;Aspistor;Both;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Ariidae;Bagre;Marine;1;0;0;1;0;0;0;1;0;0;1;0;0;0;0;0;0;
    \citet{Aguilera2013a,Lundberg2010b}
Ariidae;Cathorops;Marine;0;0;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;
    \citet{Aguilera2013a}
Ariidae;Cantarius;Marine;0;0;0;1;0;0;0;0;0;0;1;0;0;0;0;0;0;
    \citet{Aguilera2013a}
Ariidae;Notarius;Marine;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Ariidae;Sciades;Both;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Sciaenidae;Ctenosciaena;Marine;1;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Sciaenidae;Cynoscion;Both;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Sciaenidae;Equetus;Marine;1;0;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
```

```
Sciaenidae;Larimus;Marine;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Sciaenidae;Micropogonias;Marine;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Sciaenidae;Nebris;Marine;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Sciaenidae;Ophioscion;Marine;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Sciaenidae;Pachypops;Freshwater;0;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Sciaenidae;Paralonchurus;Marine;1;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Sciaenidae;Plagioscion;Freshwater;1;0;0;0;0;0;1;1;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Sciaenidae;Protosciaena;Marine;1;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Sciaenidae;Xenotolithus;NA;0;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Serranidae;Epinephelus;Marine;0;0;0;0;0;0;0;1;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Sphyraenidae;Sphyraena;Marine;1;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Scombridae;Acanthocybium;Marine;1;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
    \citet{Lundberg2010b}
Lepidosirenidae;Lepidosiren;Freshwater;0;1;1;0;0;1;0;0;0;0;0;0;0;1;1;0;0;
    \citet{Antoine2016,Lundberg2010b,Tejada-Lara2015}
```

similarity.R

```
library(vegan)

### faunal similarity analyses for the Marakaipao fish fauna
dataset <- read.delim(file = "neogeneFishOccs_withWare.tab", stringsAsFactors = FALSE)

#################### only freswater taxa
dataset <- dataset[which(dataset$Environment == "Freshwater" | dataset$Environment == "Both"), ]

### Solimões-Pebas does not seem to be a fauna but a collection of different faunas across
### the Amazon, remove it
dataset <- dataset[, -grep(pattern = "Pebas", x = colnames(dataset))]

# remove the columns family, environment, and references
comMatrix <- dataset[, -c(1, 3, 20)]
rownames(comMatrix) <- comMatrix$Taxon
comMatrix <- comMatrix[-1]
comMatrix <- t(comMatrix)

### number of occurrences per fauna
sort(x = apply(X = comMatrix, MARGIN = 2, FUN = sum), decreasing = TRUE)

### include only those faunas with at least the number of occurrences
### of Makaraipao
selectFaunas <- names(which(apply(X = comMatrix, MARGIN = 1, FUN = sum) >=
                            apply(X = comMatrix, MARGIN = 1, FUN = sum)["Makaraipao"]))
comMatrix <- comMatrix[selectFaunas, ]

### remove zero-sum species after faunal selection
selectSpp <- names(which(apply(X = comMatrix, MARGIN = 2, FUN = sum) > 0))
comMatrix <- comMatrix[, selectSpp]
```

```
### calculate the distance matrix using Bray-Curtis' method
distMatrixBray <- vegan::vegdist(comMatrix, method = "bray", binary = TRUE)
# rename labels in order to replace dots with spaces
attr(distMatrixBray, "Labels") <- gsub(pattern = "\\.", replacement = " ",
                                       x = attr(distMatrixBray, "Labels"))
# rename trans-Andean labels in order to place a leading asterisk
attr(distMatrixBray, "Labels") <- gsub(pattern = "Urumaco", replacement = "( T ) Urumaco",
                                       x = attr(distMatrixBray, "Labels"), fixed = TRUE)
attr(distMatrixBray, "Labels") <- gsub(pattern = "La Venta", replacement = "( T ) La Venta",
                                       x = attr(distMatrixBray, "Labels"), fixed = TRUE)
attr(distMatrixBray, "Labels") <- gsub(pattern = "Makaraipao", replacement = "( T ) Makaraipao",
                                       x = attr(distMatrixBray, "Labels"), fixed = TRUE)
attr(distMatrixBray, "Labels") <- gsub(pattern = "Ware", replacement = "( T ) Ware",
                                       x = attr(distMatrixBray, "Labels"), fixed = TRUE)

### plot the dendrogram
pdf(file = "faunalSim_withWare.pdf")
plot(hclust(as.dist(distMatrixBray), method = "average"), main = "Faunal similarity",
     sub = NA, xlab = NA, ylab = "Bray-Curtis coefficient")
dev.off()
```

E.5 Statistical estimation of drainage separation

E.5.1 Input divergence time estimation data

In order to reconstruct the tabular structure of the dataset, search and replace four spaces and newline (`M-S-%` search `Î ` and replace with nothing in Emacs). Field separators (commas `,`) can be replaced to tabulations with the same approach.

Divergence time estimation data were compiled from the following sources: Abe et al. (2014); Cheviron et al. (2005); Collins and Dubach (2000); Cortés-Ortiz et al. (2003); D'Horta et al. (2013); Dick et al. (2003); Elias et al. (2009); Fernandes et al. (2014); Grau et al. (2005); Gutiérrez et al. (2014); Hardman and Lundberg (2006); Hernández Torres (2015); Machado et al. (2014); Miller et al. (2008); Patané et al. (2009); Patel et al. (2011); Picq et al. (2014); Ribas et al. (2005, 2007); Ruiz-García et al. (2015); Smith et al. (2014); Voss et al. (2013); Weir and Price (2011); Řříčan et al. (2013).

```
Pair,Major group,Area (P = pacific slope),Age (Ma),CI min,CI max,Notes,Type of CI,Reference
Lepidothrix coronata cis-trans haplotype groups,Aves,CT allo,2.3,1.3,3.3,Data in Table 2
    disagree with the 7.1 Ma estimate present in the last paragraph before discussion in
    the same page,Unclear. Seems to be based on variation of sequence divergence,
    \citet{Cheviron2005}
Ateles trans-Andean clade,Mammalia,CT allo? Migration?,3.1,NA,NA,Higly speculative.
    Unsupported by Fig 3 where trans-Andeans are in a polytomy with A. belzebuth (cis),NA,
    \citet{Collins2000}
Allouata cis and trans split,Mammalia,CT allo,6.8,6.6,6.8,Calibrated using varaition in
    sequence distance. Neither CI nor HPD,NA,\citet{Cortes-Ortiz2003}
Sclerurus andinus - peruvianus,Aves,CT allo,2.55,1.92,3.31,Table 1 with data and Figure 3
    with areas/nodes,Said to be 95% CI but constructed using Beast; it might be HPD 95%,
    \citet{DHorta2013}
Symphonia globulifera ITS trans haplotypes,Angiospermae,CT allo,15,NA,NA,Fixed due to
    presence of pollen. Very inaccurate estimation,NA,\citet{Dick2003}
Napeogenes inachia - (apulia+gracilis),Lepidoptera,CT allo,2,NA,NA,Upper Amazon - Col+
```

