Evolutionary History and Taxonomy

of the Titi Monkeys (Callicebinae)

Hazel Byrne

University of
Salford
MANCHESTER

This thesis is presented to the School of Environment & Life Sciences, University of

Salford, in fulfilment of the requirements for the degree of Ph.D.

Supervised by Prof. Jean P. Boubli

June 2017
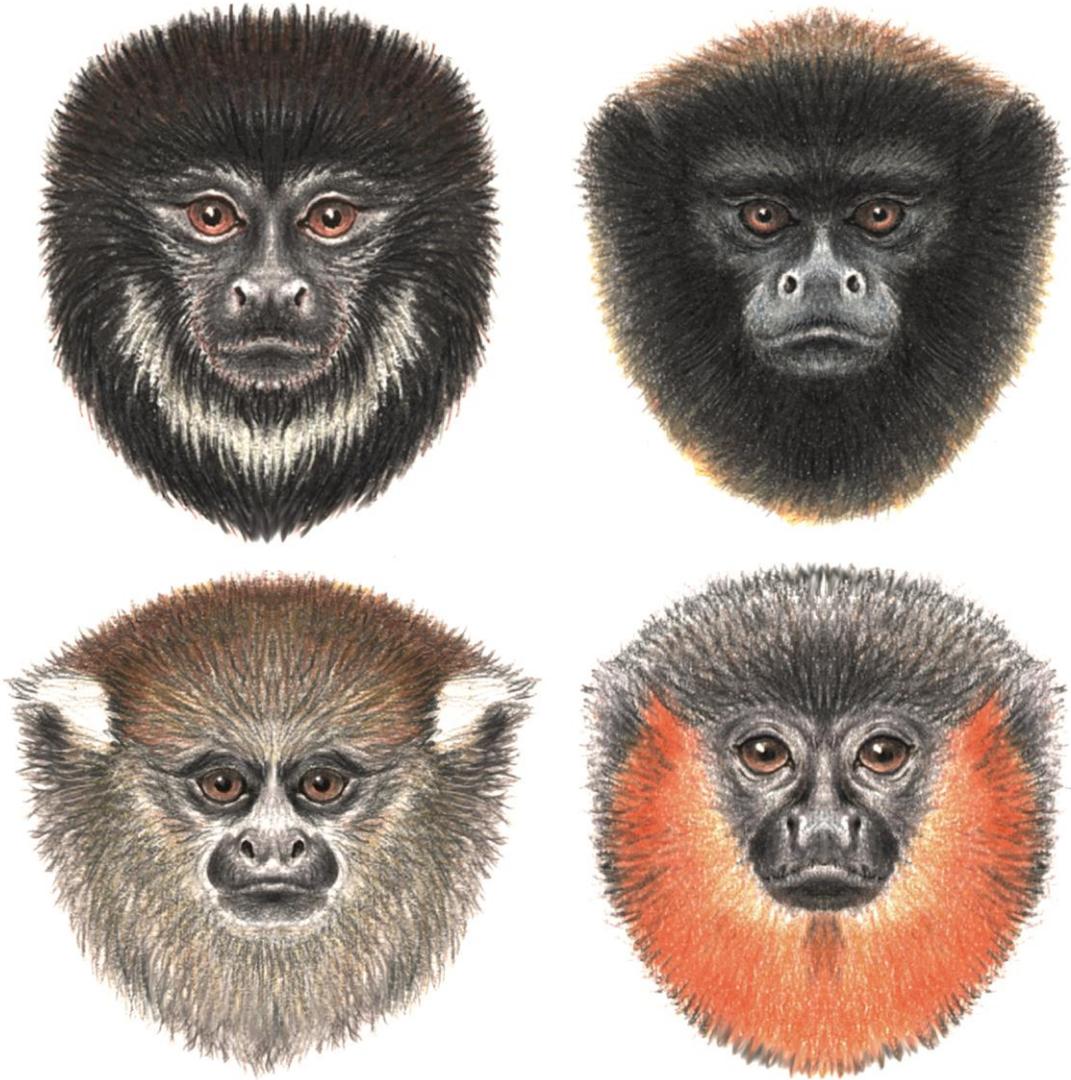
Titi monkeys (subfamily Callicebinae). Illustrations by Stephen D. Nash ©Conservation International.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to sincerely thank my supervisor, Jean P. Boubli, for his on-going support, patience and encouragement, for being a friend as well as a mentor throughout this project, and for providing the opportunity to pursue my interests and the foundations of a research career. I look forward to continuing our work together. I owe thanks to the many others at the University of Salford that have helped me at various points; Catriona Nardone, Robin Beck, Robert Jehle, Chrysoula Gubili, Felipe Ennes, Stefano Mariani, Chiara Benvenuto, Peter Shum, Gemma Lace-Costigan, to name a few.

I would like to express my gratitude to those at LEGAL, UFAM, and the Bragança Genetics Lab, UFPA, in particular, I am indebted to Tomas Hrbek, Izeni Farias, Fabricio Bertuol, Horacio Schneider, Iracilda Sampaio, Jeferson Carneiro, and Luciana Watanabe, for providing facilities, technical assistance, advice, expertise, and allowing this work to be possible. I am eternally grateful to Adrianne Freitas, Priscila Santos, Fabricio Bertuol, and Rodrigo Araújo, and everyone at UFPA and UFAM who made me feel at home. I would like to thank Jessica Lynch Alfaro for her help and support throughout this thesis.

I am grateful to Bruce Patterson and Lauren Smith at the Field Museum, Chicago, Roberto Portela Miguez at the NHM, London, and João Oliveira at the National Museum of Rio de Janeiro, as well as Maria Nazareth F. da Silva and all those at INPA, Manaus, for their assistance over the course of this project. I would like to thank Wilson Spironello at INPA for providing some additional specimens for the ddRADseq study. I am especially grateful to Anthony Rylands for his ever-willing advice and Stephen Nash for allowing us to use his beautiful illustrations.

I am most grateful of all to my parents, family and friends, in particular, Julia, Shiv and Mark, for their endless support and encouragement, Amo, Steph and Alia for sharing so much of this journey with me (especially Amo for joining me in random places around the world), John and Anna for their help with translating some old German literature, and of course, my cats.

We thank the Yanomami people and the Ribeirinho communities of the Rio Negro Basin for their help and support during JPB's surveys. We also thank the Catholic and Evangelic missionaries who provided assistance to our

Preface

The research conducted for this thesis forms part of a research collaboration between the University of Salford, UK (led by Jean P. Boubli), the Federal University of Amazonas, Brazil (led by Tomas Hrbek and Izeni Farias), and the Federal University of Pará, Brazil (led by Iracilda Sampaio and Horacio Schneider). All analyses conducted herein are my own original work.

Chapter 2 and Appendix 1 of this thesis have been published as Byrne H, Ryland AB, Carneiro JC, Lynch Alfaro JW, Bertuol F, da Silva MNF, Messias M, Groves CP, Mittermeier RA, Farias I, Hrbek T, Schneider H, Sampaio I, and Boubli JP. 2016. "Phylogenetic relationships of the New World titi monkeys (Callicebus): First appraisal of taxonomy based on molecular evidence". Frontiers in Zoology, 13(10), 1–25. JPB and I conceived the ideas; ABR, JWLA, RAM, CPG, TH, MM, MNFS, IS, HS and IF contributed to the development of the ideas, genus names, or gave samples; I performed the laboratory work and collected most of the data; JCC, FB and IF aided with collecting the mitochondrial data; I prepared the molecular datasets, determined the nature of the analyses and conducted all analyses; JPB coordinated and supervised the study; and I drafted the manuscript and led the writing, with guidance from ABR on the taxonomy aspects, and from JPB. All authors read and approved the final manuscript.

Chapter 3 and Appendix 2 of this thesis has been read and approved by the collaborative group mentioned above, as well as JWLA. I, JPB and JWLA conceived the study, TH, IS, HS and IF contributed to the development of ideas, I performed the laboratory work and collected the data, prepared the molecular datasets, determined the nature of the analyses, conducted and interpreted all analyses, drafted the manuscript and led the writing, with guidance from JWLA, JPB and TH.

Chapter 4 and Appendix 3 of this thesis have only been seen by JPB. TH conceived the laboratory design of this study. I, along with TH and FB, generated the ddRADseq data. I conceived, determined, and conducted all other components of this study, and drafted/wrote the manuscript.

Abstract

Titi monkeys (Callicebinae; Pitheciidae) are a diverse, species-rich group of New World primates with an extensive range across South America. They diverged from their sister clade (Pitheciinae) in the early Miocene, and thus, they comprise one of the oldest lineages of extant New World primates. To date, there has been no comprehensive molecular investigation of the phylogenetic relationships among Callicebinae species and, consequently, the evolutionary history of this diverse clade remains poorly studied. The overall goal of this PhD dissertation is, therefore, to provide insight into the evolutionary and biogeographic history of the subfamily Callicebinae using DNA sequence data. To infer phylogeny and estimate divergence times, we generated sequence data for 50+ wild-caught titi monkey specimens using multi-locus Sanger sequencing (22 nuclear and mitochondrial loci, > 14,500 bp) and reduced representation, genome-wide double-digest restriction-associated DNA (ddRAD) sequencing. A statistical biogeographical approach was employed to reconstruct the biogeography of Callicebinae and investigate the processes responsible for shaping present day distributions. Furthermore, the ddRAD sequence dataset was used to provide additional insight into phylogenetic relationships and genetic structure among taxa of the *moloch* group. Our phylogenetic and biogeographic results indicate that titi monkeys are divided into three distinct clades that diverged in the late Miocene through vicariance of a widespread ancestral range. Species relationships were generally recovered with strong support, and species-level diversification in the Amazonian clades was characterised by sequential founder events across river barriers in the Pleistocene. We propose a revised genus-level classification for Callicebinae that recognises three genera (*Cheracebus, Callicebus, Plecturocebus*) based on the results from the phylogenetic analyses, as well as morphological, karyological and biogeographic evidence. Overall, this study represents a major advance in our understanding of the evolutionary history of this strikingly poorly studied group, with implications for classification and research priorities.

Chapter 1: Introduction

1.1 Phylogenetic systematics

A central aim of biological research is to describe and explain the diversity of life, however, in order to communicate effectively about biological diversity, a system of classification is necessary. In the 18th century, Carl Linnaeus devised one of the first comprehensive classification systems, which systematically categorised living organisms into a ranked hierarchy based on structural similarity and standardised the use of formal binomial nomenclature. With the advent of the theory of evolution and the modern evolutionary synthesis, this pre-Darwinian Linnaean system has been gradually reworked into a system of classification based on the evolutionary relationships between organisms (Mayr & Bock, 2002; Mayr & Provine, 1998). Phylogenetic systematics is the study of evolutionary history involving the reconstruction of evolutionary relationships among taxa and consequent application to the classification of organisms (Wiley & Lieberman, 2011).

Phylogenetics operates on the principle that all living organisms have descended from a single common ancestor through patterns of branching and divergence (Hennig, 1966), resulting in a tree-like hierarchy of relatedness among taxa, with closely related lineages sharing a more recent common ancestor. Phylogenies are reconstructions of these branching patterns, typically representing evolutionary hypotheses regarding the order of divergence (branch order) and the amount of evolutionary change (branch lengths), and provide a powerful framework within which to study evolutionary history (Baum et al., 2005; Gregory, 2008). Traditional approaches to phylogenetics relied primarily on morphological data to infer evolutionary relationships, however, the limitations of morphological characters, such as complex evolutionary change and the difficulty in obtaining large mathematical datasets, became a major driving force behind the paradigm shift towards the use of molecular data (Nei & Kumar, 2000). At the end of the 20th century, advances in molecular biology and sequencing technologies led to the dominance of molecular markers for estimation of phylogeny and the rapid growth of molecular phylogenetics (San Mauro & Agorreta, 2010). The ability to generate large

and unambiguous molecular datasets amenable to statistical analysis has revolutionised our understanding of the evolutionary relationships among taxa, and thus, our ability to establish taxonomic classifications representing patterns of diversification.

Biological diversity is a continuum and the spectrum of diversity among any group of organisms cannot be fully accounted for by discrete units in a hierarchical classification system. The diversification process can also be more complex than reconstructed in a strictly bifurcating tree of life (e.g., Suh et al., 2015; Willis, 2017). Even considering such limitations, systematics is the primary language for communicating about diversity, and it is important that organisms are classified in a manner that attempts to represent, if it cannot truly reflect, the patterns and processes of diversification.

1.2 Molecular phylogenetics

Most early molecular phylogenies were inferred from a single locus, this however results in the reconstruction of a gene tree rather than an evolutionary tree of the species (Nei, 1987; Neigel & Avise, 1986; Tateno et al., 1982). Species trees can differ significantly from gene trees owing to processes such as incomplete lineage sorting and interspecific gene flow, which are especially problematic when considering lineages that are closely related or potentially hybridising (Avise et al., 1983; Edwards, 2009; Maddison, 1997; Pamilo & Nei, 1988). Phylogenetics has since progressed to a multi-gene approach that can increase the likelihood of identifying species trees through the use of multiple independent loci (Pamilo & Nei, 1988; Rokas et al., 2003; Takahata, 1989; Wu, 1991). Increasing the number of loci helps to control for the stochastic forces affecting individual genes (Maddison & Knowles, 2006), and results in phylogenies with increased resolution and confidence (e.g., Betancur-R et al., 2013; Murphy et al., 2001; Perelman et al., 2011; Yi et al., 2014).

With the advance of next-generation sequencing (NGS) technologies, phylogenetics has entered a new era where genome-wide or whole-genome data is used to reconstruct evolutionary history (phylogenomics; Delsuc et al., 2005). NGS is a more cost-effective method of producing large sequence datasets than traditional Sanger sequencing approaches and the high-coverage capacity can reduce stochastic errors and enhance resolution power (Shendure & Ji, 2008). The power of genome-

wide data to address difficult phylogenetic problems has been clearly demonstrated (e.g. Botero-Castro et al., 2013; Escudaro et al., 2014; Wagner et al., 2013). Large molecular datasets, however, can increase systematic error owing to model violation or misspecification, posing significant statistical and computational challenges to the estimation of a reliable phylogeny (Nishihara et al., 2007; Rannala & Yang, 2008; Rodríguez-Ezpeleta et al., 2007). This has led to much recent discussion about the most appropriate methods for the estimation of phylogeny from multi-locus and genomic datasets (e.g., Edwards et al., 2016; Liu et al., 2009; Liu et al., 2015; Simmons & Gatesy, 2015; Springer & Gatesy, 2016). The traditional approach involves the concatenation of all loci into a super-matrix and analysed simultaneously (Kluge, 1989; de Queiroz & Gatesy, 2007). One of the advantages of concatenation is robustness to missing data as taxa are represented by a large number of informative characters even when sequence data is incomplete (Delsuc et al., 2005). The development of partitioning methods that allow each partition to evolve under a different model, thus taking into account across gene-heterogenity, have enhanced our ability to extract the phylogenetic signal from concatenated sequences (Lanfear et al., 2012; Nylander et al., 2004).

Many strongly supported phylogenetic trees that include a large number of diverse taxa have been inferred using this approach (e.g Betancur-R et al., 2013; Perelman et al., 2011), however, differing gene histories can lead concatenation to infer well supported but inaccurate phylogenies (Degnan & Rosenberg, 2009; Edwards, 2009; Kubatko & Degnan, 2007; Roch & Steel, 2015). This has led to the recent growth of coalescent-based species tree methods that assume a multispecies coalescent model which can account for gene tree discordance (Edwards, 2009; Edwards et al., 2016; Liu et al., 2015; Ogilvie et al., 2016b). Although a wide variety of approaches are becoming available (e.g. Chifman & Kubatko, 2014; Larget et al., 2010; Mirarab et al., 2014), coalescent-based methods where gene trees and species trees are coestimated in a Bayesian framework have received the widest support (Bryant et al., 2012; Heled & Drummond, 2010; Leaché & Rannala, 2011; Liu, 2008; Liu & Pearl, 2007; Ogilvie et al., 2016a,b). As phylogenies inferred through concatenation and coalescent-based methods are often consistent, some recent studies have questioned the emerging consensus that multispecies coalescent models are consistently superior (Gatesy & Springer, 2014; Springer & Gatesy, 2016; Tonini et al., 2015). Regardless of comparable results, Liu et al. (2015; p. 26) note that

justification for the superiority of coalescent-based methods "lies in their acknowledgement of fundamental genetic processes inherent in all organisms" (i.e., that genes are stochastically independent), resulting in more realistic models for phylogenetic inference. It has also been suggested that concatentation can be viewed as a special case of the multispecies coalescent model in which all gene trees are identical and, therefore, criticism of multispecies coalescent models can be extended to concatention (Edwards et al., 2016). Phylogenetic analyses of large molecular datasets are limited still by a trade-off between model accuracy and computational constraints. Concatenation will remain a useful and important approach to phylogenetic tree reconstruction for many researchers until more computationally efficient models are available for coalescent-based species tree estimation with large datasets (Bayzid & Warnow, 2013; Liu et al., 2015).

In order to more accurately reconstruct the complex patterns of lineage diversification among organisms, the field of molecular phylogenetics is continuing to advance through novel sequencing technologies, increased computational power and improved methods of phylogenetic inference. The availability of tremendous amounts of sequence data and the consequent revision of our understanding of evolutionary history is of significance to many areas of research including systematics, evolutionary biology, historical biogeography, and conservation.

1.3 Titi monkeys (*Callicebus*)

Titi monkeys, *Callicebus* Thomas, 1903, are small to medium-sized (1–2 kg) New World primates of the family Pitheciidae (subfamily Callicebinae). They diverged from their sister clade, the Pitheciinae, in the Miocene, *c.* 20 Ma (Perelman et al., 2011; Schrago et al., 2013; Springer et al., 2012;), thus *Callicebus* comprise of one of the oldest lineages of extant New World primates. They conform to a classic pattern of social monogamy and are characterised by several unique behaviours, such as the antiphonal duet call of the pair-bonded adults, extensive male involvement in infant care, and the affiliative tail-twining behaviour (Fragaszy et al., 1982; Moynihan, 1966; Norconk, 2011; Robinson, 1979).

Callicebus is the most species rich of any primate genus, 31 were listed by Ferrari et al. (2013), and since, two new species have been described (*Callicebus miltoni*, Dalponte et al., 2014; *C. urubambensis*, Vermeer & Tello-Alvarado, 2015),

and *C. toppini* Thomas, 1914, previously considered a synonym of *C. cupreus*, has been reinstated (Vermeer & Tello-Alvarado, 2015). These species form a diverse group of primates, showing interspecific differences in body size, pelage colour, cranial and post-cranial dimensions, and chromosome number (Bueno & Defler, 2010; Hershkovitz, 1988, 1990; Kobayashi, 1995; Van Roosmalen et al., 2002). Although the first species of titis were described in the early 19th century, they were generally included in the genus *Callithrix* É. Geoffroy Saint-Hilaire, 1812. It was almost 100 years before the name *Callicebus* Thomas, 1903, was proposed, which has been in use ever since.

Hershkovitz (1963, 1988, 1990) established the basis for the present classification for the genus. In 1963, he recognised just 10 taxa across two polytypic species (*Callicebus moloch* and *C. torquatus*). This view of titi monkey diversity prevailed until Hershkovitz's revisions in 1988 and 1990. His analysis of around 1,200 museum specimens resulted in the recognition of 25 taxa across five polytypic and eight monotypic species, many of which were resurrected from previous descriptions. He also arranged the taxa in four clusters that he labelled the *modestus*, *donacophilus*, *moloch* and *torquatus* species groups. These reviews by Hershkovitz (1963, 1988, 1990) were peerless studies aiming to investigate and organise all extant forms of *Callicebus*, a difficult task in light of their extensive phenotypic variation, and his work still forms an integral part of much of our knowledge of this genus today.

To infer phylogenetic relationships, Kobayashi (1995) carried out a morphometric analysis of cranial measurements for 23 taxa and modified Hershkovitz's (1988, 1990) species groups. He maintained the *torquatus* and *donacophilus* groups, but included *C. modestus* in the latter, and split the *moloch* group into three: the *personatus* group, the *moloch* group and the *cupreus* group. As other characters, such as pelage colouration, karyotype, and geographic range, were consistent with this classification, he argued that these groups represented monophyletic clades. Kobayashi (1995) suggested that the *donacophilus*, *moloch*, and *cupreus* groups were closely related, while the *personatus* and *torquatus* groups presented a higher degree of character differentiation. The current taxonomic arrangement was established in the review by Van Roosmalen et al. (2002); they followed the species groups proposed by Kobayashi (1995) but listed all 28 recognised taxa as valid species, as proposed by Groves (2001; see also Kobayashi &

5

Langguth, 1999). Five new species have been described since 2002, and *C. toppini* Thomas, 1914, has been reinstated (Vermeer & Tello-Alvarado, 2015).

Species-level classification has focused particularly on pelage colouration (e.g., Groves, 2001, 2005; Hershkovitz, 1963, 1988, 1990; Van Roosmalen et al., 2002), but there are evident limitations to this phenotypic system in light of the considerable intraspecific and within-population variation (e.g., Aquino et al., 2008; Auricchio, 2010; Defler, 2012; Heymann et al., 2002). Intraspecific variation, particularly in traits used as diagnostic characters, has led to controversy surrounding *Callicebus* taxonomy and species identification. Auricchio (2010) analysed museum specimens for 25 species of *Callicebus* and found that *C. moloch*, *C. cupreus* and *C. hoffmannsi* possessed polymorphic phenotypes. Heymann et al. (2002) noted the inconsistency in hand colouration reported for *C. lucifer*, and similarly, Aquino et al. (2008) found phenotypic differences in the throat collar and hand colouration between populations of this species. Defler (2012) also found that individuals on both sides of the Rio Caquetá in Colombia were often indistinguishable despite apparently belonging to two species (*C. lucifer* and *C. lugens*). To comprehend the diversity of the titis, congruency is required between phenotypic traits and additional characters, such as DNA sequence data.

Some recent phylogenetic studies based on large molecular datasets have clarified high-level (genus and family) taxonomic relationships for primates (Jameson Kiesling et al., 2015; Perelman et al., 2011; Springer et al., 2012). These higher-level phylogenies reveal support for some of Kobayashi's (1995) morphological species groups, as well as surprisingly deep divergence dates (Miocene) for the major *Callicebus* clades. However, many specimens were of captive origin and few titi species were included in these studies, limiting their usefulness in inferring species-level relationships. To date, there has been no explicit molecular investigation of the phylogenetic relationships of *Callicebus* species and, consequently, the evolutionary history of titi monkeys remains poorly studied, our understanding of their diversity is limited, and the current taxonomy has yet to be tested using molecular evidence.

1.4 Biogeography of the titi monkeys

Titis have an extensive range spread across nearly all ecogeographic zones inhabitable by non-human primates in the Neotropics (Jameson Kiesling et al., 2015; except Mesoamerica), from the foothills of the northern Andes in Colombia to the tropical forests of the Amazon and Orinoco basins, the Atlantic forest region of Brazil, forest patches in the xerophytic Caatinga of northeast Brazil, and the Beni Plain in northern Bolivia, extending south as far as the Chacoan forests south and east of Santa Cruz in Bolivia and into northeast Paraguay (Defler, 2004; Hershkovitz, 1990; Martínez & Wallace, 2010; Rumiz, 2012; Stallings et al., 1989; Van Roosmalen et al., 2002).

Most of the *Callicebus* species groups show distinct distributional patterns. The *personatus* group are entirely extra-Amazonian and isolated from all other titis by over 500km of drier habitats (the Cerrado shrubby savannas). They are endemic to eastern Brazil from south of the Rio São Francisco as far as the state of São Paulo, predominantly in the Atlantic Forest biome but also in neighbouring arboreal Caatinga regions. The *torquatus* group occur in the Amazon and Orinoco basins from the eastern foothills of the Andes to the Rio Branco and the Rio Purus (north and south of the Rio Amazonas, respectively). The *moloch* and *cupreus* groups occur throughout the southern and western Amazon basin, as far east as the Rio Tocantins, as well as some isolated regions in Colombia. Sympatry among titis occurs between species of *torquatus* and *moloch/cupreus* groups in the Amazon, west of the Rio Purus. The *donacophilus* group occupy forest patches and gallery forests in wooded savannas, the Pantanal, and Chaco scrublands of Bolivia, Brazil, Peru and Paraguay (Ferrari et al., 2013; Hershkovitz, 1990; Printes et al., 2013; Van Roosmalen et al., 2002). In light of their broad and diverse distribution, it is notable that titi monkeys are absent from both Central America and the Guiana Shield (from east of the Rio Branco), and they have a large gap in their distribution in the Cerrado biome of central Brazil.

There have been multiple independent radiations into drier habitats, such as *C. barbarabrownae*, a member of the *personatus* group found in the Caatinga (Printes et al., 2013), and *C. donacophilus* and *C. pallescens*, members of the *donacophilus* group in the Guaporé grasslands and the Paraguayan Chaco scrublands (Ferrari *et al*., 2000; Rumiz, 2012). Larger rivers in Amazonia frequently delimit the geographic distribution of titi monkey species, and a recent study suggested that rivers can act as

isolating barriers for sister taxa, promoting vicariance in this group (Boubli *et al.*, 2015). In light of the above, the biogeography of the genus *Callicebus* is of particular interest in order to better understand their spatial diversification and the processes responsible for shaping present day distributions. There have been few attempts, however, to understand the biogeographic history of titi monkeys, and most existing evaluations are confounded by taxonomic uncertainty and a lack of information regarding species relationships (e.g., Hershkovitz, 1963, 1988; Kinzey, 1982; Kinzey & Gentry, 1979).

1.5 Research objectives

It is over 20 years since Kobayashi (1995) drew our attention to the complete dearth of information surrounding *Callicebus* phylogenetic relationships and evolutionary history, yet, his morphological phylogenetic analysis has remained the only species-level phylogeny available. In the same time-period, *Callicebus* has emerged as the most species-rich of all primate genera (34 species), partially owing to the elevation of all subspecies to species status. Despite the availability of genetic material for many taxa and the rapid advance of molecular phylogenetics, molecular studies have been limited to higher-level primate phylogenies, which do not provide a detailed understanding of species-level relationships within this diverse group. There has been no molecular genetic investigation with a focus on *Callicebus*, and as such, phylogenetic relationships remain contentious and an understanding of their genetic diversity and patterns of diversification is lacking. The paucity of information regarding the evolutionary history of *Callicebus* requires attention as it impedes scientific communication and raises challenges in related fields of research. For example, species-level classification relies on variation in pelage colouration to differentiate between taxa, and consequently, there is controversy surrounding species identification and validity. It is over a decade since the most recent comprehensive taxonomic reviews of the genus (Groves, 2005; Van Roosmalen et al., 2002), and these proposals have yet to be evaluated using molecular evidence. Explicit phylogenetic hypotheses are also a necessary component for understanding the patterns of spatial diversification, and thus, study of the biogeography of titis is hindered by the absence of a species-level phylogeny.

The encompassing goal of this thesis was, therefore, to employ molecular data to provide insight into the evolutionary history and biogeography of Callicebinae, one of the most strikingly poorly studied groups of primates. The following objectives were identified:

- Generate a multi-locus molecular dataset using Sanger sequencing to infer phylogeny and divergence times for *Callicebus* species and species groups, and revise the taxonomy of the genus based on molecular and morphological evidence (Chapter 2).

- Employ a statistical biogeographical approach to perform ancestral-area estimations across a time-calibrated phylogeny using the multi-locus data and reconstruct titi monkey biogeographic history (Chapter 3).

- Generate a genome-wide molecular dataset using double digest restriction-site associated DNA sequencing (ddRADseq) to infer phylogeny and divergence times, assess genetic structure, and test for interspecific gene flow (Chapter 4).

Chapter 2: Phylogenetic relationships of the New World titi monkeys (*Callicebus*): First appraisal of taxonomy based on molecular evidence

2.1 Abstract

Titi monkeys, *Callicebus*, comprise the most species-rich primate genus—34 species are currently recognised, five of them described since 2005. The lack of molecular data for titi monkeys has meant that little is known of their phylogenetic relationships and divergence times. To clarify their evolutionary history, we assembled a large molecular dataset by sequencing 20 nuclear and two mitochondrial loci for 15 species, including representatives from all recognised species groups. Phylogenetic relationships were inferred using concatenated maximum likelihood and Bayesian analyses, allowing us to evaluate the current taxonomic hypothesis for the genus. Our results show four distinct *Callicebus* clades, for the most part concordant with the currently recognised morphological species-groups—the *torquatus* group, the *personatus* group, the *donacophilus* group, and the *moloch* group. The *cupreus* and *moloch* groups are not monophyletic, and all species of the formerly recognized *cupreus* group are reassigned to the *moloch* group. Two of the major divergence events are dated to the Miocene. The *torquatus* group, the oldest lineage, diverged *c.* 11 Ma; and the Atlantic forest *personatus* group split from the ancestor of all *donacophilus* and *moloch* species at 8–9 Ma. Considering molecular, morphological and biogeographic evidence, we propose a new genus level taxonomy for titi monkeys: *Cheracebus* (Byrne et al., 2016) in the Orinoco, Negro and upper Amazon basins (*torquatus* group), *Callicebus* Thomas, 1903, in the Atlantic Forest (*personatus* group), and *Plecturocebus* (Byrne et al., 2016) in the Amazon basin and Chaco region (*donacophilus* and *moloch* species groups).

2.2 Introduction

2.2.1 Background

Titi monkeys, *Callicebus* Thomas, 1903, are small to medium-sized (1–2 kg) New World primates of the family Pitheciidae. They comprise an old platyrrhine radiation that diverged from their sister clade, the Pitheciinae, in the Miocene, *c*. 20 Ma (Perelman et al., 2011; Schrago et al., 2013; Springer et al., 2012). *Callicebus* species occur only in South America, with an extensive range from the foothills of the northern Andes in Colombia to the tropical forests of the Amazon and upper Orinoco basins, the Atlantic forest region of Brazil, forest patches in the xerophytic Caatinga of northeast Brazil, and the Beni Plain in northern Bolivia, extending south as far as the Chacoan forests south and east of Santa Cruz in Bolivia and into northeast Paraguay (Defler, 2004; Hershkovitz, 1990; Martínez & Wallace, 2010; Rumiz, 2012; Stallings et al., 1989; Van Roosmalen et al., 2002).

Callicebus is the most species rich of any primate genus; 31 were listed by Ferrari et al. (2013). Two new species have been described since then, *Callicebus miltoni* Dalponte et al., 2014 and *C. urubambensis* Vermeer & Tello-Alvarado, 2015. Vermeer & Tello-Alvarado (2015) also reinstated *C. toppini* Thomas, 1914, for long incorrectly considered a synonym of *C. cupreus*. These 34 titi species form a highly diverse group of primates, showing interspecific differences in body size, pelage colour, cranial dimensions, and chromosome number (Bueno & Defler, 2010; Hershkovitz, 1988, 1990; Kobayashi, 1995; Van Roosmalen et al., 2002). Kobayashi (1995) employed cranial morphometrics to propose the current species-group arrangement for *Callicebus* taxa, which he suggested was consistent with variation in other characters, such as pelage colouration, karyotype, and geographic range. Species-level classification, however, has focused particularly on pelage colouration (e.g., Groves, 2001, 2005; Hershkovitz, 1963, 1988, 1990; Van Roosmalen et al., 2002), but there are evident limitations to this phenotypic system in light of the considerable intraspecific and within-population variation (e.g., Aquino *et al.*, 2008; Auricchio, 2010; Defler, 2012; Heymann *et al.*, 2002). To comprehend the real taxonomic diversity of the titis, congruency is required between phenotypic traits and additional characters, such as DNA sequence data.

Some recent phylogenetic studies based on large molecular datasets have clarified high-level (genus and family) taxonomic relationships for primates (Jameson

Kiesling et al., 2015; Perelman et al., 2011; Springer et al., 2012). These higher-level phylogenies reveal surprisingly deep divergence dates (Miocene) for the major *Callicebus* clades. However, most specimens were of captive origin and rather few titi species were included in these studies, limiting their usefulness in inferring species-level relationships. To date, there has been no explicit molecular investigation of the phylogenetic relationships of *Callicebus* species and, consequently, the evolutionary history of titi monkeys remains poorly studied. The current taxonomy has yet to be tested using molecular evidence.

Here, we present a molecular phylogeny of the genus *Callicebus* based on DNA sequence data from 20 independent nuclear loci and two mitochondrial loci. In taking a molecular approach, we investigate phylogenetic relationships and divergence times among 15 species (with representatives of all species groups *sensu* Kobayashi, 1995) using concatenated Bayesian and maximum likelihood (ML) analyses. In contrast to high-level primate phylogenies (Jameson Kiesling et al., 2015; Perelman et al., 2011; Springer et al., 2012), most of the *Callicebus* species included this study are represented by multiple wild-caught specimens of known provenance and taxonomic identification. Taking into account the results from our phylogenetic analyses, as well as morphological and biogeographic evidence, we suggest a revised taxonomy that recognises three genera of titi monkey in the subfamily Callicebinae that are largely coherent with Kobayashi's (1995) morphological species groups. Below, we review changes to the taxonomy of the titis since Hershkovitz's (1963, 1988, 1990) reviews.

2.2.2 *Callicebus* taxonomy

Simia Linnaeus, C. 1758. *Syst. Nat.* 10th ed., 1: 25. In part. Humboldt, A. von. 1811. *Rec. Obs. Zool. Anat. Comp.* 1: 319. *Simia lugens* (= *Callicebus lugens*).

Cebus Erxleben, C. P. 1777. *Systema Regni Anim. Mammalia*, p. 44. In part. Hoffmannsegg, G. von. 1807. *Mag. Ges. Naturf. Freunde*, Berlin, 9: 97. *Cebus moloch* (= *Callicebus moloch*).

Callitrix Hoffmannsegg, G. von. 1807. *Mag. Ges. Naturf. Fr.*, Berlin, 10: 86. Type species by monotypy *Callitrix torquata* Hoffmannsegg. Name pre-occupied by *Callitrix* Desmarest, 1804, a junior synonym of *Cebus* Erxleben, 1777.

Callithrix Geoffroy Saint-Hilaire, É. 1812. Suite en Tableau des Quadrumanes. *Ann. Mus. Hist. Nat. Paris*, 19: 112. Included *Callithrix sciureus* (Linnaeus) (= *Saimiri*

sciureus), *Callithrix personnatus* [sic] É. Geoffroy Saint-Hilaire (= *Callicebus personatus*), *Callithrix lugens* (Humboldt), *Callithrix amictus* É. Geoffroy Saint-Hilaire, *Callithrix torquatus* (Hoffmannsegg), and *Callithrix moloch* (Hoffmannsegg). Name pre-occupied by *Callithrix* Erxleben, 1777, for the marmosets, Callitrichidae Thomas, 1903.

Saguinus Lesson, R. P. 1827. *Manuel de mammalogie*. J. B. Baillière, Paris: 56. Included all species listed by É. Geoffroy Saint-Hilaire (1812) for *Callithrix*, along with *Saguinus melanochir* (Weid-Neuwied) (= *Callicebus melanochir*), and *Saguinus infulatus* Kuhl (= *Aotus infulatus*). Name pre-occupied by *Saguinus* Hoffmannsegg, 1807, for the tamarins, Callitrichidae.

Callicebus Thomas, O. 1903. *Ann. Mag. Nat. Hist.*, 7th series, 12: 456. Type species *Simia personata* É. Geoffroy Saint-Hilaire, 1812.

In the 1800s, titis were generally included in the genus *Callithrix* É. Geoffroy Saint-Hilaire, 1812. Thomas (1903) pointed out that the name was pre-occupied by *Callithrix* Erxleben, 1777 (the currently accepted generic epithet for the marmosets) and proposed the name *Callicebus* Thomas, 1903, which has been in use ever since.

Goodman et al. (1998) suggested that members of the *torquatus* species group should be placed in a subgenus due to the last common ancestor with *Callicebus moloch* having an estimated age of more than 6 Ma. They suggested the name *Torquatus*. Groves (2001, 2005) listed *Torquatus* as a subgenus of *Callicebus* Thomas, 1903, with *Callicebus torquatus* (Hoffmannsegg, 1807), as the type species. As pointed out by Groves himself (in litt.), Goodman et al.'s (1998) suggestion of the name *Torquatus*, as proposed, does not conform to the requirements of Article 13 of the *International Code of Zoological Nomenclature* (ICZN, 1999): Names published after 1930. 13.1. "To be available, every new name published after 1930 must satisfy the provisions of Article 11 and must − 13.1.1 be accompanied by a description or definition that states in words characters that are purported to differentiate the taxon, or − 13.1.2 be accompanied by a bibliographic reference to such a public statement […], or − 13.1.3 be proposed expressly as a new replacement name (*nomen novum*) for an available name […]". Thus the name *Torquatus* is a *nomen nudum*, and unavailable.

2.2.3 Species and species groups

Elliot (1913), Cabrera (1958), and Hill (1960) listed 22–34 titi monkeys, of which 22 are considered valid taxa today. Hershkovitz (1963, 1988, 1990) subsequently established the basis for the present classification for the genus. In 1963, he recognised just 10 taxa across two polytypic species (*Callicebus moloch* and *C. torquatus*). Although the Atlantic forest *C. personatus* taxa were not included in this early review, Hershkovitz (1963) suggested that they were subspecies of *C. moloch*. This view of titi monkey diversity prevailed until Hershkovitz's revisions in 1988 and 1990. His analysis of around 1,200 museum specimens resulted in the recognition of 25 taxa across five polytypic and eight monotypic species, which he arranged in four clusters that he labelled the *modestus*, *donacophilus*, *moloch* and *torquatus* species groups (Table 2.1) (Hershkovitz, 1988, 1990).

To infer phylogenetic relationships, Kobayashi (1995) carried out a morphometric analysis of cranial measurements for 23 taxa, and modified Hershkovitz's (1988, 1990) species groups. He maintained the *torquatus* and *donacophilus* groups, but included *C. modestus* in the latter. He split the *moloch* group into three: the *personatus* group, the *moloch* group and the *cupreus* group (Table 2.1). As other characters, such as pelage colouration, karyotype, and geographic range, were consistent with this classification, he argued that these groups represented phylogenetically independent clades. Kobayashi (1995) suggested that the *donacophilus*, *moloch*, and *cupreus* groups were closely related, while the *personatus* and *torquatus* groups presented a higher degree of character differentiation. Based upon the occlusal pattern of the upper molars, the *torquatus* group was proposed as the earliest lineage (Kobayashi, 1990).

Table 2.1 The taxonomy of the titis. Taxonomic arrangement for *Callicebus* taxa as proposed by Hershkovitz (1963); Hershkovitz (1988, 1990); Kobayashi (1995) and Kobayashi & Langguth (1999); Van Roosmalen et al. (2002); Groves (2005); and the present study. Taxa included in this study are denoted with an asterisk. Classification for species not included in this study follows Groves (2005), and species described and reinstated after Groves (2005); (Dalponte et al., 2014; Defler et al., 2010; Gualda-Barros et al., 2012; Vermeer & Tello-Alvarado, 2015; Wallace et al., 2006), with the exception of *modestus* where we follow Kobayashi (1995).

Hershkovtiz (1988, 1990)	Kobayashi (1995)	Van Roosmalen et al. (2002)	Groves (2005)	Present study
Genus *Callicebus*	Genus *Callicebus*	Genus *Callicebus*	Genus *Callicebus*	Genus *Cheracebus*
--	--	--	Subgenus *Torquatus*	--
torquatus group	*torquatus* group	*torquatus* group	*torquatus* group	--
C. torquatus torquatus	*C. torquatus torquatus*	*C. torquatus*	*C. torquatus*	*C. torquatus*
C. t. lugens	*C. t. lugens*	*C. lugens*	*C. lugens*	*C. lugens**
C. t. lucifer	*C. t. lucifer*	*C. lucifer*	*C. lucifer*	*C. lucifer*
C. t. purinus	*C. t. purinus*	*C. purinus*	*C. purinus*	*C. purinus**
C. t. regulus	*C. t. regulus*	*C. regulus*	*C. regulus*	*C. regulus*
C. t. medemi	*C. t. medemi*	*C. medemi*	*C. medemi*	*C. medemi*
--	--	--	Subgenus *Callicebus*	Genus *Callicebus*
moloch group	*personatus* group	*personatus* group	*personatus* group	--
C. personatus personatus	*C. personatus*	*C. personatus*	*C. personatus*	*C. personatus**
C. p. melanochir	*C. melanochir*	*C. melanochir*	*C. melanochir*	*C. melanochir*
C. p. nigrifrons	*C. nigrifrons*	*C. nigrifrons*	*C. nigrifrons*	*C. nigrifrons**
C. p. barbarabrownae	*C. barbarabrownae*	*C. barbarabrownae*	*C. barbarabrownae*	*C. barbarabrownae*
--	*C. coimbrai*[1]	*C. coimbrai*	*C. coimbrai*	*C. coimbrai**
--	--	--	--	Genus *Plecturocebus*
--	*moloch* group	*moloch* group	*moloch* group	*moloch* group
C. moloch	*C. moloch*	*C. moloch*	*C. moloch*	*P. moloch**
C. cinerascens	*C. cinerascens*	*C. cinerascens*	*C. cinerascens*	*P. cinerascens**
C. hoffmannsi hoffmannsi	*C. hoffmannsi hoffmannsi*	*C. hoffmannsi*	*C. hoffmannsi*	*P. hoffmannsi**
C. h. baptista	*C. h. baptista*	*C. baptista*	*C. baptista*	*P. baptista*
--	--	*C. bernhardi*	*C. bernhardi*	*P. bernhardi**
C. brunneus	*C. brunneus*	*C. brunneus*	*C. brunneus*	*P. brunneus**
--	*cupreus* group	*cupreus* group	--	--
C. cupreus cupreus	*C. cupreus cupreus*	*C. cupreus*	*C. cupreus*	*P. cupreus**
C. c. discolor	*C. c. discolor*	*C. discolor*	*C. discolor*	*P. discolor*
C. c. ornatus	*C. c. ornatus*	*C. ornatus*	*C. ornatus*	*P. ornatus*
C. caligatus	*C. caligatus*	*C. caligatus*	*C. caligatus*	*P. caligatus**
C. dubius	*C. dubius*[2]	*C. dubius*	*C. dubius*	--
--	--	*C. stephennashi*	*C. stephennashi*	*P. stephennashi*
--	--	--	--	*P. aureipalatii*
--	--	--	--	*P. caquetensis*
--	--	--	--	*P. vieirai*
--	--	--	--	*P. miltoni**
--	--	--	--	*P. toppini*
donacophilus group	*donacophilus* group	*donacophilus* group	*donacophilus* group	*donacophilus* group
C. donacophilus donacophilus	*C. donacophilus donacophilus*	*C. donacophilus*	*C. donacophilus*	*P. donacophilus**
C. d. pallescens	*C. d. pallescens*	*C. pallescens*	*C. pallescens*	*P. pallescens*
C. oenanthe	--	*C. oenanthe*	*C. oenanthe*	*P. oenanthe*
C. olallae	*C. olallae*	*C. olallae*	*C. olallae*	*P. olallae*
--	--	--	--	*P. urubambensis*
modestus group	--	--	*modestus* group	--
C. modestus	*C. modestus*	*C. modestus*	*C. modestus*	*P. modestus*
25 taxa	25 taxa	28 species	28 species	33 species

[1]Kobayashi & Langguth (1999)

[2]Species group undetermined

The distinctiveness of the *torquatus* group has long been recognised; *C. torquatus* was one of the two species in Hershkovitz's first appraisal in 1963. He considered it polytypic, with three subspecies: *C. t. torquatus* (Hoffmannsegg, 1807); *C. t. lugens* (Humboldt, 1811); and *C. t. medemi* Hershkovitz, 1963. Hershkovitz (1988, 1990) subsequently resurrected three other taxa: *lucifer* Thomas, 1914; *regulus* Thomas, 1927; and *purinus* Thomas, 1927—all as subspecies of *torquatus*. As of 1990, therefore, the *torquatus* group consisted of a single species with six subspecies. Groves (2001) listed *medemi* as a species, but otherwise followed Hershkovitz in maintaining the remaining forms as subspecies of *torquatus*. Van Roosmalen et al. (2002) and Groves (2005) classified all members of the *torquatus* group as species. Taking note of the suggestion of Goodman et al. (1998), Groves (2005) placed the members of the *torquatus* group in the subgenus *Torquatus* (all other titis in the subgenus *Callicebus*), although, as mentioned, he subsequently realised that the name as suggested by Goodman et al. (1998) was a *nomen nudum*.

Hershkovitz (1988) recognised three subspecies of *C. personatus; C. p. personatus* (É. Geoffroy Saint-Hilaire, 1812); *C. p. melanochir* (Wied-Neuwied, 1820); and *C. p. nigrifrons* (Spix, 1823). He indicated that they could be considered subspecies of *C. moloch*, and placed them in his *moloch* species group (1988, 1990). In his 1990 revision, he described another subspecies from northeast Brazil, *C. p. barbarabrownae*. Kobayashi (1995) continued to recognise these four titis as subspecies but placed them in a separate species group, based on the high degree of character differentiation between *C. personatus* and other *Callicebus* taxa. Kobayashi & Langguth (1999) described *C. coimbrai*, a member of the *personatus* group from northeast Brazil, and determined that all members of the *personatus* group be considered distinct species.