```
    NEcuador+SEcuador,95% credibility interval from Figure S3,\citet{Elias2009}
Hylophylax naevioides naevioides vs. Several cis-Andean ssp of the H. naevius complex,
    Aves,CT allo,3.5,2.9,4.3,Estimated time in Fig 5 and CI in Molecular dating (p. 1097),,
    \citet{Fernandes2014}
Aburria aburri (andean) - cis-Andean Aburria,Aves,CT allo,3.1,2.4,5.2,,,\citet{Grau2005}
Marmosa robinsoni haplogroups,Mammalia,CT allo,3.65,1.94,5.74,Slightly mixed areas in
    Maracaibo. Otherwise clearly cis-trans,,\citet{Gutierrez2014}
Leiarius perruno - pictus,Actinopterygii,CT allo,NA,8,10,No node estimate. Min age is
    estimated based on fosil; max based on orogenic info. Consider remotion,,
    \citet{Hardman2006}
Rhamdia cis-trans,Actinopterygii,CT allo,10.5,6,15.3,,,\citet{HernandezTorres2015}
Rhamdia guatemalensis haplotypes,Actinopterygii,CT allo,0.7,0.1,1.4,,,
    \citet{HernandezTorres2015}
Ereomyzomys vs remaining rats,Mammalia,CT allo,3.29,0.45,6.97,Molecular dating and ancestral
    areas in p. 124,,\citet{Machado2014}
Oryzomyz-Aegliaoryzomys trans-Andean genera,Mammalia,CT allo,1.85,0.34,3.99,Molecular
    dating and ancestral areas in p. 124,,\citet{Machado2014}
Mionectes oleagineus CA+WEcuad - cis,Aves,CT allo,1.5,1.1,1.9,Supplementary material S2,,
    \citet{Miller2008}
Mionectes oleagineus Guyana+Amazonia+Brasil - Panama,Aves,CT allo,0.8,0.6,1,Supplementary
    material S2,,\citet{Miller2008}
Mionectes oleagineus Epanama - Guyana,Aves,CT allo,0.5,0.4,0.6,Supplementary material S2,,
    \citet{Miller2008}
Ramphastos a. swainsoni - Ramphastos a. ambiguus,Aves,CT allo,NA,0.33,0.81,Fig. 6; text in
    p.931,95% HPD,\citet{Patane2009}
Ramphastos sulfuratus - (R. dicolorus - ariel N),Aves,CT allo,NA,2.6,4.16,Fig. 6; text in
    p.930,95% HPD,\citet{Patane2009}
Ramphastos brevis - (R. ariel SE - R. ariel N),Aves,CT allo,NA,0.99,1.75,Fig. 6; text in
    p.931,95% HPD,\citet{Patane2009}
Ramphastos v. citrolaemus - (R. vitellinus - R. ariel N),Aves,CT allo,NA,0.2,1.4,Fig. 6;
    text in p.931,95% HPD,\citet{Patane2009}
Pteroglossus,Aves,CT allo,NA,2.05,3.28,Table 4,95% HPD,\citet{Patel2011}
Brachyhypopomus cis-trans,Actinopterygii,CT allo,11.2,NA?,NA?,Not explicit; maybe estimable
    from ,,\citet{Picq2014}
Gypopsitta cis-trans species,Aves,CT allo,6.95,6.84,7.06,,,\citet{Ribas2005}
Pionus chalcopterus - cyanescens,Aves,transGarzon allo,0.45,0.12,0.92,Separation between
    Ecord and Ecuadorian Andes through Garzon Massif,,\citet{Ribas2007}
Pionus gr. sordidus split between northern and southern Andean clade,Aves,transGarzon allo,
    2.41,1.52,3.5,Separation between northern Andean clade and Ecuador+Peru Andes through Garzon,
    ,\citet{Ribas2007}
Caquetaia,Actinopterygii,CT allo,10.9,9.4,12.4,Appendix S3 http://api.onlinelibrary.wiley.com/
    asset/v1/doi/10.1111%2Fjbi.12023/asset/supinfo%2Fjbi12023-sup-0003-AppendixS3.doc?l=j6
    %2BNsqLlmq9piJeTweJMwEP%2Flr5RFjM2UqFM4Xk2ObABeYodVg%2F0s%2FWO7ZTSUqVs8JXuusAJZpDW
    %0AfsaQcHBqyeHxEmBDJVLP,,\citet{Rican2013}
Tapirus terrestris-pinchaque,Mammalia,CT allo,3.33,NA,NA,Uncertainty not presented. “trans” is
    actually cordilleran,,\citet{Ruiz-Garcia2015}
Coendou prehensilis - (quichua+rufescens+mexicanus),Mammalia,CT allo,4.4,2.8,6.6,Fig. 4; Fig. 9;
    Table 8,95% HPD,\citet{Voss2013}
Dendrocincla homocroa - remaining spp,Aves,CT allo,3.6,NA,NA,Fig. 5,,\citet{Weir2011}
Dendrocincla anabatina - remaining spp,Aves,CT allo,2.1,NA,NA,Fig. 5,,\citet{Weir2011}
Dendrocincla fuliginosa ridgwayi - D. f. Meruloides,Aves,CT allo,0.9,NA,NA,Fig. 5,,
    \citet{Weir2011}
Salminus affinis - Salminus sp. (Amazonia),Actinopterygii,CT allo,3.6,2.1,5.1,Fig. 4; page 12,
    95% HPD,\citet{Abe2014}
Brycon moorei - remaining Brycon,Actinopterygii,CT allo,18.2,13.4,23,Fig. 4; page 12,95% HPD,
    \citet{Abe2014}
Automolus Central America - Other Automolus,Aves,CT allo?,2.9821,1.6795,4.1844,Supplementary
    material requested to Smith,95% HPD,\citet{Smith2014}
Automolus Central America - Other Automolus,Aves,CT allo?,1.184,0.6949,1.6574,Supplementary
    material requested to Smith,95% HPD,\citet{Smith2014}
```

```
Attila Mexico - Remaining ambiguous trans-then-cis topology,Aves,CT allo?,1.397,0.5288,2.6158,
    Supplementary material requested to Smith,95% HPD,\citet{Smith2014}
Brotogeris,Aves,CT allo,2.9808,1.7753,4.1164,Supplementary material requested to Smith,
    95% HPD,\citet{Smith2014}
Chlorophanes from ancestral Cis to Magdalena/Choco,Aves,CT allo,0.6237,0.3011,0.9372,
    Supplementary material requested to Smith,95% HPD,\citet{Smith2014}
Colonia Napo/Guiana - CentralAm/Choco,Aves,CT allo,2.5431,1.4168,3.932,Supplementary material
    requested to Smith,95% HPD,\citet{Smith2014}
Cyanerpes Magdalena/Esmeraldas - Remaining Cis,Aves,CT allo,1.4309,0.7441,2.1832,Supplementary
    material requested to Smith,95% HPD,\citet{Smith2014}
Cymbilaimus Choco/CentralAM - Napo/Inambari,Aves,CT allo,0.2749,0.1341,0.4459,Supplementary
    material requested to Smith,95% HPD,\citet{Smith2014}
Dendrocincla Foothills/Catatumbo/Magdalena/CentralAM/Choco - Cis,Aves,CT allo,1.1031,0.5859,
    1.5453,Supplementary material requested to Smith,95% HPD,\citet{Smith2014}
Glyphorhynchus CentralAM/Choco - Guiana/Imeri,Aves,CT allo,0.619,0.3471,0.8752,Supplementary
    material requested to Smith,95% HPD,\citet{Smith2014}
Henicorhina trans to Henicorhina cis derived,Aves,CT allo?,2.3866,1.6048,3.3292,Supplementary
    material requested to Smith,95% HPD,\citet{Smith2014}
Lepidothrix Guiana/Imeri - CentralAM/Choco,Aves,CT allo,2.21,1.5153,3.0278,Supplementary
    material requested to Smith,95% HPD,\citet{Smith2014}
Microcerculus from Trans to Cis derived,Aves,CT allo,3.937,2.6682,5.3463,Supplementary material
    requested to Smith,95% HPD,\citet{Smith2014}
Myrmotherula Choco/CentralAM - Napo/Huallaga/Inambari,Aves,CT allo,0.6266,0.3532,0.9496,
    Supplementary material requested to Smith,95% HPD,\citet{Smith2014}
Piaya Choco/CentralAM - Cis,Aves,CT allo,1.6201,1.0494,2.2605,Supplementary material requested
    to Smith,95% HPD,\citet{Smith2014}
Pteroglossus Choco/CentralAM - Cis,Aves,CT allo,1.6423,0.9096,2.4765,Supplementary material
    requested to Smith,95% HPD,\citet{Smith2014}
Pyrilia Trans to Cis,Aves,CT allo,3.9147,2.709,5.3913,Supplementary material requested to
    Smith,95% HPD,\citet{Smith2014}
Pyrilia Trans to Cis,Aves,CT allo,4.5497,3.3,6.016,Supplementary material requested to
    Smith,95% HPD,\citet{Smith2014}
Ramphastos Magdalena - Guiana/Tapajos/Inambari/Napo/Rondonia/Belem,Aves,CT allo,1.0314,0.9885,
    2.3664,Supplementary material requested to Smith,95% HPD,\citet{Smith2014}
Schiffornis Cis to Choco,Aves,CT allo,1.8281,1.039,2.4692,Supplementary material requested to
    Smith,95% HPD,\citet{Smith2014}
```

E.5.2 `datasets.R`

```
# classical model
# Diaz de Gamero (1997), Hoorn et al. (2010), Guerrero (1997)
geoAges <- c(10, 11.3, 11.8)

# expanded dataset post-qualification in appendix
ages <- c(2.3, 3.1, 6.8, 2.55, 15, 2, 3.5, 3.1, 3.65, NA, 10.5, 0.7, 3.29, 1.85, 1.5, 0.8,
          0.5, NA, NA, NA, NA, NA, 11.2, 6.95, 0.45, 2.41, 10.9, 3.33, 4.4, 3.6, 2.1, 0.9,
          3.6, 18.2, 2.9821, 1.184, 1.397, 2.9808, 0.6237, 2.5431, 1.4309, 0.2749, 1.1031,
          0.619, 2.3866, 2.21, 3.937, 0.6266, 1.6201, 1.6423, 3.9147, 4.5497, 1.0314, 1.8281)

confidences <- data.frame(
    min = c(1.3, NA, 6.6, 1.92, NA, NA, 2.9, 2.4, 1.94, 8, 6, 0.1, 0.45, 0.34, 1.1, 0.6, 0.4,
            0.33, 2.6, 0.99, 0.2, 2.05, NA, 6.84, 0.12, 1.52, 9.4, NA, 2.8, NA, NA, NA, 2.1,
            13.4, 1.6795, 0.6949, 0.5288, 1.7753, 0.3011, 1.4168, 0.7441, 0.1341, 0.5859, 0.3471,
            1.6048, 1.5153, 2.6682, 0.3532, 1.0494, 0.9096, 2.709, 3.3, 0.9885, 1.039),
    max = c(3.3, NA, 6.8, 3.31, NA, NA, 4.3, 5.2, 5.74, 10, 15.3, 1.4, 6.97, 3.99, 1.9, 1,
            0.6, 0.81, 4.16, 1.75, 1.4, 3.28, NA, 7.06, 0.92, 3.5, 12.4, NA, 6.6, NA, NA, NA,
            5.1, 23, 4.1844, 1.6574, 2.6158, 4.1164, 0.9372, 3.932, 2.1832, 0.4459, 1.5453,
            0.8752, 3.3292, 3.0278, 5.3463, 0.9496, 2.2605, 2.4765, 5.3913, 6.016, 2.3664,
            2.4692))
```

```
confidences <- confidences[complete.cases(confidences), ]
```

E.5.3 descripStatDatasets.R

```
# load the datasets
source("datasets.R")

### plot a density of point estimates with base
# modified from https://www.r-graph-gallery.com/
# 82-boxplot-on-top-of-histogram.html
pdf(file = "descriptStatDivtime.pdf")
# split screen
layout(mat = matrix(c(1, 2), 2, 1, byrow = TRUE),  height = c(1, 8))
 # Draw the boxplot and the histogram
par(mar = c(0, 3.1, 1.1, 2.1))
boxplot(ages, horizontal = TRUE, ylim = c(-0.5, 20), xaxt = "n" , col = rgb(0, 0, 1, alpha = 1/3),
    frame = FALSE, na.rm = TRUE)
par(mar = c(4, 3.1, 1.1, 2.1))
hist(ages, probability = TRUE, breaks = 15, col = rgb(0, 0, 1, alpha = 1/3), main = "",
    xlab = "Ages (Ma)", ylab = "density", xlim = c(-0.5, 20))
lines(density(ages, na.rm = TRUE), lwd = 2, lty = 2)
dev.off()
```

E.5.4 Methods based on stratigraphic confidence intervals (the stratCI R package)

alpha.R