The craniometric study of Kobayashi (1995) showed that the *donacophilus*, *moloch*, and *cupreus* groups are more closely related to each other than they are to the *torquatus* and *personatus* groups. This is reflected in the early history of their taxonomy. Hershkovitz (1963) recognised a single species with seven subspecies in his *moloch* group: *C. moloch moloch* (Hoffmannsegg, 1807); *C. m. cupreus* (Spix, 1823); *C. m. donacophilus* (d'Orbigny, 1836); *C. m. brunneus* (Wagner, 1842); *C. m. discolor* (I. Geoffroy & Deville, 1848); *C. m. ornatus* (Gray, 1866); and *C. m. hoffmannsi* Thomas, 1908. Hershkovitz's subsequent revisions (1988, 1990) resulted in the description of a new species, *dubius* Hershkovitz, 1988, and the reinstatement

of *cinerascens* Spix, 1823, *caligatus* Wagner, 1842, *modestus* Lönnberg, 1939, *olallae* Lönnberg, 1939, *baptista* Lönnberg, 1939, *pallescens* Thomas, 1907, and *oenanthe* Thomas, 1924, as valid taxa. Excluding the *C. personatus* subspecies, Hershkovitz (1990) listed 15 species and subspecies, and classified them into three species groups; the *modestus* group, the *donacophilus* group, and the *moloch* group (Table 2.1). Groves (2001) maintained the species groups of Hershkovitz (1990), but raised all the *donacophilus* and *moloch* (but not *C. personatus*) group members to species. In his review, Groves questioned the distinction between *C. cupreus*, *C. caligatus*, *C. discolor*, and *C. dubius*, and placed the latter three as synonyms of *C. cupreus*. Groves (2005), however, subsequently accepted them as valid species.

The current taxonomic arrangement was established in the review by Van Roosmalen et al. (2002). They followed the species groups proposed by Kobayashi (1995) but listed all recognised taxa as species, as proposed by Groves (2001, 2005; see also Kobayashi & Langguth, 1999). Van Roosmalen et al. (2002) described *C. bernhardi* and *C. stephennashi*, belonging to the *moloch* and *cupreus* groups, respectively. Five new species have been described since 2002; *C. aureipalatii* Wallace et al., 2006, *C. caquetensis* Defler et al., 2010, *C. vieirai* Gualda-Barros et al., 2012, *C. miltoni* Dalponte et al., 2014, in the *moloch* and *cupreus* groups, and *C. urubambensis* Vermeer & Tello-Alvarado, 2015, assigned to the *donacophilus* group. Vermeer & Tello-Alvarado (2015) also reinstated *C. toppini* Thomas, 1914, as a member of the *cupreus* group.

2.3 Methods

2.3.1 Taxon sampling

A total of 50 fresh tissue samples were collected from museum voucher specimens from the following Brazilian institutions: National Institute of Amazonian Research (INPA), Federal University of Pará (UFPA), Federal University of Rondônia (UNIR), Federal University of Amazonas (UFAM) and the Goeldi Museum (MPEG). The majority of these specimens were obtained in the context of an Amazonian-wide faunal inventory project (CNPq/SISBIOTA) carried out in accordance with the appropriate collection permits (IBAMA 483 license No. 005/2005 – CGFAU/LIC). This research adhered to the American Society of Primatologists' and American

Society of Mammalogists' principles for the ethical treatment of primates, and Brazilian laws that govern primate research.

Fifteen species of *Callicebus* were sampled, including representatives from each of the species groups of Kobayashi (1995), and five platyrrhine species were selected as outgroup taxa. A complete list of *Callicebus* and outgroup species is presented in Table 2.2. We generated novel sequence data for a total of 49 *Callicebus* and 1 outgroup sample (JPB100, *Cebus albifrons*). All samples used in this study were from wild specimens, nearly all of which are of known provenance, and morphologically identified following Hershkovitz (1988, 1990), Van Roosmalen et al. (2002), and Dalponte et al. (2014). Three of these samples are from a new species of *Callicebus* from the Alta Floresta region of Mato Grosso, Brazil (Boubli et al., in prep), that is closely related to *C. moloch* based on geographic location and pelage colouration, and is classified here as *C.* cf. *moloch*.

We retrieved additional sequences from GenBank representing six *Callicebus* and five outgroup samples from Perelman et al. (2011), and another four *Callicebus* and four outgroup individuals. A total of 59 *Callicebus* and 10 outgroup individuals were included in this study. Additional information for all samples is presented in Table 2.2.

Of the six *Callicebus* specimens retrieved from the Perelman et al. (2011) study, our molecular datasets confirm the taxonomic validity of *C. nigrifrons* (CNI-1) and show that *C. moloch* (CMH-1) and *C. caligatus* (CCG-1) are incorrectly identified. The *C. moloch* (CMH-1) specimen is most similar to our *C. hoffmannsi* individuals and *C. caligatus* (CCG-1) is very closely related to their *C. donacophilus* specimen (CDO-1). These samples are classified as *C.* cf. *hoffmannsi* and *C. donacophilus*, respectively (Table 2.2), but we note that these samples are of captive origin and could be captive hybrids.

Table 2.2 List of genetic samples used in this study including ID, source and corresponding dataset.

Species	Sample ID	Col.	Species group	Wild or captive	Geographic origin or sample source	Data	Sample notes
Callicebus moloch	MCB63	UFPA	*moloch*	Wild	Senador José Porfírio, R bank of the Rio Xingu, Pará, Brazil	nDNA, mtDNA	
Callicebus moloch	MCB64	UFPA	*moloch*	Wild	Senador José Porfírio, R bank of the Rio Xingu, Pará, Brazil	nDNA, mtDNA	
Callicebus moloch	MCB79	UFPA	*moloch*	Wild	Senador José Porfírio, R bank of the Rio Xingu, Pará, Brazil	nDNA, mtDNA	
Callicebus moloch	857	UFPA	*moloch*	Wild	Tucuruí Dam, L bank of the Rio Tocantins, Pará, Brazil	nDNA, mtDNA	
Callicebus moloch	1516	UFPA	*moloch*	Wild	Tucuruí Dam, L bank of the Rio Tocantins, Pará, Brazil	nDNA, mtDNA	
Callicebus moloch	CTGAM420	UFAM	*moloch*	Wild	Belterra, R bank of the Rio Tapajós, Pará, Brazil	nDNA, mtDNA	
Callicebus moloch	CTGAM421	UFAM	*moloch*	Wild	Belterra, R bank of the Rio Tapajós, Pará, Brazil	nDNA, mtDNA	
Callicebus cf. *moloch*	RVR22	INPA	*moloch*	Wild	Novo Horizonte Community, Alta Floresta, Mato Grosso, Brazil	nDNA, mtDNA	
Callicebus cf. *moloch*	RVR68	INPA	*moloch*	Wild	Novo Horizonte Community, Alta Floresta, Mato Grosso, Brazil	nDNA, mtDNA	
Callicebus cf. *moloch*	RVR73	INPA	*moloch*	Wild	Novo Horizonte Community, Alta Floresta, Mato Grosso, Brazil	nDNA, mtDNA	
Callicebus bernhardi	UFRO413	UNIR	*moloch*	Wild	Machadinho D'Oeste, Rondônia, Brazil	nDNA, mtDNA	
Callicebus bernhardi	42960	MPEG	*moloch*	Wild	São Francisco do Guaporé, Guaporé Biological Reserve, Rondônia, Brazil	nDNA, mtDNA	
Callicebus bernhardi	42961	MPEG	*moloch*	Wild	São Francisco do Guaporé, Guaporé Biological Reserve, Rondônia, Brazil	nDNA, mtDNA	
Callicebus bernhardi	42964	MPEG	*moloch*	Wild	São Francisco do Guaporé, Guaporé Biological Reserve, Rondônia, Brazil	nDNA, mtDNA	
Callicebus miltoni	42991	MPEG	*moloch*	Wild	Novo Aripuanã, L bank of the Rio Aripuanã, Amazonas, Brazil	nDNA, mtDNA	
Callicebus miltoni	42992	MPEG	*moloch*	Wild	Novo Aripuanã, L bank of the Rio Aripuanã, Amazonas, Brazil	nDNA, mtDNA	
Callicebus miltoni	42993	MPEG	*moloch*	Wild	Novo Aripuanã, L bank of the Rio Aripuanã, Amazonas, Brazil	nDNA, mtDNA	
Callicebus cinerascens	UFRO352	UNIR	*moloch*	Wild	Rondon II Dam, Pimenta Bueno, Rondônia, Brazil	nDNA, mtDNA	
Callicebus cinerascens	UFRO355	UNIR	*moloch*	Wild	Rondon II Dam, Pimenta Bueno, Rondônia, Brazil	nDNA, mtDNA	
Callicebus cinerascens	UFRO499	UNIR	*moloch*	Wild	Cabixi, Rondônia, Brazil	nDNA, mtDNA	
Callicebus hoffmannsi	CTGAM248	UFAM	*moloch*	Wild	Cametá Community, L bank of the Rio Tapajós, Pará, Brazil	nDNA, mtDNA	
Callicebus hoffmannsi	CTGAM290	UFAM	*moloch*	Wild	Cametá Community, L bank of the Rio Tapajós, Pará, Brazil	nDNA, mtDNA	
Callicebus hoffmannsi	01CNP	UFPA	*moloch*	---	No location data	nDNA, mtDNA	
Callicebus cf. *hoffmannsi*	CMH1	---	*moloch*	Captive	Perelman et al. (2011): Centro Nacional de Primatas.	nDNA	*C. moloch* (Perelman et al., 2011)
Callicebus cupreus	AAM15	INPA	*moloch*	Wild	RESEX Catuá-Ipixuna, Lago do Ipixuna, Coari, Amazonas, Brazil	nDNA, mtDNA	*C. cupreus* clade A
Callicebus cupreus	CTGAM210	UFAM	*moloch*	Wild	Rebio Abufari, Turiaçu, L bank of the Rio Purus, Amazonas, Brazil	nDNA, mtDNA	*C. cupreus* clade A
Callicebus cupreus	JLP15920	INPA	*moloch*	Wild	Itamarati, L bank of the Rio Jurua, Amazonas, Brazil	nDNA, mtDNA	*C. cupreus* clade A
Callicebus cupreus	4984	UFPA	*moloch*	Wild	No location data	nDNA, mtDNA	*C. cupreus* clade B
Callicebus cupreus	4988	UFPA	*moloch*	Wild	No location data	nDNA, mtDNA	*C. cupreus* clade B
Callicebus cupreus	4990	UFPA	*moloch*	Wild	No location data	nDNA, mtDNA	*C. cupreus* clade B
Callicebus cupreus	4993	UFPA	*moloch*	Wild	No location data	nDNA, mtDNA	*C. cupreus* clade B
Callicebus brunneus	UFRO541	UNIR	*moloch*	Wild	Porto Velho, R bank of the Rio Madeira, Rondônia, Brazil	nDNA, mtDNA	
Callicebus brunneus	4009	UFPA	*moloch*	Wild	Samuel Dam, L bank of the Rio Jamari, Rondônia, Brazil	nDNA, mtDNA	
Callicebus brunneus	4019	UFPA	*moloch*	Wild	Samuel Dam, L bank of the Rio Jamari, Rondônia, Brazil	nDNA, mtDNA	
Callicebus brunneus	4346	UFPA	*moloch*	Wild	Samuel Dam, L bank of the Rio Jamari, Rondônia, Brazil	nDNA, mtDNA	
Callicebus brunneus	4505	UFPA	*moloch*	Wild	Samuel Dam, R bank of the Rio Jamari, Rondônia, Brazil	nDNA, mtDNA	
Callicebus dubius	UFRO403	UNIR	*moloch*	Wild	Porto Velho, L bank of the Rio Madeira, Rondônia, Brazil	nDNA, mtDNA	
Callicebus dubius	UFRO544	UNIR	*moloch*	Wild	Porto Velho, L bank of the Rio Madeira, Rondônia, Brazil	nDNA, mtDNA	
Callicebus caligatus	CTGAM181	UFAM	*moloch*	Wild	Tapauá, Igarapé do Jacinto, R bank of the Rio Purus, Amazonas, Brazil	nDNA, mtDNA	
Callicebus caligatus	CTGAM182	UFAM	*moloch*	Wild	Tapauá, Igarapé do Jacinto, R bank of the Rio Purus, Amazonas, Brazil	nDNA, mtDNA	

Table 2.2 cont'd List of genetic samples used in this study including ID, source and corresponding dataset.

Species	Sample ID	Col.	Species group	Wild or captive	Geographic origin or sample source	Data	Sample notes
Callicebus caligatus	MVR58	INPA	*moloch*	Wild	No location data	mtDNA	
Callicebus caligatus	CCM248	INPA	*moloch*	Wild	No location data	mtDNA	
Callicebus donacophilus	CDO1	---	*donacophilus*	Captive	Perelman et al. (2011): Center for Reproduction of Endangered Species	nDNA	
Callicebus donacophilus	--	---	*donacophilus*	---	GenBank (accession number FJ785423)	mtDNA	
Callicebus donacophilus	CCG1	---	*donacophilus*	Captive	Perelman et al. (2011): Center for Reproduction of Endangered Species	nDNA	*C. caligatus* (Perelman et al. 2011)
Callicebus personatus	CLP1	---	*personatus*	Wild	Perelman et al. (2011): Espirito Santo State, Brazil.	nDNA	
Callicebus personatus	--	---	*personatus*	---	GenBank (accession number AF289988)	mtDNA	
Callicebus coimbrai	CCO1	---	*personatus*	Wild	Perelman et al. (2011): Bahia state, Brazil.	nDNA	
Callicebus nigrifrons	CNI1	---	*personatus*	Wild	Perelman et al. (2011): Carangola, Minas Gerais State, Brazil.	nDNA	
Callicebus nigrifrons	CPE04	---	*personatus*	Wild	Minas Gerais, Brazil	nDNA, mtDNA	
Callicebus nigrifrons	CPRJ1493	---	*personatus*	---	GenBank (accession number AF524884)	mtDNA	
Callicebus lugens	JPB81	INPA	*torquatus*	Wild	Mandiquie, R bank of the Rio Negro, Amazonas, Brazil	nDNA, mtDNA	
Callicebus lugens	CRB2698	---	*torquatus*	---	GenBank (DQ337708) R bank of the Rio Negro	mtDNA	
Callicebus lugens	JPB119	INPA	*torquatus*	Wild	Marari, L bank of the Rio Negro, Amazonas, Brazil	nDNA, mtDNA	
Callicebus lugens	JPB124	INPA	*torquatus*	Wild	Pé da Serra do Aracá, L bank of the Rio Negro, Amazonas, Brazil	nDNA, mtDNA	
Callicebus lugens	JPB136	INPA	*torquatus*	Wild	Igarapé Cuieiras, L bank of the Rio Negro, Amazonas, Brazil	nDNA, mtDNA	
Callicebus purinus	CTGAM154	UFAM	*torquatus*	Wild	Rebio Abufari, L bank of the Rio Purus, Amazonas, Brazil	nDNA, mtDNA	
Callicebus purinus	CTGAM195	UFAM	*torquatus*	Wild	Rebio Abufari, L bank of the Rio Purus, Amazonas, Brazil	nDNA, mtDNA	
Callicebus purinus	CTGAM209	UFAM	*torquatus*	Wild	Rebio Abufari, L bank of the Rio Purus, Amazonas, Brazil	nDNA, mtDNA	
Cebus albifrons	JPB100	INPA	Outgroup	Wild	No location data	nDNA, mtDNA	
Cebus albifrons	CEA1	---	Outgroup	Wild	Perelman et al. (2011): Instituto Nacional de Cancer.	nDNA	
Saimiri sciureus	SSC7	---	Outgroup	Captive	Perelman et al. (2011): Schwerin Zoo, Germany.	nDNA	
Saimiri sciureus	--	---	Outgroup	---	GenBank (accession number HQ644334)	mtDNA	
Cacajao calvus	CCL1	---	Outgroup	Wild	Perelman et al. (2011): Köln Zoo, Germany.	nDNA	
Cacajao calvus	--	---	Outgroup	---	GenBank (accession number NC_021967)	mtDNA	
Chiropotes israelita	CIS1	---	Outgroup	Wild	Perelman et al. (2011): Barcelos, Amazonas State, Brazil.	nDNA	
Chiropotes israelita	--	---	Outgroup	---	GenBank (accession number NC_024629)	mtDNA	
Pithecia pithecia	PPT1	---	Outgroup	Captive	Perelman et al. (2011): Centro Nacional de Primatas.	nDNA	
Pithecia pithecia	--	---	Outgroup	---	GenBank (accession number JF459229)	mtDNA	

Collection abbreviations: UFPA = Federal University of Pará; UFAM = Federal University of Amazonas; INPA = National Institute for Amazonian Research; UNIR = Federal University of Rondônia; MPEG = Goeldi Museum

2.3.2 Molecular dataset

DNA sequence data were obtained from a total of 22 loci. We selected primers for 20 independent nuclear loci from Perelman et al. (2011) based on their performance for *Callicebus*. Most of these primers were designed for the Perelman et al. (2011) study, but some originated in previous studies (Horvath et al., 2008; Murphy et al., 2001; Teeling et al., 2000; Venta et al., 1996). The nuclear regions included exons, introns, and 3'UTRs, and two loci located on the X chromosome (Table 2.3). We also obtained DNA sequence data from two mitochondrial loci; we amplified the cytochrome *b* gene (CYTB) with novel primers designed for this study, and cytochrome *c* oxidase I (COI) using previously designed primers (Ward et al., 2005).

A total of 944 new sequences (nuclear and mitochondrial) were generated for this study from three laboratories: Universidade Federal do Pará (UFPA), Pará, Brazil; University of Salford, Manchester, UK; and the Evolution and Animal Genetics Laboratory (LEGAL), Universidade Federal do Amazonas (UFAM), Amazonas, Brazil. We retrieved an additional 209 nuclear sequences for the 11 individuals sequenced for Perelman et al. (2011) from GenBank, and 12 mitochondrial sequences from GenBank (accession numbers are listed in Table 2.4).

Three datasets were compiled from subsets of loci and samples: the nuclear dataset composed of the 20 nuclear loci totalling 12,778 bp in length; the combined dataset including all 22 loci totalling 14,578 bp in length; and the mitochondrial dataset composed of the two mitochondrial loci and a length of 1,800 bp. The nuclear and combined datasets were composed of the same set of samples, containing 47 *Callicebus* and one outgroup sequenced for this study, and the 6 *Callicebus* and 5 outgroup individuals from Perelman et al. (2011). The mitochondrial dataset included all 50 newly sequenced samples, as well as an additional eight individuals retrieved from GenBank. All *Callicebus* and outgroup species are represented in each dataset, with the exception *C.* cf. *hoffmannsi* and *C. coimbrai* (nuclear and combined only). A list of samples and number of loci sequenced for the nuclear and combined datasets is presented in Table 2.5 and for the mitochondrial dataset in Table 2.6. A summary of each dataset is presented in Table 2.7.

Table 2.3 List of the 22 loci used in this study and primer information.

Locus ID	Full Name	Forward primer sequence	Reverse primer sequence	Description	Reference	Anneal. temp. (°C)
ABCA1	ATP-binding cassette, sub-family A (ABC1), member 1	CCTCCATCTTTTCAGCTCTACCTAC	ACAAGAGCCTGGAGATTGGATAAC	Intronic	Horvath et al. (2008)	56
ADORA3	adenosine A3 receptor	ACCCCCATGTTTGGCTGGAA	GATAGGGTTCATCATGGAGTT	Exonic	Murphy et al. (2001)	60
APP	amyloid beta (A4) precursor protein	TCCAAGATGCAGCAGAACG	CTAATGTGTGCACATAAAACAGG	3'UTR	Murphy et al. (2001)	60
COI	cytochrome c oxidase I	TCCATTACCAGGCCAGCTAG	GAACTTGCTGGCTTTCATATC	Exonic; mtDNA	Ward et al. (2005)	45
CREM	cAMP responsive element modulator	AGGAACTCAAGGCCCTCAAA	GGGAGGACAAATGTCTTTCAA	3'UTR	Murphy et al. (2001)	57
CYTB	cytochrome b	GCACAACCTACAGCACCACTA	CAGCTTTGGGTGTTGAYGGTRGAA	Exonic; mtDNA	This study	60
DENND5A	DENN/MADD domain containing 5A	CCAGAGTTATCATGGCCAATC	GTACCAAGCAAGAAGCTGGG	Mostly intronic	Perelman et al. (2011)	62
DMRT1	doublesex and mab-3 related transcription factor 1	ATCCCTTGTTCTGAGTGCCA	ACATTGCAAAGACCCCTGAC	Intronic	Perelman et al. (2011)	60
ERC2	ELKS/RAB6-interacting/CAST family member 2	AGCTCATCCTCCTCCTGGTTTAG	CTCCTTGAGGATCTCCAGCAAC	Mostly intronic	Horvath et al. (2008)	57
FAM123B	APC membrane recruitment protein 1 (AMER1)	CATCACTCTGGAAGAGCTGC	TGGATTTGAGGATGATTCAGG	Exonic; X-chromosome	Perelman et al. (2011)	60
FES	FES proto-oncogene, tyrosine kinase	GGGGAACTTTGGCGAAGTGTT	TCCATGACGATGTAGATGGG	Mostly intronic	Venta et al. (1996)	56
FOXP1	forkhead box P1	TCAGCATCACTAATTTTGATGAAC	TGATGCAACTCTCAAGGAAAAG	Intronic	Perelman et al. (2011)	60
MAPKAP1	mitogen-activated protein kinase associated protein 1	TGTCAGCTCCATCGCTTATAACT	GGGCTGAATGATGGTGATTT	Intronic	Perelman et al. (2011)	60
MBD5	methyl-CpG binding domain protein 5	GGCAGATAGCTACCACCACC	CTTCCAGGCAAGGTTCATTC	3'UTR	Perelman et al. (2011)	60
NEGR1	neuronal growth regulator 1	CATTATGTGGTTGGCAGCAT	TTGCAAGATGACAAACTATGTGTT	Intronic	Perelman et al. (2011)	60
NPAS3.2	neuronal PAS domain protein 3 (NPAS3)	TCAGCATTGTTGATCTGCTTTT	TGGAATATCTAACCATCTCTGAACA	Intronic	Perelman et al. (2011)	60
RAG1	recombination activating gene 1	GCTTTGATGGACATGGAAGAAGACAT	GAGCCATCCCTCTCAATAATTTCAGG	Exonic	Teeling et al. (2000)	57
RAG2	recombination activating gene 2	GATTCCTGCTAYCTYCCTCCTCT	CCCATGTTGCTTCCAAACCATA	Exonic	Teeling et al. (2000)	60
RPGRIP1	retinitis pigmentosa GTPase regulator interacting protein 1	AGATGTTGCTTATGGCACCC	ACCTGGGCTTTCTTTCGTTT	Exonic	Perelman et al. (2011)	57
SGMS1	sphingomyelin synthase 1	TCAGAATCAAACCCCATTCAG	GTGGTGGTACAGGCCATTTC	Mostly 3'UTR	Perelman et al. (2011)	57
SIM1	single-minded family bHLH transcription factor 1	GACCTACCGCAGAAAATTCG	CTGGGGCTCATCATTCATTC	Intronic	Perelman et al. (2011)	60
ZFX	zinc finger protein, X-linked	TGGAATGAAATCCCTCAAATA	ATGTCCATCAGGGCCAATAAT	Intronic; X-chromosome	Perelman et al. (2011)	52

Table 2.4 List of the GenBank accession numbers.

Species	ID	Locus										
		ABCA1	ADORA3	APP	CREM	DENND5A	DMRT1	ERC2	FAM123B	FES	FOXP1	MAPKAP1
C. cf. *hoffmannsi*	CMH1	HM765300	HM765212	HM764676	HM763036	HM759316	HM762546	HM762182	HM762131	HM761805	HM761544	HM760645
C. donacophilus	CDO1	HM765289	HM765211	HM764675	HM763035	HM759315	HM762536	HM762211	HM762130	HM761804	HM761533	HM760634
C. donacophilus	CCG1	HM765282	HM765208	HM764672	--	HM759311	HM762529	HM762201	HM762127	--	HM761526	HM760630
C. personatus	CLP1	HM765298	HM765214	--	HM763038	HM759320	HM762544	HM762187	HM762133	HM761807	HM761542	HM760643
C. coimbrai	CCO1	HM765284	HM765209	HM764673	HM763033	HM759312	HM762531	HM762175	HM762128	HM761802	HM761528	HM760631
C. nigrifrons	CNI1	--	HM765213	HM764677	HM763037	HM759318	HM762550	HM762178	HM762132	HM761806	HM761547	HM760647
Cebus albifrons	CEA1	HM765415	HM765189	HM764657	HM763047	HM759289	HM762675	HM762192	HM762106	--	HM761665	HM760760
Saimiri sciureus	SSC7	HM765401	HM765206	HM764671	HM763149	HM759309	HM762658	HM762269	HM762126	HM761799	HM761651	HM760744
Cacajao calvus	CCL1	HM765283	HM765187	HM764655	--	HM759286	HM762530	HM762297	HM762104	--	HM761527	--
Chiropotes israelita	CIS1	HM765295	HM765194	HM764662	--	HM759295	HM762542	HM762183	HM762112	HM761786	HM761539	HM760640
Pithecia pithecia	PPT1	HM765380	HM765215	HM764678	HM763130	HM759323	HM762634	HM762204	HM762135	HM761808	HM761627	HM760722
Species	**ID**	Locus										
		MBD5	NEGR1	NPAS3.2	RAG1	RAG2	RPGRIP1	SGMS1	SIM1	ZFX	COI	CYTB
C. cf. *hoffmannsi*	CMH1	HM760560	HM760295	HM759935	HM759136	HM758968	HM758700	HM758570	HM758334	HM757152	--	--
C. donacophilus	CDO1	HM760559	HM760284	HM759924	HM759135	HM758967	HM758685	HM758480	HM758323	HM757151	--	--
C. donacophilus	CCG1	HM760556	HM760277	HM759917	HM759133	HM758965	HM758677	HM758479	HM758317	HM757148	--	--
C. personatus	CLP1	HM760562	HM760293	HM759933	HM759138	HM758970	HM758697	HM758591	HM758332	HM757153	--	--
C. coimbrai	CCO1	HM760557	HM760279	HM759919	HM759134	--	HM758679	HM758518	HM758319	HM757149	--	--
C. nigrifrons	CNI1	HM760561	HM760298	HM759938	HM759137	HM758969	HM758705	HM758576	HM758335	--	--	--
Cebus albifrons	CEA1	HM760533	HM760422	HM760060	HM759115	HM758944	HM758686	HM758491	HM758451	HM757128	--	--
Saimiri sciureus	SSC7	HM760553	HM760405	HM760044	HM759131	HM758963	HM758816	HM758621	HM758437	HM757147	--	--
Cacajao calvus	CCL1	HM760531	HM760278	HM759918	HM759113	HM758942	HM758678	HM758478	HM758318	HM757126	--	--
Chiropotes israelita	CIS1	HM760539	HM760290	HM759930	HM759120	HM758950	HM758692	HM758540	HM758329	HM757133	--	--
Pithecia pithecia	PPT1	HM760564	HM760382	HM760023	HM759140	HM758971	HM758792	HM758594	HM758414	HM757155	--	--
C. personatus	--	--	--	--	--	--	--	--	--	--	--	AF289988
C. donacophilus	--	--	--	--	--	--	--	--	--	--	FJ785423	FJ785423
C. nigrifrons	--	--	--	--	--	--	--	--	--	--	--	AF524884
C. lugens	--	--	--	--	--	--	--	--	--	--	--	DQ337708
Saimiri sciureus	--	--	--	--	--	--	--	--	--	--	HQ644334	HQ644334
Cacajao calvus	--	--	--	--	--	--	--	--	--	--	NC021967	NC021967
Chiropotes israelita	--	--	--	--	--	--	--	--	--	--	NC024629	NC024629
Pithecia pithecia	--	--	--	--	--	--	--	--	--	--	JF459229	--

Table 2.5 List of sequence length and loci coverage for samples in the combined and nuclear datasets.

Species	Sample ID	Nuclear No. loci (20)	Nuclear Length (12778)	Combined No. loci (22)	Combined Length (14578)	ABCA1	ADORA3	APP	COI	CREM	CYTB	DENND5A	DMRT1	ERC2	FAM123B	FES	FOXP1	MAPKAP1	MBD5	NEGR1	NPAS3.2	RAG1	RAG2	RPGRIP1	SGMS1	SIM1	ZFX
C. moloch	MCB63	18	11,056	20	12,856																						
C. moloch	MCB64	20	12,615	22	14,415																						
C. moloch	MCB79	19	11,536	21	13,297																						
C. moloch	857	14	8,133	16	9,933																						
C. moloch	1516	15	8,755	17	10,531																						
C. moloch	CTGAM420	16	10,068	18	11,868																						
C. moloch	CTGAM421	13	8,142	15	9,942																						
C. cf. *moloch*	RVR22	20	12,379	22	14,179																						
C. cf. *moloch*	RVR68	19	11,760	21	13,560																						
C. cf. *moloch*	RVR73	20	12,603	22	14,403																						
C. bernhardi	UFRO413	17	10,845	19	12,641																						
C. bernhardi	42960	18	11,302	20	13,102																						
C. bernhardi	42961	16	10,038	17	11,178																						
C. bernhardi	42964	19	11,743	21	13,543																						
C. miltoni	42991	20	12,405	22	14,205																						
C. miltoni	42992	20	12,400	22	14,200																						
C. miltoni	42993	19	11,685	21	13,351																						
C. cinerascens	UFRO352	11	6,570	13	8,366																						
C. cinerascens	UFRO355	20	12,343	22	14,139																						
C. cinerascens	UFRO499	18	11,102	20	12,892																						
C. hoffmannsi	CTGAM248	18	11,022	20	12,764																						
C. hoffmannsi	CTGAM290	20	12,449	22	14,191																						
C. hoffmannsi	01CNP	15	9,384	17	11,184																						
C. cf. *hoffmannsi*	CMH1	20	12,456	20	12,456																						
C. cupreus	AAM15	15	8,881	17	10,663																						
C. cupreus	CTGAM210	17	11,356	19	13,141																						
C. cupreus	JLP15920	7	4,667	9	6,428																						
C. cupreus	4984	17	10,156	19	11,956																						
C. cupreus	4988	19	11,892	21	13,692																						
C. cupreus	4990	20	12,611	22	14,411																						

Table 2.5 cont'd. List of sequence length and loci coverage for samples in the combined and nuclear datasets.

Species	Sample ID	Nuclear No. loci (20)	Nuclear Length (12778)	Combined No. loci (22)	Combined Length (14578)	ABCA1	ADORA3	APP	COI	CREM	CYTB	DENND5A	DMRT1	ERC2	FAM123B	FES	FOXP1	MAPKAP1	MBD5	NEGR1	NPAS3.2	RAG1	RAG2	RPGRIP1	SGMS1	SIM1	ZFX
C. cupreus	4993	19	11,572	21	13,372																						
C. brunneus	UFRO541	18	11,076	20	12,872																						
C. brunneus	4009	19	11,999	21	13,799																						
C. brunneus	4019	17	10,388	19	12,188																						
C. brunneus	4346	20	12,475	22	14,251																						
C. brunneus	4505	20	12,560	22	14,333																						
C. dubius	UFRO403	20	12,305	22	14,098																						
C. dubius	UFRO544	20	12,511	21	13,651																						
C. caligatus	CTGAM181	20	12,344	22	14,129																						
C. caligatus	CTGAM182	20	12,540	22	14,277																						
C. donacophilus	CDO1	20	12,387	20	12,387																						
C. donacophilus	CCG1	18	11,535	18	11,535																						
C. personatus	CLP1	19	11,458	19	11,458																						
C. coimbrai	CCO1	19	11,769	19	11,769																						
C. nigrifrons	CNI1	18	10,720	18	10,720																						
C. nigrifrons	CPE04	9	5,623	11	7,419																						
C. lugens	JPB81	16	10,276	18	11,468																						
C. lugens	JPB119	20	12,610	22	14,342																						
C. lugens	JPB124	20	12,437	22	14,234																						
C. lugens	JPB136	19	11,677	21	13,400																						
C. purinus	CTGAM154	20	12,310	22	14,095																						
C. purinus	CTGAM195	16	9,299	18	11,075																						
C. purinus	CTGAM209	17	10,142	19	11,927																						
Cebus albifrons	JPB100	15	8,761	16	9,898																						
Cebus albifrons	CEA1	19	11,036	19	11,036																						
Saimiri sciureus	SSC7	20	12,116	20	12,116																						
Cacajao calvus	CCL1	17	10,737	17	10,737																						
Chiropotes israelita	CIS1	19	11,837	19	11,837																						
Pithecia pithecia	PPT1	20	12,324	20	12,324																						

Table 2.6 List of sequence length and loci coverage for samples in the mitochondrial (mtDNA) dataset.

Species	Sample ID	Mitochondrial (mtDNA) Length (1,800)	No. loci (2)	COI	CYTB
C. moloch	MCB63	1,800	2	■	■
C. moloch	MCB64	1,800	2	■	■
C. moloch	MCB79	1,761	2	■	■
C. moloch	857	1,800	2	■	■
C. moloch	1516	1,776	2	■	■
C. moloch	CTGAM420	1,800	2	■	■
C. moloch	CTGAM421	1,800	2	■	■
C. cf. *moloch*	RVR22	1,800	2	■	■
C. cf. *moloch*	RVR68	1,800	2	■	■
C. cf. *moloch*	RVR73	1,800	2	■	■
C. bernhardi	UFRO413	1,796	2	■	■
C. bernhardi	42960	1,800	2	■	■
C. bernhardi	42961	1,140	1		■
C. bernhardi	42964	1,800	2	■	■
C. miltoni	42991	1,800	2	■	■
C. miltoni	42992	1,800	2	■	■
C. miltoni	42993	1,676	2	■	■
C. cinerascens	UFRO352	1,796	2	■	■
C. cinerascens	UFRO355	1,796	2	■	■
C. cinerascens	UFRO499	1,790	2	■	■
C. hoffmannsi	CTGAM248	1,742	2	■	■
C. hoffmannsi	CTGAM290	1,742	2	■	■
C. hoffmannsi	01CNP	1,800	2	■	■
C. cupreus	AAM15	1,782	2	■	■
C. cupreus	CTGAM210	1,785	2	■	■
C. cupreus	JLP15920	1,761	2	■	■
C. cupreus	4984	1,800	2	■	■
C. cupreus	4988	1,800	2	■	■
C. cupreus	4990	1,800	2	■	■
C. cupreus	4993	1,800	2	■	■
C. brunneus	UFRO541	1,796	2	■	■
C. brunneus	4009	1,800	2	■	■
C. brunneus	4019	1,800	2	■	■
C. brunneus	4346	1,776	2	■	■
C. brunneus	4505	1,773	2	■	■
C. dubius	UFRO403	1,793	2	■	■
C. dubius	UFRO544	1,140	1		■
C. caligatus	CTGAM181	1,785	2	■	■
C. caligatus	CTGAM182	1,737	2	■	■
C. caligatus	MVR58	1,796	2	■	■
C. caligatus	CCM248	1,140	1		■
C. donacophilus	FJ785423	1,797	2	■	■
C. personatus	AF289988	998	1		■
C. nigrifrons	CPE04	1,796	2	■	■
C. nigrifrons	CPRJ1493	1,001	1		■
C. lugens	JPB81	1,192	2	■	■
C. lugens	CRB2698	1,140	1		■
C. lugens	JPB119	1,732	2	■	■
C. lugens	JPB124	1,797	2	■	■
C. lugens	JPB136	1,723	2	■	■
C. purinus	CTGAM154	1,785	2	■	■
C. purinus	CTGAM195	1,776	2	■	■
C. purinus	CTGAM209	1,785	2	■	■
Cebus albifrons	JPB100	1,137	1		■
Saimiri sciureus	HQ644334	1,800	2	■	■
Cacajao calvus	NC021967	1,797	2	■	■
Chiropotes israelita	NC024629	1,797	2	■	■
Pithecia pithecia	JF459229	657	1	■	

Table 2.7 Summary of dataset characteristics and sequence variation for *Callicebus* taxa.

Dataset ID	Description	Length (bp)	Missing (%)	Constant sites		Variable sites		Parsimony informative sites		*Callicebus* samples
				bp	%	bp	%	bp	%	
Nuclear	nDNA (20 loci)	12,778	13.6	12,387	96.9	391	3.1	293	2.3	53 samples; 47 sequenced for this study, 6 for Perelman *et al.* (2011)
Combined	nDNA + mtDNA (22 loci)	14,578	14.6	13,735	94.2	843	5.8	678	4.7	
Mitochondrial	mtDNA (2 loci)	1,800	7.1	1,312	72.9	488	27.1	420	23.3	53 samples; 49 sequenced for this study, 4 from GenBank

For all datasets, *Callicebus* sample coverage for individual gene regions varied from 74% to 100% (average sample coverage = 90%). Length of loci varied between 402 bp and 1140 bp. A list of loci characteristics is presented in Appendix 1, Table A1.1.

2.3.3 DNA isolation, amplification and sequencing

DNA was extracted from multiple tissues (blood, muscle, kidney) using the Promega Wizard Genomic Kit according to the manufacturer's protocol. We amplified all nuclear and mitochondrial gene regions using polymerase chain reaction (PCR). The PCR reactions were carried out in a total volume of 25 µL, containing approximately 30 ng of genomic DNA; 4 µL of dNTPs (200µM each); 2.5 µL 10X buffer (200 mM Tris-HCL, 500 mM KCl); 1 µL of MgCl$_2$ (25 mM); 1 µL of each forward and reverse primer (0.2 µM); and 1 Unit of Invitrogen™ *Taq* DNA polymerase. The amplification cycles were carried out under the following conditions; initial denaturation at 95°C for 5 min; followed by 35 cycles of denaturing at 94°C for 1 min, primer annealing at between 44°C and 64°C (temperature varies per primer, see Table 2.3) for 1 min, and extension at 72°C for 1 min; a final extension was carried out at 72°C for 5 min.

PCR products were analysed on 1.5% agarose gels and those that produced clear single bands were purified with polyethylene glycol (PEG) and ethanol (Paithankar & Prasad, 1991). After purification, PCR products were sequenced directly in two reactions with forward and reverse primers. Sequencing reactions were carried out using the BigDye Terminator v3.1 cycle sequencing kit (Life Technologies). For 10 µL sequencing reactions we used 0.5 µL of BigDye; 1.5 µL of 5X Sequencing buffer; 1.0 µL of each primer (0.8 µM); and 2 µL of PCR product. Sequencing reactions were performed as follows: 96°C for 2 min; followed by 35

cycles of 96°C for 15 s, 50°C for 15 s, 60°C for 2.5 min. The sequencing products were analysed using an ABI 3500xl (Life Technologies) automatic sequencer following the manufacturer's instructions. Consensus sequences for each individual were generated from sequences in forward and reverse directions using Geneious R7.1 (Biomatters).

2.3.4 Sequence alignment, data partitioning and model selection

Each locus was first aligned independently using the standard MUSCLE (Edgar, 2004) alignment plugin in Geneious R7.1 and checked visually. The loci were then concatenated into alignments reflecting the three datasets (nuclear, combined and mitochondrial).

We used the program PartitionFinder (Lanfear et al., 2012) to objectively determine the optimal model of evolution and partitioning scheme simultaneously. Best-fit models were selected using Bayesian information criteria under a "greedy" search scheme using a subset of models specific to each programme used (RAxML, MrBayes, BEAST). When specifying the alignment subsets for PartitionFinder, we defined all intronic and UTR loci as single data-blocks and split exonic sequences into three subsets reflecting codon position. All our phylogenetic analyses used a specific partitioning scheme (containing between three and nine partitions) selected for the dataset by PartitionFinder. Additional information about each specific partitioning scheme is presented in Appendix 1, Table A1.2.

2.3.5 Phylogenetic analyses

We conducted phylogenetic inference using maximum-likelihood (ML) and Bayesian methods for each dataset. All phylogenetic analyses were run on the CIPRES Science Gateway v 3.3 server (Miller et al., 2010). Our ML phylogenetic reconstructions were conducted using the program RAxML v. 8.1 (Stamatakis, 2014). For ML inferences, we used the partitioning scheme and best-fit models chosen by PartitionFinder. We estimated support for nodes using the rapid-bootstrapping algorithm (−f a -x option) for 1000 non-parametric bootstrap replicates (Stamatakis et al., 2008). Maximum-likelihood bootstrap support values (BP) greater than 70% were considered as significant support (Hillis & Bull, 1993).

Bayesian analyses were performed using MrBayes 3.2.3 (Ronquist et al., 2008) with the Metropolis coupled Markov Chain Monte Carlo (MCMC) algorithm.

The partitioning scheme and best-fit models chosen by PartitionFinder were implemented and partitions were unlinked. MCMC convergence was checked after two independent four-chain runs of 10 million generations for each Bayesian inference. We assessed convergence by examining LnL, the average standard deviation of the split frequencies between the two simultaneous runs (< 0.01), and the Potential Scale Reduction Factor (PSRF) diagnostic in MrBayes, after a burn-in of 10%. Posterior probability values (PP) higher than 0.95 were considered as significant support (Alfaro et al., 2003).

A divergence matrix for the cytochrome *b* locus was generated for selected taxa (*C. cupreus*, *C. brunneus*, *C. caligatus*, *C. dubius*) using PAUP*4.0 (Swofford, 2002), based on the model parameters selected for the alignment by jModelTest v 2.1.6 (Darriba et al., 2012; Guindon & Gascuel, 2003).

2.3.6 Divergence-time analyses

We jointly estimated phylogeny and diversification times under an uncorrelated lognormal relaxed clock in the program BEAST v. 1.8.1 (Drummond et al., 2012). The partitioning scheme and best-fit models chosen by PartitionFinder were implemented and a Yule speciation process was used for all analyses. We ran two independent analyses for 50 million generations, sampling every 5000 generations. The sampling distributions of each run were visualized using Tracer v. 1.6 to evaluate convergence and to verify that the effective sample size was > 200 for all parameters after a burn-in of 10%. We combined runs using LogCombiner v. 1.8.1 and generated the maximum credibility tree in TreeAnnotator v. 1.8.1.

To obtain the posterior distribution of the estimated divergence times, we used two calibration points with lognormal priors to set a hard minimum and soft maximum bound (Ho & Phillips, 2009). We set a minimum age of 15.7 Ma for crown Pitheciidae based on the fossil *Proteropithecia* Kay et al., 1998 (Kay et al., 1998, 1999), and a minimum age of 12.5 Ma on crown Cebinae using the fossil *Neosaimiri* Stirton, 1951 (Hartwig & Meldrum, 2002; Rosenberger et al., 1991; Takai, 1994). For both calibration points, we set a soft maximum bound at 26 Ma using the fossil *Branisella boliviana* Hoffstetter, 1969, from the Deseadan fauna of La Salla (MacFadden, 1990). We chose this maximum age based on the evidence that *Branisella boliviana* and the Miocene Patagonian fossils belong to independent stem platyrrhine radiations (Kay et al., 2008; Schrago et al., 2013), the absence of fossils

for extant lineages in South American formations from this period (Kay & Fleagle, 2010), and the wealth of molecular evidence in support of a more recent common ancestor for extant platyrrhines (e.g., Hodgson et al., 2009; Schrago et al, 2013, 2014; Springer et al., 2012). The calibration points were implemented as lognormal distributions with an offset as the hard minimum bound. We set the standard deviation and mean such that 95% of the prior distribution falls before the maximum age to create a soft maximum bound (Table 2.8).

Table 2.8. Evolutionary rate calibration constraints (in millions of years).

Divergence	Offset fossil	Offset	95% age fossil	95% prior distribution	Standard deviation	Mean
Pitheciinae-Callicebinae	*Proteropithecia*	15.7	*Branisella boliviana*	26	0.8	1.02
Cebus-Saimiri	*Neosaimiri*	12.5	*Branisella boliviana*	26	0.8	1.29

Our divergence-time analyses were run based on all 22 loci in the combined dataset, but to minimise missing data for these analyses, we concatenated sequences from two individuals for some outgroup species (*Cacajao calvus*, *Chiropotes israelita*, *Pithecia pithecia*, *Saimiri sciureus*). For comparison of node dates and topology, we also ran our BEAST analyses using the nuclear dataset.