```
# function for iteratively solving Eq. 19 in Strauss and Sadler '89.
# arguments as follows

# H = number of occurrences/horizons
# C = confidence level. Inherited from rC2p

alpha <- function(H, C) {
    alphae <- seq(from = 0, to = 30, by = 0.001)
    y <- 1 - 2*(1 + alphae)^(1 - H) + (1 + 2*alphae)^(1 - H)
    output <- alphae[which(abs(y - C) < 0.001)]
    mean(output)
}
```

funConstants.R

```
# funConstants class for inputing function constants to the estimators for fast calculations
# or in the absence of data

funConstants <- function(...) {
    constants <- list(...)
    numerics <- sapply(constants, class)
    numerics <- numerics == "numeric"
    if (sum(numerics) != length(constants)) {
        badArgs <- names(constants)[!numerics]
        badArgs <- paste("\"", badArgs, "\" ", sep = "")
        stop(paste(c("arguments ", badArgs, " are not of class \"numeric\" \n")))
    } else {
        structure(list(...), class = "funConstants")
    }
```

```
}
```

inputData.data.frame.R

```
# inputData method for data frames. This method is special as only works for data frames downloaded
# from the paleobioDB package
# x is the data frame. It must preserve the original column names and classes as the output of
# pbdb_occurrences() function in that package
# Includes additional arguments inherited from the wrapper function straussSadler89 are

# x = data frame
# species = the species name value present in the matched_name column of the data frame
# uniques = should duplicated occurrences be excluded? This is important as it biases the value of
# alpha since H is higher than should be

inputData.data.frame <- function(x, species, uniques, by_occur) {
    x <- x[x$matched_name == species, c("early_age", "late_age", "occurrence_no")]
    x <- x[complete.cases(x), ]
    if (by_occur == TRUE) {
        x <- cbind(aggregate(late_age ~ occurrence_no, data = x, min),
                   early_age = aggregate(early_age ~ occurrence_no, data = x, max)[, 2])
    }
    x <- cbind(x, midpoint = apply(x, 1, mean))
    if (uniques == TRUE) {
        H <- length(unique(x$midpoint)) # there's something weird here as H is giving the double of
                                        #the value in Alroy's tool
    } else {
        H <- nrow(x)
    }
    min <- min(x$late_age)
    max <- max(x$early_age)
    R <- max - min
    list(H = H, min = min, max = max, R = R)
}
```

inputData.funConstants.R

```
# funConstants method for the inputData function
# x is the output of calling the function funConstants on the

inputData.funConstants <- function(x) {
    if(sum(sort(names(x)) == sort(c("R", "min", "max", "H"))) != length(names(x))) {
        stop("constant names must follow the \"H\", \"min\", \"max\", \"R\" convention")
    }
    x
}
```

inputData.numeric.R

```
# this method for numeric vectors (usually occurrences for a single species) take the argument
# and calculates the needed constants
# the only argument x is then converted to a list for further use of the parameter estimator
# functions
inputData.numeric <- function(x) {
    H <- length(x)
    min <- min(x)
    max <- max(x)
    R <- max - min
```

```
    list(H = H, min = min, max = max, R = R)
}
```

inputData.R

```
# generic function for accepting three different clases: R's numeric vector, data frames, and
# the funConstants class herein defined
# takes as arguments the following:
# x = dataset

inputData <- function(x, ...) {
    UseMethod("inputData")
}
```

print.funConstants.R

```
# funConstants method for print that does not print the part below with the attr(*, "class") information
print.funConstants <- function(x) {
    attributes(x)$class <- NULL
    print(x)
}
```

rC.R

```
# one-parameter case for the estimator of early age or min age. alpha comes from the Eq. 20 and
# formation of the confidence interval from Eq. 12
# arguments are as follows
# x = inputData(x) this applies the appropriate method for the original x data in th wrapper function
# and converts x in a list
# C = the confidence level. defaults to 0.95
# endpoint = first or last appearance to be estimated

# outputs to the minimum age (youngest) in the data and the min - alpha, the confidence interval
# extension

rC <- function(x, C, endpoint) {
    a <- x$R*((1 - C)^(-1/(x$H - 1)) - 1)
    if (endpoint == "first") {
        c(minObs = x$min, maxObs = x$max, maxEst = x$max + a, H = as.integer(x$H), alphaR = a,
            alpha = a/x$R)
    } else {
        c(minEst = x$min - a, minObs = x$min, maxObs = x$max, H = as.integer(x$H), alphaR = a,
            alpha = a/x$R)
    }
}
```

rC2p.R

```
# two-parameter case for the estimator of early age or min age. alpha comes from the Eq. 19 and
# formation of the
# confidence interval from Eqs. 12 and 13 coupled. It calls the alpha() function that iteratively
# calculates alpha from Eq. 19.
# arguments are as follows
# x = inputData(x) this applies the appropriate method for the original x data in th wrapper function and
# converts x in a list
# C = the confidence level. defaults to 0.95

# outputs to the min - alpha and the max + alpha, that jointly have the C confidence interval extension
```

```
rC2p <- function(x, C) {
    a <- alpha(x$H, C)
    c(minEst = x$min - a*x$R, minObs = x$min, maxObs = x$max, maxEst = x$max + a*x$R, H = as.integer(x$H),
      alphaR = a*x$R, alpha = a)
}
```

straussSadler89.R

```
# straussSadler89 is a wrapper for the functions with one or two parameters
# takes as input
# x = dataset
# case = whether one or two parameter case
# C = confidence level
# endpoint = used only for method = "one.par". Whether first or last occurrence are to be
# estimated
# This function depends on three other functions in the package: rC, rC2p and inputData.
# The first two are
# implementations of
# Strauss and Sadler's estimators for one or two parameter cases, and inputData is a helper
# function that
# homogenizes the data inputed to the functions
# ... = additional arguments to be passed for the inputData.data.frame method based on
# PBDB data frames,
# specially species and uniques

straussSadler89 <- function(x, method, C, endpoint, ...) {
    x <- inputData(x, ...)
    if (method == "one.par") {
        return(rC(x, C, endpoint))
    }
    if (method == "two.par") {
        return(rC2p(x, C))
    } else {
        stop("Method must be one- or two-parameter cases of Strauss & Marshall '89")
    }
}
```

transpTest.R

```
#### Approach at obtaining all combinations and their transpositions in a given vector

### DON'T USE FOR N > 7!!!!
# Takes horrors calculating for n = 8, also, for all cases the function seems to be
# symmetric with repect to number
# of transpositions, maybe a functional approach can be worth trying.

# iterative list generator
# Outputs to a list because expand.grid requires a list of elements for combination
iterList <- function(series) {
    length <- length(series)
    output <- vector(length = length, mode = "list")
    for (i in seq_len(length)) {
        output[i] <- list(series)
    }
    return(output)
}

# the following transpositions function applies the transposition counter function over
```

```
# the rows of a
# data frame and outputs a vector of transpositions
transpositions <- function(x) {
    output <- apply(x, 1, FUN = counter)
    return(output)
}

# This function counts the number of transpositions in a given vector making progressive
# pairwise comparisons
# from first to last element and testing whether the first is larger than the second
# (i.e., a transposition).
counter <- function(x) {
    transp <- vector()
    for (i in 1:(length(x)-1)) {
        transp <- c(transp, x[i] > x[(i+1):length(x)])
    }
    return(sum(transp))
}

# this function generates the combinations and cleans in order to provide only cases where
# all values are
# uniques. Its sole argument is the number of elements to combine
# Performance above n = 8 is due to this function. It runs roughly 10 times slower than
# gtools::permutations

combiner <- function(x) {
    combined <- expand.grid(iterList(1:x))
    uniques <- apply(combined, 1, FUN = unique)
    lengths <- sapply(uniques, FUN = length)
    output <- combined[lengths == x, ]
    return(output)
}
```

marshall94.R

```
####### Distribution-free non-random CIs (Marshall 1994) ##########

# redefinition of the algorithm can be useful in order to avoid calculating all th values and
# concatenating
# the vectors with results. These two improvements will speed up the functions.
# For the search algorithm try a while() loop that stops when the condition is false so that only
# one element
# is saved in memory, instead of calculating all of the values that will be useless in the end.

# Lower bound correctly gives the values in table 1
# testing eq 1.
# C is the (x + 1)th smallest gap out of N gaps
# gamma  = 2*alpha
# N = number of gaps
# x = largest integer that satisfies Eq 1
lowerBound <- function(confidence, N, quantile) {
    gamma <- (1 - confidence)/2
    leftside <- NULL
    for(i in 0:N) {
        x <- i
        leftside <- c(leftside, (choose(N, x) * quantile^x * (1 - quantile)^(N - x))) # for x <= N
    }
    sums <- NULL
    for(i in seq_along(leftside)) {
        sums <- c(sums, sum(leftside[1:i]))
```

```
    }
    tests <- sums < gamma
    output <- sum(tests)
    if(output >= 1) {
        return(output)
    } else {
        print("Impossible to calculate lower bound")
    }
}

# Upper bound correctly gives the values in table 1
upperBound <- function(confidence, N, quantile) {
    gamma <- (1 - confidence)/2
    rightside <- NULL
    for(i in 0:N) {
        x <- i
        rightside <- c(rightside, (choose(N, x) * quantile^x * (1 - quantile)^(N - x))) # for x <= N
    }
    sums <- NULL
    for(i in seq_along(rightside)) {
        sums <- c(sums, sum(rightside[1:i]))
    }
    tests <- sums > (1 - gamma)
    index <- 1:length(tests)
    output <- index[tests == TRUE][1]
    if(output %in% c(NA, N)) {
        warning("Impossible to calculate lower bound")
    } else {
        return(output)
    }
}

marshall94 <- function(intervals, confidence, quantile) {
    lower <- lowerBound(confidence, length(intervals), quantile)
    upper <- upperBound(confidence, length(intervals), quantile)
    output <- c(intervals[lower], intervals[upper])
    return(output)
}
```

pbdb.R

```
# pbdb_range is a copy of the first part of the script for the pbdb_temp_range function in the
# paleobioDB package
pbdb_range <- function(dataset, species, by_occurr = FALSE) {
    selection = dataset[dataset$matched_name == species, ]
    if(by_occurr == TRUE) {
        max = tapply(selection$early_age, as.character(selection$occurrence_no), max)
        min = tapply(selection$late_age, as.character(selection$occurrence_no), min)
    } else {
        max = tapply(selection$early_age, as.character(selection$matched_name), max)
        min = tapply(selection$late_age, as.character(selection$matched_name), min)
    }
    temporal_range = data.frame(max, min, row.names = NULL)
    return(temporal_range)
}
```

E.5.5 Inference based on stratigraphic CIs

```
# load package
```

```
library(stratCI)

# load datasets
source("datasets.R")

# remove missing data
ages <- ages[complete.cases(ages)]
# remove outliers
ages <- ages[which(ages < 10)]

# calculate the older theta that is an extension beyond the oldest occurrence
strausSadlerAndes <- straussSadler89(x = ages, method = "one.par", endpoint = "first", C = 0.95)

#calculate the confidence interval under th distribution-free approach

marshallAndes <- c(max(ages), max(ages) + marshall94(ages, confidence = 0.95, quantile = 0.95)[1])

# plot the estimations
pdf(file = "divtimeStratCI.pdf")
hist(ages, probability = TRUE, col = rgb(red = 0, green = 0, blue = 1, alpha = 0.3), xlim = c(0, 10),
    main = "CIs based on stratigraphic intervals", xlab = "Age (Ma)")
# plot the lines  for the classical CI estimator of Strauss and Sadler
arrows(x0 = strausSadlerAndes["maxObs"], y0 = 0.025, x1 = strausSadlerAndes["maxEst"], y1 = 0.025,
    code = 3, angle = 90, length = 0.1, lwd = 3, col = "darkblue")
# plot the lines for the CI estimator of Marshall 94
arrows(x0 = marshallAndes[1], y0 = 0.05, x1 = marshallAndes[2], y1 = 0.05, code = 3, angle = 90,
    length = 0.1, lwd = 3, col = "darkgreen")
# plot a legend
legend(x = "topright", legend = c("Strauss and Sadler \'89", "Marshall \'94"),
    col = c("darkblue", "darkgreen"),  lty = 1, lwd = 3)
dev.off()
```

E.5.6 Method based on the x-intercept

Description of the method

```
# load packages
library(Hmisc)
# load the datasets
source("datasets.R")

ages <- ages[complete.cases(ages)] # remove missing data
# remove extreme values for illustrative purposes,
# later this action gets justified
ages <- ages[which(ages < 10)]
# calculate the empirical ecdf and save it to an object for future plotting
ecdfAges <- Hmisc::Ecdf(ages, pl = FALSE)
norRegAges <- lm(I(1 - y) ~ x, data = ecdfAges)
### plot a density of point estimates with base along with a regression
# modified from https://www.r-graph-gallery.com/
# 82-boxplot-on-top-of-histogram.html
pdf(file = "cumulativePlot.pdf", height = 10, width = 7, pointsize = 17)
par(oma = c(0, 0, 0, 0))
split.screen(c(2, 1))
screen(1)
par(mar = c(2, 4.5, 3, 2))
plot(x = ecdfAges$x, y = (1 - ecdfAges$y), type = "s", xlim = c(0, 8), ylab = "Cumulative density")
screen(2)
par(mar = c(4.5, 4.5, 2, 2))
```