2.4 Results

2.4.1 Group-level topology

All analyses across the mitochondrial, nuclear and combined datasets yielded an identical topology for the *Callicebus* species groups (Figure 2.1). Our results support the division of *Callicebus* into four reciprocally monophyletic groups; the *torquatus* clade, here including *C. lugens* and *C. purinus*; the *personatus* clade with *C. personatus*, *C. coimbrai*, and *C. nigrifrons*; the *donacophilus* clade with *C. donacophilus*; and the *moloch* clade containing all remaining taxa (*C. hoffmannsi, C. cinerascens, C. miltoni, C. bernhardi, C. moloch, C.* cf. *moloch, C. brunneus, C. cupreus, C. dubius*, and *C. caligatus*). The *torquatus* group is strongly supported as the earliest radiation to diverge. It is followed by the separation of the *personatus* group from the *donacophilus+moloch* clade, with the final group-level split occurring

between the *donacophilus* group and the *moloch* group. These major diversification events receive significant support across all analyses (bootstrap percentage, BP > 70%; posterior probability, PP > 0.95), and thus, our results suggest a highly resolved topology for the *Callicebus* species groups (Figure 2.1). As Kobayashi's *moloch* and *cupreus* groups were not monophyletic, we adopt Groves' (2005) classification and include all *cupreus* group species (*sensu* Kobayashi, 1995) in the *moloch* group. A summary of node support per analysis is presented in Table 2.9.
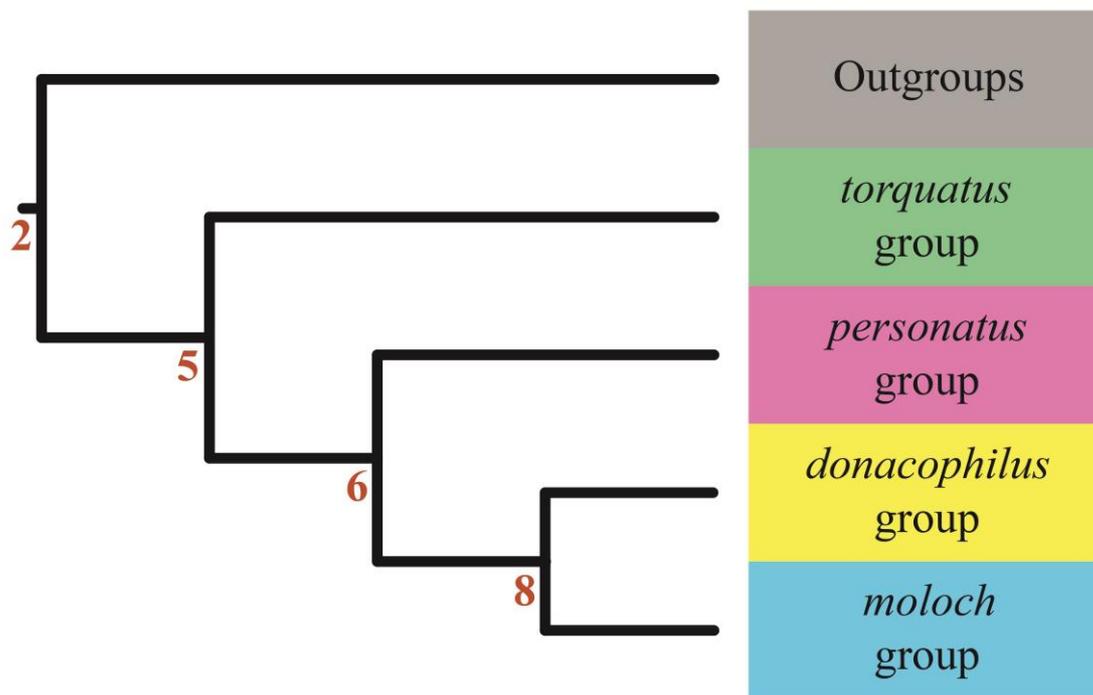
Figure 2.1 Phylogenetic reconstruction showing *Callicebus* species-group level topology. All nodes were significantly supported in all analyses (BP ≥ 70% and PP ≥ 0.95). Node numbers correspond to those in Figure 2, 3.

2.4.2 Species-level topology

Within each dataset, ML (RAxML) and Bayesian (MrBayes, BEAST) inference trees all presented similar species-level topologies. Individual trees with node support values for each analysis are found in Appendix 1, Figure A1.1–A1.3 (combined), A1.4–A1.6 (nuclear), A1.7, A1.8 (mitochondrial). Node support values for all phylogenetic analyses are listed in Table 2.9.

The phylogenetic relationships among taxa in the *torquatus* and *personatus* clades are identical for all three datasets (Figure 2.2). All nodes have significant support (BP > 70%, PP > 0.95) with the exception of the sister-relationship between

C. lugens from the left and right banks of the Rio Negro, which is not supported for the mitochondrial dataset (BP = 64%, PP = 0.78). *Callicebus donacophilus* is consistently supported as an independent lineage, sister to the *moloch* species group. Species-level relationships within the *moloch* group, however, vary according to each dataset. The principal differences were found between the combined and nuclear dataset topologies in the phylogenetic position of *C. cinerascens* and *C. miltoni*, as well as the phylogenetic relationships of *C. cupreus* and other closely related species (Figure 2.2). The mitochondrial trees largely reflect those inferred from the combined dataset except in the phylogenetic position of *C. hoffmannsi* (see Figure A1.7, A1.8), discussed below.

2.4.3 The *moloch* group

In contrast to morphological hypotheses (Kobayashi, 1995; Van Roosmalen et al., 2002), the *moloch* and *cupreus* groups were not monophyletic; *C. hoffmannsi* does not share a most recent common ancestor with other species of *moloch* group (*sensu* Kobayashi, 1995); and *C. brunneus* of the *moloch* group (*sensu* Kobayashi, 1995) is nested in the *cupreus* species group clade.

There is little molecular evidence for the separation of specimens identified as *C. caligatus* and *C. dubius*. The mitochondrial dataset supports *C. caligatus* and *C. dubius* as a monophyletic group (BP = 91%, PP = 1.00); however, the two *C. dubius* do not form a clade, and branch off independently at the base of the *C. caligatus* clade. Most of the nodes within this clade are not well supported (BP < 70%, PP < 0.95), and the topology may suggest that these taxa form one, not two, species. For the nuclear and combined datasets, *C. dubius* is monophyletic and is a minimally diverged sister taxon of *C. caligatus*. A divergence matrix based on the 1140bp cytochrome *b* locus (Appendix 1, Table A1.3) shows genetic distance values of 0.01–0.06 between the six *C. dubius* and *C. caligatus* specimens. These values are comparable to the divergence between specimens of *C. brunneus* (0.0–0.08) or of *C. cupreus* (0.02–0.19), rather than the genetic distances found between *C. brunneus*, *C. cupreus* and the *C. caligatus*/*C. dubius* complex (0.24–0.38).

Table 2.9 Node support values for all phylogenetic analyses. Node numbers correspond to those on Figure 2.1, 2.2, 2.3, A1.1 – A1.10. Bold indicates an unsupported node. Asterisk indicates the node is not represented in that topology.

Node	Split or clade	Combined dataset			Nuclear dataset			Mitochondrial dataset	
		BP RAxML	PP MrBayes	PP BEAST	BP RAxML	PP MrBayes	PP BEAST	BP RAxML	PP MrBayes
1	Pitheciidae vs. Cebinae	root	root	1.00	root	root	1.00	root	root
2	Pitheciinae vs. Callicebinae	100	1.00	1.00	100	1.00	1.00	83	1.00
3	*Saimiri* vs. *Cebus*	100	1.00	1.00	100	1.00	1.00	83	1.00
4	*Pithecia* vs. *Cacajao* + *Chiropotes*	100	1.00	1.00	100	1.00	0.98	75	1.00
5	*torquatus* group vs. all other *Callicebus*	100	1.00	1.00	100	1.00	0.98	100	1.00
6	*personatus* group vs. *donacophilus* + *moloch* groups	100	1.00	1.00	100	1.00	0.98	72	0.95
7	*Cacajao* vs. *Chiropotes*	100	1.00	1.00	100	1.00	1.00	98	0.96
8	*donacophilus* group vs. *moloch* group	100	1.00	1.00	100	1.00	0.98	100	1.00
9	*C. hoffmannsi* vs. remaining *Callicebus* [1]	85	1.00	1.00	70	1.00	1.00	**29**	**Polytomy**
10	west- vs. east-Amazonian *moloch* taxa [2]	**53**	1.00	1.00	**67**	1.00	1.00	71	**Polytomy**
11	*C. lugens* vs. *C. purinus*	100	1.00	1.00	100	1.00	1.00	100	1.00
12	*C. nigrifrons* vs. other *personatus* group taxa.	100	1.00	1.00	100	1.00	1.00	100	1.00
13	*C. cinerascens* + *C. miltoni* vs. *C. bernhardi* + *C. moloch* [2]	**44**	**0.76**	1.00	*	*	*	80	1.00
14	*C. cupreus* vs. remaining west-Amazonian taxa [3]	99	1.00	1.00	*	*	*	84	1.00
15	*C. bernhardi* vs. *C. moloch*	100	1.00	1.00	**56**	0.99	1.00	89	1.00
16	*C. coimbrai* vs. *C. personatus*	100	1.00	1.00	100	1.00	1.00	*	*
17	*C. brunneus* vs. remaining west-Amazonian taxa [3]	99	1.00	1.00	**58**	**Polytomy**	1.00	91	1.00
18	*C. cinerascens* vs. *C. miltoni* [2]	72	1.00	1.00	*	*	*	100	1.00
19	*C. hoffmannsi* vs. *C.* cf. *hoffmannsi*	85	1.00	1.00	82	1.00	1.00	*	*
20	*C. lugens* L bank vs. *C lugens* R bank	99	1.00	1.00	97	1.00	0.99	**64**	**0.78**
21	*C. moloch* vs. *C.* cf. *moloch*	98	1.00	1.00	79	1.00	1.00	94	1.00
22	*C. cupreus* A vs. *C. cupreus* B [3]	99	1.00	1.00	*	*	*	98	1.00
23	*C. dubius* vs. *C. caligatus* [4]	99	1.00	1.00	**64**	0.99	1.00	*	*
24	*C. cinerascens* (independent radiation) [2]	*	*	*	76	1.00	0.99	*	*
25	*C. miltoni* (independent radiation) [2]	*	*	*	70	0.98	0.99	*	*
26	*C. cupreus* A vs. remaining west-Amazonian taxa [3]	*	*	*	95	1.00	1.00	*	*
27	*C. cupreus* B vs. *C. caligatus* + *C. dubius* [3]	*	*	*	**23**	**Polytomy**	0.38	*	*

[1] The mitochondrial position of *C. hoffmannsi* differs from the combined and nuclear.

[2] Nuclear topology: *C. cinerascens* and *C. miltoni* form independent lineages.

[3] The relationship between *C. brunneus*, *C. cupreus* and *C. dubius/caligatus* differs between the nuclear and combined/mitochondrial topology.

[4] For the mitochondrial topology, *C. dubius* is not monophyletic.
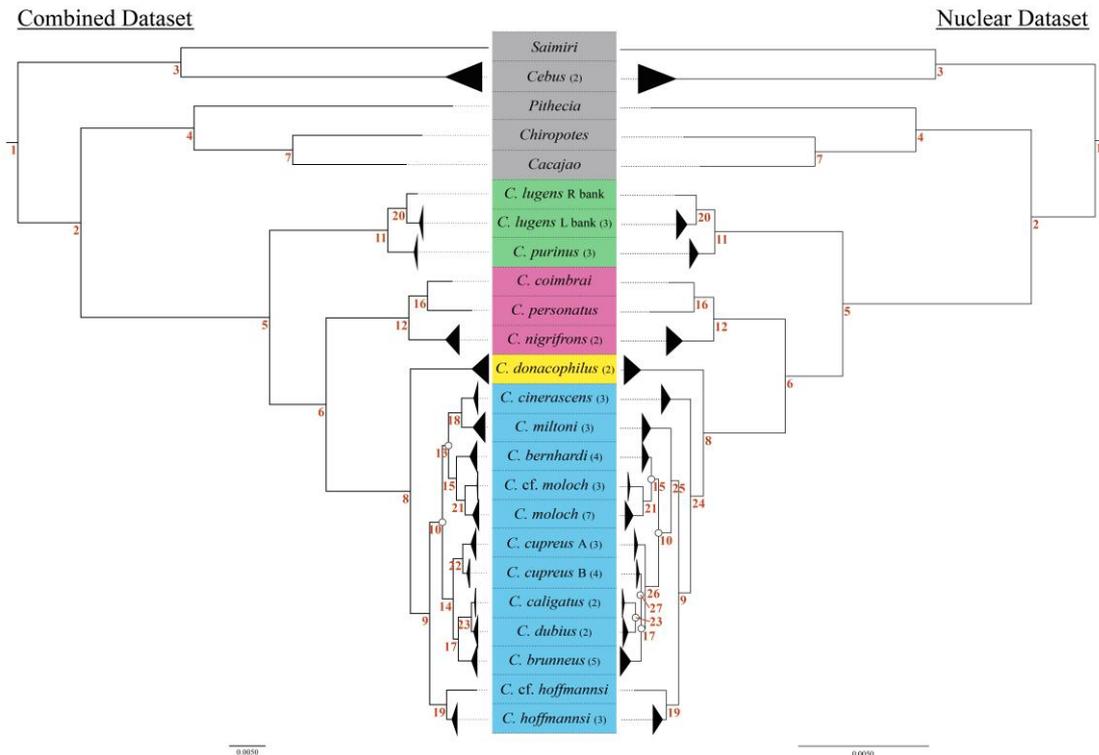
Figure 2.2 Molecular phylogeny showing relationships among *Callicebus* taxa based on 53 *Callicebus* and 6 outgroup individuals. Shown are maximum likelihood trees inferred from the combined dataset (left) and the nuclear dataset (right), with branches collapsed to represent clades of interest. Numbers in parenthesis indicate number of individuals represented in the collapsed clade. See Figure A1.1–A1.6 for the expanded ML (RAxML) and Bayesian (MrBayes, BEAST) trees with node support values. Unmarked nodes were significantly supported in all analyses (BP ≥ 70% and PP ≥ 0.95), while nodes marked with white circles received low support (BP < 70% and/or PP < 0.95). Red numbers represent nodes of interest listed with support values for all methods of analysis in Table 2.9. Background colours reflect species group; green for the *torquatus* group, pink for the *personatus* group, yellow for *donacophilus* group, blue for the *moloch* group; and grey indicates the outgroup species.

All datasets support a west-Amazonian species complex that comprises *C. brunneus, C. cupreus, C. caligatus* and *C. dubius*, and is subdivided into four distinct clades: *C. brunneus*; *C. cupreus* A; *C. cupreus* B; and *C. caligatus/C. dubius*. The sister group relationship of *C. cupreus* (*C. cupreus* A, *C. cupreus* B) to the group comprising *C. brunneus* and *C. caligatus/dubius* is consistently supported in the combined/mitochondrial phylogeny (BP > 84%, PP = 1.00). In the nuclear dataset, *C. cupreus* is paraphyletic and *C. cupreus* A is supported as the first diverging member of the group (BP = 95%, PP = 1.00). The RAxML and BEAST topologies show that *C. brunneus* is the next taxon to diverge (BP = 58%, PP = 1.00), with *C. cupreus* B being sister to *C. caligatus/dubius* (BP = 23%, PP = 0.38). However, the MrBayes

tree inferred from the nuclear dataset shows a polytomy among *C. brunneus, C. caligatus/dubius* and *C. cupreus* B.

Callicebus hoffmannsi is strongly supported as an early diverging lineage in the nuclear (between the *C. cinerascens* and *C. miltoni* radiations) and combined (as sister-group to all other species of the *moloch* group) dataset analyses. The phylogenetic relationship of *C. hoffmannsi* differs in the mitochondrial dataset (see Figure A1.7, A1.8), but has no statistical support (RAxML, BP = 28%; MrBayes, unresolved polytomy).

All analyses support a clade that contains *C. moloch*, *C.* cf. *moloch* and *C. bernhardi*, with a sister-species relationship between *C. moloch* and *C.* cf. *moloch*. All nodes within this group are significantly supported (BP > 70%, PP > 0.95) with the exception of the split between *C. bernhardi* and *C. moloch/C.* cf. *moloch* for the ML nuclear phylogeny (BP = 56%, PP > 0.99). *Callicebus cinerascens* + *C. miltoni* are a sister group to this clade in the mitochondrial (with significant support) and combined (supported only in the BEAST analysis, PP = 1.00) datasets. In the nuclear dataset, *C. cinerascens* and *C. miltoni* find significant support as independent early diverging lineages, along with *C. hoffmannsi*. Thus, there is a conflict in the phylogenetic signals of the nuclear and mitochondrial datasets, which is reflected by low support in combined datasets, but high support in independent mitochondrial and nuclear analyses. The phylogenetic position of *C. cinerascens* and *C. miltoni*, therefore, remains unresolved.

2.4.4 Divergence-time estimates

From the combined dataset (Figure 2.3, A1.9, Table 2.10), we estimated the origin of crown Pitheciidae at *c.* 21.47 Ma (95% HPD = 17.82–25.78) and the origin of crown *Callicebus* to be in the early Miocene, *c.* 18.71 Ma (95% HPD = 15.97–22.6). The most recent common ancestor of extant *Callicebus* lineages is estimated to have lived in the late Miocene (10.98 Ma; 95% HPD = 8.36–14.25); this ancestor gave rise to the progenitor of the *torquatus* species group (Amazon and Orinoco) and the progenitor of all other *Callicebus* clades. Next to diverge was the Atlantic forest *personatus* group at around 8.34 Ma (95% HPD = 6.18–10.86), also in the late Miocene. The final group-level divergence is estimated to have occurred in the Pliocene, around 4.39 Ma (95% HPD = 2.99–6.08), between *C. donacophilus* (representative of the *donacophilus* group) and the *moloch* group. In the *moloch* group, *C. hoffmannsi*

diverged at an estimated 3.44 Ma (95% HPD = 2.39–4.74), followed by the divergence of an east-Amazonian clade (*C. cinerascens, C. miltoni, C. bernhardi, C. moloch, C.* cf. *moloch*) and a west-Amazonian clade (*C. cupreus, C. brunneus, C. dubius, C. caligatus*) at around 2.81 Ma (95% HPD = 1.95–3.8).

Table 2.10 Divergence time estimates for the combined and nuclear datasets. Node numbers correspond to those on Figure 2.1, 2.2, 2.3, A1.9, A1.10. Asterisk indicates the node is not represented in that topology.

Node	Split or clade	Combined dataset			Nuclear dataset		
		Mean age (Ma)	95% HPD Lower	95% HPD Upper	Mean age (Ma)	95% HPD Lower	95% HPD Upper
1	Pitheciidae vs. Cebinae	21.47	17.82	25.78	22.89	17.82	28.92
2	Pitheciinae vs. Callicebinae	18.71	15.97	22.6	19.13	15.93	23.8
3	*Saimiri* vs. *Cebus*	15.02	12.85	17.92	14.99	12.79	18.08
4	*Pithecia* vs. *Cacajao + Chiropotes*	11.89	8.82	15.5	11.99	7.33	16.67
5	*torquatus* group vs. all other *Callicebus*	10.98	8.36	14.25	12.03	7.78	16.72
6	*personatus* group vs. *donacophilus + moloch* groups	8.34	6.18	10.86	8.94	5.52	13.07
7	*Cacajao* vs. *Chiropotes*	6.77	4.48	9.27	6.23	2.68	10.15
8	*donacophilus* group vs. *moloch* group	4.39	2.99	6.08	5.33	2.58	8.78
9	*C. hoffmannsi* vs. remaining *Callicebus*	3.44	2.39	4.74	3.57	1.55	6.29
10	west- vs. east-Amazonian *moloch* taxa [1]	2.81	1.95	3.8	2.41	1	4.33
11	*C. lugens* vs. *C. purinus*	2.6	1.57	3.77	3.15	1.03	6.22
12	*C. nigrifrons* vs. other *personatus* group taxa.	2.53	1.52	3.79	4.14	1.85	7.3
13	*C. cinerascens + C. miltoni* vs. *C. bernhardi + C. moloch* [1]	2.33	1.65	3.2	*	*	*
14	*C. cupreus* vs. remaining west-Amazonian taxa [2]	1.95	1.33	2.72	*	*	*
15	*C. bernhardi* vs. *C. moloch*	1.72	1.16	2.39	1.73	0.63	3.23
16	*C. coimbrai* vs. *C. personatus*	1.55	0.81	2.44	2.43	0.87	4.48
17	*C. brunneus* vs. remaining west-Amazonian taxa [2]	1.5	0.96	2.1	1.36	0.51	2.55
18	*C. cinerascens* vs. *C. miltoni* [1]	1.33	0.85	1.93	*	*	*
19	*C. hoffmannsi* vs. *C.* cf. *hoffmannsi*	1.23	0.53	2.07	1.91	0.6	3.72
20	*C. lugens* L bank vs. *C lugens* R bank	1.16	0.65	1.82	1.65	0.52	3.4
21	*C. moloch* vs. *C.* cf. *moloch*	1.05	0.64	1.52	1.18	0.39	2.27
22	*C. cupreus* A vs. *C. cupreus* B [2]	1	0.58	1.47	*	*	*
23	*C. dubius* vs. *C. caligatus*	0.5	0.26	0.79	0.84	0.28	1.66
24	*C. cinerascens* (independent radiation) [1]	*	*	*	4.38	1.96	7.63
25	*C. miltoni* (independent radiation) [1]	*	*	*	3.08	1.33	5.37
26	*C. cupreus* A vs. remaining west-Amazonian taxa [2]	*	*	*	1.69	0.63	3.17
27	*C. cupreus* B vs. *C. caligatus + C. dubius* [2]	*	*	*	1.16	0.42	2.16

[1] For the nuclear topology, *C. cinerascens* and *C. miltoni* form independent lineages.
[2] The relationship between *C. brunneus*, *C. cupreus* and *C. dubius/caligatus* differs in the nuclear and combined topology.

Sister species divergences are estimated at 1–3 Ma for all *Callicebus* taxa included in the dating analyses. These are especially recent for species of the *moloch* group, with all sister-species splits occurring 1−2 Ma with the exception of the *C. dubius* and *C. caligatus* divergence, which occurred more recently at *c.* 0.5 Ma (95% HPD = 0.26–0.79). Our dating analyses also suggest relatively divergent lineages within some taxa that diverged *c.* 1.0 − 1.2 Ma; *C. lugens* from the left and right bank of the Rio Negro; *C. moloch* and *C.* cf. *moloch*; and *C. cupreus* A and B.
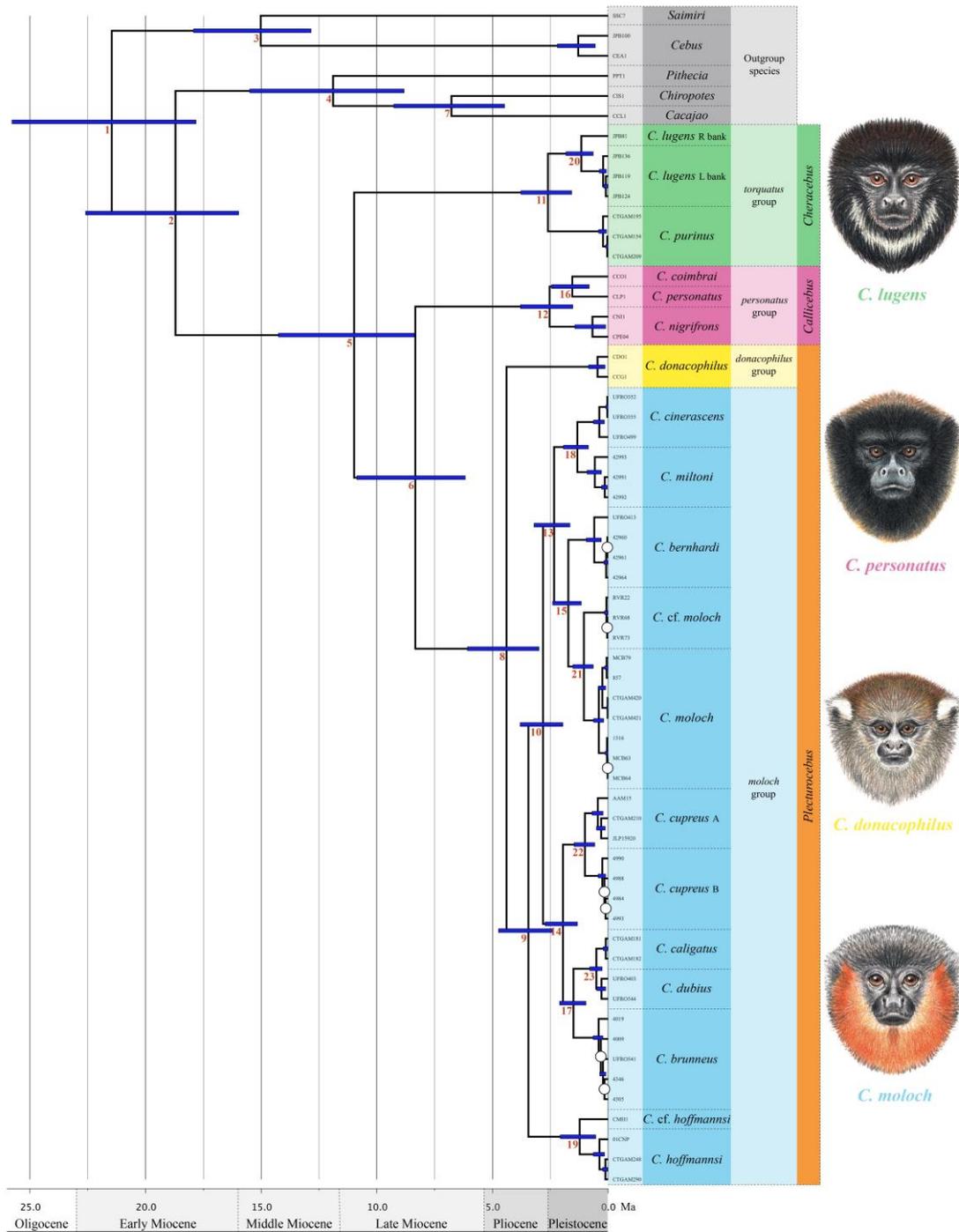
Figure 2.3 A time-calibrated phylogeny showing estimated divergence ages among *Callicebus* individuals based on the combined dataset. Unmarked nodes were strongly supported (PP ≥ 0.99), nodes marked with white circles received low support (PP < 0.95). Node bars indicate the 95% highest posterior density. Red numbers represent nodes of interest listed with specific support values and estimated divergence times in Tables 2.9, 2.10. Nodes numbered 2 and 3 were used for calibration. A time scale in million years and the geological periods are given. Background colours reflect species-group; green for the *torquatus* group, pink for the *personatus* group, yellow for *donacophilus* group, blue for the *moloch* group; and grey indicates the outgroup species. Illustrations by Stephen D. Nash ©Conservation International.

We also dated the phylogeny based on nuclear loci only (Table 2.10, Figure A1.10). Importantly, we estimated the age of divergence of *C. cinerascens* and *C. miltoni* from their sister clades in the *moloch* group at *c.* 4.38 Ma (95% HPD = 1.96–7.63) and 3.08 Ma (95% HPD = 1.33–5.37), respectively. Note that *C. cinerascens* and *C. miltoni* are weakly supported as a sister group to the *C. bernhardi*/*C. moloch*/*C.* cf. *moloch* clade in the combined dataset analyses due to mitochondrial DNA signal. Divergence dates inferred for the combined dataset BEAST analyses are consistently slightly younger across *Callicebus* than for the dating analyses based on the nuclear loci. A summary of divergence date estimates and 95% HPD intervals for the combined and nuclear dataset BEAST analyses is presented in Table 2.10.

2.5 Discussion

2.5.1 A proposal for a new taxonomy of the titi monkeys at genus-level

In this study, we assembled one of the largest molecular datasets for any group of platyrrhine primates, sequencing 20 nuclear and two mitochondrial loci totalling over 14,000 base pairs, and including representatives of all the major callicebine lineages. Using this dataset, we provide the first comprehensive review of Callicebinae using molecular data to assess phylogenetic relationships and divergence dates among the major lineages and to test morphological taxonomical hypotheses. Our analyses show that *Callicebus* is divided into three principal clades of Miocene origin, corresponding to Kobayashi's (1995) *torquatus* and *personatus* groups, and a clade containing the *donacophilus*, *moloch* and *cupreus* species groups. All phylogenetic analyses yielded identical relationships among these three clades with estimated divergence times being in the late Miocene. Based on the results from our phylogenetic analyses, and also morphological, ecological, karyological and biogeographical evidence (see below), we suggest the division of titi monkeys into three genera in the subfamily Callicebinae (Table 2.1).

Cheracebus new genus

LSID: urn:lsid:zoobank.org:act:DE67E93E-89A3-47C1-BAF3-E183F3448520

Type species: *Simia lugens* Humboldt, A. von. 1811. *Rec. Obs. Zool. Anat. Comp.* 1: 319.

We did not suggest the earlier named *Callitrix* [*sic*] *torquata* Hoffmannsegg, 1807, as the type species, because the type locality given by Schlegel (p. 235, 1876) is outside the range of *torquatus* as defined by Hershkovitz (1990), and there is a certain, as yet unresolved, confusion concerning the diagnostic phenotypic traits for the species' identification (see Lönnberg, 1939; Spix, 1823). There is, as such, a lack of clarity regarding its diagnostic characteristics, its distribution, and even its validity as a taxon. Humboldt's anecdote about *Simia lugens* was the inspiration for the name *Cheracebus* (see below).

Etymology: "*Chera*" is the Latin form of $\chi\eta\rho\alpha$; Greek for "widow". "*Cebus*" comes from the Greek "*kebos*", which means "long-tailed monkey". Humboldt (1811, 1852) referred to it as the "viudita" of the Orinoco and recounted that missionaries called it the widow monkey because of its pelage colouration—a pale face, white collar, and white hands contrasting with an overall blackish pelage—that was reminiscent of the white veil, neckerchief, and gloves of a widow in mourning. The name persevered (Tate, 1939) and in French it has been called the "veuve" (Hill, 1960; Humboldt, 1852). A synonym of *Simia lugens* is *Saguinus vidua* Lesson, 1840: 165. "*Vidua*" is Latin for widow.

> *"The saimiri, or titi of the Orinoco, the atele, the sajou, and other quadrumanous animals long known in Europe, form a striking contrast, both in their gait and habits, with the macavahu, called by the missionaries viudita, or 'widow in mourning'. The hair of this little animal is soft, glossy, and of a fine black. Its face is covered with a mask of a square form and a whitish colour tinged with blue. This mask contains the eyes, nose, and mouth. The ears have a rim: they are small, very pretty, and almost bare. The neck of the widow presents in front a white band, an inch broad, and forming a semicircle. The feet, or rather the hinder hands, are black like the rest of the body; but the fore paws are white without, and of a glossy black within. In these marks, or white spots, the missionaries think they recognize the veil, the neckerchief, and the gloves of a widow in mourning. The character of this little monkey, which sits up on its hinder extremities only when eating, is but little indicated in its appearance."* (p. 212, Humboldt, 1852).

Distinguishing characters: *Cheracebus* comprises the *torquatus* group titis as defined by Hershkovits (1963, 1988, 1990), Kobayashi (1995), and Groves (2001) (Figure 2.4). Hershkovitz's (1990) review contains detailed descriptions of the dental, cranial and post-cranial characters which distinguish the *torquatus* group, and hence, now the genus *Cheracebus*, from all other titi monkeys. He described the diagnostic characters as follows: "Average size larger than that of other species except *C. personatus* (tables 11, 13), ethmoturbinal I larger, projecting farther behind than the maxilloturbinal bone […] average cerebral index high (table 9) [29% of greatest skull length], diploid chromosome number = 20 (subspecies unknown) [see below], forehead, forearms, sideburns, feet, and tail blackish; crown reddish, reddish brown, mahogany, or blackish; sideburns little projecting; throat collar whitish or buffy, sometimes not well defined or absent; hands blackish, buffy, yellowish, or orange; upper parts from crown to tail base reddish brown, conspicuously to faintly banded or uniformly colored; chest, belly uniformly reddish, reddish brown, or blackish" (p. 78, 1990).
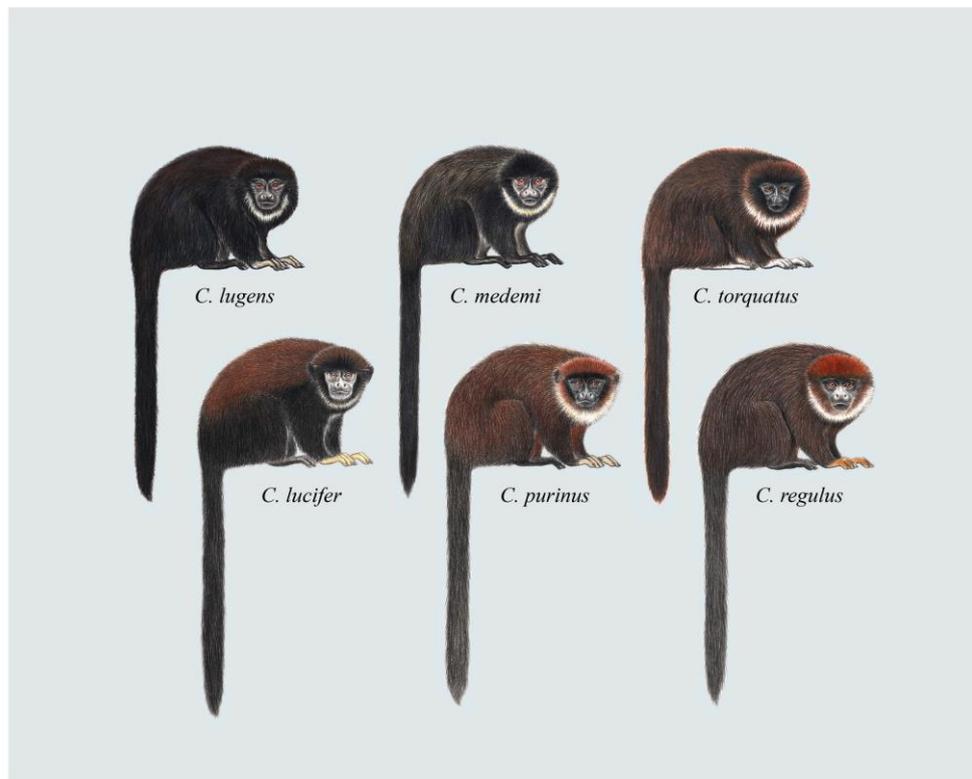
Figure 2.4. Titi monkeys, genus *Cheracebus*. Illustrations by Stephen D. Nash ©Conservation International.

Jones & Anderson (1978) summarised the diagnostic characters in a taxonomic key distinguishing *Callicebus personatus* from *Callicebus torquatus* and *Callicebus moloch*, based on Hershkovitz (1963): "Color of body reddish to black, venter either not or slightly defined from dorsum; hind feet and tail to tip black; forearms black above and below; upper surface of forefeet either whitish or blackish like the wrists".

According to Kobayashi (1990), the occlusal pattern of the upper molars is relatively smooth and simple in the *torquatus* group.

Groves (p. 176−177. 2001) added that the mesostyle and distostyle on the upper premolars are well defined, whereas in the other species-groups they are absent on P^2 and weak or absent on P^{3-4}; an entepicondylar foramen is present that is lacking in all other species; and the limbs are very long: arm 67−73% of trunk length, leg 90%. Groves (2001) did not agree with Hershkovitz's (1990) assertion that the *torquatus* group titis are unusually large.

Geographic range: Titis of the genus *Cheracebus* occur in the Amazon and Orinoco basins, in Brazil, Colombia, Ecuador, Peru, and Venezuela (Figure 2.5). North of the Solimões-Amazonas, they occur east as far as the Rio Branco in Brazil, extending into Venezuela as far north as the Rio Orinoco, west of the Río Caroni to the foothills of the Eastern Cordillera of the Andes, south of the upper Río Guaviare, Colombia, through Ecuador, north of the Río Aguarico, and into Peru to the north of the rios Amazonas and Tigre. South of the Solimões-Amazonas, they extend eastward from the Rio Javari in Brazil, across the lower and middle rios Juruá and Purus (Aquino & Encarnación, 1994; Aquino et al., 2008; Defler, 2004; Hershkovitz, 1990; Linares, 1998; Van Roosmalen et al., 2002). In Ecuador, Peru, and Brazil (primarily south of the Rio Amazonas-Solimões), titis of this genus are sympatric with a number of the smaller titis of Hershkovitz's (1988, 1990) *moloch* group.

Cheracebus lugens (Humboldt, 1811). Widow monkey, White-chested titi
Simia lugens Humboldt, A. von. 1811. *Rec. Obs. Zool. Anat. Comp.* 1: 319.
Type locality: Near San Francisco de Atabapo, at the confluence of the ríos Orinoco and Guaviare, Amazonas, Venezuela.
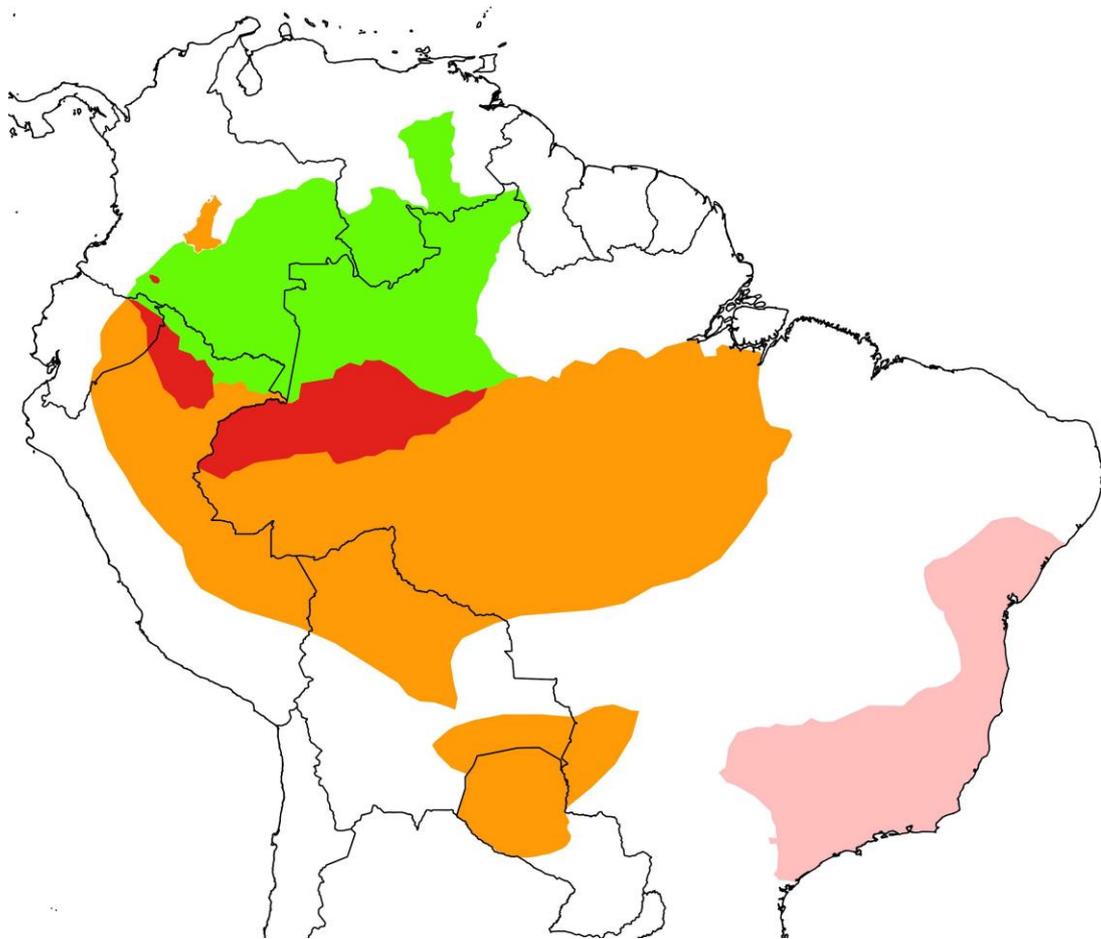
Figure 2.5 Approximate geographic distribution of *Cheracebus* (green), *Callicebus* (pink) and *Plecturocebus* (orange). The area of sympatry between *Cheracebus* and *Plecturocebus* is shown in red.

Cheracebus medemi (Hershkovitz, 1963). Black-handed titi, Medem's titi

Callicebus torquatus medemi. Hershkovitz, P. 1963. *Mammalia* 27(1): 52.

Type locality: Río Meceya, near mouth, right bank Río Caquetá, Putumayo, Colombia: altitude approximately 180 m.

Cheracebus torquatus (Hoffmannsegg, 1807). Collared titi, white-collared titi

Callitrix [sic] *torquatus* Hoffmannsegg, G. von. 1807. *Mag. Ges. Naturf. Fr.*, Berlin 10: 86.

Type locality: Codajás, north bank Rio Solimões upstream the mouth of the Rio Negro, Amazonas, Brazil (Hershkovitz, 1963; see also Lönnberg, 1939).

Cheracebus lucifer (Thomas, 1914). Yellow-handed titi

Callicebus lucifer Thomas, O. 1914. *Ann. Mag. Nat. Hist.*, 8th ser. 13: 345.

Type locality: Yahuas, N. of Loreto, about 2°40'S, 70°30'W, Alt. 500 ft. (Thomas, 1914). Yahuas territory, near Pebas, Loreto, Peru, about 125 m (Hershkovitz, 1990).

Cheracebus purinus (Thomas, 1927). Rio Purus titi
Callicebus purinus Thomas, O. 1927. *Ann. Mag. Nat. Hist*. 9th ser. 19: 509.
Type locality: Ayapuá, lower Rio Purus, southern affluent of Rio Solimões, Brazil.

Cheracebus regulus (Thomas, 1927). Juruá collared titi
Callicebus regulus Thomas, O. 1914. *Ann. Mag. Nat. Hist*. 9th ser. 19: 510.
Type locality: Fonte Boa, upper Rio Solimões, Amazonas, Brazil.

Callicebus Thomas, 1903

Thomas, O. 1903. *Ann. Mag. Nat. Hist*., 7th series, 12: 456. Type species. *Simia personata* É. Geoffroy Saint-Hilaire, 1812.

Type species. *Simia personata* Geoffroy Saint-Hilaire, É. 1812. In: Humboldt, 1812. *Rec. Obs. Zool.*, p. 357.

Etymology: "Calli" is from the Greek *kalos*, which means "beautiful". "Cebus" is from the Greek *kebos*, which means "a long-tailed monkey".

Distinguishing characters: The genus *Callicebus* is here restricted to the Atlantic forest titis that were listed as subspecies of *C. personatus* in the *moloch* group by Hershkovitz (1990), and as members of a distinct *C. personatus* group by Kobayashi (1995) and Groves (2001) (Figure 2.6). Groves (2001) also included *C. coimbrai* Kobayashi & Langguth, 1999. Hershkovitz's (1990) review contains detailed descriptions of the dental, cranial and post-cranial skeletal characters which distinguish *C. personatus* from all other titi monkeys [see also Kobayashi (1995) for craniometric differences]. Hershkovitz (p. 70−71, 1990) diagnosed *C. personatus* as follows: "Average size largest […]; cranial characters essentially as in *moloch* group except average cerebral index greater, average brain case index less […]; pelage coarse, shaggy with full coat of hidden brownish wool hairs; color of trunk variable, cover hairs with 2 or 4 pheomelanic bands sharply defined to shadowy, or uniformly,

43

pheomelanin; cheiridia blackish, the blackish often extending proximally as a tapered band to mid-arm or mid-foreleg, remainder of limbs grayish, buffy, yellowish or orange, the hairs banded or unbanded; facial hairs long, often comparatively thick but not concealing skin; forehead blackish with or without fine buffy banding; sideburns and ear tufts blackish; tail orange, reddish, mahogany, or mixed with blackish, never entirely blackish."
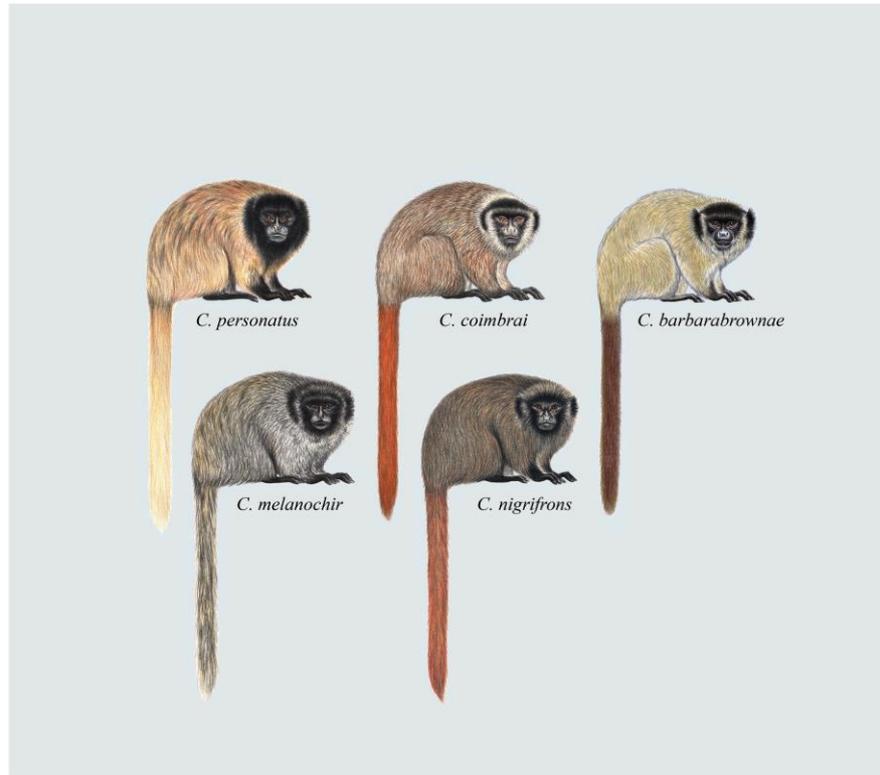
Figure 2.6. Titi monkeys, genus *Callicebus*. Illustrations by Stephen D. Nash ©Conservation International.