```
plot(x = ecdfAges$x, y = (1 - ecdfAges$y), type = "p", col = "black", xlim = c(0, 8), xlab = "Age (Ma)",
    ylab = "Cumulative density")
abline(h = 0)
abline(norRegAges, lty = 2)
points(x = -(coef(norRegAges)[1] / coef(norRegAges)[2]), y = 0, pch = 23, col = "black",
    bg = "red", cex = 1.2)
dev.off()
```

Confidence interval for the x-intercept of (Draper and Smith, 1998) and bootstrap

```
# load packages
library(boot) # for the bootstrap calculation method
library(Hmisc) # for the calculation of empirical CDFs
library(Rfit) # robust lm

# Draper-Smith CI method
# x = object of class lm or rfit
CI_DraperSmith_X0 <- function(x, alpha = 0.05, ...) {
    if (class(x) == "rfit") {
        x$model <- data.frame(y = x$y,
                              x = x$x[,"x"],
                              stringsAsFactors = FALSE)
    }
    intercept <- coef(x)[1]
    slope <- coef(x)[2]
    meanX <- mean(x$model[,2])
    n <- length(x$model[,2])
    tstar <- qt(alpha/2, n-2)
    sxx <- sum(x$model[,2]^2) - sum(x$model[,2])^2 / n
    SSresidual <- (1-cor(x$model[,1], x$model[,2])^2) *
                  (sum(x$model[,1]^2)-sum(x$model[,1])^2/n)
    S <- sqrt(SSresidual/(n-2))
    SEslope <- S / sqrt(sxx)
    Xintercept <- - intercept / slope
    y0 <- 0
    g <- (tstar / (slope/SEslope))^2
    left <- (Xintercept - meanX) * g
    bottom <- 1 - g
    Right <- (tstar * S / slope) * sqrt( ((Xintercept - meanX)^2/sxx) + bottom/n)
    lower <- Xintercept + (left + Right) / bottom
    upper <- Xintercept + (left - Right) / bottom
    return(c(lower,upper))
}

# bootstrap-based method
# x,y = the x and cumulative-y values for fitting the robust model rfit
CI_Boot_X0 <- function(x, y, p = c(0.025, 0.975), R = 1000, robust = FALSE) {
    d <- data.frame(x, y)
    bootObj <- boot::boot(d,
                          function(d, i) {
                              if(robust) {
                                  fit <- Rfit::rfit(y ~ x, data = d[i,])
                              } else {
                                  fit <- lm(y ~ x, data = d[i,])
                              }
                          -coef(fit)[1]/coef(fit)[2]
                      },
                      R = R)
```

```
    return(quantile(bootObj$t, p))
}

# load datasets
source("datasets.R")

### bootstrap-based CIs

# when calculating  the cumulative function, mind the sign as we are interested in the
# negative cumulative,
# that is, the -vector of ages in order to think in a geo-time scale
ages <- ages[complete.cases(ages)] # remove NAs
# outliers create a very weird situations in which lm and rfit give the same
# estimated coefficients, this is clearly wrong, outlier remotion improves the situation
$ but in the end al this
# illustrates how dangerous is to lean on the
# CDF for regression
ages <- ages[which(ages < 10)]

ecdfAges <- Hmisc::Ecdf(ages, pl = FALSE)
# in order to represent properly in time this information, we need the complement of
# cumulative probability
ecdfAges$y <- 1 - ecdfAges$y

## using classical regression (misleading due to non-normality of residuals)
# classical regression
norRegLogEcdfAges <- lm(y ~ x, data = ecdfAges)
## using robust regression methods
robRegLogEcdfAges <- Rfit::rfit(y ~ x, data = ecdfAges)

# CI based on the method of Draper and Smith
draperSmithNormalX0 <- CI_DraperSmith_X0(norRegLogEcdfAges)
draperSmithRobustX0 <- CI_DraperSmith_X0(robRegLogEcdfAges)

# CI from bootstraping
bootstrapX0 <- CI_Boot_X0(x = ecdfAges$x, y = ecdfAges$y)

### plot the estimations
# plot the estimations
pdf(file = "divtimeCDF_CI.pdf")
hist(ages, probability = TRUE, col = rgb(red = 0, green = 0, blue = 1, alpha = 0.3),
    xlim = c(0, 10),  main = "CDF-based on confidence intervals", xlab = "Age (Ma)")
# plot the lines for the estimator of Draper and Smith using lm
arrows(x0 = draperSmithNormalX0[2], y0 = 0.025, x1 = draperSmithNormalX0[1], y1 = 0.025,
    code = 3, angle = 90, length = 0.1, lwd = 3, col = "darkblue")
# plot the lines for the estimator of Draper and Smith using rfit
arrows(x0 = draperSmithRobustX0[2], y0 = 0.05, x1 = draperSmithRobustX0[1], y1 = 0.05,
    code = 3, angle = 90, length = 0.1, lwd = 3, col = "darkgreen")
# plot the lines for the estimator based on bootstrap
arrows(x0 = bootstrapX0[1], y0 = 0.075, x1 = bootstrapX0[2], y1 = 0.075, code = 3,
    angle = 90, length = 0.1, lwd = 3, col = "darkred")
# plot a legend
legend(x = "topright", legend = c("Draper and Smith with lm", "Draper and Smith with rfit",
                                  "Bootstrap on x0"),
    col = c("darkblue", "darkgreen", "darkred"),  lty = 1, lwd = 3)
dev.off()
```

E.5.7 Inference based on resampling

resampbasedInference.R

```
# load the dataset
source("datasets.R")

# age simulations using resampling from confidence regions
toc <- Sys.time()

N <- 1000000
meanConf <- vector(mode = "numeric", length = N)
medianConf <- vector(mode = "numeric", length = N)

# for each divtime point, sample a randon number from the uncertainty region, then calculate the mean and
# median of these values and save them to their respective vectors
for (i in 1:N) {
    sampleIndex <- sample(1:nrow(confidences), size = 15)
    df <- confidences[sampleIndex, ]
    estimates <- vector(mode = "numeric", length = nrow(df))
    for (j in 1:nrow(df)) {
        estimates[j] <- runif(n = 1, min = df$min[j], max = df$max[j])
    }
    meanConf[i] <- mean(estimates)
    medianConf[i] <- median(estimates)
}

toc <- Sys.time()
# evaluation time
toc-tic

# build a dataset with results
simDataset <- data.frame(means = meanConf, medians = medianConf, stringsAsFactors = FALSE)