Jones & Anderson (1978) summarised the diagnostic characters in a taxonomic key distinguishing *Callicebus personatus* from *Callicebus torquatus* and *Callicebus moloch*, based on Hershkovitz (1963): "Distal portion of limbs (at least forefeet and hind feet) black and in sharp contrast to the gray or rufous of wrists and other proximal parts". According to Kobayashi (1990), the *personatus* group shows the most uneven and variable occlusal pattern in the upper molars, with the largest number of small cusps and conules.

Callicebus coimbrai, not included by Hershkovitz (1990), conforms. It has a black forehead, crown, and ears, and a buffy body; pale cheek whiskers, the colour

extending to the nape; hands and feet blackish, tail orange, and zebra stripes on the upper back (Groves, 2001). The diagnostic features of the *personatus* group given by Groves (p. 175, 2001) summarised Hershkovitz (1990).

Geographic range: Endemic to Brazil (Figure 2.5). These titis are known from north-eastern Brazil, south of the Rio São Francisco in forest patches in the Caatinga (*barbarabrownae*) and Atlantic forest (*coimbrai*), south through the Atlantic forest of the states of Bahia, Espírito Santo, and Rio Janeiro, west as far as the rios Paraná and Paranaíba, and south to the Rio Tieté in the state of São Paulo (Chagas & Ferrari, 2010; Hershkovitz, 1990; Printes et al., 2013; Van Roosmalen et al., 2002).

Callicebus personatus (É. Geoffroy Saint-Hilaire, 1812). Masked titi
Simia personata Geoffroy-Saint Hilaire, É. 1812. In: Humboldt, 1812. *Rec. Obs. Zool.*, p. 357.
Type locality: Brazil. Restricted by Hershkovitz (1990) to the lower Rio Doce, Espírito Santo, Brazil.

Callicebus coimbrai Kobayashi & Langguth, 1999. Coimbra-Filho's titi
Callicebus coimbrai Kobayashi, S. & Langguth, A. 1999. *Revta. Bras. Zool*. 16(2): 534.
Type locality: Proximity of the small village of Aragão, in the region of Santana dos Frades about 11.0 km SW of Pacatuba, south of the estuary of the Rio São Francisco, state of Sergipe, Brazil. 10°32'S, 36°41'W, altitude 90 m.

Callicebus barbarabrownae Hershkovitz, 1990. Blond titi
Callicebus personatus barbarabrownae Hershkovitz, P. 1990. *Fieldiana, Zool., n.s.,* (55): 77.
Type locality: Lamarão, Bahia, Brazil, altitude about 300 m above sea level.

Callicebus melanochir (Wied-Neuwied, 1820). Southern Bahian titi
Callithrix melanochir Wied-Neuwied, M. A. P. von. 1820. *Reise nach Brasilien in den Jahren 1815 bis 1817.* Vol. 1. H. L. Bronner, Frankfurt am Main, p. 258 and fn.
Type locality: Morro d'Árara or Fazenda Arara, state of Bahia, Brazil (Hershkovitz, 1990).

Callicebus nigrifrons (Spix, 1823). Black-fronted titi

Callithrix nigrifrons Spix, J. B. von. 1823. *Sim. Vespert. Brasil.*, p. 21.

Type locality: Brazil. Restricted by Hershkovitz (1990) to the Rio Onças, municipality of Campos, Rio de Janeiro, Brazil.

Plecturocebus new genus

LSID: urn:lsid:zoobank.org:act:1E86C672-5008-4DB6-8776-53595C157FEA

Type species. *Plecturocebus moloch* (Hoffmannsegg, 1807) Red-bellied titi

Cebus moloch Hoffmannsegg, G. von. 1807. *Mag. Ges. Naturf. Fr.*, Berlin, 9: 97.

Etymology: "Plect-" comes from the Greek *plektos*, which means plaited or twisted. In Latin, *Plecto* and *plexus* refer to a braid, plait, or interweave. "Uro-" comes from the Greek word *oura*, which means "tail". "Cebus" is from the Greek *kebos*, which means "a long-tailed monkey". The name refers to the tail-twining behaviour of the Callicebinae. Titis, adults and juveniles, frequently intertwine their tails when they sit side-by-side; sometimes looped quite loosely, sometimes wound around very tightly, making several turns. The behaviour is affiliative (Moynihan, 1966).

Diagnostic characters: Hershkovitz's (1990) review contains detailed descriptions of the dental, cranial and post-cranial characters of the titi species recognized at the time, and presents summaries of the key characteristics of his *modestus* (included here in the *donacophilus* group), *donacophilus* (Figure 2.7) and *moloch* (Figure 2.8) groups. Groves' (2001) taxonomy, with some exceptions, followed that of Hershkovitz, and the distinguishing features he provided, and that we record here, are from Hershkovitz's comprehensive 1990 review.

Groves (p. 171, 2001) summarized the *modestus* group as follows: "Externally resembles the *moloch* group, but cranially primitive according to Hershkovitz (1990), with an elongate, low-slung cranium, very small cranial capacity, only 20% of greatest skull length, and short occiput, condylobasal length averaging 86% of greatest skull length. Median pterygoids very large; mandibular angle large. Postcranial skeleton unknown; chromosomes unknown".

Characteristics of species of the *donacophilus* group (*donacophilus*, *pallescens*, *olallae* and *oenanthe*) were summarized as follows by Groves (p. 171, 2001): "Cranial capacity 21−25% of greatest skull length, condylobasal length 81−84% of greatest skull length. Arm (radius plus humerus) 52−58% of trunk length, leg (tibia plus femur) 71−78%. Chromosomes 2n = 50".

Characteristics of the *moloch* group, including the species *cinerascens*, *hoffmannsi*, *baptista*, *moloch*, *brunneus*, *cupreus* (synonyms *caligatus*, *discolor*, *toppini*, and *dubius*), and *ornatus*, were summarized by Groves (p. 172−173, 2001) as follows: "Cranial capacity 26−29% of greatest skull length; condylobasal length 78−82%. Forelimb (known only for *C. cupreus*) 53-61% of trunk length, hind limb 72−81%. Chromosomes 2n = 48 (*C. moloch*, *C. brunneus*) or 46 (*C. cupreus*, *C. ornatus*)".

Figure 2.7. Titi monkeys, the *donacophilus* group of *Plecturocebus*. Illustrations by Stephen D. Nash ©Conservation International.

Geographic range: Brazil, Colombia, Ecuador, Peru, Bolivia, Paraguay (Figure 2.5). The northernmost limit is the upper reaches of the Río Meta in Colombia (*Plecturocebus ornatus*) extending south to the upper Río Guaviare. *Plecturocebus*

caquetensis occurs in a small portion of the upper Caquetá basin in Colombia. All other representatives of this genus occur throughout the greater part of the Amazon basin, south of the ríos Içá-Putumayo and Amazonas-Solimões, east of the Andes, extending south through Ecuador, Peru, Brazil, and Bolivia into Paraguay to the confluence of the ríos Pilcomayo and Paraguai. In Brazil, they occur east as far as the Rio Tocantins-Araguaia, south of the Rio Amazonas (Aquino & Encarnación, 1994; Defler, 2004; Hershkovitz, 1990; Martínez & Wallace, 2010; Stallings et al., 1989; Tirira, 2007; Van Roosmalen et al., 2002).
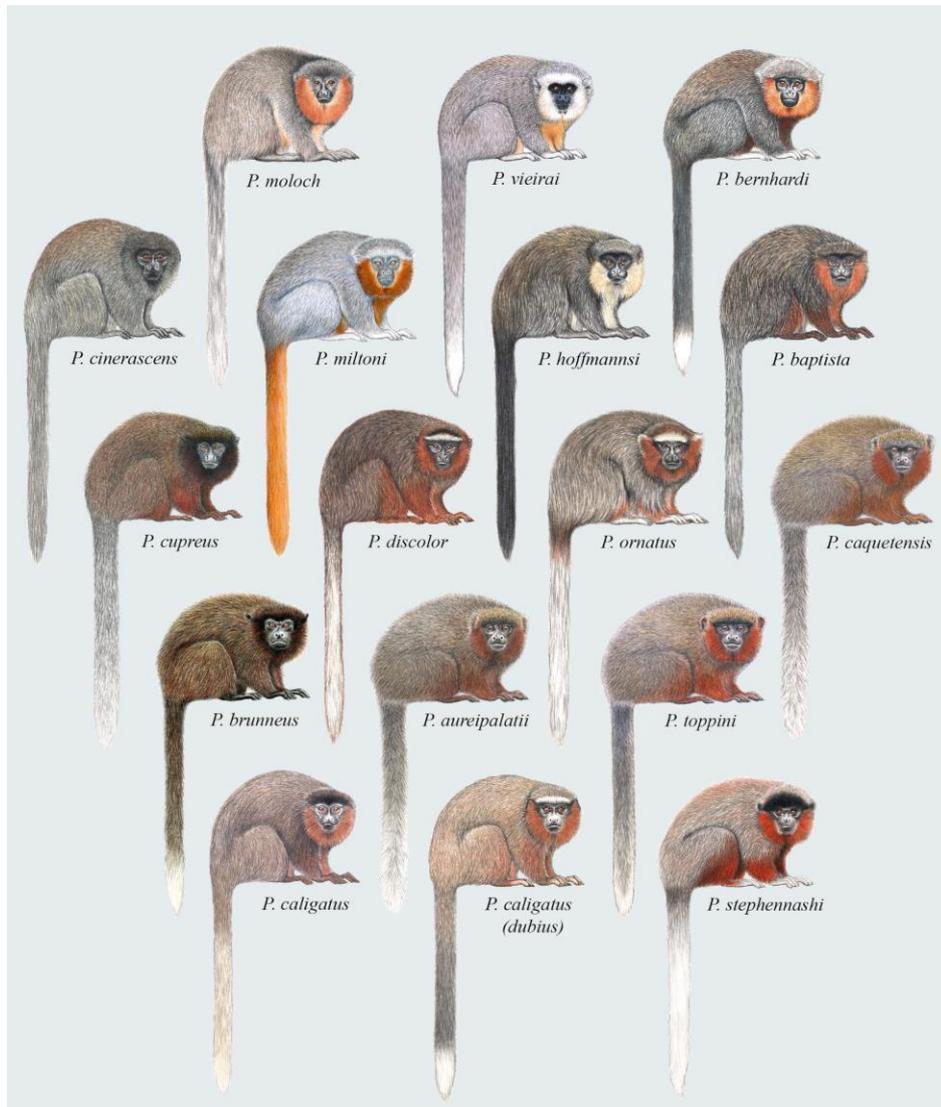
Figure 2.8. Titi monkeys, the *moloch* group of *Plecturocebus*. Illustrations by Stephen D. Nash ©Conservation International.

Plecturocebus donacophilus group

Plecturocebus donacophilus (D'Orbigny, 1836). White-eared titi

Callithrix donacophilus D'Orbigny, M. A. D. 1836. *Voy. Am. Merid., Atlas Zool*., pl. 5.

Type locality: Rio Mamoré basin, Beni, Bolivia.

Plecturocebus pallescens (Thomas, 1907). White-coated titi

Callicebus pallescens Thomas, O. 1907. *Ann. Mag. Nat. Hist*., 7th ser., 20: 161.

Type locality: Thirty miles north of Concepción, Chaco, Paraguay.

Plecturocebus oenanthe (Thomas, 1924). Río Mayo titi

Callicebus oenanthe Thomas, O. 1924. *Ann. Mag. Nat. Hist*. 9th ser., 14: 286.

Type locality: Moyobamba, San Martín, Peru, altitude *c.* 840 m above sea level.

Plecturocebus olallae (Lönnberg, 1939). Olalla Brother's titi

Callicebus olallae Lönnberg, E. 1939. *Ark. f. Zool.,* 31A, 13: 16.

Type locality: La Laguna, 5 km from Santa Rosa, Beni, Bolivia, altitude *c.* 200 m above sea level.

Plecturocebus modestus (Lönnberg, 1939). Rio Beni titi

Callicebus modestus Lönnberg, E. 1939. *Ark. f. Zool.,* 31A, 13: 17.

Type locality: El Consuelo, Río Beni, Beni, Bolivia, altitude 196 m above sea level.

Plecturocebus urubambensis (Vermeer & Tello-Alvarado, 2015). Urubamba brown titi

Callicebus urubambensis Vermeer, J. & Tello-Alvarado, J. C. 2015. *Primate Conserv.* (29): 19.

Type locality: Peru, near the Colonia Penal del Sepa, on the southern bank of the Río Sepa, a western tributary of the Río Urubamba (10°48'50"S, 73°17'80"W). Altitude 280 m.

<p align="center">*Plecturocebus moloch* group</p>

Plecturocebus moloch (Hoffmannsegg, 1807) Red-bellied titi

Cebus moloch Hoffmannsegg, G. von. 1807. *Mag. Ges. Naturf. Fr.*, Berlin, 9: 97.

Type locality: Near the town of Belém, Pará, Brazil. Hill (1960) gives the type locality as the banks of the Rio Pará (= terminal part of the Rio Tocantins), Pará, Brazil. Redetermined by Hershkovitz (1963) as the right bank of the lower Rio Tapajós, municipality of Santarém, Pará, Brazil.

Plecturocebus vieirai (Gualda-Barros, Nascimento & Amaral, 2012). Vieira's titi

Callicebus vieirai Gualda-Barros, J., Nascimento, F. O. do & Amaral, M. K. do. 2012. *Pap. Avuls. Zool., São Paulo* 52(53): 263.

Type locality: Rio Renato, tributary of Rio Teles Pires (right bank), nearby the city of Cláudia, state of Mato Grosso, Brazil (11°33'00.15"S, 55°10'59.98"W); around 370 m above sea level.

Plecturocebus bernhardi (M. G. M. Van Roosmalen, T. Van Roosmalen & Mittermeier, 2002). Prince Bernhard's titi

Callicebus bernhardi Van Roosmalen, M. G. M., Van Roosmalen, T. and Mittermeier, R. A. 2002. *Neotrop. Primates* 10(suppl.): 24.

Type locality: West bank of the lower Rio Aripuanã, at the edge of the settlement of Nova Olinda, 41 km southwest of the town of Novo Aripuanã, Amazonas state, Brazil (05°30'63"S, 60°24'61"W); altitude 45 m above sea level.

Plecturocebus cinerascens (Spix, 1823). Ashy titi

Callithrix cinerascens Spix, J. B. von. 1823. *Sim. Vespert. Brasil.*, p. 20, pl.14.

Type locality: Unknown. Spix indicated the Río Putumayo-Içá in the vicinity of the Peru-Brazil border, but, as indicated by Hershkovitz (1990), there is no evidence that it was ever collected there. This species occurs on right bank of the Rio Aripuanã, a tributary of the Rio Madeira (Noronha et al., 2007; Van Roosmalen et al., 2002).

Plecturocebus miltoni (Dalponte, Silva & Silva-Júnior, 2014). Milton's titi

Plecturocebus miltoni Dalponte, J. C., Silva, F. E. & Silva-Júnior, J. de S. 2014. *Pap. Avuls. Zool., São Paulo* 54(32): 462.

Type locality: Curva do Cotovelo, region of the mouth of the Pombal stream, Reserva Extrativista Guariba-Roosevelt, right bank of the upper Roosevelt River, municipality of Colniza, Mato Grosso, Brazil (08°59'45.21"S, 60°43'42.72"W).

Plecturocebus hoffmannsi (Thomas, 1908). Hoffmanns's titi
Callicebus hoffmannsi Thomas, O. 1908. *Ann. Mag. Nat. Hist*., 8th series, 2: 89.
Type locality: Urucurituba, Santarém, Rio Tapajós, Pará, Brazil.

Plecturocebus baptista (Lönnberg, 1939). Lake Baptista titi
Callicebus baptista Lönnberg, E. 1939. *Ark. f. Zool.,* 31A, 13: 7.
Type locality: Determined by Hershkovitz (p. 29, 1963) as the Lago do Baptista, right bank of the Rio Madeira, north of the Paraná Urariá and east of the town of Nova Olinda do Norte, Amazonas, Brazil (Van Roosmalen et al., 2002). Syntypes collected from the Lago Tapaiuna.

Plecturocebus cupreus (Spix, 1823) Coppery titi
Callithrix cuprea Spix, J. B. von. 1823. *Sim. Vespert. Brasil*., p. 23, pl. 17.
Type locality: Rio Solimões, Brazil, near the Peruvian boundary. Restricted to Tabatinga by Hershkovitz (p. 36, 1963), but should be opposite Tabatinga because the species does not occur on the north bank or Tabatinga side of the Solimões (p. 61, Hershkovitz, 1990).

Plecturocebus discolor (I. Geoffroy Saint-Hilaire & Deville, 1848). Red-crowned titi
Callithrix discolor Geoffroy Saint Hilaire, I. & Deville, É. 1848. *C. R. Acad. Sci. Paris*, 27: 498.
Type locality: Sarayacu, Río Ucayali, Ucayali, Peru.

Plecturocebus ornatus (Gray, 1866). Ornate titi
Callithrix ornata Gray J. E. 1866. *Ann. Mag. Nat. Hist*., 4th ser., 17: 57.
Type locality: "Nouvelle Grenade", now Colombia, restricted to the Villavicencio region, Río Meta, Meta, Colombia, by Hershkovitz (p. 44, 1963).

Plecturocebus caquetensis (Defler, Bueno & Garcia, 2010). Caquetá titi

Callicebus caquetensis Defler, T. R., Bueno. M. L. & García, J. 2010. *Primate Conserv.* (25): 2.

Type locality: Vereda El Jardin, east of Valparaiso, municipality of Puerto Milan, Department of Caquetá, Colombia (01°8'24.61"N, 75°32'34.04"W); 251 m above sea level.

Plecturocebus brunneus (Wagner, 1842). Brown titi

Callithrix brunea Wagner, J. A. 1842. *Arch. Naturgesch.*, 8(1): 357.

Type locality: Brazil, subsequently specified by Pelzeln (p. 20, 1883) as Rio Mamoré, Cachoeira da Bananeira, Rondônia, Brazil.

Plecturocebus aureipalatii (Wallace, Gómez, A. M. Felton & A. Felton, 2006). Madidi titi

Callicebus aureipalatii Wallace et al. 2006. *Primate Conserv.* (20): 31.

Type locality: Campamento Roco Roco, Río Hondo, Madid National Park and Natural Area of Integrated Management, La Paz Department, Bolivia (14°37'30"S, 67°43'06"W).

Plecturocebus toppini (Thomas, 1914). Toppin's titi

Callicebus toppini Thomas, O. 1914. *Ann. Mag. Nat. Hist.*, ser. 8, 13: 480.

Type locality: Rio Tahuamanu, northeast Peru [sic] near Bolivian boundary. About 12°20'S, 68°45'W. The Rio Tahuamanu and the Bolivian border are in fact in southeast Peru, not northeast; evidently a *lapsus calami*.

Plecturocebus caligatus (Wagner, 1842). Chestnut-bellied titi

Callithrix caligata Wagner, J. A. 1842. *Arch. Naturgesch.*, 8(1): 357.

Type locality: Restricted by Thomas (p. 90, 1908) to Borba, Rio Madeira, Amazonas Brazil.

Plecturocebus dubius (Hershkovitz, 1988). Doubtful titi

Callicebus dubius Hershkovitz, P. 1988. *Proc. Acad. Nat. Sci. Philadelphia* 140(1): 264.

Type locality: Said to be Lago de Aiapuá (= Ayapuá), west bank, lower Rio Purus, more likely on the east bank of the lower Rio Purus, probably opposite of the Lago do

Aiapuá (Hershkovitz, 1990). Röhe & Silva-Júnior (2009) recorded that the species had crossed from the Mucuim-Ituxi interfluvium to the right bank of the Rio Mucium using a man-made bridge. Here considered a junior synonym of *P. caligatus*.

Plecturocebus stephennashi (M. G. M. Van Roosmalen, T. Van Roosmalen & Mittermeier, 2002). Stephen Nash's titi

Callicebus stephennashi Van Roosmalen, M. G. M., Van Roosmalen, T. and Mittermeier, R. A. 2002. *Neotrop. Primates* 10(suppl.): 15.

Type locality: Unknown, holotype and paratypes said to be have been caught somewhere along the middle to upper Rio Purus, Amazonas, Brazil.

2.5.2 Genus-level topology

Our proposal to divide *Callicebus* into three distinct genera gains support from previous molecular phylogenetic analyses (e.g., Canavez et al., 1999; Perelman et al., 2011; Springer et al., 2012). Our divergence-time estimates for the genus-level splits (*Cheracebus c.* 11 Ma; *Callicebus c.* 8.3 Ma), are comparable to those reported by Springer et al. (2012) (*Cheracebus c.* 7.8 Ma; *Callicebus c.* 7.2 Ma) and Perelman et al. (2011) (*Callicebus c.* 9.9 Ma). Based on phylogenomic evidence, Jameson Kiesling et al. (2015) estimated the divergence time of *Callicebus* and *Plecturocebus* at 6.7 Ma, and suggested that these two species groups required the designation of separate genera based on the time-classification criteria proposed by Goodman et al. (1998).

The phyletic groups proposed by Kobayashi (1995) using cranial morphometrics correspond with the arrangement found using molecular evidence in the present study. Kobayashi (1995) noted that the *torquatus* group (*Cheracebus*) and the *personatus* group (*Callicebus*) presented a high degree of character differentiation, while the *donacophilus*, *moloch* and *cupreus* groups (*Plecturocebus*) were more closely related. In discordance with his proposal, we found support for the division of *Plecturocebus* into two, not three, species groups. The *donacophilus* group is indeed a distinct early diverging lineage but Kobayashi's (1995) *moloch* and *cupreus* groups are better described as a single group, which began diversifying *c.* 3.4 Ma. To account for paraphyly in the current group arrangement, we propose that all Amazonian titis of the *cupreus* and *moloch* groups (*sensu* Kobayashi, 1995) should be assigned to a single *moloch* group, conforming to the *moloch* group identified by Groves (2001).

We argue that increased resolution of the species-level relationships among these species is required to justify erecting any additional species group.

Body size and pelage colouration also support our taxonomic hypothesis. The *moloch* species group of *Plecturocebus* is composed of medium-sized "typical" titis characterised by the greyish or brownish dorsum with a contrasting whitish, orange or reddish belly (except *P. cinerascens* and *P. brunneus*; see Figure 2.8), while the *donacophilus* clade taxa are the smallest species, generally showing a buffy to dark grey pelage that lacks contrast (Figure 2.7) (Hershkovitz, 1988; Kobayashi, 1995). *Callicebus* are distinguished by their large size and overall appearance (Figure 2.6), distinct from other callicebine taxa (see Groves, 2001). Hershkovitz (1988) indicated that *Cheracebus* species are larger than the species of *Plecturocebus*, but Groves (p. 176, 2001) found that this was not borne out by the available measurements. They are distinguishable from all other titis, however, by their uniform dark reddish to blackish pelage with contrasting whitish throat collar (Figure 2.4) and also their postcranial skeleton.

Our conclusions based on molecular evidence are further supported by karyological data. The subfamily Callicebinae presents extensive karyotypic variation that corresponds closely to the present genera derived from molecular and morphological data. *Cheracebus* is characterised by low chromosome number; 2n = 20 in *C. torquatus* (Benirschke & Bogart, 1976) and *C. lucifer* (Bueno & Defler, 2010), and 2n = 16 in *C. lugens*, the lowest diploid chromosome number and most derived karyotype known among all primates (Stanyon et al., 2003). *Callicebus nigrifrons* and *C. personatus*, show intermediate chromosome numbers of 2n = 42 and 2n = 44, respectively (Rodrigues et al., 2004). *Plecturocebus* taxa have the highest chromosome numbers, ranging from 2n = 44 (*P. ornatus*) (Bueno et al., 2006) to 2n = 50 (*P. hoffmannsi*, *P. donacophilus*) (de Boer, 1974; Rodrigues et al., 2001).

Wood & Collard (1999) argued that the designation of a genus should include "an ecological situation, or adaptive zone, that is different from that occupied by the species of another genus". Our three genera satisfy these conditions with each having distinct geographic distributions (Figure 2.5) and habitat preferences (Ferrari et al., 2013). *Callicebus* species are entirely extra-Amazonian and geographically well separated from all other titis. They are found in the Atlantic Forest region of eastern Brazil, as far south and west as the Tietê-Paraná-Parnaíba river system, and as far

north as the Rio São Francisco (Chagas & Ferrari, 2010). This includes the range of *C. barbarabrownae*, which occupies the Caatinga biome of northeast Brazil.

Cheracebus is the northern-most genus, occurring in the Amazon Basin to the west of the rios Branco and Negro (north of the Rio Amazonas) and west of the Rio Purus (south of the Rio Amazonas), with the geographic range of *C. lugens* extending north of the Rio Negro into Venezuela and Colombia (Ferrari et al., 2013). In the southern part of their range, *Cheracebus* species are sympatric with species of the *moloch* group of *Plecturocebus*, which occur throughout the southern and western Amazon basin (Figure 2.5). However, it is unlikely that this has resulted in extensive niche overlap. *Cheracebus* species prefer open-canopy forests, with tall trees and well-drained soils, and make use of higher levels of the canopy, whereas *moloch* group species occupy the dense understoreys of vegetation, thick with lianas (Johns, 1991; Defler, 2004). Where they are sympatric, it has been reported that *Cheracebus* species often inhabit areas of poor vegetation, outcompeted by the *moloch* group species for more favourable habitats (Kinzey, 1981; Kinzey & Gentry, 1979) Although still little studied, *Cheracebus* and sympatric *Plecturocebus* undoubtedly have different dietary preferences, with *Cheracebus* species consuming more insects, seeds and tougher fruits, while the diets of the *moloch* group species contain more leaves (Bicca-Marques & Heymann, 2013; Ferrari et al., 2013; Heymann & Nadjafzadeh, 2013; Palacios et al., 1997; Palacios & Rodríguez, 2013).

The range of the *donacophilus* group species of *Plecturocebus* extends far south of the Amazon basin and they have the most disjunct set of species distributions of the titi monkey clades. They occupy forest patches and gallery forests in the savannah floodplains of Bolivia, Paraguay and Brazil, with the range of *P. pallescens* extending into the Chaco scrublands and Pantanal swamps in Paraguay and Brazil (Ferrari et al., 2013; Rumiz, 2012; Silva-Júnior et al., 2013; Stallings, 1985)

As we have sequence data for only one species of the *donacophilus* clade, we are limited in our ability to make novel inferences about this group. The morphological, molecular and ecological differences between these two species groups may justify a new classification for taxa of the *donacophilus* clade, pending increased taxonomic sampling and sequence data.

For the taxa not included in this study we will continue to follow the arrangement proposed by Groves (2005) (Table 2.1), with the exception of *P. modestus*. Only a single adult specimen has been collected to date. Hershkovitz (1988,

1990) noted the unusual elongated skull of *P. modestus* and regarded it as the most primitive titi monkey species. Because of this, he created the *modestus* group, a proposal followed by Groves (2001, 2005). Kobayashi (1995) moved *P. modestus* to the *donacophilus* group, but stated "the phylogenetic position of *P. modestus* is morphometrically debatable" (p.119) and that a sufficient number of samples need to be collected to clarify placement. Although new observations have been made in the wild (Felton et al., 2006), no further adult *P. modestus* specimens have been collected and thus we follow Kobayashi (1995) in maintaining *P. modestus* in the *donacophilus* group.

2.5.3 Species-level topology

Our phylogenetic analyses showed strong support for most of the nodes in the Callicebinae phylogeny. At species-level, phylogenetic relationships among taxa of *Cheracebus* and *Callicebus* are identical in all analyses, however they varied among species of the *moloch* group of *Plecturocebus*.

Based on the analysis of museum specimens, Auricchio (2010) suggested that the pelage colouration of *P. bernhardi* is consistent with polymorphic variation found in *P. moloch* specimens, and considered *P. bernhardi* as a junior synonym of *P. moloch*. He states that a mitochondrial phylogeny also supports the classification of all "*moloch*" phenotypes as polymorphic variants of the same species, including *P. bernhardi* and a specimen from the Alta Floresta region (likely *P.* cf. *moloch*). However, the molecular data and phylogenetic trees were not presented in the study. This classification is in conflict with the results from our molecular datasets, showing support for three distinct taxa, with a sister-clade relationship between *P. bernhardi* and *P. moloch*/*P.* cf. *moloch*. Divergence time analyses date the split between *P. bernhardi* and *P. moloch*/*P.* cf. *moloch* at *c.* 1.7 Ma, representing one of the oldest speciation events within the *moloch* group and providing support for the validity of *P. bernhardi* as a distinct species. *Plecturocebus moloch* and *P.* cf. *moloch* are highly supported as distinct sister-taxa across all datasets, and divergence time analyses date the split at *c.* 1.1 Ma, comparable to other speciation times within the *moloch* group. Seven *P. moloch* specimens from three different localities (see Table 2.2) are included in this study, however, in contrast, the earliest diversification event within *P. moloch* is estimated at *c.* 0.4 Ma. The molecular evidence presented here provides support for the designation of *P.* cf. *moloch* as a valid species. This taxon occurs in the Alta

Floresta region of Mato Grosso, Brazil and our group is currently working on this new species description (Boubli et al. in prep.).

Our results cast doubt on the validity of the species status of *P. dubius*. For the nuclear and combined datasets, *P. dubius* is a minimally diverged sister taxon of *P. caligatus*, and for the mitochondrial dataset, *P. dubius* is paraphyletic and most of the nodes within the *P. caligatus*/*P. dubius* clade show low support. Pairwise genetic distances estimated for the cytochrome *b* locus between *P. caligatus* and *P. dubius* (see Table A1.3; 0.01–0.06) indicate that the two taxa show very low genetic differentiation. *Plecturocebus caligatus* occurs in the interfluve delineated by the rios Purús/Solimões/Madeira/Ipixuna, and to the southwest *P. dubius* is found between the rios Purús/Mucuím/Madeira (southern limit unknown). The pelage colouration of *P. caligatus* and *P. dubius* is also highly similar; Hershkovitz (1988) noted that the only distinguishing feature between *P. caligatus* and *P. dubius* was the whitish frontal band found in the latter, and suggested that rather than indicating two distinct species, forehead colouration could be a variable feature in *P. caligatus*. This white frontal blaze is poorly developed in some *P. dubius* specimens (pers. obs.). Considering the morphological, molecular, and geographical affinities between *P. caligatus* and *P. dubius*, we propose the designation of *P. dubius* (Hershkovitz, 1988) as a junior synonym of a polymorphic *P. caligatus*. We suggest that the phenotypic differences found between these taxa represent geographic variation in pelage colouration.

Based on cranial morphometrics, Kobayashi (1995) suggested that *P. brunneus* was closely related to his *moloch* group species, however, the skulls of *P. brunneus* studied were of two species, *P. urubambensis* and *P. brunneus*, which may have affected the results. Our analyses support a western Amazonian species-complex composed of *P. brunneus*, and Kobayashi's *cupreus* group species, *P. cupreus* and *P. caligatus*. *Plecturocebus cupreus* is the earliest diverging lineage within this clade, and *P. brunneus* is the sister taxon to *P. caligatus* (*P. dubius*). Although the relationships between these west-Amazonian species are well resolved in the mitochondrial and combined datasets, the nuclear topology differs but with low support across most of the nodes. We consistently find two distinct *P. cupreus* clades, with an estimated divergence time of 1 Ma. These two clades are not sister in the nuclear dataset phylogenies. The *P. cupreus* clade A samples are from museum specimens with known localities in the Amazon basin, whereas those of *P. cupreus*

clade B come from a collection of blood samples with no available skins, skulls or geographical data.

Plecturocebus cinerascens has an overall grey agouti pelage, lacking the contrasting colours characteristic of the *moloch* group, leading Hershkovitz (1988) to suggest that *P. cinerascens* is the most primitive member. In this study, the nuclear dataset supports *P. cinerascens* as the earliest diverging lineage, forming a sister-clade to all other species of the *moloch* group, followed by the divergence of *P. hoffmannsi* and then *P. miltoni* and the rest of the moloch group. However, the mitochondrial dataset supports an alternative topology where *P. cinerascens* and *P. miltoni* form a sister-group to the *P. bernhardi* and *P. moloch* clade. Analyses based on combined data show the same topology as mitochondrial phylogenies, but with low support for the *P. cinerascens*/*P. miltoni* and *P. bernhardi*/*P. moloch* sister group relationship, likely as a result of strong conflict between the nuclear and mitochondrial phylogenetic signals. Using mitochondrial loci alone does not resolve the phylogenetic position of *P. hoffmannsi*, however, the combined phylogenetic signal from nuclear and mitochondrial markers supports *P. hoffmannsi* as an early diverging lineage. All taxonomic reviews to date infer a close relationship with *P. baptista*, and thus our results suggest that *P. hoffmannsi* and *P. baptista* are a sister clade to all remaining *moloch* group taxa, with the exception of *P. cinerascens* and *P. miltoni* (position unresolved).

The *P. caligatus* and *P. moloch* specimens sequenced by Perelman et al. (2011) were incorrectly identified and our results indicate that their *P. caligatus* sample is likely *P. donacophilus*. The identity of the *P. moloch* specimen of Perelman et al. (2011) is unknown, however, it is sister to our *P. hoffmannsi* individuals in all analyses and so we labelled it *P.* cf. *hoffmannsi*. In our divergence time analyses, we estimate that these taxa diverged *c.* 1.2 Ma, thus *P.* cf. *hoffmannsi* may be a distinct species. Further investigation is required to confirm whether *P.* cf. *hoffmannsi* is one of the known species of *Plecturocebus* that have not been analysed. It is also possible both these specimens from Perelman et al. (2011) are captive hybrids.

2.5.4 Age estimates

Our time-calibrated phylogeny suggests that the callicebine lineages began to radiate in the late Miocene, with the origin of *Cheracebus* at around 11 Ma, followed by the divergence of *Callicebus* and *Plecturocebus* at around 8.3 Ma. The timescale for titi monkey evolution estimated here is compatible with the fossil record of the platyrrhines and with other recent molecular analyses (see Table 2.11; Jameson Kiesling et al., 2015; Perelman et al., 2011; Schrago et al., 2013; Springer et al., 2012). Within *Plecturocebus*, we find support for divergent lineages leading to *P. donacophilus*, *P. hoffmannsi*, and the remaining taxa that date to the Pliocene, *c.* 4.4 Ma and 3.4 Ma, respectively. Within the *moloch* group, we find a sister-clade relationship between east- and west-distributed Amazonian species, which diverged *c.* 2.8 Ma. Nearly all the *moloch* group sister-species divergences in this study occurred 2–1 Ma, pointing to a rapid Pleistocene diversification of this group.

Table 2.11 Comparison of estimated divergence times (combined dataset) with other recent studies.

Clade or Split	Mean age (Ma)				
	Perelman et al. (2011)	Springer et al. (2012)	Schrago et al. (2013)	Kiesling et al. (2015)	Present study
Crown Pitheciidae	24.82	23.3	21.9	25.51	21.47
Pitheciinae vs. Callicebinae	20.24	20.7	19.6	18.08	18.71
Cheracebus vs. *Callicebus* + *Plecturocebus*	n/a	7.81	n/a	n/a	10.98
Callicebus vs. *Plecturocebus*	9.86	7.16	n/a	6.65	8.34
Plecturocebus: *donacophilus* group vs. *moloch* group	4.69	3.22	n/a	n/a	4.39

The three Callicebinae genera proposed here are isolated from each other by major biogeographical barriers: the Amazonian *Plecturocebus* titis are largely separated from the northernmost genus, *Cheracebus*, by the Rio Amazonas, and from the Atlantic Forest genus, *Callicebus*, by the Cerrado and Caatinga biomes of central Brazil (Figure 2.5). At the species level, larger rivers in Amazonia frequently delimit the geographic distribution of titi monkeys, and recent evidence suggests that they can act as isolating barriers for sister taxa, promoting vicariant divergence (Boubli et al., 2015). Together, these characteristics make the subfamily Callicebinae of particular interest for the study of Amazonian biogeographical history.

2.6 Conclusions

In this study, we provide the first molecular review of the subfamily Callicebinae, and our phylogenetic analyses help to clarify a number of issues on the taxonomic relationships among its species and genera. We provide evidence for an early divergence of three major Callicebinae lineages, and infer a highly supported phylogeny for all species included, with the exception of *P. miltoni* and *P. cinerascens*, which require further investigation. The three Callicebinae genera identified here can be clearly separated on biogeographical, morphological and molecular grounds, and together, these factors provide strong evidence in support of our taxonomic proposal. Recent taxonomic revisions using molecular, ecological and morphological evidence have argued for the separation at the generic level of the robust and the gracile capuchins (Lynch Alfaro et al., 2012) and, likewise, saddleback and black-mantled tamarins from the remaining species of the genus *Saguinus* (Rylands et al., 2016). As with the tamarins and capuchins, this new classification will undoubtedly make for a taxonomy that reflects more clearly titi monkey evolutionary history. It is evident that questions remain regarding the species-level taxonomy of the Callicebinae, and thus phylogenetic hypotheses will be modified with the availability of sequence data for remaining titi species. Taken together, our work illustrates the value of a molecular phylogenetic approach to taxonomic classification and here provides a basis for future studies on the evolutionary history and taxonomy of titi monkeys.

Chapter 3: Biogeography of the titi monkeys (Callicebinae)

3.1 Abstract

Titi monkeys (Callicebinae; Pitheciidae) are a diverse group of platyrrhine primates with an extensive range across South America. There have been few attempts to understand the biogeographic history of Callicebinae and most evaluations have been limited by taxonomic and phylogenetic uncertainty. Here, we reconstructed a time-calibrated molecular phylogeny for Callicebinae under Bayesian inference using two mitochondrial and five nuclear loci. Statistical biogeographic methods implemented in BioGeoBEARS were employed to estimate ancestral areas and to reconstruct Callicebinae biogeographic history using 12 biogeographic models. Our results indicate that the most recent common ancestor to all extant titi monkeys was widespread from the present-day Andean foothills in the Colombian Amazon, through the savannas of Bolivia and Brazil, to the Atlantic Forest of eastern Brazil. Genus-level divergences were characterised by vicariance of ancestral range in the late Miocene resulting in the isolation of *Cheracebus* in north-western Amazon, *Callicebus* in the Atlantic Forest, and *Plecturocebus* in the wet and dry savanna regions. Species-level diversification in both Amazonian clades occurred as they spread across the Amazon in the Pleistocene and were largely characterised by long-distance dispersal from a narrow area of origin (Napo, *Cheracebus*; Rondônia, *P. moloch* group) through sequential founder-events across rivers. These founder-events were sufficiently rare to allow diversification in isolation after dispersal, supporting the role of major Amazonian rivers as strong barriers to gene flow. Overall, our biogeographic reconstruction is most consistent with the "Dynamic Young Amazon" model, suggesting that the diversification of Callicebinae lineages was influenced by the evolution of the Pebas wetland system of western Amazon. This study comprises one of the first large-scale investigations of the evolutionary history of titi monkeys in the context of Amazonian and South American historical biogeography, and sheds light on the processes that generated the great diversity found among Callicebinae taxa.

3.2 Introduction

3.2.1 Background

Titi monkeys (Callicebinae; Pitheciidae) are a diverse group of New World primates found throughout much of South America. They were considered monogeneric (*Callicebus* Thomas, 1903) for much of their taxonomic history, however, the most recent classification (Chapter 2) recognises 33 species across three genera; *Cheracebus* Byrne et al. (2016) for the widow titis; *Callicebus* Thomas, 1903, for the titis of the Atlantic Forest and Caatinga; and *Plecturocebus* Byrne et al. (2016), comprised of the *donacophilus* and *moloch* species groups. Further support for this classification is found in other recent molecular phylogenetic studies (Carneiro et al., 2016; Hoyos et al., 2016), and this classification is followed throughout this chapter.

Titi monkeys have an extensive range spread across nearly all ecogeographic zones inhabitable by non-human primates in the Neotropics (Jameson Kiesling et al., 2015; except Mesoamerica), with each genus or group showing a distinct distributional pattern (Figure 2.5, Chapter 2). The Atlantic forest *Callicebus* are entirely extra-Amazonian and isolated from all other titis by over 500km of drier habitats (the Cerrado shrubby savannas). They are endemic to eastern Brazil from south of the Rio São Francisco as far as the state of São Paulo, predominantly in the Atlantic Forest biome but also in neighbouring arboreal Caatinga regions. The widow titis (genus *Cheracebus*) occur in the Amazon and Orinoco basins from the eastern foothills of the Andes to the Rio Branco and the Rio Purus (north and south of the Rio Amazonas, respectively). The *moloch* group of *Plecturocebus* occur throughout the southern and western Amazon basin, as far east as the Rio Tocantins, as well as some isolated regions in Colombia (*P. ornatus* and *P. caquetensis*). Sympatry among titis occurs between species of *Cheracebus* and *Plecturocebus* in the Amazon, west of the Rio Purus. The *donacophilus* group of *Plecturocebus* occupy forest patches and gallery forests in wooded savannas, the Pantanal, and Chaco scrublands of Bolivia, Brazil, Peru and Paraguay (Ferrari et al., 2013; Hershkovitz, 1990; Printes et al., 2013; Van Roosmalen et al., 2002). In light of their broad and diverse distribution, it is notable that titi monkeys are absent from both Central America and the Guiana Shield (from east of the Rio Branco), and they have a large gap in their distribution in the Cerrado biome of central Brazil (which separates *Callicebus* from all other titis).

A recent review of the biogeography of New World primates (Lynch Alfaro et

al., 2015b) highlighted the family Pitheciidae (which is comprised of the subfamilies Callicebinae and Pitheciinae) as the group most urgently in need of further biogeographic research. As with all New World primate lineages, little information regarding the biogeography of Callicebinae can be derived from the fossil record. The oldest fossil with a definite resemblance to modern titi monkeys (*Miocallicebus villaviega* Takai et al., 2001) comes from La Venta fauna of the middle Miocene (*c*. 12–11 Ma), in the modern day Tatacoa Desert of Colombia (Kay et al., 2013). It is the only fossil to document the callicebine lineage; however, the material is limited (consisting of one maxillary fragment) with little else known about *Miocallicebus*. There have been few attempts to understand the biogeographic history of Callicebinae, and most existing evaluations are confounded by taxonomic uncertainty and a lack of information regarding species relationships (e.g., Hershkovitz, 1963, 1988; Kinzey, 1982; Kinzey & Gentry, 1979).

Titi monkey evolutionary history has only recently been elucidated using molecular evidence (see Chapter 2; Carneiro et al., 2016; Hoyos et al., 2016), providing new insights into the relationships among Callicebinae lineages, and thus an appropriate phylogenetic framework to investigate the biogeographic history of the clade. In placing the northernmost genus *Cheracebus* as the earliest lineage to diverge, interesting biogeographic patterns emerge such as the sister clade relationship of the Atlantic Forest *Callicebus* and *Plecturocebus* of the Amazon and the wet and dry savanna ecosystems. As such, the extant distribution of titi monkeys makes their biogeographic history of particular interest, especially when interpreted in light of phylogenetic relationships and estimated lineage divergence times. The subfamily Callicebinae diverged from their sister clade (Pitheciinae) in the early Miocene, *c*. 18–20 Ma (Chapter 3), thus titi monkeys comprise of one of the oldest lineages of extant New World primates. Based on recent molecular dating analyses, the ancestors of current titi genera appeared in the late Miocene, with initial diversification of *Plecturocebus* occurring in the Pliocene, and extant species diverging mostly in the Pleistocene (Chapter 3; Hoyos et al., 2016). Based on nuclear data, Perelman et al. (2011) inferred slightly older Plio-Pleistocene species divergences, however they are difficult to interpret except for *Callicebus* because of the misidentification of *Plecturocebus* specimens and lack of taxonomic coverage (see Chapter 2). In light of their evolutionary history, a deeper understanding of the spatial diversification of extant titi monkey taxa has the potential to provide insight

into the biogeographic history of South America since the late Miocene.

3.2.2 Callicebinae biogeography

The earliest hypotheses for Callicebinae biogeography were proposed by Hershkovitz (1963, 1988), and focused on the upper Amazon Basin as the centre of origin for each major lineage. From this area, he suggested that ancestral stock from each clade dispersed downstream to lowland areas through newly available habitats following receding floodplains during the Pleistocene, with fluvial dynamics important in driving and maintaining species-level diversification. His biogeographic scenarios are elaborate and contain details about the historical distribution and spatial diversification of each lineage (Hershkovitz, 1988). However, they are based on species relationships inferred from Hershkovitz's metachromism hypothesis (the evolution of pelage colouration via specific unidirectional pathways), which has not been corroborated by molecular phylogenetic hypotheses (e.g., see Jacobs et al., 1995). An evaluation of facial colouration in Neotropical primates found no support for the metachromism hypothesis and much higher support for a model assuming no constraints on colour change (Santana et al., 2012). Owing to Hershkovitz's strict adherence to this hypothesis, it is difficult to extract further details from his biogeographic reconstructions that are interpretable in light of our current understanding of titi monkey evolutionary history. The first biogeographic scenario derived from an explicit phylogenetic hypothesis is that of Hoyos et al. (2016) who suggested that the central Amazon region was the ancestral area of origin for *P. cupreus* and closely related taxa, with subsequent dispersal westwards following the southern bank of the Rio Amazonas towards the foothills of the Andes, and then northwards along the eastern Andes into Colombia. Since Hoyos et al. (2016) focuses solely on select species of the genus *Plecturocebus*, broader biogeographic patterns remain unclear.

A number of biogeographic models have been proposed for the Amazon basin, and many of these models have the potential to explain the diversification of Callicebinae. The most frequently invoked of these is the riverine barrier hypothesis (Wallace, 1852), which identifies river dynamics influenced by tectonic activity as primarily responsible for the isolation and diversification of Amazonian biota. Larger rivers frequently delineate the distributions of Callicebinae species in the Amazon and river dynamics have been proposed as an important speciation force in the clade (e.g.,

Hershkovitz, 1988). Different timescales have been proposed for the formation of the current Amazonian drainage system. The "Old Amazon" geologic model suggests that by *c*. 7 Ma a transcontinental drainage system and major rivers in the Amazon had established, the Pebas lake and floodplain system of the western Amazon had disappeared, and from then on this region bore the key geographic features of the current landscape (Hoorn et al., 2010). In contrary, the "Dynamic Young Amazon" model infers a more recent Plio-Pleistocene origin for the transition from a lacustrine system to the current drainage system in the western Amazon (Campbell et al., 2006; Latrubesse et al., 2010; Rosetti et al., 2005). This model implies that different regions in Amazonia may have undergone distinct rates of landscape change, the most dynamic area being western Amazon where lowland forest was only established following the Pliocene to Pleistocene transition and recession of the Pebas system. As such, diversification of Amazonian biota is associated with tectonically mediated fluvial dynamics as well as the availability of suitable habitat to colonise the western Amazon (Aleixo & Rosetti, 2007). Based on patterns of diversification in the widespread Amazonian avian genus *Psophia,* Ribas et al. (2012) proposed a timescale for the drainage of the wetlands and formation of major rivers in the Amazon within the last three million years. Similar patterns found in other avian (e.g., Fernandes et al., 2012; d'Horta et al., 2013) and primate taxa (Boubli et al., 2015; Buckner et al., 2015; Lynch Alfaro et al., 2015a) have also been hypothesized to be related to the establishment of the current drainage system in the Plio-Pleistocene.

Jameson Kiesling et al. (2015) reconstructed a biogeographic scenario for New World primates using a statistical biogeographical analysis based on a genomic dataset and identified the Amazon as the area of origin for the most recent common ancestor of extant New World primates, each of the families (Pitheciidae, Atelidae, Cebidae), and most genera. Rather than diversification in geographic isolation, these authors argue that the Amazon rainforest was the key generator of diversity, providing such a rich environment for niche exploitation that diversification of the major lineages could have occurred largely in sympatry (Jameson Kiesling et al., 2015). Following the divergence of genera in the Amazon, members of some lineages colonised other subregions occupied by New World primates (Atlantic Forest, Central Grasslands, Caatinga, Cerrado). Although Callicebinae was represented as a monogeneric clade in this study, under the current classification this implies that the progenitors of the Atlantic Forest genus *Callicebus* and the *Plecturocebus*

donacophilus group originated in the Amazon and subsequently dispersed to their respective biomes.

All New World primate genera found in the Atlantic Forest have closely related sister taxa in the Amazon (Lynch Alfaro et al., 2015b) and it has been suggested based on the timing of divergences that there were intermittent periods of increased connectivity between the two regions during the middle-late Miocene. Similar patterns among avian sister taxa in the Amazon and Atlantic Forest have been associated with a middle-late Miocene corridor along the southern Cerrado in Mato Grosso or along the transition towards the Chaco of Bolivia and Paraguay, while younger Plio-Pleistocene divergences correspond to connections between the north-eastern Atlantic Forest and the eastern extreme of the Amazon through the Caatinga (Batalha-Filho et al., 2013). Batalha-Filho et al. (2013) suggested that the older southern connections were driven by geological events associated with the uplift of the Andes and that the more recent northern connections were influenced by climatic changes that promoted the intermittent expansion of gallery forest through the Cerrado and Caatinga. Hershkovitz (1988) proposed a similar scenario to that of Batalha-Filho et al. (2013) for the spread of titi monkeys to the Atlantic Forest. He suggested that riparian forest along Cerrado river systems facilitated the dispersal of the progenitor of the genus *Callicebus* to the Atlantic Forest, and that these forests largely disappeared during the climatic changes in the Pleistocene. The late Miocene origin for this clade, however, would suggest that titis arrived in the Atlantic Forest biome via the southern connections, which is further supported by the extant distributions of members of their sister clade (*Plecturocebus*) through the wet and dry savannas along this corridor.

3.2.3 Biogeographic hypotheses for Callicebinae diversification

In this study, through a statistical biogeographical approach, we reconstructed the biogeographic history of Callicebinae to better understand their spatial diversification and the processes responsible for shaping present day distributions. Our primary aim was to infer the origin of Callicebinae and the origin of the major titi clades, and the history of their colonisation of South America since the late Miocene. Geographical areas were based upon vertebrate centres of endemism in the Amazon as well as major biomes (Figure 3.1); Pantepui, Napo, Inambari, Rondônia, Pará, Wet & Dry Savannas, Atlantic Forest (see 3.3.3 Methods: Biogeographical analyses). Statistical

methods based on maximum likelihood were applied to a time-calibrated molecular phylogeny, allowing the comparison of models to determine the relative importance of vicariance and dispersal in forming the current distributions of Callicebinae taxa. This approach has been used to better understand the biogeographic history of other Neotropical primates (e.g. tamarins and marmosets, Buckner et al., 2015; capuchins, Lima et al., 2017; squirrel monkeys, Lynch Alfaro et al., 2015a). The current hierarchy of the subfamily Callicebinae corresponds to geological periods, such that genus- and species-level divergences shed light on diversification dynamics in the late Miocene or the Plio-Pleistocene, respectively. In light of this, we discuss the predictions regarding patterns of genus- and species-level diversification separately below.
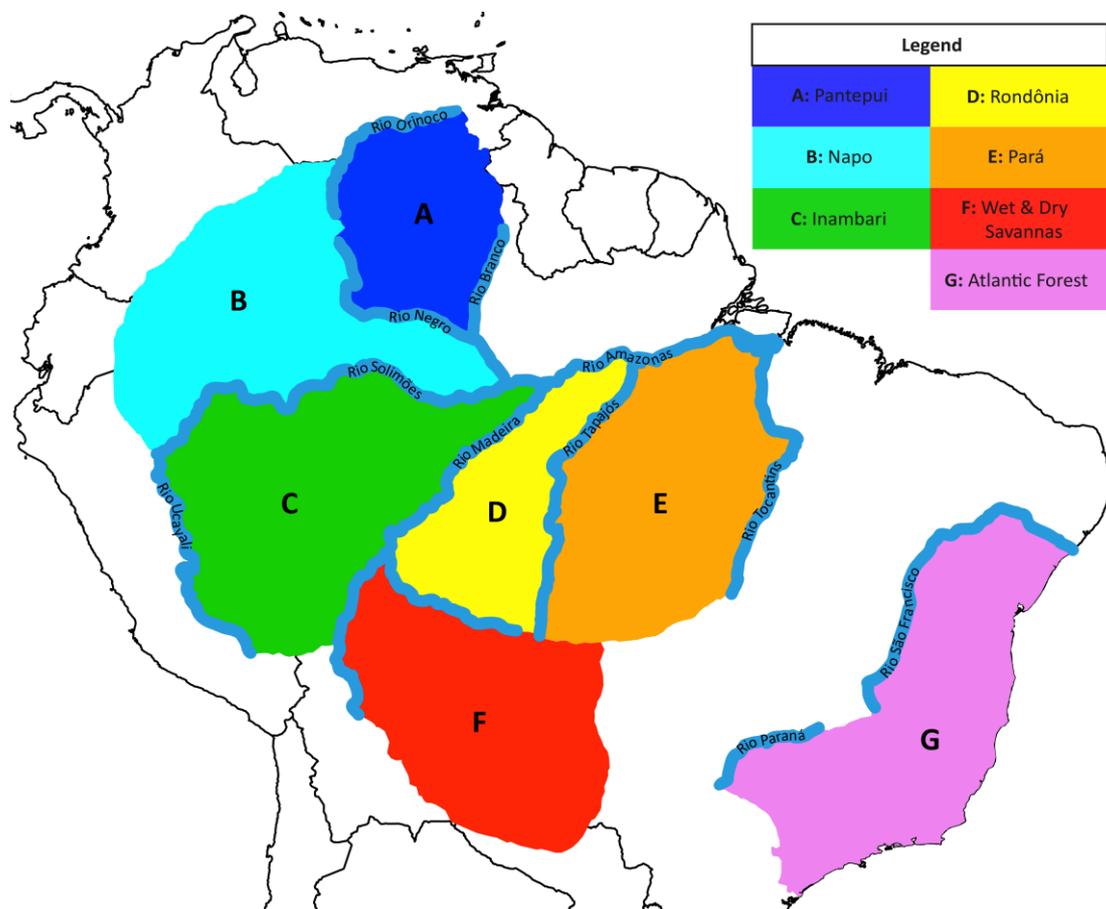
Figure 3.1 Map showing the seven biogeographic regions/centres of endemism used in this study: (A) Pantepui (& Imeri), (B) Napo, (C) Inambari, (D) Rondônia, (E) Pará (Tapajós & Xingu), (F) Wet & Dry Savannas, (G) Atlantic Forest. Major Amazonian rivers are also shown.

3.2.4 Genus-level diversification

We tested whether the ancestral populations to all modern titi monkeys were widespread across South America (scenario 1a) or had a more regionally restricted ancestral range (1b), with the aim to shed light on the patterns and processes of genus-level diversification and the present day sympatry between *Cheracebus* and *Plecturocebus* in the Amazon. We also tested whether titis colonised the Atlantic Forest biome through the "southern" Miocene pathway via the Chaco, or the more recent "northern" pathway to the eastern Amazon via the Caatinga in the Plio-Pleistocene (Batalha-Filho et al., 2013) (2a–b).

If Callicebinae genus-level divergences represent the fragmentation of a widespread ancestral range owing to a series of ecological or geological vicariant events, we expect our biogeographic model and reconstruction to emphasise vicariance and the ancestral range should span across multiple regions, from the Napo to the Atlantic Forest (1a). Vicariance may have been driven by geological events associated with the accelerated uplift of the Andes since the late Miocene. Under the "Old Amazon" geologic model, a transcontinental fluvial pathway had developed by *c*. 10 Ma and the Rio Amazonas was fully established by 7 Ma (Hoorn et al., 2010). If the formation of the drainage system promoted diversification, this model suggests that the development of this transcontinental waterway may have led to vicariance of ancestral range within the Amazon, isolating the *Cheracebus* ancestor to the north of the Rio Amazonas (Napo + Pantepui) and the *Callicebus + Plecturocebus* ancestor to the south (Inambari + other southern areas) (1a.i) (Figure 3.2). Under this model, present day sympatry between *Plecturocebus* and *Cheracebus* is explained by dispersal over the Rio Amazonas, northwards and southwards, respectively.

Alternatively the "Dynamic Young Amazon" scenario (see above) suggests that Miocene divergences in Amazonian biota were driven by the evolution of the extensive Pebas wetlands system of the western Amazon (Aleixo & Rosetti, 2007). This model proposes that the western Amazon was largely inhospitable to upland forest taxa throughout the late Miocene, resulting in vicariance of ancestral range across more geologically stable terra firme regions: the Brazilian Shield, the eastern foothills of the Andes, and the Guiana Shield/northern and north-eastern periphery of the Pebas system. Thus, we expect initial vicariance to result in isolation of the *Cheracebus* ancestor in the Andean foothills (within the Napo region) and the *Callicebus + Plecturocebus* ancestor in the Brazilian Shield and/or surrounding

regions (i.e., Rondônia, Pará, Wet & Dry Savannas, Atlantic Forest), and the absence of ancestral range in Inambari until the Pleistocene (1a.ii) (Figure 3.2). In this scenario, present day sympatry is explained by the spread of *Cheracebus* and *Plecturocebus* from opposing sides of the Amazon following the recession of the wetlands in the Pleistocene. In both scenarios (1ai + 1a.ii), we predict that vicariance between the Atlantic Forest and the Amazon/Wet & Dry Savanna regions resulted in the divergence of *Callicebus* and *Plecturocebus*.
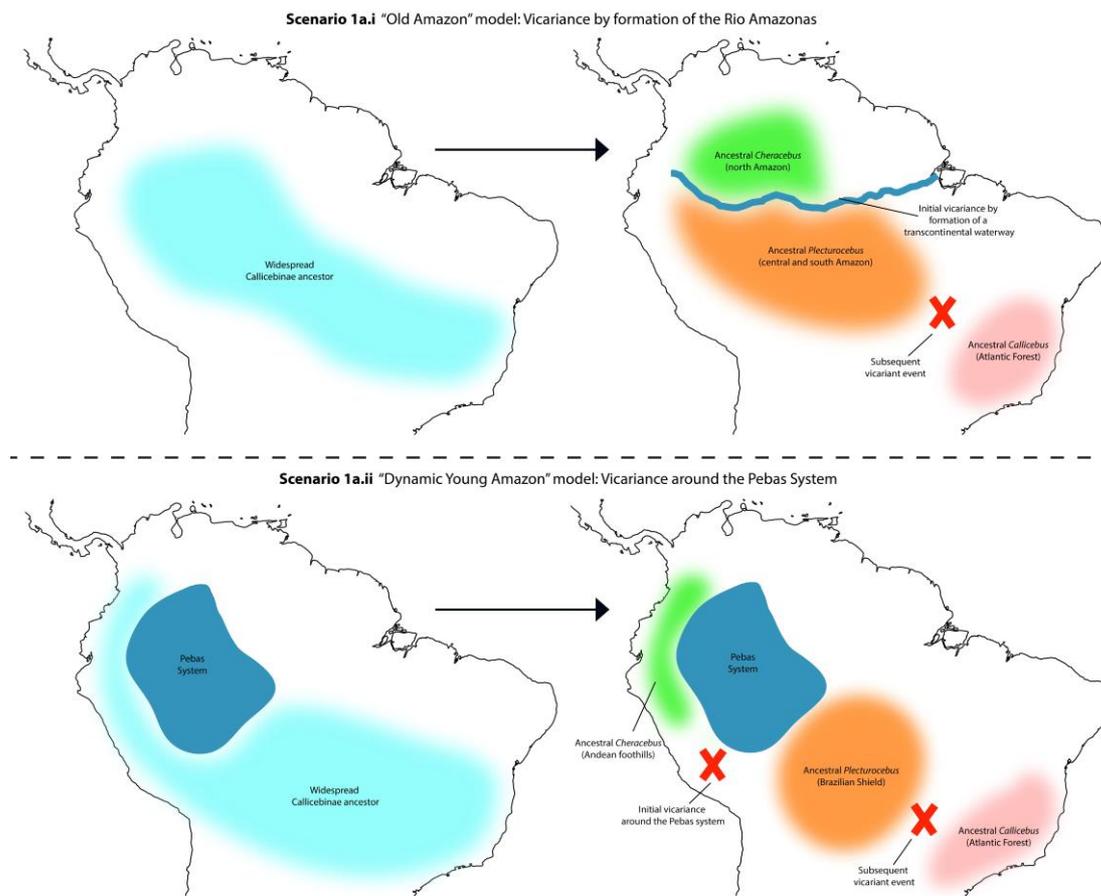
Figure 3.2 Graphical summary of alternative vicariance scenarios. Shown are the predicted patterns of genus-level diversification through vicariance of a widespread Callicebinae ancestor in the late Miocene under the "Old Amazon" model (1a.i) and "Dynamic Young Amazon" model (1a.ii).

In the case of a narrow ancestral range, dispersal should be emphasised and the ancestral areas should be restricted to one or two regions (1b). If most Neotropical primate genera originated in the Amazon, as suggested by Jameson Kiesling et al. (2015), we expect the ancestral range for each genus to be restricted to Amazonian subregions, with subsequent dispersal of *Callicebus* to the Atlantic Forest and the

69

donacophilus group of *Plecturocebus* to the Wet & Dry Savannas region (1b.i)
(Figure 3.3).
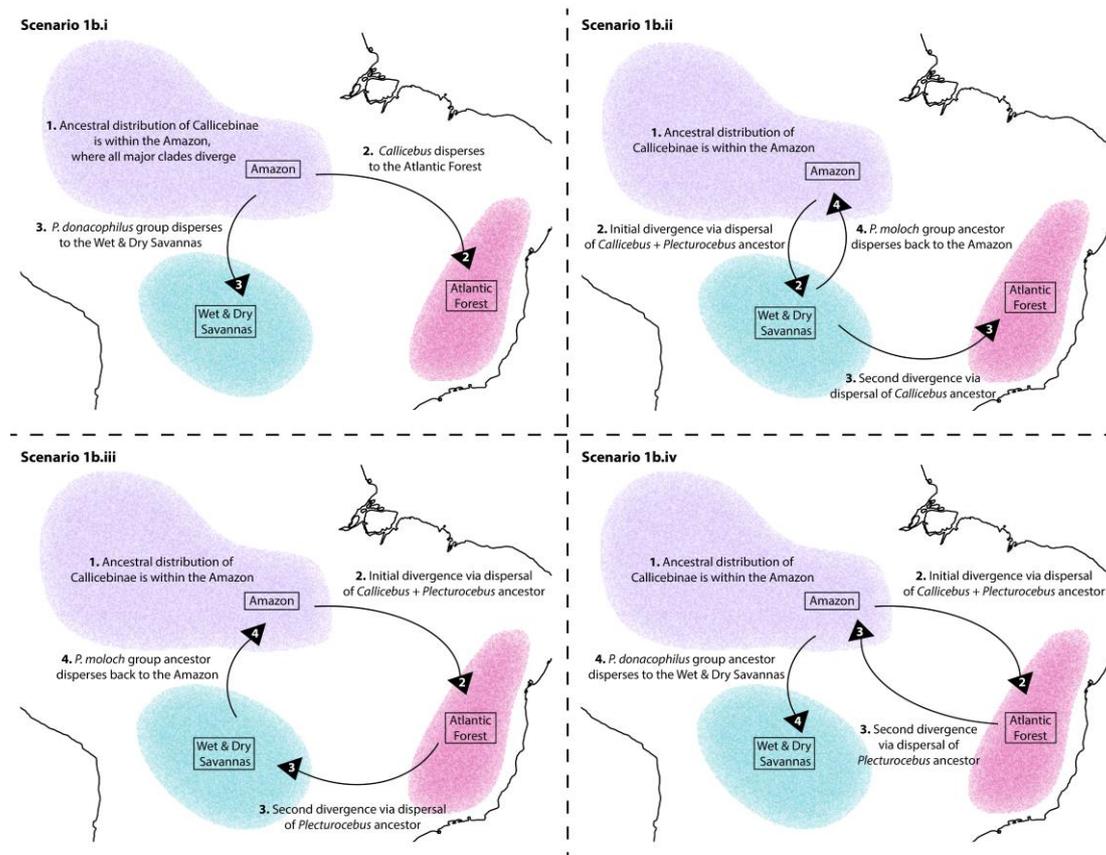
Figure 3.3 Graphical summary of alternative dispersal scenarios. Shown are possible patterns of dispersal between major biomes characterising the diversification of Callicebinae lineages, beginning with a geographically restricted ancestor in the late Miocene. In scenario 1b.i, all major divergences occur within the Amazon, with subsequent dispersal of *Callicebus* and the *P. donacophilus* group to their respective biomes prior to the Pleistocene. In the pure dispersal scenarios (1b.ii – iv), major divergences among Callicebinae lineages are driven by dispersal between biomes. Arrows represent dispersal events and shaded regions depict major biomes: Amazon (purple), Wet & Dry Savannas (green), Atlantic Forest (pink).

There are two alternative scenarios based on pure dispersal between major biomes that infer the origin of *Callicebus* and *Plecturocebus* outside the Amazon. The first plausible scenario involves dispersal of *Callicebus* + *Plecturocebus* ancestor from the Amazon to the Wet & Dry Savannas region, followed by dispersal of *Callicebus* to the Atlantic Forest, and subsequently, reinvasion of the Amazon by *Plecturocebus* from the Wet & Dry Savannas (1b.ii) (Figure 3.3). Alternatively, the *Callicebus* + *Plecturocebus* ancestor dispersed first from the Amazon to the Atlantic Forest, followed by dispersal of *Plecturocebus* either to the Wet & Dry Savannas

region and back to the Amazon, or vice versa (1b.iii + iv) (Figure 3.3). In these latter dispersal scenarios, ancestral *Plecturocebus* originated outside the Amazon and subsequently reinvaded, thus explaining the present day sympatry with *Cheracebus*.

3.2.5 Species-level diversification

Titi monkeys diversified relatively rapidly in the Pleistocene, and as such, species-level divergences in the Amazonian clades (*Cheracebus* and the *Plecturocebus moloch* group) can shed light on the processes and mechanisms of speciation in Amazonia in the last two million years. Below we provide alternative hypotheses for the biogeographic history of these Callicebinae lineages (3a–b).

The "Old Amazon" geologic model suggests that major rivers in the Amazon were established by 7 Ma (Hoorn et al., 2010), before the divergence of extant Callicebinae species, and as such, these lineages are too young to be driven by vicariance owing to river formation. However, assuming that rivers are barriers to the dispersal of titis, this model predicts that current species distributions are best explained by a pattern of rare dispersal events across rivers (pure dispersal). In this scenario, we expect the biogeographic model and reconstruction to emphasize founder-event speciation, and dispersal should have occurred between adjacent regions delineated by major rivers (3a).

The "Dynamic Young Amazon" scenario proposes that species-level diversification in some currently widespread Amazonian taxa was associated with the dynamic geologic history of the western Amazon and the relatively recent establishment of terra firme forest in this region (Aleixo & Rosetti, 2007). This model implies that *Cheracebus* and the *P. moloch* group were isolated in the more stable Andean foothills and Brazilian Shield, respectively, until dispersal was facilitated by the availability of suitable 'upland' forest habitat in the western Amazon. This landscape change was driven by tectonically mediated fluvial dynamics and the formation of key features of the current drainage system in the Plio-Pleistocene (Aleixo & Rosetti, 2007; Campbell et al., 2006; Rosetti et al., 2005), however, the exact timing of these events remains uncertain. If major rivers were in place early in the Plio-Pleistocene prior to species divergences, then we expect diversification to be explained by rare dispersal events across rivers and the biogeographic model and reconstruction should emphasize founder-event speciation, similar to the "Old Amazon" model. On the other hand, there are specific constraints on the pattern of

dispersal which are viewed as support (in addition to consideration of genus-level divergences) for the "Dynamic Young Amazon" model; we expect diversification to be associated with recent dispersal into the western Amazon (Inambari) by both Amazonian clades, which were previously isolated in the Brazilian Shield (Rondônia/Pará; *P. moloch* group) or Andean foothills (within Napo; *Cheracebus*) (3b.i).

Alternatively, according to the temporal scale proposed by Ribas et al. (2012), major Amazonian tributaries were formed within the last two million years, concurrent with the diversification of extant Callicebinae species. In this scenario, ancestral populations expanded between 2–3 Ma following drainage of the wetlands and establishment of terra firme forest in the western Amazon, with subsequent vicariance owing to river formation. Thus, we expect ancestral range to be reconstructed as widespread, through the western Amazon, by *c.* 2 Ma and major divisions by vicariance following the order of river formation (from Ribas et al., 2012); first across the Rio Amazonas/Solimões (2–3 Ma), then the Rio Madeira (1–2 Ma), Rio Tapajós (1.3–0.8 Ma), and Rio Negro (*c.* 1 Ma) (3b.ii).

3.3 Methods

3.3.1 Molecular dataset

A molecular sequence dataset was assembled comprising five nuclear (FES, MAPKAP1, RAG1, RAG2, ZFX) and two mitochondrial loci (CYTB, COI). These loci were chosen to maximise taxonomic coverage while minimising missing data across the dataset, as only mitochondrial sequences were available for some Callicebinae species. Twenty-one individuals from 19 Callicebinae species were represented in the dataset, including the two distinct *Cheracebus lugens* lineages from opposing banks of the Rio Negro, Brazil (Boubli et al., 2015), and also one taxon in the process of description (*P.* cf. *moloch*, Boubli et al., in prep). A total of 223 sequences were included; 115 sequences belonging to Callicebinae taxa, most of which were generated for Chapter 2; and 108 sequences retrieved from GenBank belonging to 16 species of Platyrrhini (5) and Catarrhini (11) selected as outgroups in order to include nodes with reliable fossil calibrations. GenBank accession numbers for the retrieved sequences are provided in Table 3.1.

Table 3.1 List of the GenBank accession numbers.

Species	COI	CYTB	FES	MAPKAP1	RAG1	RAG2	ZFX
Plecturocebus ornatus	---	KX353784	---	---	---	---	---
Plecturocebus caquetensis	---	KX353781	---	---	---	---	---
Plecturocebus discolor	---	KX353786	---	---	---	---	---
*Plecturocebus donacophilus**	FJ785423	FJ785423	HM761804	HM760634	HM759135	HM758967	HM757151
*Callicebus nigrifrons**	---	---	---	---	HM759137	HM758969	---
*Callicebus personatus**	---	---	HM761807	HM760643	HM759138	HM758970	HM757153
Callicebus coimbrai	---	---	HM761802	HM760631	HM759134	---	HM757149
*Pithecia pithecia**	JF459229	KR902424	HM761808	HM760722	HM759140	HM758971	HM757155
*Cacajao calvus**	NC021967	NC021967	---	---	HM759113	HM758942	HM757126
*Chiropotes israelita**	NC024629	NC024629	HM761786	HM760640	HM759120	HM758950	HM757133
*Cebus albifrons**	AJ309866	KU694249	KU694628	KU694723	HM759115	KU694935	KU695108
*Saimiri sciureus**	HQ644334	HQ644334	HM761799	HM760744	HM759131	HM758963	HM757147
*Trachypithecus obscurus**	AY863425	AY863425	HM761732	HM760754	HM759066	HM758893	HM757077
*Colobus guereza**	AY863427	AY863427	HM761695	HM760637	HM759029	HM758852	HM757038
*Chlorocebus aethiops**	NC007009	NC007009	HM761691	HM760620	HM759026	HM758848	HM757034
*Macaca fascicularis**	NC012670	NC012670	HM761702	HM760689	---	HM758859	HM757045
*Papio anubis**	KC757406	KC757406	HM761717	HM760709	HM759049	HM758875	HM757061
*Theropithecus gelada**	FJ785426	FJ785426	HM761728	---	HM759062	HM758888	HM757077
*Hylobates lar**	HQ622766	HQ622766	HM761737	HM760669	HM759071	HM758898	HM757082
*Pongo pygmaeus**	NC001646	NC001646	HM761749	HM760724	HM759081	HM758910	HM757094
*Gorilla gorilla**	KF914214	KF914214	HM761744	HM760662	HM759077	HM758905	HM757089
*Pan troglodytes**	EU095335	EU095335	HM761747	HM760726	HM759080	HM758908	HM757092
*Homo sapiens**	EF061150	EF061150	HM761735	HM760672	HM759069	HM758896	HM757080

*composite individuals

Six new sequences were obtained from museum voucher specimens to include Callicebinae species (*Cheracebus lucifer, Plecturocebus vieirai, Callicebus personatus*) for which no or little molecular sequence data was available. Laboratory work was carried out at the University of Salford, Manchester, UK. DNA was extracted from blood and muscle tissues using the Qiagen DNeasy Blood & Tissue Kit according to manufacturer's protocol. Six new sequences were generated for COI (3), CYTB (2) and RAG1 (1) (see primer information in Chapter 2, Table 2.3). The PCR reactions were carried out in a total volume of 50 μL, containing approximately

30 ng of genomic DNA, 4 µL of dNTPs (200µM each), 5 µL 10X PCR buffer (100 mM Tris-HCL, 500 mM KCL, 15 mM Mg2+), 1 µL of each forward and reverse primer (0.2 µM), and 0.25 µL of TaKaRa *Taq* DNA polymerase (1 Unit). The amplification cycles were carried out under the following conditions; initial denaturation at 95 °C for 5 min; followed by 35 cycles of denaturing at 94 °C for 1 min, primer annealing for 1 min, and extension at 72 °C for 1 min; a final extension was carried out at 72 °C for 5 min. PCR products were analysed on 1.5 % agarose gels and then Sanger sequenced commercially by Source Bioscience (Cambridge, UK). Consensus sequences were generated from forward and reverse reads using Geneious R7.1 (Biomatters).

Table 3.2 Locus coverage and ID for Callicebinae taxa.

Species	Sample ID	COI	CYTB	FES	MAPKAP1	RAG1	RAG2	ZFX
Plecturocebus hoffmannsi	CTGAM290	■	■	■	■	■	■	■
Plecturocebus cinerascens	UFRO355	■	■	■	■	■	■	■
Plecturocebus miltoni	42991	■	■	■	■	■	■	■
Plecturocebus bernhardi	42964	■	■	■	■	■	■	■
Plecturocebus cf. moloch	RVR73	■	■	■	■	■	■	■
Plecturocebus vieirai	2594	■	■					
Plecturocebus moloch	MCB64	■	■	■	■	■	■	■
Plecturocebus brunneus	4505	■	■	■	■	■	■	■
Plecturocebus caligatus	CTGAM182	■	■	■	■	■	■	■
Plecturocebus ornatus	ZP01		■					
Plecturocebus caquetensis	ICN19439		■					
Plecturocebus discolor	ZP03		■					
Plecturocebus cupreus	CTGAM210	■	■	■	■	■		■
Plecturocebus donacophilus*	NA	■	■	■	■	■	■	■
Callicebus nigrifrons*	NA	■	■	■	■	■	■	
Callicebus personatus*	NA	■	■	■	■	■	■	■
Callicebus coimbrai	CCO1	■		■	■	■	■	■
Cheracebus lugens (RN)	JPB81	■	■	■	■	■	■	■
Cheracebus lugens (LN)	JPB119	■	■	■	■	■	■	■
Cheracebus lucifer	CTGAM703	■	■			■		■
Cheracebus purinus	CTGAM154	■	■	■	■	■	■	■

*composite individuals

Each locus was aligned independently using the MUSCLE algorithm in Geneious R7.1 (Biomatters) and subsequently combined in a matrix resulting in a total alignment length of 5,233 bp. PartitionFinder (Lanfear et al., 2012) was used to determine the optimal partitioning scheme and the best-fit substitution models for each partition under the Bayesian information criterion (BIC). The analysis was run using the complete search algorithm ("all") and linked branch lengths (see Table 3.3 for selected partitioning scheme).

Table 3.3 Partitioning scheme and substitution models selected by PartitionFinder.

Partition	Model	Loci
1	K80+G	FES
2	K80+I	MAPKAP1, RAG1
3	HKY+G	RAG2, ZFX
4	GTR+I+G	COI, CYTB

3.3.2 Phylogenetic analysis and molecular dating

Phylogeny and diversification times were jointly estimated under an uncorrelated lognormal relaxed clock in the program BEAST v. 1.8.2 (Drummond et al., 2012). The partitioning scheme and best-fit models chosen by PartitionFinder were implemented and a Yule speciation process was used. We conducted two replicate runs of 50 million generations, sampling every 5,000 generations. The sampling distributions were visualized using Tracer v. 1.6 to evaluate convergence, performance, and burn-in. We combined the runs using LogCombiner v. 1.8.2 and generated the maximum credibility tree using a burn-in of 10% in TreeAnnotator v. 1.8.2. To obtain the posterior distribution of the estimated divergence times, we used six calibration points (Table 3.4) with lognormal priors to set hard minimum and soft maximum bounds; (i) Callicebinae/Pitheciinae (95%: 15.7–26.0 Ma); (ii) *Cebus/Saimiri* (95%: 12.5–26.0 Ma); (iii) Hominoid/Cercopithecoid (95%: 21.0–30.0 Ma); (iv) *Homo/Pongo* (95%: 12.5–18.0 Ma); (v) *Homo/Pan* (95%: 5.0–10.0 Ma); and (vi) *Theropithecus/Papio* (95%: 3.5–6.5 Ma). Standard deviation was set at 0.5 for all nodes.

Table 3.4 Evolutionary rate calibration constraints (in millions of years).

Divergence	Offset	95% Prior Distribution	Mean	Fossil	Reference	Age
Callicebinae-Pitheciinae	15.7	26.0[a]	5.13	*Proteropithecia*	Kay et al. (1998)	≈15.7
Cebus-Saimiri	12.5	26.0[a]	6.72	*Neosaimiri*	Hartwig & Meldrum (2002)	≈12.1
Homo-Pan	5	10	2.49	*Ardipithecus*	Haile-Selassie (2001)	5.2
				Orrorin	Senut et al. (2001)	6
				Sahelanthropus	Brunet et al. (2002); Vignaud et al. (2002)	6.0 – 7.0
Homo-Pongo	12.5	18	2.74	*Sivapithecus*	Kelley (2002)	≈12.5
Hominoid-Cercopithecoid	21	30	4.48	*Morotopithecus*	Young & MacLatchy (2004)	>20.6
				Victoriapithecus	Pilbeam & Walker (1968);	≈19.0
					Benefit & McCrossin (2002)	
Papio-Theropithecus	3.5	6.5	1.49	*Theropithecus*	Leakey (1993)	≈3.5

[a] Based on the fossil *Branisella boliviana* Hoffsetter, 1969.

3.3.3 Biogeographic analyses

We divided the distribution of titi monkeys into seven geographic regions (Pantepui, Napo, Inambari, Rondônia, Pará, Wet & Dry Savannas, Atlantic Forest) and coded each taxon for presence/absence in each of these regions (see Table 3.5). The geographic regions were broadly defined by major biomes or centres of endemism in the Amazon, primarily following Cracraft (1985) and Silva et al. (2002) (see Figure 3.1, Table 3.6).

The Pará region (Cracraft, 1985) represents both the Tapajós and Xingu areas of endemism (Silva et al., 2002), as delineating these areas separately is not of central importance to the present study. Our Pantepui region covers the highlands of the Guiana Shield (from west of the Branco-Essequibo divide/Rupununi graben) in the Duida and parts of the Gran Sabana subcentres of the Pantepui area of endemism and the lowlands of the Imeri area of endemism (Cracraft, 1985; Silva et al., 2002). It is delineated by the Negro, Orinoco and Branco rivers and contains the geographic distribution of the left bank Rio Negro *C. lugens* lineage. Geographic regions were assigned for the Atlantic Forest biome and for the wet and dry savanna ecosystems. The Wet & Dry Savannas region is comprised largely of the seasonally flooded Pantanal and Llanos de Moxos (Beni) savannas, and the dry wooded Chaco and Chiquitano savannas, as well as the southernmost tip of the Amazon basin. It broadly

follows the Central Grasslands area of Jameson Kiesling et al. (2015) but is refined to the known limits of extant titi monkey distributions. These open habitats are of importance to the biogeography of Callicebinae, which are one of the most diverse groups of New World primates found across these regions. Four species of the *P. donacophilus* group are known to occur here (Martínez & Wallace, 2007, 2013; Rumiz, 2012), only *P. donacophilus*, however, was represented in this dataset owing to the lack of available sequence data for the other taxa.

Table 3.5 Biogeographic regions used to reconstruct the biogeographic history of Callicebinae. Single letter codes correspond to those in Figure 3.1, 3.5, A2.2 –A2.5.

Area code	Corresponding area of endemism/biome	Taxa
A	Pantepui (& Imeri)	*C. lugens* (left bank Rio Negro lineage)
B	Napo	*C. lugens* (right bank Rio Negro lineage), *C. lucifer, P. ornatus, P. caquetensis, P. discolor*
C	Inambari	*C. purinus, P. cupreus, P. caligatus*
D	Rondônia	*P. hoffmannsi, P. cinerascens, P. miltoni, P. bernhardi, P. brunneus*
E	Pará (Tapajós & Xingu)	*P. moloch, P.* cf. *moloch, P. vieirai*
F	Wet & Dry Savannas	*P. donacophilus*
G	Atlantic Forest	*C. personatus, C. coimbrai, C. nigrifrons*

Our main biogeographic analyses were based on time-calibrated trees from the BEAST analysis with the outgroup taxa removed. Two approaches were employed to perform ancestral-area estimations across the phylogeny of Callicebinae. We first used the R package BioGeoBEARS (Matzke, 2013a) to reconstruct the biogeographic history of Callicebinae under alternative models implemented in a likelihood framework: DIVALIKE, DIVALIKE+J, DEC, DEC+J, BAYAREALIKE, BAYAREALIKE+J. Each model allows for a different subset of biogeographic processes, such as dispersal, vicariance and extinction (see Figure 1 in Matke 2013b), which are implemented as free parameters that are estimated from the data. The "J" parameter corresponds to founder-event speciation ("jump dispersal", see Matzke, 2014).

Table 3.6 Description of the geographic limits of the biogeographic regions.

Region code	Corresponding area of endemism/biome	Boundaries
A	Pantepui (& Imeri)	Bordered on the east by the Rio Branco, north-west by the Rio Orinoco, and south-west by the Rio Negro.
B	Napo	Bordered on the west by the Andes, north-east by the Rios Meta, Orinoco and Negro, and south-east by the Rio Solimões.
C	Inambari	Bordered on the west by the Rio Ucayali, north by the Rio Solimões-Amazonas, and south-east by the Rio Madeira-Madre De Dios.
D	Rondônia	Bordered on the west by the Rio Madeira, north by the Rio Amazonas, east by the Rio Tapajós-Juruena, and south by the Rio Itenez O Guapore.
E	Pará (Tapajós & Xingu)	Bordered on the west by the Rio Tapajós-Juruena, north by the Rio Amazonas, east by the Rio Tocantins, and south by the Amazon watershed and Cerrado.
F	Wet & Dry Savannas	Bordered on the west by the Andes, north-west by the Rio Madre De Dios, north-east by the Rio Itenez O Guapore and the Amazon watershed, east by the Cerrado, and south by the known limits to extant Callicebinae distributions. It contains the Beni Savanna, the Chiquitano Savanna, the dry Chaco and the Pantanal ecoregions.
G	Atlantic Forest	Bordered on the east by the Atlantic coastline, south by the Rio Tietê, and north and west by the Rios Paraná and São Francisco and the Cerrado.

To account for the influence of distance on dispersal, all BioGeoBEARS analyses were also conducted under a distance-based dispersal model where dispersal probability is multiplied by distance to the power of x ("X" parameter; Van Dam & Matzke, 2016): DIVALIKE+X, DIVALIKE+J+X, DEC+X, DEC+J+X, BAYAREALIKE+X, BAYAREALIKE+J+X. Distances were calculated as the physical distance in kilometres between the centre of the geographic areas and then scaled to the smallest distance (Table 3.7).

We set the maximum number of ancestral areas at a given node to four for all biogeographic analyses given that extant Callicebinae species have relatively restricted distributions and to avoid intractability of the analyses. We compared the twelve different BioGeoBEARS models for statistical fit using the corrected Akaike Information Criterion (AICc). We performed Biogeographic Stochastic Mapping (BSM) simulations in BioGeoBEARS to estimate the overall probability of the different biogeographic processes under the specified phylogeny, parameters and models (Matzke, 2016). The BSM simulations were conducted across 1000 stochastic maps (simulated histories) under the best-fit BioGeoBEARS models based on AICc scores, and checked for convergence. We also ran all 12 BioGeoBEARS models on time-calibrated trees that included two Pitheciinae outgroup taxa (*Cacajao calvus*,

Inambari + Napo; *Chiropotes israelita*, Pantepui) to test the impact on the ancestral areas estimated for the origin of Callicebinae.

The second approach to ancestral-area estimations was using the Bayesian framework implemented in the software RASP 3.2 (Yu et al., 2015), which reconstructs ancestral states over a posterior distribution of trees. The biogeographic history of titi monkeys was reconstructed under the Bayesian DIVA (S-DIVA), Bayesian DEC (S-DEC), and Bayesian Binary MCMC (BBM) models. All analyses were run on a random subset of 1,000 trees from the BEAST analysis. The BBM chains were run for 5 million generations (sampling every 1000 generations, a temperature of 0.1, and 10% burn-in), and state frequencies were estimated under the F81+G model.

Table 3.7. Distance-dependent dispersal matrix for the "+X" models in BioGeoBEARS.

Areas		Pantepui	Napo	Inambari	Rondônia	Pará	Wet & Dry Savannas	Atlantic Forest
	Area Code	A	B	C	D	E	F	G
Pantepui	A	0	1	1.81	1.86	2.31	2.99	4.4
Napo	B	1	0	1.15	1.82	2.66	2.7	4.7
Inambari	C	1.81	1.15	0	1.16	2.22	1.64	4.01
Rondônia	D	1.86	1.82	1.16	0	1.07	1.18	2.91
Pará	E	2.31	2.66	2.22	1.07	0	1.73	2.09
Wet & Dry Savannas	F	2.99	2.7	1.64	1.18	1.73	0	2.69
Atlantic Forest	G	4.4	4.7	4.01	2.91	2.09	2.69	0

3.4 Results

3.4.1 Phylogeny

Our phylogenetic results showed monophyly of the three Callicebinae genera (Figure 3.4, see Figure A2.1 for the full timetree with outgroup taxa) and are broadly concordant with previous molecular genetic analyses (Chapter 2; Carneiro et al., 2016; Hoyos et al., 2016). Phylogenetic relationships within the subfamily Callicebinae were largely resolved with most nodes showing strong support (posterior probability/PP = 1.00), and only posterior probability values of less than 1.00 are discussed below. A summary of PP and age estimates with 95% HPD intervals are reported in Table 3.8. The genus *Cheracebus* was strongly supported as sister taxon to remaining titis, diverging in the late Miocene *c.* 10.25 Ma, followed by the divergence of *Callicebus* and *Plecturocebus c.* 8.27 Ma, and then the division of *Plecturocebus*

into *P. donacophilus* (representative of the *P. donacophilus* group) and the *P. moloch* group in the Pliocene *c*. 4 Ma.

Among *Cheracebus* taxa, the earliest divergence occurred between *C. lugens* and *C. lucifer* + *C. purinus* in the Pleistocene *c*. 1.87 Ma, while the sister-species *C. lucifer* and *C. purinus* diverged *c*. 1.26 Ma, and the *C. lugens* lineages diverged *c*. 1 Ma. *Callicebus nigrifrons* was supported as the sister taxon to a clade containing *C. personatus* and *C. coimbrai*, diverging *c*. 2.85 Ma (the oldest estimated speciation age). The divergence of *C. personatus* and *C. coimbrai* was recovered with low support (PP = 0.61), likely as a result of missing data (no mitochondrial loci for *C. coimbrai*). Our results, however, are concordant with Chapter 2 and Perelman et al. (2011).
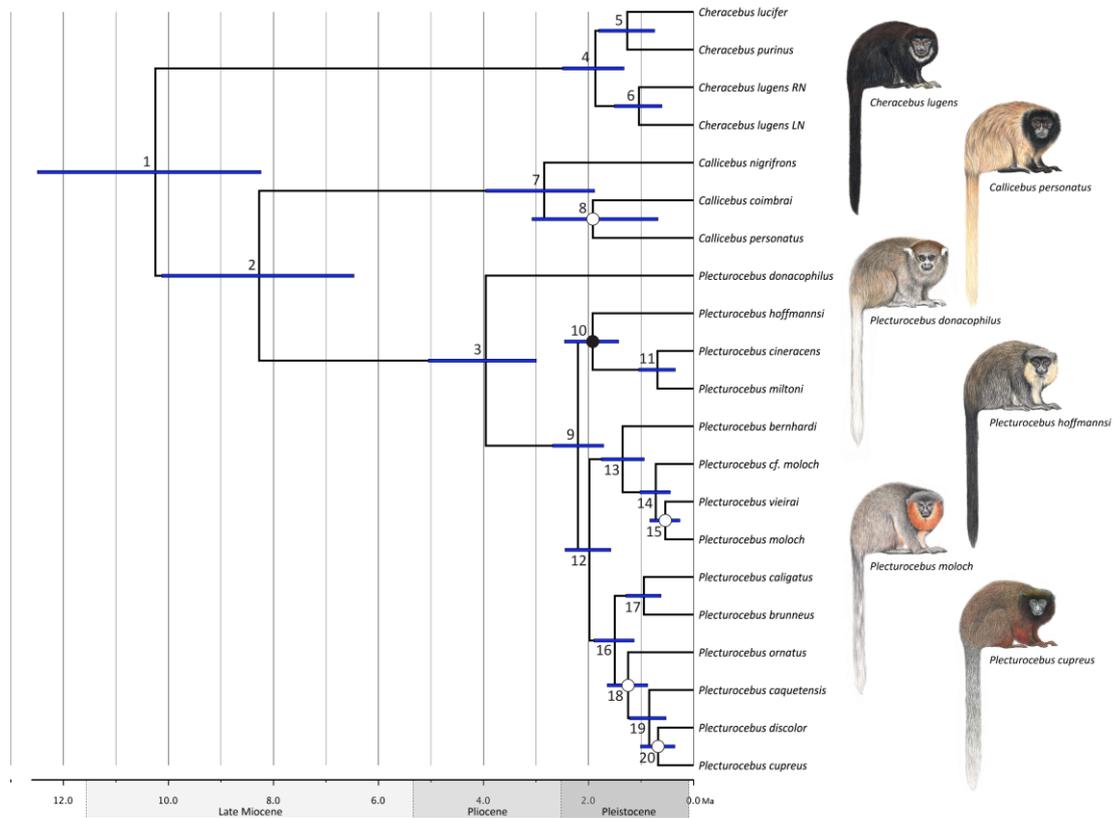
Figure 3.4 A time-calibrated phylogeny of Callicebinae. Unmarked nodes were strongly supported (PP = 1.00), the node marked with a black circle received good support (PP = 0.96), and nodes marked with white circles were recovered without significant support (PP < 0.95). Node bars indicate the 95% highest posterior density (HPDs). Support values, estimated divergence ages, and HPDs are listed according to node numbers in Table 3.8. Illustrations by Stephen D. Nash ©Conservation International.