# to write a backup of simulated data:
write.table(simDataset, paste("millionRunsAugmentedData", format(Sys.time(), "%d-%m-%Y"), ".csv", sep = ""),
    sep = ",", row.names = FALSE, col.names = TRUE)

# generate composite histogram
pdf("divtimeResampling.pdf", width = 6, heigh = 5)
hist(ages, breaks = 13, probability = TRUE, ylim = c(0, 0.8), xlim = c(0, 20), col = "black",
    xlab = "Age (Ma)", ylab = "Density", main = "Parameter resampling (1.000.000 iterations)")
hist(medianConf, probability = TRUE, col = rgb(1, 0, 0, 1/3), add = TRUE)
hist(meanConf, probability = TRUE, col = rgb(0, 0, 1, 1/3), add = TRUE)
abline(v = median(ages, na.rm = TRUE), col = "red", lwd = 3)
abline(v = mean(ages, na.rm = TRUE), col = "blue", lwd = 3)
abline(v = mean(geoAges), col = "black", lty = "dashed", lwd = 3)
box()
legend(x = "topright", legend = c("Nodes", "Median", "Mean"), fill = c("black", rgb(1, 0, 0, 1/3),
    rgb(0, 0, 1, 1/3)))
dev.off()
```

Bibliography

Abe, K.T., Mariguela, T.C., Avelino, G.S., Foresti, F., and Oliveira, C. 2014. Systematic and historical biogeography of the Bryconidae (Ostariophysi: Characiformes) suggesting a new rearrangement of its genera and an old origin of Mesoamerican ichthyofauna. *BMC evolutionary biology*, 14:152. 214

Aguilera, O.A., Lundberg, J.G., Birindelli, J.L.O., Sabaj Pérez, M.H., Jaramillo, C.A., and Sánchez-Villagra, M.R. 2013a. Palaeontological evidence for the last temporal occurrence of the ancient Western Amazonian river outflow into the Caribbean. *PLoS ONE*, 8:1–17. 203, 211

Aguilera, O.A., Moraes-Santos, H., Costa, S.A.R.F., Ohe, F., Jaramillo, C.A., and Nogueira, A. 2013b. Ariid sea catfishes from the coeval Pirabas (Northeastern Brazil), Cantaure, Castillo (Northwestern Venezuela), and Castilletes (North Colombia) formations (early Miocene), with description of three new species. *Swiss Journal of Palaeontology*, 132:45–68. 203, 211

Antoine, P.O., Abello, M.A., Adnet, S., Altamirano Sierra, A.J., Baby, P., Billet, G., Boivin, M., Calderón, Y., Candela, A., Chabain, J., Corfu, F., Croft, D.A., Ganerød, M., Jaramillo, C.A., Klaus, S., Marivaux, L., Navarrete, R.E., Orliac, M.J., Parra, F., Pérez, M.E., Pujos, F., Rage, J.C., Ravel, A., Robinet, C., Roddaz, M., Tejada-Lara, J.V., Vélez-Juarbe, J., Wesselingh, F.P., and Salas-Gismondi, R. 2016. A 60-million-year Cenozoic history of western Amazonian ecosystems in Contamana, eastern Peru. *Gondwana Research*, 31:30–59. 203, 211

Azpelicueta, M.d.l.M. and Cione, A.L. 2016. A southern species of the tropical catfish genus *Phractocephalus* (Teleostei: Siluriformes) in the Miocene of South America. *Journal of South American Earth Sciences*, 67:221–230. 203, 211

Ballen, G.A. and Moreno-Bernal, J.W. 2019. New records of the enigmatic Neotropical fossil fish *Acregoliath rancii* (Teleostei *incertae sedis*) from the middle Miocene Honda group of Colombia. *Ameghiniana*, 56:431–440. 203, 211

Bogan, S., Sidlauskas, B., Vari, R.P., and Agnolin, F.L. 2012. *Arrhinolemur scalabrinii* Ameghino, 1898, of the late Miocene - a taxonomic journey from the Mammalia to the

Anostomidae (Ostariophysi: Characiformes). *Neotropical Ichthyology*, 10:555–560. 203, 211

Cheviron, Z.A., Hackett, S.J., and Capparella, A.P. 2005. Complex evolutionary history of a Neotropical lowland forest bird (*Lepidothrix coronata*) and its implications for historical hypotheses of the origin of Neotropical avian diversity. *Molecular Phylogenetics and Evolution*, 36:338–357. 214

Cione, A.L. and Azpelicueta, M.d.l.M. 2013. The first fossil species of *Salminus*, a conspicuous South American freshwater predatory fish (Teleostei , Characiformes), found in the Miocene of Argentina. *Journal of Vertebrate Paleontology*, 33:1051–1060. 203, 211

Cione, A.L., Azpelicueta, M.d.l.M., Bond, M., Carlini, A.A., Casciotta, J.R., Cozzuol, M.A., de la Fuente, M., Gasparini, Z., Goin, F.J., Noriega, J., Scillato-Yané, G.J., Soibelzon, L., Tonni, E.P., Verzi, D., and Vucetich, M.G. 2000. The Miocene vertebrates from Paraná, eastern Argentina Miocene vertebrates from Entre Ríos province, eastern Argentina. *Correlación Geológica*, 14:191–237. 203, 211

Cione, A.L., Dahdul, W.M., Lundberg, J.G., and Machado-Allison, A. 2009. *Megapiranha paranensis*, a new genus and species of Serrasalmidae (Characiformes, Teleostei) from the upper Miocene of Argentina. *Journal of Vertebrate Paleontology*, 29:350–358. 203, 211

Collins, A.C. and Dubach, J.M. 2000. Biogeographic and ecological forces responsible for speciation in Ateles. *International Journal of Primatology*, 21:421–444. 214

Cortés-Ortiz, L., Bermingham, E., Rico, C., Rodríguez-Luna, E., Sampaio, I., and Ruiz-García, M. 2003. Molecular systematics and biogeography of the Neotropical monkey genus, *Alouatta*. *Molecular Phylogenetics and Evolution*, 26:64–81. 214

D'Horta, F.M., Cuervo, A.M., Ribas, C.C., Brumfield, R.T., and Miyaki, C.Y. 2013. Phylogeny and comparative phylogeography of *Sclerurus* (Aves: Furnariidae) reveal constant and cryptic diversification in an old radiation of rain forest understorey specialists. *Journal of Biogeography*, 40:37–49. 214

Dick, C.W., Abdul-Salim, K., and Bermingham, E. 2003. Diversification of a Widespread Tropical Rain Forest Tree. *The American Naturalist*, 162:691–703. 214

Draper, N.R. and Smith, H. 1998. *Applied Regression Analysis*. Wiley Interscience, New York. 224