The most notable result for the *P. moloch* group is the strongly supported division into three clades; the Aripuanã-Tapajós clade containing *P. hoffmannsi, P. miltoni* and *P. cinerascens;* the Eastern clade containing *P. bernhardi, P. cf. moloch, P. vieirai and P. moloch*; and the Western clade comprised of *P. brunneus, P. caligatus, P. ornatus, P. caquetensis, P. discolor* and *P. cupreus*. These clades diversified rapidly in the early Pleistocene, with the Aripuanã-Tapajós clade diverging from the ancestor of the Eastern and Western clades *c.* 2.2 Ma, and the Eastern and Western clades diverging at *c.* 2 Ma. Within the Aripuanã-Tapajós clade, *P. hoffmannsi* was the earliest diverging lineage (PP = 0.96), diverging from *P. cinerascens + P. miltoni* at *c.* 1.92 Ma. For the Eastern Amazonian taxa, *P. bernhardi* and *P. cf. moloch* were recovered as successive sister taxa to the *P. moloch + P. vieirai* clade, with estimated divergence times of 1.35 Ma and 0.72 Ma, respectively. The sister-species relationship of *P. moloch* and *P. vieirai* received low support (PP= 0.52), and the divergence was dated at 0.54 Ma, which is the youngest cladogenetic event across the phylogeny. For the Western Amazonian species, our results mirror those of Hoyos et al. (2016), but with addition of *P. brunneus* which was recovered as the sister taxon to *P. caligatus*. The earliest diversification event within this clade was the division between *P. brunneus + P. caligatus* and remaining taxa at *c.* 1.5 Ma. *Plecturocebus ornatus* and *P. caquetensis* were recovered as successive sister taxa to the *P. discolor + P. cupreus* clade, with estimated divergence times of 1.25 Ma and 0.84 Ma. Nodes representing the *P. ornatus* split and the divergence of *P. discolor* and *P. cupreus* were recovered with low support (0.79 and 0.59, respectively). Overall, these age estimates are slightly younger than those from Chapter 2. A summary of PP and age estimates with 95% HPD intervals are reported in Table 3.8.

3.4.2 Biogeographic reconstruction

In total, we performed 15 different biogeographic analyses using BioGeoBEARS (12) and RASP (3). Of the 12 BioGeoBEARS models evaluated, the DIVALIKE+J+X model (ΔAICc = 0) produced the best statistical fit to the data (Table 3.9). Figure 3.5 depicts the ancestral areas with the greatest probability at each node under this model, graphic depictions of state probabilities are presented in Figure A2.2.

Table 3.8 Summary of posterior probability and ages estimates from BEAST, and ancestral areas and probabilities estimated under best-fit models in BioGeoBEARS. States with a probability ≥ 0.05 are listed, bold indicates the most probable state. Node numbers correspond to those on Figure 3.4.

Node	Divergence	PP	Mean age (Ma)	95% HPD Lower	95% HPD Upper	Ancestral states & probabilites (DIVALIKE+J+X)	Ancestral states & probabilites (DEC+J+X)
1	*Cheracebus* vs. *Callicebus* + *Plecturocebus*	1.00	10.25	8.24	12.5	**BFG (0.19)**, BDFG (0.12), AFG (0.09), ABFG (0.08), BDG (0.07), BCFG, (0.07), ACFG (0.06)	**BFG (0.26)**, BDG (0.12), AFG (0.12), BDFG (0.11), CFG (0.08), ADG (0.06)
2	*Callicebus* vs. *Plecturocebus*	1.00	8.27	6.47	10.13	**FG (0.56)**, DG (0.23), DFG (0.12)	**FG (0.45),** DG (0.21), DFG (0.16)
3	*P. donacophilus* group vs. *P. moloch* group	1.00	3.96	2.99	5.05	**F (0.57),** D (0.23), DF (0.13)	**F (0.47),** D (0.22), DF (0.16)
4	*C. purinus* + *C. lucifer* vs. *C. lugens*	1.00	1.87	1.32	2.5	**B (0.38)**, A (0.17), BC (0.15), AB (0.11), C (0.10), AC (0.08)	**B (0.57)**, A (0.26), C (0.16)
5	*C. purinus* vs. *C. lucifer*	1.00	1.26	0.74	1.81	**B (0.58)**, C (0.35), BC (0.08)	**B (0.72)**, C (0.27)
6	*C. lugens* left bank vs. *C. lugens* right bank	1.00	1.04	0.6	1.51	**A (0.57)**, B (0.41)	**B (0.51)**, A (0.49)
7	*C. nigrifrons* vs. *C. personatus* + *C. coimbrai*	1.00	2.85	1.89	3.96	**G (1.00)**	**G (1.00)**
8	*C. personatus* vs. *C. coimbrai*	0.61	1.92	0.68	3.08	**G (1.00)**	**G (1.00)**
9	Aripuanã-Tapajós clade vs. Eastern + Western *P. moloch* clades	1.00	2.2	1.71	2.68	**D (0.91)**, C (0.05)	**D (0.84)**, C (0.08)
10	*P. hoffmannsi* vs. *P. cinerascens* + *P. miltoni*	0.96	1.92	1.43	2.46	**D (1.00)**	**D (1.00)**
11	*P. cinerascens* vs. *P. miltoni*	1.00	0.69	0.35	1.04	**D (1.00)**	**D (1.00)**
12	Eastern vs. Western *P. moloch* clades	1.00	1.99	1.58	2.45	**D (0.69)**, C (0.23), E (0.08)	**D (0.62)**, C (0.25), E (0.08)
13	*P. bernhardi* vs. *P.* cf. *moloch* + *P. vieirai* + *P. moloch*	1.00	1.35	0.94	1.76	**D (0.84),** E (0.16)	**D (0.82),** E (0.17)
14	*P.* cf. *moloch* vs. *P. vieirai* + *P. moloch*	1.00	0.72	0.44	1.02	**E (1.00)**	**E (1.00)**
15	*P. vieirai* vs. *P. moloch*	0.52	0.54	0.26	0.83	**E (1.00)**	**E (1.00)**
16	*P. brunneus* + *P. caligatus* vs. *P. ornatus* + *P. caquetensis* + *P. discolor* + *P. cupreus*	1.00	1.5	1.13	1.9	**C (0.73),** D (0.25)	**C (0.71),** D (0.25)
17	*P. brunneus* vs. *P. caligatus*	1.00	0.95	0.62	1.29	**C (0.71)**, D (0.29)	**C (0.70)**, D (0.30)
18	*P. ornatus* vs. *P. caquetensis* + *P. discolor* + *P. cupreus*	0.79	1.25	0.87	1.65	**B (0.59)**, C (0.41)	**B (0.60)**, C (0.40)
19	*P. caquetensis* vs. *P. discolor* + *P. cupreus*	1.00	0.84	0.52	1.22	**B (0.54)**, C (0.46)	**B (0.54)**, C (0.45)
20	*P. discolor* vs. *P. cupreus*	0.59	0.68	0.36	1.01	**C (0.60)**, B (0.40)	**C (0.60)**, B (0.40)

The second best model, DEC+J+X (ΔAICc = 3.94; see Appendix 2, Figure A2.3, A2.4), estimated identical most likely states (ancestral areas) at all nodes with similar support values as the best-fitting model (DIVALIKE+J+X) with one exception (*C. lugens*; discussed below). The concordance between the ancestral areas and diversification patterns recovered under these two best-fit models are viewed as strong support for the inferred biogeographic scenario, detailed below. Together these two best-fit models comprise 0.98 of the relative likelihood according to corrected Akaike weights (AICc), with the other ten BioGeoBEARS models combined accounting for just 0.02 of the relative likelihood (Table 3.9). The results from each of the RASP models are comparable to one of the corresponding BioGeoBEARS analyses, with the same most likely states recovered at all nodes; BBM to BAYAREALIKE+J, S-DEC to DEC, and S-DIVA to DIVALIKE (except for the most recent common ancestor of *Cheracebus*).

Our results support a widespread ancestral population to all titi monkeys occurring through the Napo, Wet & Dry Savannas and Atlantic Forest regions early in the late Miocene and genus-level diversification events were characterised by vicariance of this ancestral range. An initial major vicariant event at *c.* 10.25 Ma resulted in the isolation of *Cheracebus* in the Napo and the *Callicebus* + *Plecturocebus* ancestor in the Atlantic Forest and Wet & Dry Savannas. A subsequent vicariant event at *c.* 8.27 Ma led to the isolation of *Plecturocebus* in the Wet & Dry Savannas region and *Callicebus* in the Atlantic Forest, where all species-level divergences within this clade occurred. The *Plecturocebus* species groups diverged in the Pliocene, *c.* 4 Ma, through a founder-event when the progenitor of the *P. moloch* group dispersed from the Wet & Dry Savannas into Rondônia, while the *P. donacophilus* group ancestor remained in the Wet & Dry Savannas region. All diversification events within the Amazonian clades, *Cheracebus* and the *P. moloch* group, were characterised by founder-event speciation (jump dispersal) or occurred within one geographic area ("narrow sympatry", see Matzke, 2013b). The initial divergence within the *P. moloch* group occurred within Rondônia between the ancestors of the Aripuanã-Tapajós clade and the Eastern + Western clade. The divergence between the Eastern and Western clades was explained by jump dispersal of the progenitor of the Western clade from Rondônia into Inambari at *c.* 2 Ma.

Table 3.9 Comparison of likelihood values (LnL) and corrected Akaike's information criterion (AICc, ΔAICc, AICc weight) for each of the BioGeoBEARS analyses.

Model	No. params	LnL	AICc	ΔAICc	AICc weight	*p*-value*
DIVALIKE	2	-40.69	86.05	29.2	3.95E-07	1.80E-06
DIVALIKE+J	3	-29.28	65.98	9.13	9.00E-03	
DIVALIKE+X	3	-32.66	72.72	15.87	3.10E-04	1.30E-05
DIVALIKE+J+X	4	-23.17	56.85	0	0.86	
DEC	2	-45.78	96.24	39.39	2.42E-09	4.70E-08
DEC+J	3	-30.86	69.13	12.28	1.86E-03	
DEC+X	3	-38.75	84.91	28.06	6.98E-07	1.80E-07
DEC+J+X	4	-25.14	60.79	3.94	0.12	
BAYAREALIKE	2	-55.88	116.4	59.55	1.01E-13	1.70E-11
BAYAREALIKE+J	3	-33.24	73.88	17.03	1.73E-04	
BAYAREALIKE+X	3	-51.18	109.8	52.95	2.75E-12	2.20E-11
BAYAREALIKE+J+X	4	-28.78	68.07	11.22	3.17E-03	

*chi-squared test between LnL

Among the Eastern clade taxa, *P. bernhardi* diverged when the ancestor of *P. cf. moloch* + *P. vieirai* + *P. moloch* dispersed into the Pará region, and subsequent divergence of these latter taxa occurred within Pará. The species of the Western clade of the *P. moloch* group showed the most complex pattern of spatial diversification characterised by four cladogenetic dispersal events. The initial divergence among Western clade taxa occurred when the ancestral population to the clade containing *P. ornatus, P. caquetensis, P. discolor,* and *P. cupreus*, dispersed from Inambari into Napo. Subsequent diversification of these taxa was largely explained by founder-events between Napo and Inambari. The *P. caligatus* + *P. brunneus* ancestor remained in Inambari and subsequently diverged through a founder-event when the ancestor of *P. brunneus* dispersed back to Rondônia. Among the *Cheracebus* taxa, the DIVALIKE+J+X model inferred that the earliest divergence between *C. lugens* and *C. purinus* + *C. lucifer* occurred when the ancestor of *C. lugens* dispersed from the Napo into the Pantepui region, and subsequently, the two *C. lugens* lineages diverged when the right bank Rio Negro lineage moved back into the Napo region. The DEC+J+X model, however, recovered a slightly different pattern where the initial divergence between *C. lugens* and *C. purinus* + *C. lucifer* was within Napo, and the *C. lugens* lineages divided when the left bank Rio Negro ancestor dispersed into the Pantepui region. Under both models, *C. purinus* and *C. lucifer* diverged via a founder-event when the ancestor of *C. purinus* colonised Inambari.
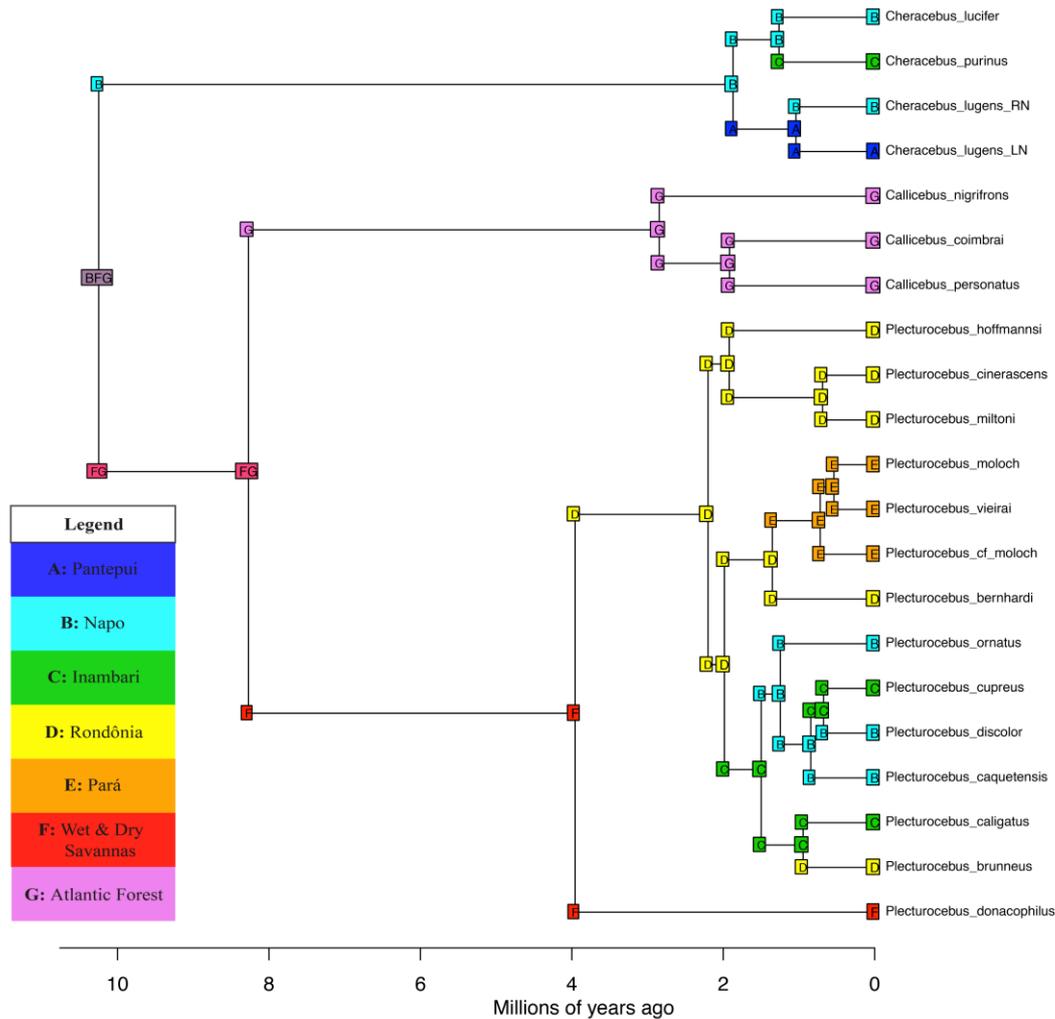
Figure 3.5 DIVALIKE+J+X reconstruction of ancestral areas for Callicebinae.

The above reconstruction of Callicebinae spatial diversification is based on the ancestral states with the greatest probability at each node under the best-fit models (DIVALIKE+J+X and DEC+J+X). The most probable ancestral areas recovered at most nodes under these models showed good support (see graphic depictions of state probabilities in Appendix 2, Figure A2.2, A2.4, and a summary of all probable states at each node in Table 3.8). The greatest uncertainty is found at the root node, which is often characteristic of biogeographic methods due to a lack of direct information about historical distributions and extinct taxa (Landis, 2017). Despite this ambiguity, all analyses within the DEC and DIVA model sets (including the RASP S-DIVA and S-DEC models) inferred the most likely ancestral area of origin for Callicebinae across Napo, Wet & Dry Savannas, and Atlantic Forest, and some included Rondônia. Furthermore, under the best-fit models, the second most probable state at the ancestral node included Rondônia along with the original areas. Taken together, these results

85

provide stronger support for an ancestrally widespread population to all modern titis across the Napo, Wet & Dry Savannas, and Atlantic Forest regions, and potentially Rondônia. For the BioGeoBEARS test analyses that included *Cacajao calvus* and *Chiropotes israelita*, the same models produced the best statistical fit to the data (DIVALIKE+J+X and DEC+J+X) and identical ancestral states were estimated for the origin of Callicebinae under these models as for the main analyses based on only ingroup taxa. Biogeographic studies focusing on the family Pitheciidae (with good taxonomic coverage) will likely increase our understanding of the geographic origin of Callicebinae owing to the added information at the root node derived from the estimated ancestral distributions of Pitheciinae.

For each set of analyses (DEC, DIVALIKE, BAYAREALIKE), the addition of both the X and the J parameters conferred large improvements to likelihood compared to all the simpler analogous models. Importantly, DIVALIKE+J is highly comparable to the best-fit model, DIVALIKE+J+X, which is nested within it. Under the DIVALIKE+J model (where no assumptions are made about the relationship between distance and dispersal), slightly different ancestral areas were recovered at a small number of nodes, the most significant of which was the inclusion of Rondônia in the root ancestral states, as noted above (see Figure A2.5). The biogeographic processes recovered under both models, however, are identical: early vicariant events marking divergence of the genera, while all further cladogenetic events were explained by founder-event speciation or occurred within one geographic area.

The impact of the J parameter on the likelihood across all analyses is significant (Table 3.9). It is evident that founder-event speciation is an important process in explaining current species distributions from both the statistical fit of the +J models and the proportion of jump dispersal events in the reconstructions. BSM simulations that estimate the overall probability of different biogeographic processes under the specified phylogeny, parameters and model, lend further support to jump dispersal (J) as a strong contributor to explaining the data. Event counts derived from 1000 simulated histories under both the DIVALIKE+J+X and DEC+J+X models (where the J parameter is ~ 0.5) indicate that about 47–48% of the cladogenetic events were founder-events, 40% were within-area sympatry, and only around 11–12% were vicariance, while there were zero anagenetic dispersal events (Table 3.10).

Table 3.10 Biogeographic stochastic mapping event counts under DIVALIKE+J+X and DEC+J+X models in BioGeoBEARS.

		Founder-event	Vicariance	Narrow sympatry	Subset sympatry	Anagenetic dispersal
DIVALIKE+J+X	**Mean event counts**	9.43	2.5	8.07	0	0
	Standard deviation	1.07	0.52	1.01	0	0
	% of cladogenetic events	47.15%	12.49%	40.36%	---	---
DEC+J+X	**Mean event counts**	9.59	2.21	8.09	0.1	0
	Standard deviation	1.22	0.52	1.03	0.44	0
	% of cladogenetic events	47.95%	11.07%	40.48%	0.50%	---

3.5 Discussion

3.5.1 Phylogenetic inference

The most notable phylogenetic result is the strongly supported division of the *P. moloch* group into three clades, the Aripuanã-Tapajós, Eastern, and Western clades. The phylogenetic relationships between *P. miltoni, P. cinerascens* and all other *P. moloch* group taxa were previously unresolved, finding support as independent early diverging lineages based on nuclear loci or as the sister clade to the Eastern Amazonian taxa based on mitochondrial sequences (see Chapter 2). The species relationships and age estimates recovered are concordant with the proposal that all current *P. moloch* group taxa should continue to be recognized as a single species group (Chapter, 2; Carneiro et al., 2016; Hoyos et al., 2016). The descriptive names for each of the three major *P. moloch* group clades recovered in this study are suggested to aid communication given the absence of justification to erect additional species groups, and thus the necessity for formal classification. They are based upon the geographic centre of the distribution of each clade in the Amazon basin. The Eastern and Western *P. moloch* clades were estimated to have diverged at a similar time to *P. cinerascens + P. miltoni* and *P. hoffmannsi* (2 Ma vs. 1.92 Ma), however, it is helpful to denote the Eastern and Western clades separately given their taxonomic history and their general phenotypic and geographic differences.

Although we are missing a small number of described species from the *P. moloch* group, namely *P. baptista, P. toppini* and *P. aureipalatii* (which may be a

synonym of *P. toppini*, see Vermeer & Tello-Alvarado, 2015), our dataset largely covers the spectrum of species diversity known within this group. We suggest that *P. baptista* is likely a member of the Aripuanã-Tapajós clade, while *P. toppini* and *P. aureipalatii* likely are members of the Western clade. We are also missing sequences for *P. stephennashi*, however, doubt surrounds the validity and species status of this taxon given the morphological and geographical affinities to *P. caligatus*. For the biogeographical analyses, *P. dubius* is best represented as *P. caligatus*, regardless of whether it is considered a synonym, geographic variant or subspecies (see Byrne et al., 2016; Carneiro et al., 2016; Hoyos et al., 2016; Serrano-Villavicencio et al., 2017), and thus it was not included in this dataset. Vermeer & Tello-Alvarado (2015) placed *P. urubambensis* in the *P. donacophilus* group based on the metachromism hypothesis of Hershkovitz (1988), however as discussed, this is not generally a reliable predictor of species relationships. These authors also indicated that *P. urubambensis* could be most closely related to *P. brunneus* based on phenotypic similarities, and thus it may also be a member of the Western *P. moloch* clade.

3.5.2 Historical biogeography

The patterns of diversification reconstructed across both the genera and species are consistent with the predictions of the "Dynamic Young Amazon" model (Aleixo & Rosetti, 2007; Campbell et al., 2006; Latrubesse et al., 2010; Rosetti et al., 2005). Here we summarise our interpretation of the biogeography of Callicebinae under this model (Figure 3.6, 3.7).

Prior to genus divergences, the ancestral population to all modern titis was widespread between the Napo, Wet & Dry Savannas and Atlantic Forest regions in the late Miocene (Figure 3.6i). Despite occurring in surrounding regions, titis were absent from the western Amazon (Inambari), which is consistent with the proposal that the western Amazon was inhospitable to upland forest lineages in the Miocene owing to the extensive Pebas wetlands system. These wetlands would have also extended into parts of Napo, thus this interpretation assumes that titi monkeys were restricted to the region around the lake in the current Andean foothills of the Napo region. The existence of Callicebinae towards the western extreme of the Napo region in the late Miocene is supported by the fossil *Miocallicebus* Takai et al., 2001, which is dated to *c.* 11–12 Ma (Kay et al., 2013). *Miocallicebus* belongs to the La Venta fauna of the modern day Tatacoa Desert, Magdalena Valley, Colombia, which would

have been contiguous with the Napo region in the late Miocene prior to the formation of the eastern Andes. Similar patterns in other Neotropical primates have been associated with a broad ancestral distribution in an arc along the west bank of the Pebas wetlands (Buckner et al., 2015), suggesting that Napo was at least intermittently connected to the southern regions through the Andean foothills west of the Inambari region. This is consistent with the proposal that land connections existed between the Andean foothills and Brazilian Shield around the southern rim of the Pebas system during the Miocene. Furthermore, the absence of titis from the Guiana Shield until the radiation of *C. lugens* in the Pleistocene suggests that connectivity around north of the Pebas system between the Guiana Shield/Pantepui region and Napo had ceased by the late Miocene (Wesselingh & Salo, 2006).
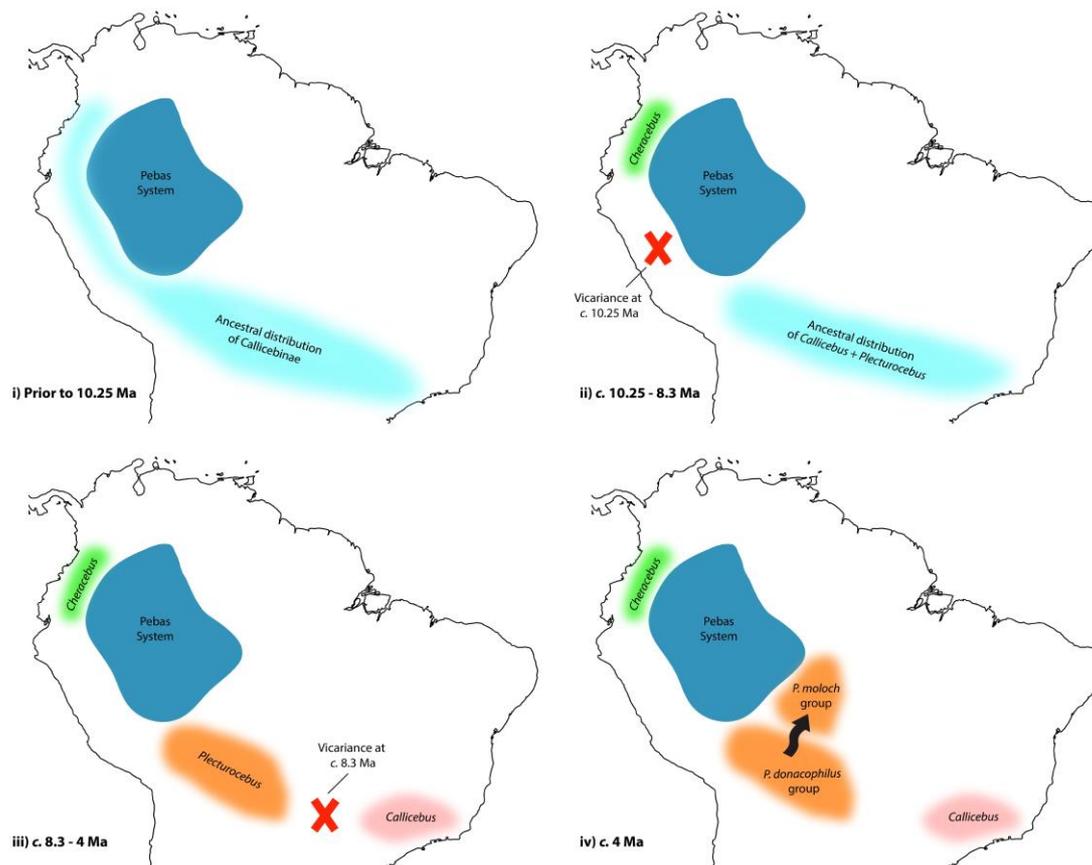
Figure 3.6 Graphical depiction of approximate Callicebinae historical distributions and spatial diversification pattern in the late Miocene and Pliocene under the "Dynamic Young Amazon" model.

The earliest divergence within the subfamily Callicebinae at *c.* 10.25 Ma was characterized by a vicariant event that disrupted all connectivity between the Napo and southern regions, consistent with vicariance across relatively geologically stable terra firme centres of the Andean foothills (of the Napo) and the Brazilian Shield (parts of the Wet & Dry Savannas region) as predicted under this model (Figure 3.6ii). This vicariance may have been associated with the period of strong and widespread Andean uplift that begun around the late Miocene (e.g., Garzione et al., 2006, 2008; Ghosh et al., 2006; Gregory-Wodzicki, 2000). Once the link around the southern rim of the Pebas system disappeared, the region of the Napo occupied by the *Cheracebus* ancestor remained disconnected from all other areas until the floodplains of the western Amazon receded in the Pleistocene. The origin of *Cheracebus* in the Andean foothills of the Napo is consistent with the observation that older Amazonian avian taxa that diverged in the Miocene are mainly found in the north-western Amazon basin (and the Guianas), with the highest concentration found around the tropical Andes of Ecuador, Colombia and Peru (see Figure 1a in Batalha-Filho et al., 2013).

The *Callicebus* + *Plecturocebus* ancestor was widespread across the Wet & Dry Savannas and Atlantic Forest, which remained in connection whether continuously or intermittently, until another vicariant event at *c.* 8.3 Ma led to their disjunction, isolating the ancestor of *Callicebus* in the Atlantic Forest and *Plecturocebus* in the Wet & Dry Savannas region (Figure 3.6iii). As such, our results indicate that titis spread to the Atlantic Forest via a southern pathway in the Miocene, rather than in the Plio-Pleistocene from the eastern Amazon, providing support for the proposal that the Amazon and the Atlantic Forest were connected along the western edge of the Brazilian Shield and the transition towards the Chaco of Bolivia and Paraguay during the Miocene (Batalha-Filho et al., 2013). Our reconstruction suggests that this corridor disappeared in two stages in the late Miocene: first the Wet & Dry Savannas lost connectivity with the Napo region *c.* 11–10 Ma, and around 2 million years later, the Atlantic Forest became isolated following the closure of the link along the modern-day Chaco/southern Cerrado. Lineages of plants found in the Cerrado began to diversify around 9 Ma (Simon et al., 2009), and it is possible that ecological changes early in the transition to the Cerrado biome disrupted this connection to the Atlantic Forest and played a role in the divergence of *Callicebus* and *Plecturocebus*. Furthermore, this scenario suggests that the north-western Amazon was connected to

the southern Atlantic Forest, consistent with diversification patterns seen for avian taxa where most upland lineages found in the southern Atlantic Forest originated in the Miocene and have closely related taxa distributed around the Andes (see Fig 2. Batalha-Filho et al., 2013; see also Percequillo et al., 2011).

After these major vicariant events, no further change to ancestral distributions occurred until the progenitor of the *P. moloch* group dispersed from the Wet & Dry Savannas region into Rondônia at *c.* 4 Ma (Figure 3.6iv). An alternative scenario (according to some models in the DEC and DIVA sets and the second most probable states under the best-fit models, see section 3.4.2) indicates that ancestral range extended further within the Brazilian Shield through the Wet & Dry Savannas into Rondônia since initial divergence in the late Miocene, which is still consistent with the "Dynamic Young Amazon" model. Here, the divergence of *Plecturocebus* species groups was characterised by vicariance between Rondônia and the Wet & Dry Savannas, which are separated by the Rio Madeira watershed suggesting a potential role for the formation of this river system.

Species-level diversification in the Amazonian clades, *Cheracebus* and the *P. moloch* group, occurred in an explosive manner from the early Pleistocene to *c.* 0.5 Ma (Figure 3.7). The *P. moloch* group showed a complex pattern of diversification with a non-monophyletic assemblage of taxa endemic to Rondônia, the area of origin for this group. The initial major division among *P. moloch* group taxa occurred within Rondônia at *c.* 2.2 Ma and current distributions correlate with the Rio Roosevelt-Aripuanã such that the Aripuanã-Tapajós clade is restricted to the right bank and the Rondônian members of the Western and Eastern clade (*P. brunneus* and *P. bernhardi*) are found on the left bank (Figure 3.7i). Based on similar patterns in some avian taxa, Fernandes (2013) proposed several mini interfluvial areas of endemicity for Rondônia that correspond to the subdivisions found among *P. moloch* group taxa. The distributions of *P. brunneus* and *P. bernhardi* are contained within the regions denoted R1 and R2, on the left bank of the Rio Roosevelt-Aripuanã, while the regions denoted R3 and R4 correspond to the distribution of the Aripuanã-Tapajós clade from the right bank of the Rio Roosevelt-Aripuanã (see Figure 2 in Fernandes, 2013), although the range of *P. cinerascens* also extends upstream of the headwaters of the Rio Roosevelt.
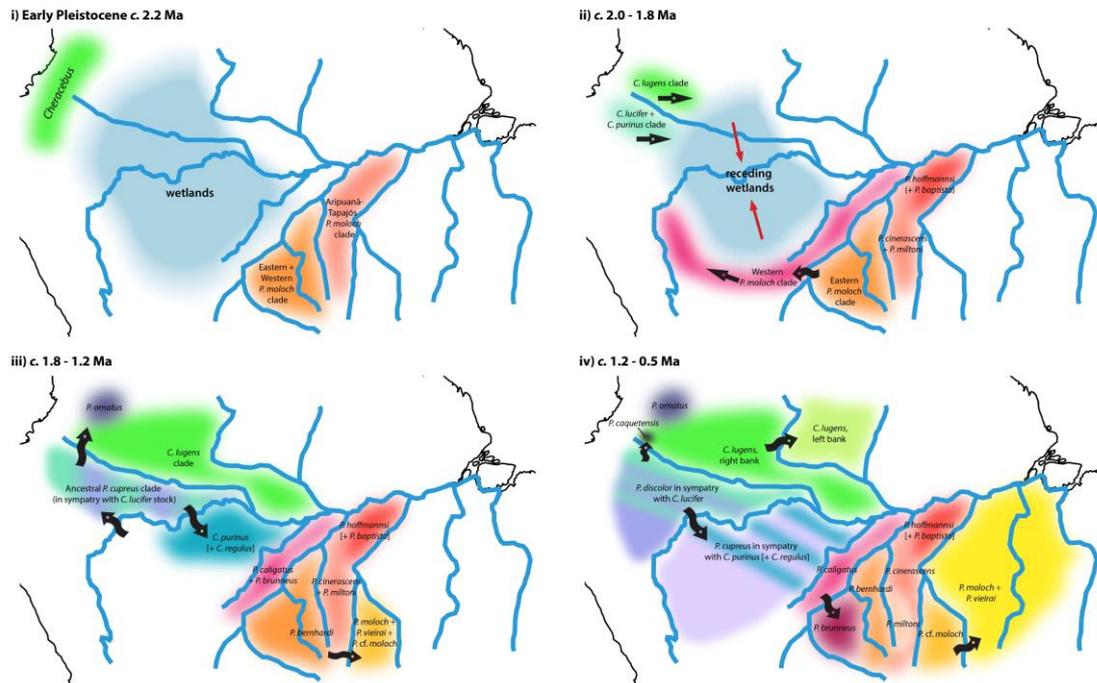
Figure 3.7 Graphical depiction of the spatial diversification pattern for Amazonian titis of the genus *Cheracebus* and the *P. moloch* group in the Pleistocene under "Dynamic Young Amazon" model. Black arrows indicate dispersal events. Major Amazonian rivers are also shown.

One proposed explanation for the complex diversification patterns seen in this region is changing river dynamics in the Plio-Pleistocene, whereby current tributaries of the Rio Madeira such as the Rio Roosevelt-Aripuanã may have captured a significant part of the Madeira or even Tapajós drainages (Fernandes, 2013; Muniz et al., in revision). Furthermore, evidence for two large mega-fans from the late Pleistocene involving the Rios Roosevelt-Aripuanã and Jiparaná indicates the existence of a much wider and more complex drainage system in this interfluvium in the Pleistocene (Latrubesse, 2002). This complex history of river system evolution is likely to have played an important role in driving historical distributions in Rondônia. We reason that the initial division within the *P. moloch* group was associated with the Rio Roosevelt-Aripuanã, whether through a rare founder-event over this barrier or vicariance owing to river formation, restricting the ancestor to the Aripuanã-Tapajós clade on the right bank to the northern part of Rondônia. Species-level diversification of the Aripuanã-Tapajós clade taxa occurred within this region with a notably early divergence between *P. hoffmannsi* and *P. cinerascens + P. miltoni* dated at *c.* 1.9 Ma. The current distribution of this clade suggests that the lower Rio Amazonas, as well as the Rios Madeira and Tapajós were largely established by *c.* 2.2 Ma, providing barriers to the dispersal of this clade. We further suggest that the occurrence of *P.*

92

cinerascens upstream of the headwaters of the Rio Roosevelt-Aripuanã represents recent range expansion that could be related to fluvial dynamics associated with this river since the Late Pleistocene. Rondônia is the only geographic area with members from more than one of the *P. moloch* clades and is of evident importance to understanding the biogeographic history of this group. This pattern of shared lineages in the Rondônia region appears to be general as similar patterns are observed in birds (Fernandes, 2013; Thom & Aleixo, 2015) and lizards (Oliveira et al., 2016). Studies with a specific focus on the biogeography of the *P. moloch* group will provide insight into diversification dynamics within Rondônia at a finer scale.

After the initial major divergence, the ancestral population to the Western + Eastern clade was isolated briefly in the southern section of Rondônia on the left bank of the Rio Roosevelt-Aripuanã. At around 2 Ma, the Western and Eastern clades diverged when the progenitor of Western taxa dispersed to Inambari via a founder-event across the upper Rio Madeira (Figure 3.7ii). The Western clade subsequently spread westward across Inambari, reaching the Rio Solimões and dispersing into the Napo by *c*. 1.5 Ma. Western clade taxa had likely reached the Andean foothills in Colombia at the north-western extreme of current titi monkey distributions by around 1.25 Ma based on the divergence of *P. ornatus*. This diversification pattern suggests that the wetlands had receded and suitable upland forest habitat had begun to be established in the western Amazon by the early Pleistocene allowing the rapid colonisation of Inambari and Napo. Species divergences within the Western clade were largely characterised by sequential "island hopping" between Inambari and Napo across the Rio Solimões-Amazonas, although a founder-event back across the upper Rio Madeira, from Inambari to southern Rondônia, led to the divergence of *P. brunneus*. Further details about the exact pattern of spatial diversification, however, should be interpreted with caution given the important taxa missing from this clade in this study, such as *P. toppini*, *P. aureipalatii* and potentially *P. urubambensis* (see Vermeer & Tello-Alvarado, 2015).

The progenitor of the Eastern clade remained in Rondônia to the left bank of the Rio Roosevelt-Aripuanã where *P. bernhardi* diverged when the ancestor of the remaining Eastern clade taxa dispersed eastwards over or around the headwaters of the Rio Tapajós-Juruena into the Pará region (Tapajós area of endemism) at *c*. 1.35 Ma. *Plecturocebus* cf. *moloch* is restricted to the Alta Floresta region between the Rios Juruena and São Manuel, and likely diverged from the ancestor of *P. vieirai + P.*

moloch as they dispersed eastwards from Rondônia. The phylogenetic and biogeographic relationship between *P. moloch* and *P. vieirai* requires further investigation given the low support for this node and the current lack of clear limits to their geographic distributions; *Plecturocebus moloch* is broadly distributed between the Rios Tapajós and Tocantins while the whiter phenotype classified as *P. vieirai* is known from various localities around the left bank of the Rio Xingu.

The two best-fitting BioGeoBEARS models inferred slightly different scenarios for the spatial diversification of the genus *Cheracebus*; the DEC+J+X reconstruction, however, is more consistent with extant species distributions and thus discussed here. Under this model, the initial divergence between *C. lugens* and *C. lucifer* + *C. purinus* at *c.* 1.87 Ma occurred within the Napo region and current distributions suggest that it was associated with the Rio Caquetá-Japurá. As such, we propose that ancestral *Cheracebus* began to expand out of the Andean foothills of the north-western Amazon following the receding wetlands in the early Pleistocene. This population was divided when it reached and dispersed over the Rio Caquetá (in either direction), or by vicariance owing to river dynamics, leaving the ancestor to *C. lugens* to the north and *C. lucifer* + *C. purinus* to the south. These clades diversified as they spread eastwards on either side of the Rios Caquetá-Japurá and Solimões, with the *C. lugens* lineages diverging via a founder-event across Rio Negro into the Pantepui region (Figure 3.7ii–iv). The divergence of *C. lucifer* and *C. purinus* occurred when the ancestor to *C. purinus* dispersed across the Rio Solimões into Inambari at *c.* 1.26 Ma. For the *Cheracebus* species for which no sequence data are available, this pattern suggests that *C. regulus* (Inambari) is the sister taxon to *C. purinus*, while *C. medemi* (Napo) is either sister to *C. lucifer* or the earliest diverging lineage within the clade south of the Rio Caquetá-Japurá.

Amazonian titis (*Cheracebus* and the *P. moloch* group) are found in terra firme regions and occasionally in seasonally inundated black-water forests (Igapó), but they are generally absent from white-water flooded forest habitats (Varzéa). It is likely that the extensive lake and floodplains that existed in the Western Amazon were largely inhospitable to titi monkeys and colonisation of this region would have depended upon the availability of suitable lowland forest habitats. This is supported by the absence of ancestral range from Inambari throughout the late Miocene and Pliocene. Our reconstruction indicates that both *Cheracebus* and *Plecturocebus* taxa dispersed to Inambari from other regions between 2–1 Ma, thus suggesting that the

western Amazon transitioned from a lacustrine system to floodplain to lowland forest in the Pleistocene ("Dynamic Young Amazon" model).

The largely cosmopolitan distribution of the *P. moloch* group across the Amazon is notable: from east of the Rio Tocantins in eastern Amazon, to the Rio Beni in northern Bolivia, and northwest as far as the Andean foothills of the Meta department in Colombia, entirely absent only from the Pantepui region. Differences in ecology/habitat preference and dispersal ability may have allowed the *P. moloch* group to spread much more widely across the Amazon than *Cheracebus* within a similar time period. Rather than soil type, the structure of forest vegetation is thought to best explain differences in habitat preference (Defler, 1994). *Cheracebus* species occur in forests that are tall and well-stratified, mostly in undisturbed highland terra firme regions but they have also been recorded in well-developed Igapó habitats. Taxa of the *P. moloch* group are frequently found in poorly stratified low forests including disturbed, secondary and liana-rich habitats in terra firme and seasonally inundated regions where they are often located at the edge of large gaps in the canopy (Defler, 1994; Haugaasen & Peres, 2005; Peres, 1997; Van Roosmalen et al., 2002). As poorly stratified gallery forests are often in low-lying areas along waterways, rare founder-events by passive transfer across riverbanks may occur more frequently among taxa of the *P. moloch* group, facilitating passive dispersal across major Amazonian rivers and the rapid colonisation of the greater part of the Amazon within the last 2 million years. This scenario is consistent with the complex diversification pattern based on sequential founder-events reconstructed in this study for the *P. moloch* group clade, including several dispersal events across major rivers such as the Solimões and Madeira. Additionally, owing to their ability to thrive at edge and in disturbed habitats, it is possible that *P. moloch* group taxa have been able to disperse more broadly through the use of temporary, secondary, or developing lowland forests and forests at the edge of their range, while the dispersal of *Cheracebus* species may depend upon the establishment of well-developed forests and connectivity between them. Our reconstruction implies that the Western *P. moloch* clade invaded Inambari and Napo slightly before *Cheracebus* dispersed out of the Napo region, which may also be explained by these differences in habitat preference, however the time difference is not significant when confidence intervals are considered (see 95% HPD intervals in Table 3.8).

This biogeographic reconstruction sheds light on the present-day sympatry

between *Cheracebus* and *Plecturocebus* species across the Napo and Inambari regions. The ancestors to these clades were likely isolated at opposing extremes of the Amazon in the terra firme centres of the Andean foothills and Brazilian Shield by the Pebas system in the late Miocene, existing in isolation in these distinct regions for 8–9 million years. Range expansion from the east (*Plecturocebus*) and west (*Cheracebus*) in the Pleistocene following the establishment of lowland forest in the western Amazon has led to sympatry (see Figure 2.5) with distinct ecological niches.

Our biogeographic reconstruction supports a sequential, long-distance dispersal model of speciation by "island hopping" across pre-existing river barriers. These founder-events were sufficiently rare to allow divergence in isolation after dispersal, pointing to major rivers as relatively strong barriers to gene flow (riverine barrier hypothesis). Founder-events (jump dispersal) simply indicate that dispersal and speciation were coincident, but whether this could have been associated with the classic but controversial "founder effect" (Mayr, 1954) would depend on the number of individuals that dispersed, it is important to note that these two terms are not equivalent.

In support of this mode of speciation, evidence also exists to indicate that dispersal across major rivers, or around the headwaters, is an on-going process albeit infrequent. A specimen classified as *P. cupreus* (FR 62) based on morphology and collection locality has a cytochrome *b* gene with the closest genetic affinity to *P. hoffmannsi* and *P. moloch* (Hoyos et al., 2016). Although no nuclear data was generated, this supports at least one relatively recent dispersal event by a member of either the Eastern or Aripuanã-Tapajós clade across the Rio Madeira, and potentially subsequent introgression with *P. cupreus*. A specimen collected near the UHE Rondon II in Rondônia (UFRO 354) identified as *P. cinerascens* based on pelage colouration, confirmed by genome-wide nuclear data (discussed in Chapter 4), has a *P. bernhardi* mitochondrial genome (e.g., it was classified as *P. bernhardi* based on mtDNA in Carneiro et al., 2016). This locality is situated at the upper reaches of the Rio Roosevelt, which forms the major barrier delimiting their geographic distributions, and may indicate more recent gene flow and a potential contact/hybrid zone in south-western Rondônia between these taxa. Considering only a relatively small number of specimens have been sequenced (nuclear or mitochondrial loci) for Callicebinae as a whole, two possible cases of gene flow between taxa divided by major rivers is significant and also particularly relevant to phylogenetic inferences

based solely on mitochondrial data. Dispersal of the Western *P. moloch* clade across the Rio Madeira and Rio Solimões was likely along the upper sections in higher regions of Amazon basin where titis are found closer to the edge of the river, rather than the lower reaches where these rivers are fringed with impenetrable Varzéa. Additionally, it is likely that other rivers not delineated in this study, such as the Rios Purus, Juruá and Caquetá, also played an important role in the diversification of the Amazonian clades.

Our reconstruction suggests that major Amazonian tributaries were largely established prior to the beginning of species-level diversification in Amazonian clades, and supports no role for major rivers as vicariant agents potentially owing to the narrow ancestral ranges of both clades prior to Pleistocene range expansion. Although our results give little indication whether major elements of the current drainage system had formed by late Miocene or in the Plio-Pleistocene, they suggest a Pleistocene transition from floodplain to lowland forest in the Western Amazon, in support of the "Dynamic Young Amazon" geologic model, which itself suggests that the current Amazonian drainage system was established in the Plio-Pleistocene.

The study of the biogeography of titis is limited by the absence of a species-level phylogeny with all Callicebinae taxa represented. Explicit phylogenetic hypotheses are a necessary component of understanding the spatial patterns of diversification and the processes driving extant and historical distributions. The *P. donacophilus* group is particularly poorly represented in the current study owing to the lack of available sequence data for most taxa, and little information about their spatial diversification can be inferred from our reconstruction apart from the putative origin of this clade in the Wet & Dry Savannas region, which comprises the majority of their current geographic distribution (except *P. oenanthe*). Given the unusual and disjunct distributions of these species, in particular, *P. oenanthe* in the Andean foothills of Peru, a phylogenetic framework including these species will be essential for a comprehensive reconstruction of the biogeographic history of Callicebinae at species-level and to assess these biogeographic hypotheses. Until then, we advocate caution over strict interpretation of the results, however, we believe that the overall patterns of diversification recovered here are significant, i.e., major divergences between the genera occurred by vicariance of widespread ancestral populations, the connection between the northwestern Amazon and Atlantic Forest, and the divergence of Amazonian species through sequential founder-events from a narrow area of origin

(likely the northwestern Amazon for *Cheracebus*, and Rondônia for the *P. moloch* group).

Similarly, further inferences regarding the Atlantic Forest *Callicebus* depend upon the inclusion of multiple distinct regions within the Atlantic Forest biome. Our biogeographic results suggest that Callicebinae spread to the Atlantic Forest prior to 10.25 Ma via the present-day Chaco likely to the southern Atlantic Forest and subsequently northwards through this region. As such, they indicate that *Callicebus* originated in this region, rather than the Amazon (see Jameson Kiesling et al., 2015), diverging from the progenitor of *Plecturocebus* by vicariance of a geographically widespread ancestor. Age estimates suggest that diversification among *Callicebus* species represented in this study begun at *c.* 2.85 Ma, earlier than among *Cheracebus* or *P. moloch* group taxa, and it is likely that the pattern of diversification differs from the Amazonian clades. Forests in the southern Atlantic Forest are cooler than lowland forests in the northern Atlantic Forest and studies have shown that the distinction between these regions is important for many primates (Kinzey, 1982), corresponding to the limit between the geographic distributions of *C. nigrifrons* and *C. personatus* + *C. coimbrai* (see Figure 5 in Printes et al., 2013). Future studies focusing specifically on each of the major titi monkey clades, with increased sampling and further geographic delineations within each region, will allow a more in-depth investigation of their biogeographic history and the processes promoting their diversification at a finer scale.

3.6 Conclusions

In this study, we reconstructed a well-supported phylogeny for Callicebinae that supports the diversification of the *P. moloch* group into three major clades (Aripuanã-Tapajós, Eastern, Western). Our phylogenetic results clarify several questions regarding Callicebinae evolutionary history and species relationships, such as the relationship between *P. cinerascens* + *P. miltoni* and the other taxa of the *P. moloch* group. Our biogeographic reconstruction indicates that the diversification of titi monkey genera initiated in the late Miocene via the fragmentation of a widespread ancestor distributed across the north-western Amazon (*Cheracebus*), Wet & Dry Savannas (*Plecturocebus*), and Atlantic Forest (*Callicebus*). Species-level diversification of taxa of the Amazonian clades, *Cheracebus* and the *P. moloch* group,

occurred in an explosive manner from the early Pleistocene to *c.* 0.5 Ma, and was largely characterised by sequential founder-events across rivers. These founder-events were sufficiently rare to allow divergence in isolation after dispersal, supporting the role of major Amazonian rivers as relatively strong barriers to gene flow. Our biogeographic reconstruction suggests that the evolution of the Pebas system in the western Amazon may have influenced the diversification and distribution of extant Callicebinae lineages, which were absent from the western Amazon until the recession of these wetlands and the establishment of suitable forest habitat starting in the Pleistocene ("Dynamic Young Amazon" model).

This study comprises one of the first reconstructions of the biogeography of titis based on explicit phylogenetic hypotheses and investigations into their evolutionary history in the context of Amazonian and South American biogeography. Although this research provides only a large-scale reconstruction of callicebine biogeography and should be interpreted with caution, it represents a critical starting point for future research investigating the biogeographic history of Callicebinae and of each of the major clades, and the processes promoting their diversification at a finer scale. Increased taxonomic sampling (particularly for the *P. donacophilus* group) and geographic delineations within each region (especially within the Atlantic Forest biome and Rondônia centre of endemism) will allow more in-depth investigation of titi monkey biogeography and testing of the biogeographic scenarios recovered in this study.

Chapter 4: Phylogenomics of titi monkeys (Callicebinae) using ddRADseq data with a focus on the *Plecturocebus moloch* group

4.1 Abstract

Recent molecular phylogenetic studies have revealed an increasingly comprehensive picture of the evolutionary history of titi monkeys (Callicebinae), one of the most species-rich groups of New World primates. Across these studies, however, conflicting phylogenetic relationships were recovered among some species of the *Plecturocebus moloch* group. Restriction-site associated DNA sequencing (RADseq) has become an important method of generating genome-wide molecular data for non-model organisms in order to address difficult phylogenetic questions. Here, reduced representation genome-wide data were generated for 12 Callicebinae species (45 specimens) using a double digest RADseq (ddRADseq) approach to infer phylogenetic relationships and to test for introgressive hybridisation. Phylogenetic analyses recovered a strongly supported topology for Callicebinae with species of the *P. moloch* group divided into three major clades; Aripuanã-Tapajós, Eastern, and Western clades. D-statistic tests detected a pattern of genetic introgression between *P. cinerascens* and *P. bernhardi* that was supported by independent evidence of mitochondrial introgression, and together, strongly suggested that *P. cinerascens* individuals from the left bank of upstream Rio Roosevelt, Rondônia, are admixed with *P. bernhardi*. We discuss putative sources of topological incongruence across loci and across previous studies, and the complicated nature of lineage diversification in the *P. moloch* group. Despite the strong resolution among the recovered species relationships, we advocate that strict interpretation of the phylogenetic results should conducted with caution until assessed using a multispecies coalescent-based model with the ddRADseq data.

4.2 Introduction

The evolutionary history of titi monkeys (Callicebinae) has only very recently received renewed scientific attention, with a number of molecular phylogenetic studies advancing our understanding of the relationships among the many species (Byrne et al., 2016; Carneiro et al., 2016; Hoyos et al., 2016;). Although genetic data are still lacking for many taxa, an increasingly comprehensive picture of the evolutionary history of this diverse group is being uncovered. All molecular evidence presented in these studies supports the monophyly and distinction of the three genera (*Cheracebus, Callicebus, Plecturocebus*), and molecular dating analyses suggest that these genera diverged in the late Miocene (Chapter 2; Hoyos et al., 2016). Species relationships recovered among taxa of the *Plecturocebus moloch* group, however, vary across the phylogenies inferred to date.

Phylogenetic conflict is most often associated with the relationship between *P. hoffmannsi*, as well as the sister species, *P. cinerascens* and *P. miltoni*, and the remaining members of the *P. moloch* group. The combined phylogenetic signal from nuclear and mitochondrial loci typically recovers *P. hoffmannsi* as the earliest diverging lineage and *P. cinerascens + P. miltoni* as sister to a clade containing *P. moloch, P. bernhardi* and closely related taxa (e.g., see Chapter 2, Figure 2.2 ; Carneiro et al., 2016). However, combined nuclear and mitochondrial data matrices have also recovered a monophyletic clade containing *P. hoffmannsi*, *P. cinerascens* and *P. miltoni* (Aripuanã-Tapajós clade, Chapter 3). Based on 20 nuclear loci, *P. cinerascens* and *P. miltoni* form independent early diverging lineages within the *P. moloch* group (Chapter 2, Figure 2.2), and when only mitochondrial loci are included, *P. hoffmannsi* is sister to the western Amazonian species (i.e., *P. cupreus* and closely related taxa; Chapter 2), or eastern Amazonian species with low support (represented by *P. moloch*; Hoyos et al., 2016). Difficulties in resolving the phylogenetic relationships among these species may be associated with an insufficient quantity of phylogenetically informative nuclear data, or topological incongruence could be a result of discordant phylogenetic signals owing stochastic processes such as incomplete lineage sorting and interspecific gene flow.

The generation of large genome-wide molecular datasets for non-model organisms has been revolutionised by the advent of next generation sequencing (NGS) technologies and genome complexity reduction methods (Davey et al., 2011;

Pukk et al., 2015). Restriction-site associated DNA sequencing (RADseq) is one of the most broadly used and cost-effective methods for generating reduced representation libraries, targeting DNA sequences adjacent to specific restriction enzyme recognition sites throughout the genome (Baird et al., 2008). The utility of RADseq data to address difficult phylogenetic problems was clearly demonstrated with Lake Victoria's cichlids, an adaptive radiation of over 500 species which diversified within the last 15,000 years (Wagner et al., 2013). Phylogenetic analyses employing up to 5.8 million base pairs of RADseq data recovered well-supported species relationships and reciprocal monophyly for the cichlid species included. Since then, RADseq data has been applied to questions in phylogenetics at interspecific level for many groups of organisms (e.g., Cruaud et al., 2014; Díaz-Arce et al., 2016; Hipp et al., 2014; Manthey et al., 2016;), and provided insight into species limits (e.g., Herrara & Shank, 2016; Pante et al., 2015; Razkin et al., 2016) and evidence for introgressive hybridisation (e.g., Chattopadhyay et al., 2016; Combosch & Vollmer, 2015; Eaton et al., 2015; Rheindt et al., 2014).

In the present study, double digest restriction-site-associated DNA sequencing (ddRADseq) (Peterson et al., 2012) was employed to sample unlinked genomic regions across 45 Callicebinae samples (12 species). Concatenated data matrices were generated by assembling the ddRADseq loci *de novo* using the pyRAD pipeline (Eaton & Ree, 2013; Eaton, 2014). ddRADseq data were then used to reconstruct phylogeny using maximum likelihood and Bayesian methods, generate a fossil-calibrated phylogeny to estimate divergence times, assess whether genetic structure among *P. moloch* group taxa corresponds with taxonomic classification using Bayesian clustering analyses (Pritchard et al., 2000), and test for introgression between *P. moloch* group lineages using the D-statistic test (Green et al., 2010). A multispecies coalescent model as implemented in StarBEAST2 (Ogilvie et al., 2016a) was employed to infer a species tree and estimate divergence times for Callicebinae from multiple gene trees based on the multi-locus Sanger sequenced dataset generated in Chapter 2. To assess putative mitochondrial introgression, a molecular dataset comprising two mitochondrial loci was assembled for some species of the *P. moloch*, and a maximum-likelihood tree was inferred. The primary aims of this study were to add to our understanding of the evolutionary history of Callicebinae using genome-wide data, increase the resolution of species relationships among the *P. moloch* group taxa, and assess potential factors influencing discordance across studies and across

loci (e.g., combined vs. mitochondrial vs. nuclear topologies). Parts of this study were inspired by the applications of the pyRAD pipeline and the investigation of introgression among the American live oaks by Eaton et al. (2015).

4.3 Methods

4.3.1 Taxon sampling

A total of 45 fresh tissue samples were collected from museum voucher specimens from the following Brazilian institutions: National Institute of Amazonian Research (INPA), Federal University of Pará (UFPA), Federal University of Rondônia (UNIR), Federal University of Amazonas (UFAM), the Goeldi Museum (MPEG), and the Rio de Janeiro Primate Center (CNRJ). The majority of these specimens were obtained in the context of an Amazonian-wide faunal inventory project (CNPq/SISBIOTA) carried out in accordance with the appropriate collection permits (IBAMA 483 license No. 005/2005 – CGFAU/LIC). This research adhered to the American Society of Primatologists' and American Society of Mammalogists' principles for the ethical treatment of primates, and Brazilian laws that govern primate research.

Twelve currently recognised Callicebinae species (45 specimens) were sampled from the *Plecturocebus moloch* group (eight), *Cheracebus* (three), and *Callicebus* (one) (Table 4.1). Individuals from diverse lineages within these species, most of which are identifiable in the phylogenetic trees in Chapter 2, were sampled; *P. cupreus* (clade A + B), *P. cinerascens* (clade A + B), *P. bernhardi* (clade A + UFRO + CCM), and *C. lugens* from the left and right banks of the Rio Negro (LN + RN), as well as a newly collected *C. lugens* specimen from the left bank of the Rio Japurá (LJ). Following Serrano-Villavicencio et al. (2017), *P. caligatus caligatus* and *P. caligatus dubius* are labelled as subspecies of *P. caligatus*, and samples from both taxa were included in the dataset. In addition, one sample included in the dataset represents *P.* cf. *moloch*, a new taxon (Boubli et al., in prep) of the *P. moloch* group from the Alta Floresta region of Mato Grosso, Brazil (see Chapter 2; Carneiro et al., 2016). All of the samples used in this study were from wild-caught specimens, nearly all of which are of known provenance (see Table 4.1).

Table 4.1 Genetic samples included in the ddRADseq analyses including ID, museum collection, and geographic origin.

Species (clade)	Sample ID	Col.	Geographic origin
Plecturocebus hoffmannsi	CTGAM248	UFAM	Cametá Community, L bank of the Rio Tapajós, Pará, Brazil
Plecturocebus hoffmannsi	CTGAM249	UFAM	Cametá Community, L bank of the Rio Tapajós, Pará, Brazil
Plecturocebus hoffmannsi	CTGAM290	UFAM	Cametá Community, L bank of the Rio Tapajós, Pará, Brazil
Plecturocebus miltoni	42991	MPEG	Novo Aripuanã, L bank of the Rio Aripuanã, Amazonas, Brazil
Plecturocebus miltoni	42992	MPEG	Novo Aripuanã, L bank of the Rio Aripuanã, Amazonas, Brazil
Plecturocebus cinerascens (B)	UFRO352	UNIR	Rondon II Dam, Pimenta Bueno, Rondônia, Brazil
Plecturocebus cinerascens (B)	UFRO354	UNIR	Rondon II Dam, Pimenta Bueno, Rondônia, Brazil
Plecturocebus cinerascens (B)	UFRO355	UNIR	Rondon II Dam, Pimenta Bueno, Rondônia, Brazil
Plecturocebus cinerascens (A)	UFRO499	UNIR	Cabixi, Rondônia, Brazil
Plecturocebus cinerascens (A)	WRS03	INPA	Apuí, Apuí, Amazonas, Brazil
Plecturocebus cinerascens (A)	WRS04	INPA	Apuí, Apuí, Amazonas, Brazil
Plecturocebus bernhardi (A)	42961	MPEG	São Francisco do Guaporé, Guaporé Biological Reserve, Rondônia, Brazil
Plecturocebus bernhardi (A)	42964	MPEG	São Francisco do Guaporé, Guaporé Biological Reserve, Rondônia, Brazil
Plecturocebus bernhardi	UFRO413	UNIR	Machadinho D'Oeste, Rondônia, Brazil
Plecturocebus bernhardi	CCM173	INPA	Rio Mariepauá, R bank tributary of the Rio Madeira, Amazonas, Brazil
Plecturocebus cf. *moloch*	RVR73	INPA	Novo Horizonte Community, Alta Floresta, Mato Grosso, Brazil
Plecturocebus moloch	CTGAM420	UFAM	Belterra, R bank of the Rio Tapajós, Pará, Brazil
Plecturocebus moloch	CTGAM421	UFAM	Belterra, R bank of the Rio Tapajós, Pará, Brazil
Plecturocebus brunneus	4505	UFPA	Samuel Dam, R bank of the Rio Jamari, Rondônia, Brazil
Plecturocebus brunneus	UFRO327	UNIR	Cujubim, Fazenda Manoa, Rondônia, Brazil
Plecturocebus brunneus	UFRO541	UNIR	Porto Velho, R bank of the Rio Madeira, Rondônia, Brazil
Plecturocebus cupreus (A)	AAM15	INPA	RESEX Catuá-Ipixuna, Lago do Ipixuna, Coari, Amazonas, Brazil
Plecturocebus cupreus (A)	JLP15920	INPA	Itamarati, L bank of the Rio Juruá, Amazonas, Brazil
Plecturocebus cupreus (A)	CTGAM210	UFAM	Rebio Abufari, Turiaçu, L bank of the Rio Purus, Amazonas, Brazil
Plecturocebus cupreus (B)	4987	UFPA	No location data
Plecturocebus cupreus (B)	4988	UFPA	No location data
Plecturocebus cupreus (B)	4990	UFPA	No location data
Plecturocebus c. caligatus	CTGAM181	UFAM	Tapauá, Igarapé do Jacinto, R bank of the Rio Purus, Amazonas, Brazil
Plecturocebus c. caligatus	CTGAM182	UFAM	Tapauá, Igarapé do Jacinto, R bank of the Rio Purus, Amazonas, Brazil
Plecturocebus c. caligatus	CCM248	INPA	No location data
Plecturocebus c. caligatus	MVR58	INPA	No location data

Table 4.1 cont'd Genetic samples included in the ddRADseq analyses including ID, museum collection, and geographic origin.

Species (clade)	Sample ID	Col.	Geographic origin
Plecturocebus c. dubius	UFRO403	UNIR	Porto Velho, L bank of the Rio Madeira, Rondônia, Brazil
Plecturocebus c. dubius	UFRO427	UNIR	L bank of the Rio Mucuim, Canutama, Amazonas, Brazil
Plecturocebus c. dubius	UFRO544	UNIR	Porto Velho, L bank of the Rio Madeira, Rondônia, Brazil
Plecturocebus c. dubius	2804	CNRJ	No location data
Callicebus personatus	2466	CNRJ	Aracruz, Espirito Santo, Brazil
Cheracebus lugens (LN)	JPB160	INPA	São Gabriel da Cachoeira, L bank of the Rio Negro, Amazonas, Brazil
Cheracebus lugens (LN)	JPB161	INPA	São Gabriel da Cachoeira, L bank of the Rio Negro, Amazonas, Brazil
Cheracebus lugens (RN)	JPB81	INPA	Igarapé Mandiquie, R bank of the Rio Negro, Amazonas, Brazil
Cheracebus lugens (LJ)	CTGAM733	UFAM	L bank of the Rio Japurá, Amazonas, Brazil
Cheracebus lucifer	CTGAM703	UFAM	R bank of the Rio Japurá, Amazonas, Brazil
Cheracebus lucifer	CTGAM726	UFAM	R bank of the Rio Japurá, Amazonas, Brazil
Cheracebus purinus	CTGAM154	UFAM	Rebio Abufari, Turiaçu, L bank of the Rio Purus, Amazonas, Brazil
Cheracebus purinus	CTGAM195	UFAM	Rebio Abufari, Turiaçu, L bank of the Rio Purus, Amazonas, Brazil
Cheracebus purinus	CTGAM209	UFAM	Rebio Abufari, Turiaçu, L bank of the Rio Purus, Amazonas, Brazil
Pithecia mittermeieri	CTGAM215	UFAM	L bank of the Rio Tapajós, Aveiro, Pará, Brazil
Chiropotes albinasus	CTGAM213	UFAM	L bank of the Rio Tapajós, Aveiro, Pará, Brazil
Chiropotes israelita	CTGAM5713	UFAM	Marari, Carauari, Amazonas, Brazil
Chiropotes sagalatus	CTGAM515	UFAM	Floresta Nacional de Saracá-Taquera, Pará, Brazil
Cacajao calvus	INPA5241	INPA	Tarauacá, Acre, Brazil
Cacajao melanocephalus	CTGAM0065	UFAM	Rio Daraá, L bank of the Rio Negro, Amazonas, Brazil
Cacajao ayresi	CTGAM5667	UFAM	L bank of the Rio Acará, Barcelos, Amazonas Brazil
Cacajao hosomi	CTGAM5698	UFAM	Serra do Imeri, Xamata, Amazonas, Brazil

Collection abbreviations: UFPA = Federal University of Pará; UFAM = Federal University of Amazonas; INPA = National Institute for Amazonian Research; UNIR = Federal University of Rondônia; MPEG = Goeldi Museum; CNRJ = Rio de Janeiro Primate Center

4.3.2 Library preparation and sequencing

Laboratory procedures were performed at the Evolution and Animal Genetics Laboratory (LEGAL) in the Federal University of Amazonas (UFAM), Manaus, Brazil. Total genomic DNA was extracted from blood and muscle tissues using the standard phenol-chloroform extraction protocol of Sambrook et al. (1989). The concentration of the extracted DNA was quantified using a Nanodrop 2000 spectrophotometer (Thermo Scientific), and the DNA was diluted to 50 ng/µL.

A reduced representation genomic library was constructed using the double digest restriction-site-associated DNA sequencing protocol (ddRADseq) (Peterson et al., 2012). The ddRADseq library preparation protocol was modified by Tomás Hrbek and LEGAL (UFAM) to allow simultaneous digestion and ligation and for sequencing on the Ion Torrent PGM (https://github.com/legalLab). ddRAD sequencing allows extreme genome complexity reduction by double digesting DNA with two restriction enzymes, a "common-cutter" and a "rare-cutter", followed by strict size selection of the library to standardize the size of the sequenced fragments. Here, the 8-base pair cutter SdaI (recognition site CCTGCA^GG), and the 4-base pair cutter Csp6I (recognition site G^TAC), were used as the rare and common cutting restriction enzymes, respectively. These enzymes create cohesive ends on the digested DNA fragments allowing ligation of complementary IonTorrent P and A adapters. The P1 adapter is common to all samples and binds to the sticky-end created by the SdaI restriction enzyme. The A adapter contains a unique molecular barcode for identification of individuals (to allow post-sequencing demultiplexing), it binds to the sticky-end created by the Csp6I restriction enzyme and is a divergent "Y" adapter to ensure that only fragments with one P1 and one A adapter are enriched. The SdaI and Csp6I restriction enzymes were chosen because we expected to observe up to 12,000 ddRADseq fragments in the range of 320 to 400 bp based on the *in-silico* digestion of complete primate genomes deposited in GenBank (Boubli et al., in prep). This information was used to optimize the number of individuals to be analysed in one run of the IonTorrent PGM.

DNA digestion and ligation were carried out simultaneously in the same reaction in a final volume of 50 uL: 4 μL of DNA (200ng) was digested with 0.1 μL (1 U) of restriction enzymes SdaI and Csp6I (Thermo Scientific), and ligated with 2 µL of the IonTorrent adapters, P1 (0.1 µM) and A (5 µM; individual barcode adapter), 0.5 µL of T4 DNA ligase (5 U), 0.5 µL of ATPs (5 mmol), and 5 µL of 10X Tango Buffer. The digestion and ligation step was carried out at 37 °C for 180 minutes, followed by heat-inactivation at 68 °C for 15 minutes. A PCR test was then carried out to check the performance of the digestion. The PCR test was carried out in a final volume of 15 µL: 1.5 µL of 10X NH_4SO_4 buffer (Thermo Scientific), 1.5 µL of primer P1 (2 mM), 1.5 µL of primer A-amp (2 mM), 1.2 µL of dNTPs (10 mM), 1.2 µL MgCl2 (25 mM), 0.35 µL of DNA *Taq* polymerase (1 U), and 1 µL of the digested

adapter-ligated DNA. PCR conditions were as follows: 2 minutes at 94 °C; followed by 35 cycles of 15 s at 94 °C, 35 s at 55 °C, and 90 s at 68 °C. The PCR tests were checked by electrophoresis on 1% agarose gel stained with GelRed^{TM} (Biotium, Inc.).

The next step involves enrichment of the digested adapter-ligated DNA fragments, which were amplified in five separate PCRs for each sample. The enrichment PCRs were carried out at final volume of 25 µL: 2.0 µL of MgCl2 (25 mM), 2.0 µL of dNTPs (10 mM), 2.5 µL of 10X NH$_4$SO$_4$ buffer (Thermo Scientific), 2.5 µL of primer P1 (2 mM), 2.5 µL of primer A-amp (2 mM), 0.1 µL of KlenTaq (0.5 U KlenTaq DNA Polymerase Technology), and 1 µL of the digested adapter-ligated DNA. Enrichment PCR conditions were as follows: 1 minute at 68°C; followed by 18 cycles of 10 s at 93°C, 35 s at 52°C, and 90 s at 68°C; and a final cycle of 7 minutes at 68°C. Each of the five enriched PCR products for each sample were then combined in a single tube to a total volume of 100 µL (20 µL each), and purified using 0.8-fold volume of solid-phase reversible immobilization (SPRI) bead solution (AMPure). The concentration of the enriched DNA samples were then measured using a Qubit 2.0 Fluorometer (Invitrogen) and all samples were pooled together equimolarly in a single tube. DNA fragments in the range of 320 to 400 bp were selected using the Pippin Prep (2% agarose cartridge; Sage Science), owing to the ability of the IonTorrent PGM to sequence fragments up to 400 bp. The ddRADseq library was purified again using AMPure beads (0.7-fold volume) and sequenced on an Ion Torrent PGM (Life Technologies) using the 400-bp PGI 318 Ion PGM sequencing kit following manufacturers' recommendations.

4.3.3 ddRADseq assembly

Raw sequence data were demultiplexed, quality filtered and assembled into *de novo* loci using the pyRAD v3.0.63 pipeline (Eaton & Ree, 2013; Eaton, 2014). pyRAD assembles ddRADseq data into clusters of similar sequences, which are considered different loci (i.e. orthologs), without the use of a reference genome. The pyRAD pipeline was used as it was developed to search for homologies among divergent samples/taxa and allows the presence of insertions and deletions (indel variation) owing to use of global alignment clustering methods. It is also suitable for processing ddRADseq data generated by the Ion Torrent PGM, which has an increased indel error rate relative to other sequencing platforms (Laehnemann et al., 2016) and otherwise high-quality sequence data can be discarded when using pipelines where

indels are not considered [e.g., Stacks (Catchen et al., 2011), which was developed for population-level analyses].

ddRADseq data from eight Pitheciinae species (eight specimens) that were sampled by Bertuol et al. (in prep.) to investigate species relationships among black uakaris were added to the dataset after demultiplexing. These specimens were sequenced at the same laboratory (LEGAL, UFAM) using the same protocol for ddRADseq library preparation as described above (see section 4.3.2 Library preparation and sequencing) with one exception; the purification step was carried out with the GFX™ PCR DNA and Gel Band Purification Kit (GE Healthcare Life Sciences, USA) rather than SPRI bead purification.

Sequence data from the 53 individuals were separated by pyRAD using the sample-specific molecular barcodes that were attached during library preparation. IonTorrent adapters, individual barcodes and restriction sites were removed with the filter setting "1". Bases with a Phred quality score of less than 15 were turned into undetermined sites (Ns), and reads with more than ten undetermined sites (Ns) were discarded. This minimum quality score was chosen to account for the slight systematic underestimation of base call accuracy (Phred score) by the IonTorrent PGM (e.g., Bragg et al., 2013). The maximum number of ten undetermined sites (Ns) in a read was set in consideration of the 300–400 bp length of the sequenced fragments. The optimal value for the clustering threshold parameter (the minimum similarity required to consider sequences as orthologs) depends on various study-specific factors such as the amount of polymorphism and sequencing error (Catchen et al., 2013). In order to assess the impact on the assembled data matrices and downstream phylogenetic inference, two different clustering thresholds were tested for within- and across-sample clustering (85% and 92%). Quality-filtered reads were clustered within samples at 85% or at 92% sequence similarity using the VSEARCH algorithm (Rognes et al., 2016) and then aligned using MUSCLE (Edgar, 2004). Heterozygosity and error rate were estimated for each individual from the aligned clusters and used to make base calls in the consensus sequence for each within-sample cluster. Clusters were retained if the minimum depth of coverage was at least 5X, and if the consensus sequence contained no more than six heterozygous sites, ten undetermined sites (Ns), and two alleles after error correction (as all taxa in the study are diploid). Consensus sequences were then clustered across samples at 85% or at 92% sequence similarity using the VSEARCH algorithm and aligned using

MUSCLE. In the final filtering step, loci were discarded as putative paralogs if more than five individuals shared a heterozygous site, as excessive shared heterozygosity across species may represent fixed differences among paralogs rather than shared heterozygosity within orthologs.

Steps one to five in pyRAD (the generation of quality-filtered consensus sequences for within-sample clusters) were performed once for each clustering threshold (85% and 92%) for all sequenced individuals. Final datasets with different combinations of samples/taxa for downstream analyses were then constructed by running steps six and seven of pyRAD (clustering of consensus sequences across samples at 85% or 92% and subsequent filtering) with only the target individuals included. A locus was only represented in a dataset if it was recovered for at least ~50% of the ingroup individuals in that dataset. The minimum number of individuals required at a locus for each dataset is listed in Table 4.2.

Table 4.2 List of the data matrices assembled in pyRAD including clustering threshold and minimum coverage parameter settings, sample information and dataset usage.

Dataset	Clust. thresh.	Description	No. ingroup	Min. cov.	Outgroup addon	Usage
A85	85%	All samples	45	22	8 Pitheciidae	RAxML, MrBayes
A92	92%	All samples	45	22	8 Pitheciidae	RAxML, MrBayes
B85	85%	One sample per taxon	19	9	8 Pitheciidae	BEAST, RAxML, MrBayes
P85	85%	*Plecturocebus*	33	16	--	STRUCTURE
Pi85	85%	Aripuanã-Tapajós clade	11	5	--	STRUCTURE
Pii85	85%	Eastern clade	7	4	--	STRUCTURE
Piii85	85%	Western clade	16	8	--	STRUCTURE
PD85	85%	*Plecturocebus + Cheracebus*	32	16	9 *Cheracebus*	D-statistics, RAxML, MrBayes

A total of eight ddRADseq datasets were assembled (Table 4.2). Initially, data matrices including all sequenced individuals were assembled using both 85% and 92% clustering thresholds (referred to as the A85 and A92 datasets). All further datasets used a clustering threshold of 85%. A dataset including one well-sequenced individual from each Callicebinae taxon or lineage, as well as all Pitheciinae outgroup taxa, was constructed (B85 datasets) primarily for use in divergence dating analyses (see section 4.3.4). Four datasets containing different combinations of individuals

from the *Plecturocebus moloch* group were assembled for downstream clustering analyses with STRUCTURE (Pritchard et al., 2000) (see section 4.3.5): a) all taxa from the *Plecturocebus moloch* group (P85); b) Aripuanã-Tapajós clade taxa only (Pi85); c) Eastern clade taxa only (Pii85); and d) Western clade taxa only (Piii85). A dataset was also constructed for the investigation of introgression between species of the *P. moloch* group using D-statistics (see section 4.3.6) that included all *P. moloch* group individuals as ingroup taxa (except three poorly sequenced individuals that were excluded) and all *Cheracebus* individuals as outgroups (PD85).

In order to assess the distribution and source of missing data between samples, a heatmap of shared ddRADseq loci between samples across the two datasets including all individuals was generated, and a Mantel correlation test was performed with 9999 permutations. The mantel test measured Spearman's rank correlation between the Jaccard's distance of the proportion of shared loci between samples, pairwise phylogenetic distance, and the number of raw input reads (Eaton et al., 2015). Pairwise genetic distances between individuals were calculated (K80 model) using the R package "adegenet" (Jombart, 2008).

4.3.4 Phylogenetic analysis of ddRADseq data

Phylogenetic inference was conducted using maximum-likelihood (ML) and Bayesian methods for four of the concatenated ddRADseq loci data matrices assembled in pyRAD: A85 and A92 (all individuals); B85 (select individuals from each taxon); and PD85 (only individuals of the *P. moloch* group and *Cheracebus*). Maximum-likelihood trees were inferred using RAxML v. 8.2.10 (Stamatakis, 2006, 2014) with the GTR + G (gamma) substitution model and 1,000 bootstrap replicates integrated with 200 searches for the optimal tree. Bayesian analyses were performed using MrBayes 3.2.3 (Ronquist et al., 2012) with the GTR + G substitution model. MCMC (Markov Chain Monte Carlo) convergence was checked after two independent four-chain runs of 2 million generations for each Bayesian inference. Convergence was assessed by examining LnL, the average standard deviation of the split frequencies between the two simultaneous runs (< 0.01), and the Potential Scale Reduction Factor (PSRF) diagnostic in MrBayes, after a burn-in of 10%.

Phylogeny and diversification times were jointly estimated for the B85 dataset under an uncorrelated lognormal relaxed clock model in the program BEAST v. 1.8.2 (Drummond et al., 2012). A Yule speciation process and the GTR + G substitution

model were used, and the ucld.mean prior was set to a gamma distribution (shape = 0.001; scale = 1000). Two replicate runs of 100 million MCMC generations, sampling every 10,000 generations, were conducted. The sampling distributions of each run were visualized using Tracer v. 1.6 to evaluate convergence and to verify that the effective sample size was > 200 for all parameters after a burn-in of 10%. Independent runs were combined with the first 1000 (10%) samples of the posterior distribution discarded as burn-in using LogCombiner v. 1.8.2 and the maximum credibility tree was generated in TreeAnnotator v. 1.8.2. To obtain the posterior distribution of the estimated divergence times, one calibration point on the root node (Callicebinae/Pitheciinae) was implemented with a lognormal distribution to set hard minimum and soft maximum bounds. A minimum age of 15.7 Ma was used based on the fossil *Proteropithecia* Kay et al., 1998, (Kay et al., 1998, 1999) and a soft maximum bound was set at 26 Ma based on the fossil *Branisella boliviana* Hoffstetter, 1969, from the Deseadan fauna of La Salla (McFadden, 1990). The standard deviation (= 0.5) and mean (= 1.51) were set such that 95% of the prior distribution falls before the maximum age to create the soft maximum bound. All phylogenetic analyses (using RAxML, MrBayes, and BEAST) were run on the CIPRES Science Gateway v 3.3 server (Miller et al., 2010).

4.3.5 Bayesian clustering analyses

A Bayesian model-based clustering method was applied to investigate genetic structure among members of the *P. moloch* group, as implemented in the software STRUCTURE v. 2.3.2 (Falush et al., 2003; Pritchard et al., 2000). STRUCTURE allocates individuals into clusters (K) as to minimize deviations from Hardy–Weinberg equilibrium and maximise linkage equilibrium, independent of population information. Individuals are assigned probabilistically into each of the K clusters based on their membership coefficient (Q value), and joint membership in two or more clusters may be an indication of admixture. STRUCTURE analyses were initially performed using the P85 dataset (all taxa of the *P. moloch* group), and subsequently using the subsampled Pi85, Pii85, and Piii85 datasets to evaluate the existence of finer structure which may be obscured by major axes of structure (e.g., among the major clades) in the overall P85 dataset. One SNP (single nucleotide polymorphism) per locus was selected at random and data matrices of unlinked SNPs (assuming that SNPs of different ddRADseq loci are effectively unlinked) in coded

SNP STRUCTURE-format files were generated for each dataset by pyRAD. All STRUCTURE analyses were conducted with the admixture model, correlated allele frequencies, and no putative origins specified for individuals. For the P85 dataset including all *P. moloch* group taxa, six runs at each value of *K* (ranging from one to sixteen) were performed with a burn-in of 200,000 steps and MCMC length of 800,000 steps. Between eight and thirteen clusters were predicted *a priori* since nine described taxa of the *P. moloch* group are included in the dataset, as well as individuals from divergent lineages for at least four species (*P. moloch, P. bernhardi, P. cinerascens, P. cupreus*). For the subsampled datasets (Pi85, Pii85, and Piii85), five runs at each value of *K* (ranging from one to six or seven) were performed with a burn-in of 100,000 steps and MCMC length of 500,000. Between three and five clusters were suspected *a priori* for each of the subsampled *P. moloch* group datasets.

The model choice criterion implemented in STRUCTURE to infer the most probable number of clusters is an estimate of the posterior probability of the data for a given *K*; the number of clusters that provides the highest likelihood, LnP(D), across runs is considered the most likely *K* value (Pritchard et al., 2000). As the primary goal of these analyses was to detect genetic structure that was suspected *a priori* based on taxonomic classification and phylogenetic results, the optimum number of clusters for each dataset was primarily assessed by LnP(D). If several *K* values had similar LnP(D) scores, the assignment of the additional clusters was evaluated to check if they were informative or assigned equally to the putative populations (Pritchard & Wen, 2004). The ad hoc statistic ΔK (Evanno et al., 2005) implemented in the program STRUCTUREHARVESTER (Earl & vonHoldt, 2012) was also considered to attempt to detect the most appropriate number of clusters, in particular for the subsampled datasets. The Evanno method chooses the optimum number of clusters based on the second order rate of change in the log probability of data between successive *K* values, however, when there is strong hierarchical structure it often returns only the top level of stratification (Evanno et al., 2005). For the full P85 dataset, it is likely that ΔK will strongly detect the major axes of structure representing deeper divergences within the *P. moloch* group, possibly obscuring finer structure within or even among species, and resulting in a smaller value (e.g., three) than expected (eight or above) for the optimum number of clusters using this method.

STRUCTUREHARVESTER (Earl & vonHoldt, 2012) was used to examine LnP(D) and ΔK for each possible number of clusters (*K*) for each dataset. The

program CLUMPP (Jakobsson & Rosenberg 2007) was used to combine and average individual assignments probabilities across all replicates, and individual Q values were plotted and visualised using DISTRUCT 1.1 (Rosenberg, 2004).

4.3.6 Introgression analyses

The D-statistic (Durand et al., 2011; Green et al., 2010) was used to evaluate whether ancestral admixture has occurred between species of the *P. moloch* group. This test was first employed to assess introgression between *Homo sapiens sapiens* and Neanderthals (Green et al., 2010) based on whole genome data, and has since been applied to non-model organisms using partial representation genome-wide data (e.g., Eaton & Ree, 2013; Eaton et al., 2015). Based on the assumption of a true four-taxon tree with the topology (((P1, P2) P3) Outgroup), the four-taxon D-statistic evaluates the occurrence of two biallelic site patterns, ABBA and BABA (A = ancestral allele, B = derived allele), which are incongruent with this species tree. ABBA and BABA represent sites in which an allele is derived in the P3 lineage, and in either of the P2 (ABBA) or the P1 (BABA) sister lineages, but not in both (i.e., BBBA, which agrees with the species tree). These discordant allele patterns are expected to occur at equal frequencies (D = 0) if they arise through stochastic processes such as incomplete lineage sorting, whereas one discordant pattern is expected to occur more frequently than the other if introgression has occurred between P3 and either P1 or P2 (measured by the significant deviation of D from 0).

For the four-taxon D-statistic test, taxa were assigned at species-level and selected as follows: P1 and P2 were set as species of the same *P. moloch* group clade (Western, Eastern, or Aripuanã-Tapajós); P3 was set as each of the remaining *P. moloch* group species, providing the hypothesised species tree was not violated (e.g., when P1 and P2 were set as *P. hoffmannsi* and *P. miltoni*, P3 could not be *P. cinerascens*); and all *Cheracebus* individuals (*n* = 9) were used as the outgroup. The hypothesised species tree followed the topology of the ddRADseq phylogenetic trees, however, given the absence of strong support across the phylogenetic analyses the only assumption made species relationships among the Western clade taxa was that *P. brunneus* and *P. cupreus* couldn't form a clade to the exclusion of *P. caligatus*. D was calculated over all possible combinations of species assigned to P1, P2 and P3 that satisfied these conditions (tests 1 – 39). A select set of four-taxon D-statistic tests were performed in which the two distinct *P. cinerascens* lineages (clade A and clade

B) were differentiated (tests 40 – 44), as clade B individuals were suspected to have mixed ancestry (see section 4.3.7). A final set of tests were performed in which the two *P. cupreus* clades (A and B) and the *P. caligatus* subspecies (*P. c. dubius* and *P. c. caligatus*) were defined separately to assess introgression between these lineages.

All ingroup taxa contain multiple individuals in the four-taxon D-statistic tests and D was calculated separately for all possible combinations of different individuals for P1, P2 and P3 in each of the tests. All samples were pooled for the outgroup taxon and a locus can be used if the three ingroup taxa and at least one outgroup individual are represented. Input files for the D-statistic tests were generated with modified python scripts and all tests were performed in pyRAD v. 3.0.63 (Eaton, 2014; Eaton et al., 2015) using the ddRADseq loci generated for the PD85 dataset (*P. moloch* group and *Cheracebus* taxa). The standard deviation of the D-statistic was calculated from 1000 bootstrap replicates. The measured D was converted to a Z-score (the number of standard deviations from zero) and significance was assessed from a two-tailed *p*-value using $\alpha = 0.01$ as a conservative cutoff for significance after Holm-Bonferoni correction for multiple testing. A significant Z-score indicates that gene flow may have occurred between P3 and P1 or P2.

The partitioned D-statistic (Eaton & Ree, 2013) is an extension to the four-taxon D-statistic test (Durand et al., 2011) based on a five-taxon species tree, $(((P1, P2), (P3_1, P3_2))$ Outgroup), with two lineages from the P3 clade ($P3_1$, $P3_2$). The partitioned D-statistic test evaluates the occurrence of derived alleles (B) which are present in P2 or P1 (but not both) and in $P3_1$ (D_1), or $P3_2$ (D_2), or both P3 sublineages (D_{12}). Three D-statistics are estimated in this test, one for each pair of discordant biallelic site patterns: D_{12} for ABBBA/BABBA; D_1 for ABBAA/BABAA; and D_2 for ABABA/BAABA. In the partitioned D-statistic, P3 is defined as the donor lineage and the D_{12} statistic indicates which direction gene flow occurred; if introgression occurred from P3 into P2 or P1, then both P3 lineages ($P3_1$ and $P3_2$) will share derived alleles with the recipient taxon, regardless of which P3 lineage was involved, because it is likely that some derived alleles arose in the ancestor to the P3 clade. A significant Z-score for D_{12} indicates that introgression occurred from the P3 clade into P2 or P1, and a significant Z-score for D_1 or D_2 signals whether the $P3_1$ or $P3_2$ lineage (or both) was involved. A significant Z-score for D_1 or D_2, but not D_{12}, suggests that gene flow occurred in the opposite direction, i.e., from P1 or P2 into P3. Thus, partitioned D-statistic tests assess the direction of gene flow and which lineages were

involved, and they are particularly useful when considering older species trees, interspecific hybridisation, and diverse species that include distinct intraspecific lineages which may have admixed independently (Eaton et al., 2015).

For the four-taxon D-statistic tests with a significant Z-score, partitioned D-statistic tests were performed in order to infer the direction of gene flow and assess which intraspecific clades were involved for species represented by multiple divergent lineages. Partitioned D-statistics were employed to evaluate two scenarios, admixture involving *P. cinerascens* and *P. bernhardi* + *P. moloch* (I), and involving *P. cinerascens* and *P. c. caligatus* (II). In the latter case (II), *P. cinerascens* clade A and another taxon from the Aripuanã-Tapajós clade (*P. cinerascens* clade B*, P. miltoni,* or *P. hoffmannsi*) were selected for P1 and P2, while P3$_1$ and P3$_2$ were set as *P. c. caligatus* and another taxon from the Western clade (*P. c. dubius, P. brunneus,* or *P. cupreus*), based on the combination of individuals with a significant Z-score in the four-taxon test. Partitioned D-statistic tests were performed with all possible combinations of these taxa (tests 60 – 68).