Elias, M., Joron, M., Willmott, K., Silva-Brandão, K.L., Kaiser, V., Arias, C.F., Piñerez, L.M.G., Uribe, S., Brower, A.V.Z., Freitas, A.V.L., and Jiggins, C.D. 2009. Out of the Andes: patterns of diversification in clearwing butterflies. *Molecular Ecology*, 18:1716–1729. 214

Fernandes, A.M., Wink, M., Sardelli, C.H., and Aleixo, A. 2014. Multiple speciation across the Andes and throughout Amazonia: The case of the spot-backed antbird species complex (*Hylophylax naevius/Hylophylax naevioides*). *Journal of Biogeography*, 41:1094–1104. 214

Grau, E.T., Pereira, S.L., Silveira, L.F., Höfling, E., and Wajntal, A. 2005. Molecular phylogenetics and biogeography of Neotropical piping guans (Aves: Galliformes): Pipile Bonaparte, 1856 is synonym of Aburria Reichenbach, 1853. *Molecular Phylogenetics and Evolution*, 35:637–645. 214

Gutiérrez, E.E., Anderson, R.P., Voss, R.S., Ochoa-G., J., Aguilera, M., and Jansa, S.A. 2014. Phylogeography of *Marmosa robinsoni*: insights into the biogeography of dry forests in northern South America. *Journal of Mammalogy*, 95:1175–1188. 214

Hardman, M. and Lundberg, J.G. 2006. Molecular phylogeny and a chronology of diversification for "phractocephaline" catfishes (Siluriformes: Pimelodidae) based on mitochondrial DNA and nuclear recombination activating gene 2 sequences. *Molecular Phylogenetics and Evolution*, 40:410–418. 214

Hernández Torres, C.L. 2015. *Species Delimitation, Phylogenetics, and Biogeography of the Catfish Genus Rhamdia Bleeker (Heptapteridae) of Central America and the Trans-Andean Region of Colombia.* Phd thesis, The University of Southern Mississippi. 214

Lundberg, J.G., Sabaj Pérez, M.H., Dahdul, W.M., and Aguilera, O.A. 2010. The Amazonian Neogene fish fauna. In C. Hoorn and F.P. Wesselingh (eds.), *Amazonia, Landscape and Species Evolution: A Look Into the Past*, pp. 281–301. Blackwell Publishing. 203, 211

Machado, L.F., Leite, Y.L.R., Christoff, A.U., and Giugliano, L.G. 2014. Phylogeny and biogeography of tetralophodont rodents of the tribe Oryzomyini (Cricetidae: Sigmodontinae). *Zoologica Scripta*, 43:119–130. 214

Miller, M.J., Bermingham, E., Klicka, J., Escalante, P., do Amaral, F.S.R., Weir, J.T., and Winker, K. 2008. Out of Amazonia again and again: episodic crossing of the Andes promotes diversification in a lowland forest flycatcher. *Proceedings of the Royal Society B*, 275:1133–42. 214

Patané, J.S.L., Weckstein, J.D., Aleixo, A., and Bates, J.M. 2009. Evolutionary history of *Ramphastos* Toucans: Molecular phylogenetics, temporal diversification, and biogeography. *Molecular Phylogenetics and Evolution*, 53:923–934. 214

Patel, S., Weckstein, J.D., Patané, J.S.L., Bates, J.M., and Aleixo, A. 2011. Temporal and spatial diversification of *Pteroglossus* Araçaris (Aves: Ramphastidae) in the neotropics: Constant rate of diversification does not support an increase in radiation during the Pleistocene. *Molecular Phylogenetics and Evolution*, 58:105–115. 214

Picq, S., Alda, F., Krahe, R., and Bermingham, E. 2014. Miocene and Pliocene colonization of the Central American Isthmus by the weakly electric fish *Brachyhypopomus occidentalis* (Hypopomidae, Gymnotiformes). *Journal of Biogeography*, 41:1520–1532. 214

Ribas, C.C., Gaban-Lima, R., Miyaki, C.Y., and Cracraft, J. 2005. Historical biogeography and diversification within the Neotropical parrot genus *Pionopsitta* (Aves: Psittacidae). *Journal of Biogeography*, 32:1409–1427. 214

Ribas, C.C., Moyle, R.G., Miyaki, C.Y., and Cracraft, J. 2007. The assembly of montane biotas: linking Andean tectonics and climatic oscillations to independent regimes of diversification in *Pionus* parrots. *Proceedings. Biological sciences / The Royal Society*, 274:2399–2408. 214

Ruiz-García, M., Vásquez, C., Sandoval, S., Kaston, F., Luengas-Villamil, K., and Shostell, J.M. 2015. Phylogeography and spatial structure of the lowland tapir (*Tapirus terrestris* , Perissodactyla: Tapiridae) in South America. *Mitochondrial DNA*, 1736:1–9. 214

Smith, B.T., Mccormack, J.E., Cuervo, A.M., Hickerson, M.J., Aleixo, A., Cadena, C.D., Pérez-Emán, J., Burney, C.W., Xie, X., Harvey, M.G., Faircloth, B.C., Glenn, T.C., Derryberry, E.P., Prejean, J., Fields, S., and Brumfield, R.T. 2014. The drivers of tropical speciation. *Nature*, pp. 1–8. 214

Tejada-Lara, J.V., Salas-Gismondi, R., Pujos, F., Baby, P., Benammi, M., Brusset, S., de Franceschi, D., Espurt, N., Urbina, M., and Antoine, P.O. 2015. Life in proto-Amazonia: Middle Miocene Mammals from the Fitzcarrald Arch (Peruvian Amazonia). *Palaeontology*, 58:341–378. 203, 211

Toledo-Piza, M. 2000. The Neotropical fish subfamily Cynodontinae (Teleostei: Ostariophysi: Characiformes): A phylogenetic study and a revision of *Cynodon* and *Rhaphiodon*. *American Museum Novitates*, 3286:1–88. 210

Voss, R.S., Hubbard, C., and Jansa, S.A. 2013. Phylogenetic Relationships of New World Porcupines (Rodentia, Erethizontidae): Implications for Taxonomy, Morphological, and Biogeography. *American Museum Novitates*, pp. 1–36. 214

Weir, J.T. and Price, M. 2011. Andean uplift promotes lowland speciation through vicariance and dispersal in *Dendrocincla* woodcreepers. *Molecular Ecology*, 20:4550–4563. 214

Říčan, O., Piálek, L., Zardoya, R., Doadrio, I., and Zrzavý, J. 2013. Biogeography of the Mesoamerican Cichlidae (Teleostei: Heroini): colonization through the GAARlandia land bridge and early diversification. *Journal of Biogeography*, 40:579–593. 214