For scenario I, four-taxon tests show evidence of admixture involving *P. cinerascens* and both *P. bernhardi* and *P. moloch*, and thus it was of interest to test whether these taxa hybridised independently and also to assess which lineages were involved (e.g., *P. cinerascens* clade A or clade B*; P. bernhardi* clade A or UFRO*; P. moloch* or *P*. cf. *moloch*). For these tests, P1 and P2 were set as one of the *P. cinerascens* lineages and another member of the Aripuanã-Tapajós clade, while taxa from the Eastern clade (*P. moloch, P*. cf *moloch, P. bernhardi* clade A, or *P. bernhardi* UFRO) were selected as P3$_1$ and P3$_2$. Partitioned D-statistic tests were performed with all possible combinations of these taxa (tests 45 – 59).

As in the four-taxon tests, the outgroup taxa is represented by the pooled *Cheracebus* samples for all partitioned tests, however, the ingroup taxa (terminals P1, P2, P3$_1$ and P3$_2$) were defined as a single individual per taxon to reduce redundancy and divergent intraspecific lineages of interest were also defined separately. There are six incongruent allele patterns measured in the five-taxon test and each site must be represented across the five taxa included (versus two patterns across four taxa for the four-taxon D-statistic). Because fewer sites will meet these conditions than for the four-taxon test, the individual representing each taxon was chosen based on the number of loci recovered in the ddRADseq dataset to maximise the statistical power of the partitioned D-statistic test. The standard deviation of the D-statistics were

calculated from 1000 bootstrap replicates. The measured D (D_1, D_2, D_{12}) were converted to Z-scores and significance was assessed from a two-tailed *p*-value using α = 0.01 as a conservative cutoff for significance (Z-score > 2.55).

4.3.7 Mitochondrial phylogeny

A recent study investigating the evolutionary history of Callicebinae taxa (Carneiro et al., 2016) classified one of the *P. cinerascens* individuals included in this ddRADseq dataset (UFRO354) as a *P. bernhardi*. This specimen along with two other individuals (UFRO352, UFRO355) collected from the same locality form a sister clade (labeled clade B) to other *P. cinerascens* samples in the ddRADseq phylogenies. Photos obtained of the UFRO354 specimen confirms that it phenotypically resembles *P. cinerascens*, however, these specimens were collected on the left bank of upstream Rio Roosevelt at the Rondon II dam, near Pimenta Bueno, Rondônia, which is outside the known geographic distribution for *P. cinerascens*, within an interfluvium where *P. bernhardi* is found. The phylogeny reconstructed in Carneiro et al. (2016) is based on two mitochondrial loci (COI and CYTB) as well as *alu* markers, and it is possible that the phylogenetic signal from the mitochondrial sequences overwhelmed the information contained in the nuclear sequences owing to the higher mutation rate and significantly greater number of informative sites. In this case, the UFRO354 individual may truly have a *P. bernhardi* mitochondrial genome, which is especially interesting in light of the nuclear (ddRADseq) and phenotypic evidence that UFRO354 is more closely related to *P. cinerascens*.

To independently verify that UFRO354 has a *P. bernhardi* mitochondrial genome, and also because the molecular data used by Carneiro et al. (2016) were not published, new mitochondrial sequences (for COI and CYTB) were generated for this specimen. To infer phylogeny and compare sequence identity, sequence data were also generated for three other *P. cinerascens* and two *P. bernhardi* individuals. These six tissue samples were collected from museum voucher specimens (Table 4.3) DNA was extracted from muscle tissues using the Qiagen DNeasy Blood & Tissue Kit according to manufacturer's protocol. Eleven new sequences were generated for COI (5), and CYTB (6) (see primer information for COI and CYTB in Table 2.3, Chapter 2). The PCR reactions were carried out in a total volume of 50 μL, containing approximately 30 ng of genomic DNA, 4 μL of dNTPs (200μM each), 5 μL 10X PCR buffer (100 mM Tris-HCL, 500 mM KCL, 15 mM Mg2+), 1 μL of each forward and

reverse primer (0.2 µM), and 0.25 µL of TaKaRa *Taq* DNA polymerase (1 Unit). The amplification cycles were carried out under the following conditions; initial denaturation at 95 °C for 5 min; followed by 35 cycles of denaturing at 94 °C for 1 min, primer annealing for 1 min at 45°C (COI) or 60°C (CYTB), and extension at 72 °C for 1 min; a final extension was carried out at 72 °C for 5 min. PCR products were analysed on 1.5 % agarose gels and then Sanger sequenced commercially. Consensus sequences were generated from forward and reverse reads using Geneious R7.1 (Biomatters).

The complete cytochrome b (CYTB) locus (1140 bp) and 660 bp of the cytochrome c oxidase subunit I (COI) locus were aligned using the MUSCLE algorithm in Geneious R7.1 and subsequently concatenated. Twelve further individuals sampled in the multi-locus dataset in Chapter 2 were also added, including some *P. miltoni, P. moloch* and *P.* cf. *moloch* specimens (see Table 4.3). Sequences for COI and CYTB were extracted from a whole mitochondrial genome sequence that was retrieved from GenBank for one *P. donacophilus* specimen in order to root the tree (accession number = FJ785423). All specimens were represented at both loci except *P. cinerascens* FR31 and *P. bernhardi* 42961, which are missing data for COI. A maximum-likelihood tree was reconstructed using RAxML v. 8.1 (Stamatakis, 2014) with the GTR + G (gamma) substitution model. Node support was estimated using the rapid-bootstrapping algorithm (−f a -x option) for 1000 non-parametric bootstrap replicates (Stamatakis et al., 2008).

Table 4.3 Genetic samples included in the mtDNA phylogenetic analysis including ID, museum collection, data source, and geographic origin.

Species (clade)	Sample ID	Col.	Data	Geographic origin
Plecturocebus miltoni	42991	MPEG	Ch. 2	Novo Aripuanã, L bank of the Rio Aripuanã, Amazonas, Brazil
Plecturocebus miltoni	42992	MPEG	Ch. 2	Novo Aripuanã, L bank of the Rio Aripuanã, Amazonas, Brazil
Plecturocebus cinerascens (B)	UFRO352	UNIR	Ch. 2	Rondon II Dam, Pimenta Bueno, Rondônia, Brazil
Plecturocebus cinerascens (B)	UFRO354	UNIR	New	Rondon II Dam, Pimenta Bueno, Rondônia, Brazil
Plecturocebus cinerascens (B)	UFRO355	UNIR	Ch. 2	Rondon II Dam, Pimenta Bueno, Rondônia, Brazil
Plecturocebus cinerascens (A)	UFRO499	UNIR	Ch. 2	Cabixi, Rondônia, Brazil
Plecturocebus cinerascens (A)	FR31	INPA	New	Novo Aripuanã, R bank of the Rio Aripuanã, Amazonas, Brazil
Plecturocebus cinerascens (A)	FR50	INPA	New	Sucunduri, Apuí, Amazonas, Brazil
Plecturocebus cinerascens (A)	FR123	INPA	New	Novo Aripuanã, R bank of the Rio Aripuanã, Amazonas, Brazil
Plecturocebus bernhardi (A)	42960	MPEG	Ch. 2	São Francisco do Guaporé, Guaporé Biological Reserve, Rondônia, Brazil
Plecturocebus bernhardi (A)	42961	MPEG	Ch. 2	São Francisco do Guaporé, Guaporé Biological Reserve, Rondônia, Brazil
Plecturocebus bernhardi (A)	42964	MPEG	Ch. 2	São Francisco do Guaporé, Guaporé Biological Reserve, Rondônia, Brazil
Plecturocebus bernhardi	UFRO413	UNIR	Ch. 2	Machadinho D'Oeste, Rondônia, Brazil
Plecturocebus bernhardi	FR26	INPA	New	Novo Aripuanã, L bank of the Rio Aripuanã, Amazonas, Brazil
Plecturocebus bernhardi	CCM173	INPA	New	Rio Mariepauá, R bank tributary of the Rio Madeira, Amazonas, Brazil
Plecturocebus cf. *moloch*	RVR73	INPA	Ch. 2	Novo Horizonte Community, Alta Floresta, Mato Grosso, Brazil
Plecturocebus moloch	CTGAM420	UFAM	Ch. 2	Belterra, R bank of the Rio Tapajós, Pará, Brazil
Plecturocebus moloch	MCB63	UFPA	Ch. 2	Senador José Porfírio, R bank of the Rio Xingu, Pará, Brazil

Collection abbreviations: UFPA = Federal University of Pará; UFAM = Federal University of Amazonas; INPA = National Institute for Amazonian Research; UNIR = Federal University of Rondônia; MPEG = Goeldi Museum.

4.3.8 Coalescent-based species tree analysis: StarBEAST2

A multispecies coalescent model as implemented in StarBEAST2 (Ogilvie et al., 2016a) was employed to infer a species tree for Callicebinae from multiple gene trees based on the multi-locus Sanger sequenced dataset generated in Chapter 2. One of the goals of this analysis was to assess whether the species relationships inferred using

the multi-locus sequences under a multi-species coalescent model are congruent with the results from the ddRADseq phylogenetic analyses.

StarBEAST2 is a newly developed version of *BEAST (Heled & Drummond, 2010) with several improvements such as novel MCMC operators to allow more accurate inference of species trees, divergence times, and substitution rates (Ogilvie et al., 2016a). As information for each locus is used to infer the species tree, it is important to minimise missing data as it may affect the estimation of the gene and species trees (e.g., Townsend et al., 2011). Thus, thirteen loci (11 nuclear and 2 mitochondrial; Table 4.4) were chosen that were represented by at least one individual for all Callicebinae taxa included, with two exceptions (*C. coimbrai* was missing mitochondrial data and the RAG2 locus; and *C. nigrifrons* was not represented at the ZFX locus). Chapter 2 contains information on each of the loci (Table 2.3). Between one and three individuals were included for each Callicebinae taxon, depending on the number available from the original dataset and sequencing coverage, and three Pitheciinae taxa were included as outgroups. Most of the sequences included were generated for Chapter 2, although some were retrieved from GenBank (see Table 4.4 and accession numbers in Table 4.5).

Each locus was aligned using the MUSCLE algorithm in Geneious R7.1 (Biomatters) and substitution models were set according to the model selected for each alignment using Bayesian information criterion (BIC) in jModelTest v 2.1.6 (Darriba et al., 2012) as follows; HKY+G for CYTB and ABCA1; HKY+I for FES; K80 for DENND5A; K80+I for RAG1 and RAG2; TRN+G for COI; and HKY for the other six loci. The multispecies coalescent model was applied as implemented in StarBEAST2, an extension of BEAST v 2.4.4 (Bouckaert et al., 2014). Callicebinae individuals were grouped into 16 terminals representing 14 described species (*P. c. caligatus / P. c. dubius* and *P. moloch / P.* cf. *moloch* were defined separately), and Pitheciinae taxa were set as an outgroup terminal. The clock, site and tree models were unlinked across loci, except the two mitochondrial loci which shared a gene tree. All clock models were set as lognormal uncorrelated relaxed clocks and a Yule speciation process was applied. The Analytical Population Size Integration model was set as the population size model. A lognormal distribution (default settings) was set for the species-population mean prior.

Table 4.4 Locus coverage for the samples included in the StarBEAST2 analysis. Colour indicates source of the data: grey = Chapter 2; orange = Perelman et al. (2011); and purple = other sequences from GenBank.

Species	Sample ID	No. loci (13)	COI	CYTB	ABCA1	DENND5A	DMRT1	ERC2	FES	FOXP1	MAPKAP1	NPAS3.2	RAG1	RAG2	ZFX
P. hoffmannsi	CTGAM248	11	G	G	G	G		G	G	G	G	G		G	G
P. hoffmannsi	CTGAM290	13	G	G	G	G	G	G	G	G	G	G	G	G	G
P. miltoni	42991	13	G	G	G	G	G	G	G	G	G	G	G	G	G
P. miltoni	42992	13	G	G	G	G	G	G	G	G	G	G	G	G	G
P. miltoni	42993	12	G	G	G	G	G	G	G	G	G	G	G	G	
P. cinerascens	UFRO355	13	G	G	G	G	G	G	G	G	G	G	G	G	G
P. cinerascens	UFRO499	12	G	G	G	G		G	G	G	G	G	G	G	G
P. bernhardi	UFRO413	12	G	G	G	G		G	G	G	G	G	G	G	G
P. bernhardi	42960	12	G	G	G	G	G	G	G	G	G	G	G	G	
P. bernhardi	42964	13	G	G	G	G	G	G	G	G	G	G	G	G	G
P. moloch	MCB64	13	G	G	G	G	G	G	G	G	G	G	G	G	G
P. moloch	MCB79	12	G	G	G	G	G	G	G	G	G	G	G	G	
P. moloch	CTGAM420	11	G	G	G			G	G	G	G	G	G	G	G
P. cf. moloch	RVR22	13	G	G	G	G	G	G	G	G	G	G	G	G	G
P. cf. moloch	RVR68	12	G	G		G	G	G	G	G	G	G	G	G	G
P. cf. moloch	RVR73	13	G	G	G	G	G	G	G	G	G	G	G	G	G
P. brunneus	4009	13	G	G	G	G	G	G	G	G	G	G	G	G	G
P. brunneus	4346	13	G	G	G	G	G	G	G	G	G	G	G	G	G
P. brunneus	4505	13	G	G	G	G	G	G	G	G	G	G	G	G	G
P. cupreus	AAM15	10	G	G	G	G		G	G	G	G	G		G	
P. cupreus	CTGAM210	12	G	G	G	G	G	G	G	G	G	G		G	G
P. cupreus	4988	13	G	G	G	G	G	G	G	G	G	G	G	G	G
P. cupreus	4990	13	G	G	G	G	G	G	G	G	G	G	G	G	G
P. cupreus	4993	12	G	G	G	G	G	G	G	G	G	G	G		G
P. c. caligatus	CTGAM181	13	G	G	G	G	G	G	G	G	G	G	G	G	G
P. c. caligatus	CTGAM182	13	G	G	G	G	G	G	G	G	G	G	G	G	G
P. c. dubius	UFRO403	13	G	G	G	G	G	G	G	G	G	G	G	G	G
P. c. dubius	UFRO544	12		G	G	G	G	G	G	G	G	G	G	G	G
P. donacophilus	CDO*	13	P	P	O	O	O	O	O	O	O	O	O	O	O
C. lugens	JPB119	13	G	G	G	G	G	G	G	G	G	G	G	G	G
C. lugens	JPB124	13	G	G	G	G	G	G	G	G	G	G	G	G	G
C. lugens	JPB136	13	G	G	G	G	G	G	G	G	G	G	G	G	G
C. lugens	JPB81	12	G	G	G	G		G	G	G	G	G	G	G	G
C. purinus	CTGAM154	13	G	G	G	G	G	G	G	G	G	G	G	G	G
C. purinus	CTGAM195	11	G	G	G	G	G	G	G	G		G		G	G
C. purinus	CTGAM209	11	G	G	G	G	G	G		G		G	G	G	G
C. personatus	CLP1*	12		P	O	O	O	O	O	O	O	O	O	O	O
C. personatus	CNRJ2466	2	G	G											
C. coimbrai	CCO1	10			O	O	O	O	O	O	O	O	O		O
C. nigrifrons	CNI1*	10		P		O	O	O	O	O	O	O	O	O	
C. nigrifrons	CPE04	8	G	G	G	G		G	G			G			
Cacajao calvus	CCL1*	11	P	P	O	O	O	O		O		O	O	O	O
Chiropotes israelita	CIS1*	13	P	P	O	O	O	O	O	O	O	O	O	O	O
Pithecia pithecia	PPT1*	13	P	P	O	O	O	O	O	O	O	O	O	O	O

* = composite individual

Table 4.5 Accession numbers for the GenBank sequences included in the StarBEAST2 analysis.

Species	Sample ID	Locus				
		COI	CYTB	ABCA1	DENND5A	DMRT1
P. donacophilus	CDO1*	FJ785423	FJ785423	HM765289	HM759315	HM762536
C. personatus	CLP1*	---	AF289988	HM765298	HM759320	HM762544
C. coimbrai	CCO1	---	---	HM765284	HM759312	HM762531
C. nigrifrons	CNI1*	---	AF524884	---	HM759318	HM762550
Pithecia pithecia	PPT1*	JF459229	KR902424	HM765380	HM759323	HM762634
Cacajao calvus	CCL1*	NC021967	NC021967	HM765283	HM759286	HM762530
Chiropotes israelita	CIS1*	NC024629	NC024629	HM765295	HM759295	HM762542

Species	Sample ID	Locus			
		ERC2	FES	FOXP1	MAPKAP1
P. donacophilus	CDO1*	HM762211	HM761804	HM761533	HM760634
C. personatus	CLP1*	HM762187	HM761807	HM761542	HM760643
C. coimbrai	CCO1	HM762175	HM761802	HM761528	HM760631
C. nigrifrons	CNI1*	HM762178	HM761806	HM761547	HM760647
Pithecia pithecia	PPT1*	HM762204	HM761808	HM761627	HM760722
Cacajao calvus	CCL1*	HM762297	---	HM761527	---
Chiropotes israelita	CIS1*	HM762183	HM761786	HM761539	HM760640

Species	Sample ID	Locus			
		NPAS3.2	RAG1	RAG2	ZFX
P. donacophilus	CDO1*	HM759924	HM759135	HM758967	HM757151
C. personatus	CLP1*	HM759933	HM759138	HM758970	HM757153
C. coimbrai	CCO1	HM759919	HM759134	---	HM757149
C. nigrifrons	CNI1*	HM759938	HM759137	HM758969	---
Pithecia pithecia	PPT1*	HM760023	HM759140	HM758971	HM757155
Cacajao calvus	CCL1*	HM759918	HM759113	HM758942	HM757126
Chiropotes israelita	CIS1*	HM759930	HM759120	HM758950	HM757133

* = composite individual

The species tree was calibrated by applying a lognormal distribution on the root node with a hard minimum bound of 15.7 Ma based on the fossil *Proteropithecia* Kay et al., 1998, and a soft maximum bound of 26 Ma based on the fossil *Branisella boliviana* Hoffstetter, 1969, implemented as described for the ddRADseq divergence dating analysis in section 4.3.4. Each of the nuclear gene trees were loosely calibrated based on the average substitution rates found for Callicebinae taxa across the 54 nuclear loci employed by Perelman et al. (2011) (including the 11 loci used in this analysis). The clock rate calibrations were applied as a uniform distribution with broad upper and lower bounds (3.06×10^{-4} – 12.5×10^{-4} substitutions per site per million years) set based on the average upper and lower 95% HPD values estimated for the substitutions rate across the eight Callicebinae species included in the

Perelman et al. (2011) phylogeny. Clock rates for the mitochondrial loci were estimated relative to the this substitution rate under a broad exponential distribution (default settings). Note that the broad substitution rate calibrations were applied to the gene tree clock rates whereas the root node fossil calibration was applied to the species tree.

The StarBEAST2 analysis was conducted for 200 million MCMC generations, sampling every 10,000 generations, and run on the CIPRES Science Gateway v 3.3 server (Miller et al., 2010). The sampling distributions were visualized using Tracer v. 1.6 to evaluate convergence and to verify that the effective sample size was > 200 for all parameters after a burn-in of 10%. The maximum clade credibility species tree and gene trees were generated in TreeAnnotator v. 2.4.4. DensiTree (Bouckaert, 2010) was used to visualise uncertainty in the species tree.

4.4 Results

4.4.1 ddRADseq data assembly

The samples sequenced for this study (i.e., Callicebinae taxa) had an average of 155K raw reads (total = 7 million), which were reduced to an average 80K quality filtered reads per sample. These clustered into an average of 2342 and 2407 stacks (85% and 92% clustering threshold, respectively), with a mean depth of around 12.8X, and subsequently filtered to an average 2199 (85%) and 2250 (92%) consensus sequences per sample (Table 4.6).

Table 4.6 Summary of the ddRADseq data assembly: comparison of average sequencing effort for Callicebinae vs. Pitheciinae samples.

Samples	Clust. thresh.	Reads	Reads passed	Clusters[1]	Avg. depth[1]	Cons. loci	No. sites	H[2]
Callicebinae	85%	154863	79762	2342	12.80	2199	655204	0.0022
Callicebinae	92%	--	--	2407	12.77	2250	670320	0.0022
Pitheciinae	85%	86138	40606	1124	8.27	1030	294005	0.0012
Pitheciinae	92%	--	--	1144	8.29	1049	299304	0.0012
All	85%	144490	73852	2159	12.11	2022	600684	0.0020
All	92%	--	--	2217	12.09	2068	614317	0.0020

[1]After excluding loci with depth <5

[2] Heterozygosity measured as the proportion of called sites

Sequencing effort was comparatively low for the outgroup samples (Pitheciinae taxa) that were added to the dataset. These had an average of 41K quality filtered reads per sample (86K raw reads/sample) which clustered into an average of 1124 (85%) and 1144 (92%) clusters, with a mean depth of around 8.3X, and filtered to an average of 1030 (85%) and 1049 (92%) consensus sequences per sample (Table 4.6).

The generation of consensus sequences per sample (steps 1 – 5 in pyRAD) was highly comparable across both clustering thresholds (85% and 92%) in relation to the number of clusters, average depth, the number of filtered putative loci, and the heterozygosity of each individual (Tables 4.7, A3.1). Clustering of consensus sequences across samples (steps 6 + 7 in pyRAD) at both thresholds also resulted in a similar number of total loci included in the final datasets (1178 versus 1129 loci in A85 and A92 datasets, respectively). However, the outgroup samples are represented at fewer loci in the A92 dataset, with around a 40% decrease (compared to A85) in the average number of loci sequenced for the outgroup taxa as a percentage of the total number of loci (e.g., see Table 4.8).

This trend is also found for *Cheracebus* and *Callicebus* individuals (around a 5% decrease), while *Plecturocebus* samples have around the same ratio of sequenced to total loci in both datasets. Given that all taxa had similar numbers of consensus sequences per sample under both thresholds (Tables 4.7, A3.1), it is likely that clustering across samples at 92% similarity resulted in fewer loci included for the outgroup taxa owing to the relatively old divergence (early Miocene) and thus, higher sequence divergence between Callicebinae and Pitheciinae. In light of this, all further datasets were assembled using a clustering threshold of 85% to maximise sequence coverage across all clades, each of which contained between 1805 and 3048 loci (545K – 932K total bp). Information about each assembled dataset is listed in Table 4.9.

Table 4.7 Summary of the ddRADseq data assembly (85% clustering threshold): sequencing information per sample.

Species	Sample ID	Reads	Reads passed	Clusters[1]	Avg. depth[1]	Cons. loci	No. sites	H[2]
P. hoffmannsi	CTGAM248	306337	151852	3500	16.85	3302	978010	0.0028
P. hoffmannsi	CTGAM249	111856	55727	2643	12.51	2489	742761	0.0029
P. hoffmannsi	CTGAM290	102270	52813	2636	12.12	2462	734999	0.0028
P. miltoni	42991	103221	52457	2472	12.05	2344	698258	0.0019
P. miltoni	42992	26936	13386	765	7.40	696	209065	0.0020
P. cinerascens	UFRO352	273113	139449	3487	18.42	3284	967058	0.0030
P. cinerascens	UFRO354	340623	206741	4004	22.96	3820	1162308	0.0026
P. cinerascens	UFRO355	345291	184024	3168	15.74	2945	890986	0.0031
P. cinerascens	UFRO499	290049	175012	3825	21.36	3651	1117138	0.0024
P. cinerascens	WRS03	282987	178158	3897	21.96	3737	1141566	0.0025
P. cinerascens	WRS04	289893	174354	3722	21.72	3567	1087395	0.0026
P. bernhardi	42961	54657	27783	1621	9.17	1494	446649	0.0019
P. bernhardi	42964	162111	84364	3170	15.19	2996	891006	0.0022
P. bernhardi	UFRO413	276283	133737	2159	12.05	2000	591162	0.0023
P. bernhardi	CCM173	38910	20194	959	8.18	872	258247	0.0026
P. cf. moloch	RVR73	131902	66449	2710	13.36	2552	761102	0.0012
P. moloch	CTGAM420	150107	73122	2733	12.79	2553	768096	0.0015
P. moloch	CTGAM421	253243	120890	3326	17.71	3136	926124	0.0015
P. brunneus	4505	136376	68082	2731	13.72	2580	765038	0.0018
P. brunneus	UFRO327	403145	199843	2709	13.26	2518	736335	0.0019
P. brunneus	UFRO541	58000	29213	1584	10.18	1475	441644	0.0021
P. cupreus	AAM15	239043	119942	3654	16.38	3423	1011058	0.0016
P. cupreus	JLP15920	165045	85552	3147	15.02	2999	890722	0.0021
P. cupreus	CTGAM210	53988	25065	698	6.84	619	175367	0.0019
P. cupreus	4987	126201	60676	2508	12.95	2355	700502	0.0018
P. cupreus	4988	39680	19363	1168	8.36	1071	321155	0.0022
P. cupreus	4990	146403	75103	2840	13.31	2688	799467	0.0018
P. c. caligatus	CTGAM181	73412	33052	753	7.15	679	191748	0.0019
P. c. caligatus	CTGAM182	203884	104355	3279	16.24	3105	919765	0.0023
P. c. caligatus	CCM248	64071	30994	1830	9.65	1686	504487	0.0023
P. c. caligatus	MVR58	241597	118306	3430	16.41	3234	951910	0.0020
P. c. dubius	UFRO403	252706	124259	3284	16.34	3109	920152	0.0023
P. c. dubius	UFRO427	134817	62740	1208	8.93	1098	309948	0.0023
P. c. dubius	UFRO544	107090	47568	1895	11.78	1777	531692	0.0025
P. c. dubius	2804	127611	63963	2744	12.86	2585	767292	0.0023
C. personatus	2466	13162	6331	155	7.33	135	40451	0.0023
C. lugens	JPB160	85568	41459	1112	7.81	1025	290127	0.0017
C. lugens	JPB161	111957	51591	1022	8.43	930	260906	0.0016
C. lugens	JPB81	160013	78654	3024	14.41	2834	843338	0.0017
C. lugens	CTGAM733	101580	46833	2192	10.29	2018	616928	0.0019
C. lucifer	CTGAM703	50854	24646	1421	8.51	1299	403499	0.0025
C. lucifer	CTGAM726	47586	24161	1490	8.63	1379	413427	0.0027
C. purinus	CTGAM154	29785	13879	736	6.93	654	202980	0.0024
C. purinus	CTGAM195	82079	37058	969	7.55	882	250041	0.0019
C. purinus	CTGAM209	173405	86094	3032	15.04	2876	852287	0.0019
Pithecia mittermeieri	CTGAM215	49088	22116	1068	8.15	975	302059	0.0012
Cacajao ayresi	5667	75721	35587	1159	8.02	1054	298552	0.0007
Cacajao calvus	5241	109827	52274	1035	8.40	955	267997	0.0010
Cacajao hosomi	5698	94259	44687	1220	8.76	1122	316673	0.0011
Cacajao melanocephalus	0065	75938	35650	1007	7.79	921	260177	0.0013
Chiropotes albinasus	CTGAM213	82121	38938	1009	8.01	920	259987	0.0011
Chiropotes israelita	5713	105778	50461	1331	8.70	1241	349903	0.0013
Chiropotes sagalatus	CTGAM515	96369	45135	1166	8.37	1051	296689	0.0018

[1] After excluding loci with depth <5

[2] Heterozygosity measured as the proportion of called sites

Table 4.8 Summary of ddRADseq locus coverage per sample for each dataset.

Species	Sample ID	Dataset // total loci									
		A85	**A92**	**B85**	**PD85**	**P85**	**Pi85**	**Pii85**	**Piii85**	**PW85**	**PS85**
		1178	1129	1987	1931	1944	3048	1805	1959	2259	2306
P. hoffmannsi	CTGAM248	827	792	1396	1291	1306	1931	--	--	--	--
P. hoffmannsi	CTGAM249	971	938	--	1445	1448	1894	--	--	--	1674
P. hoffmannsi	CTGAM290	966	932	--	1415	1418	1852	--	--	--	1638
P. miltoni	42991	941	897	1424	1371	1365	1760	--	--	--	1583
P. miltoni	42992	382	359	--	461	465	575	--	--	--	532
P. cinerascens	UFRO352	877	849	1452	1377	1386	1965	--	--	--	--
P. cinerascens	UFRO354	937	900	--	1451	1457	2399	--	--	--	1692
P. cinerascens	UFRO355	535	515	--	808	808	1497	--	--	--	--
P. cinerascens	UFRO499	958	931	--	1419	1426	2332	--	--	--	--
P. cinerascens	WRS03	976	942	1502	1480	1490	2407	--	--	--	1708
P. cinerascens	WRS04	935	903	--	1417	1425	2338	--	--	--	--
P. bernhardi	42961	700	670	--	970	966	--	1030	--	--	--
P. bernhardi	42964	1002	956	1642	1567	1567	--	1598	--	--	1846
P. bernhardi	UFRO413	500	476	818	776	779	--	858	--	--	914
P. bernhardi	CCM173	409	394	--	546	--	--	609	--	--	--
P. cf. *moloch*	RVR73	941	897	1463	1408	1417	--	1498	--	--	1641
P. moloch	CTGAM420	905	881	--	1359	1367	--	1451	--	--	--
P. moloch	CTGAM421	924	884	1563	1466	1479	--	1519	--	--	1745
P. brunneus	4505	957	933	1492	1448	1457	--	--	1503	1676	1677
P. brunneus	UFRO327	532	508	--	798	812	--	--	846	953	--
P. brunneus	UFRO541	647	622	--	905	907	--	--	973	1058	1030
P. cupreus	AAM15	964	936	1588	1522	1533	--	--	1605	1812	1788
P. cupreus	JLP15920	1010	970	--	1580	1589	--	--	1666	1860	--
P. cupreus	CTGAM210	199	188	--	--	--	--	--	--	--	--
P. cupreus	4987	949	921	--	1405	1415	--	--	1490	1649	--
P. cupreus	4988	544	527	--	714	716	--	--	757	814	--
P. cupreus	4990	989	953	1552	1519	1527	--	--	1584	1769	1781
P. c. caligatus	CTGAM181	172	162	--	--	250	--	--	309	--	--
P. c. caligatus	CTGAM182	951	921	1578	1506	1522	--	--	1587	1783	1802
P. c. caligatus	CCM248	790	759	--	1101	1102	--	--	1157	1259	--
P. c. caligatus	MVR58	921	884	--	1476	1490	--	--	1562	1763	1757
P. c. dubius	UFRO403	953	926	--	1488	1501	--	--	1567	1762	1764
P. c. dubius	UFRO427	232	222	--	--	341	--	--	407	--	--
P. c. dubius	UFRO544	801	783	1173	1134	1138	--	--	1234	1334	1305
P. c. dubius	2804	981	941	--	1468	1479	--	--	1558	1734	--
C. personatus	2466	48	38	58	--	--	--	--	--	--	--
C. lugens	JPB160	174	156	280	208	--	--	--	--	--	--
C. lugens	JPB161	167	145	--	198	--	--	--	--	--	--
C. lugens	JPB81	704	657	1111	879	--	--	--	--	--	--
C. lugens	CTGAM733	530	486	777	630	--	--	--	--	--	--
C. lucifer	CTGAM703	334	306	--	386	--	--	--	--	--	--
C. lucifer	CTGAM726	460	407	647	527	--	--	--	--	--	--
C. purinus	CTGAM154	180	164	--	210	--	--	--	--	--	--
C. purinus	CTGAM195	188	163	--	208	--	--	--	--	--	--
C. purinus	CTGAM209	691	636	1063	856	--	--	--	--	--	--
Pithecia mittermeieri	CTGAM215	151	103	201	--	--	--	--	--	--	--
Cacajao ayresi	5667	101	74	172	--	--	--	--	--	--	--
Cacajao calvus	5241	82	47	117	--	--	--	--	--	--	--
Cacajao hosomi	5698	116	77	180	--	--	--	--	--	--	--
Cacajao melanocephalus	0065	109	58	167	--	--	--	--	--	--	--
Chiropotes albinasus	CTGAM213	100	47	145	--	--	--	--	--	--	--
Chiropotes israelita	5713	110	63	167	--	--	--	--	--	--	--
Chiropotes sagalatus	CTGAM515	102	60	150	--	--	--	--	--	--	--

Sequencing coverage for some ingroup individuals was low. In particular, *Callicebus personatus* had only 155 clusters and 135 consensus sequences before across sample clustering and was represented in < 5% of the total loci in the final datasets where it is included (Table 4.8). Based on this, it was tempting to exclude this specimen entirely from the phylogenetic analyses, however, it is the only representative of the genus *Callicebus* sequenced in this study and the extremely low sequencing coverage appeared to have little impact on the inferred phylogenetic relationships (see section 4.4.2). Three *Plecturocebus* samples also had notably low sequencing coverage (CTGAM210, CTGAM181, UFRO427) in comparison to other individuals from this clade (e.g., see Figure 4.1), and they were excluded from some datasets (e.g., PD85, assembled for introgression analyses).

Table 4.9 Summary of the final assembled ddRADseq datasets, including number of loci, concatenated length, and percentage missing data.

Dataset	Clust. thresh.	No. Loci	Length (bp)	Variable (bp)	Pis (bp)	Missing data
A85	85%	1178	360499	25394	13138	49.8%
A92	92%	1129	345045	21889	11166	50.3%
B85	85%	1987	602518	32263	15076	55.9%
P85	85%	1944	591179	21011	8692	37.5%
Pi85	85%	3048	932098	14475	5655	37.9%
Pii85	85%	1805	545483	5142	1608	32.5%
Piii85	85%	1959	590357	9821	3080	37.2%
PD85	85%	1931	588527	30509	15364	44.6%

In the final datasets including Pitheciinae, each outgroup sample is represented in only around 6 – 9 % of the total loci (although a much greater % of loci contain at least one outgroup sample). The low number of shared loci recovered between the outgroups and Callicebinae (Figure 4.1) is influenced by two main factors; the number of raw reads and the relatively deep divergence between these clades (Eaton et al., 2017). The average number of raw and quality filtered reads, within-sample clusters and consensus sequences for the Pitheciinae individuals was much lower (close to half) than for the Callicebinae taxa, as discussed above (see Tables 4.6, 4.7), and the mean number of raw reads was a significant predictor of the number of shared loci between samples across the three datasets tested (Mantel r⇢ = 0.44 – 0.5, *p*-value < 0.0006; see Figure 4.2 and Table 4.10).
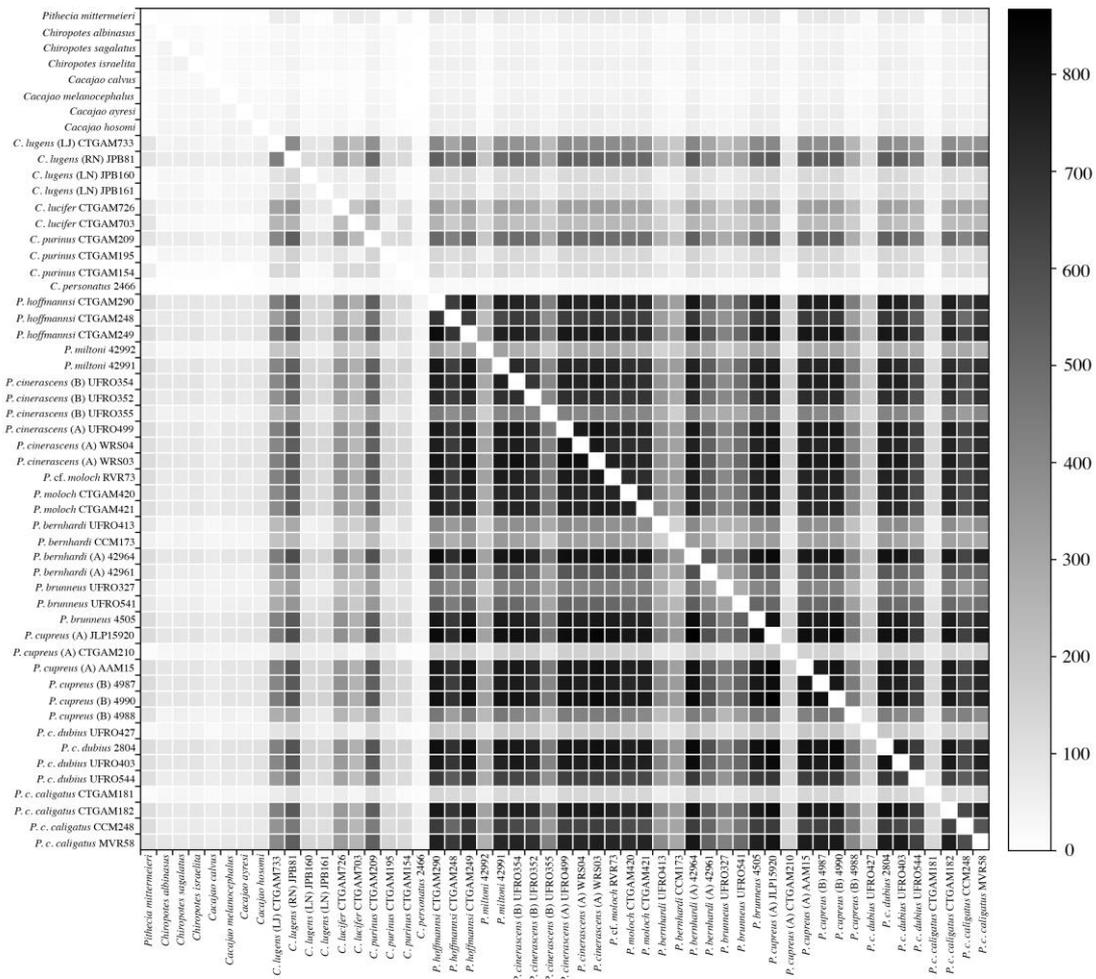
Figure 4.1 Heatmap showing the number of shared ddRADseq loci between samples across two datasets assembled with different clustering thresholds (85% or 92%). The A85 dataset (total ddRADseq loci = 1178) is below the diagonal and the A92 dataset (total ddRADseq loci = 1129) is above the diagonal.

The number of shared loci between Pitheciinae and Callicebinae may have been influenced by the slight modification to the sequencing protocol (SPRI bead versus kit purification), which could account for some of the difference in the mean number of raw reads across these clades. Variation between ddRADseq libraries can also be attributed to the use of fragment size selection to adjust the quantity of loci, however, this is minimised with size selection tools such as the Pippin Prep (Sage Science), which was used in both protocols (Puritz et al., 2014).

127
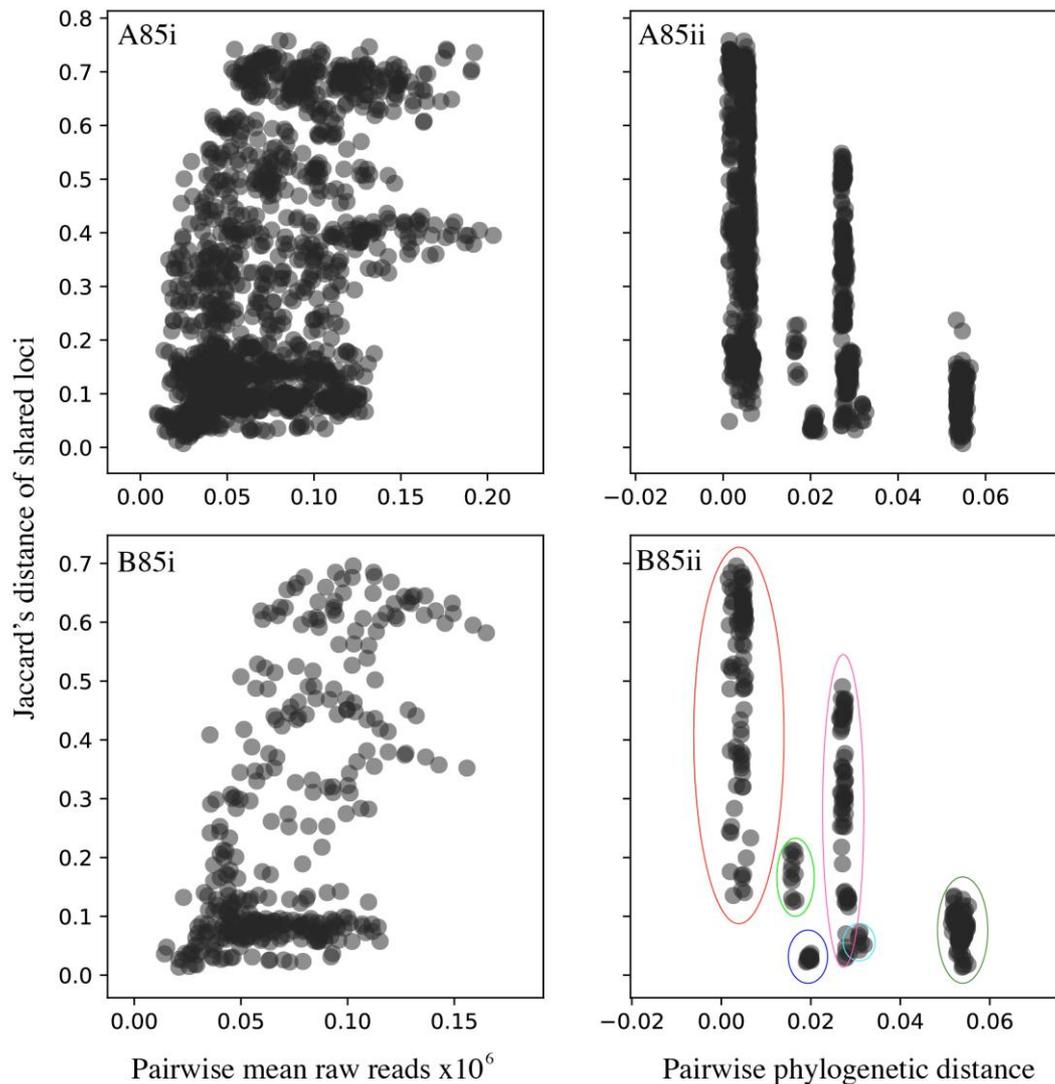
Figure 4.2 The correlation between Jaccard's distance of the proportion of shared loci between samples and (i) the number of raw input reads or (ii) pairwise phylogenetic distance. The A85 dataset (all individuals) is on top and the B85 dataset (one individual per taxon) is below. Coloured circles on the B85ii plot indicate loci shared across taxa, as follows: closely related species (red); *Cheracebus* and *Plecturocebus* (pink); *C. personatus* and *Plecturocebus/Cheracebus* (dark blue); *Cacajao* and *Chiropotes* (light green); *Pithecia* and *Cacajao/Chiropotes* (light blue); and taxa from different subfamilies (dark green).

The second factor is related to the decrease in shared RADseq loci between taxa with increasing phylogenetic distance as a result of the disruption of restriction recognition sites owing to mutations ("locus dropout"). Mantel correlation tests indicate that there was strong hierarchical structure in the distribution of missing data (Mantel r···→ = -0.73 – -0.76, *p*-value = 0.0001; see Table 4.10), regardless of the clustering threshold used. The occurrence of locus dropout could be overestimated if

the number of shared loci between the Pitheciinae and Callicebinae libraries is impacted by other sources of variation. Variation between these libraries, however, could not account entirely for the strong hierarchical structure given that pairwise phylogenetic distance predicts the number of shared loci at all phylogenetic scales. This is clearly demonstrated in Figure 4.2; closely related species from the same genus share the most loci (red circle on the B85ii plot, Figure 4.2) and the least loci are generally shared across taxa from the different subfamilies (dark green) [but also for *C. personatus* vs. *Plecturocebus / Cheracebus* (dark blue) owing to the extremely low sequencing coverage for *C. personatus*]. Intermediate amounts of shared loci are typically found across the genera [*Cacajao* vs. *Chiropotes* (light green); *Cheracebus* vs. *Plecturocebus* (pink); *Pithecia* vs. *Cacajao / Chiropotes* (light blue)].

Table 4.10 Mantel correlation tests.

Dataset	No. of raw reads		Phylogenetic distance	
	Mantel r⇢	***p*-value**	**Mantel r⇢**	***p*-value**
A85	0.447	0.0003	-0.743	0.0001
A92	0.442	0.0003	-0.764	0.0001
B85	0.5	0.0006	-0.734	0.0001

Because *Plecturocebus* taxa have better sampling in terms of number of individuals and number of raw reads, there is a bias towards loci recovered across *Plecturocebus* samples (see Figure 4.1), whereas loci found only in *Cheracebus*, *Callicebus*, or Pitheciinae will not be represented. Thus, nearly all *Plecturocebus* individuals have a higher number of loci sequenced in the final datasets (Table 4.8). The potential loss of loci as a result of locus dropout formed part of the justification for constructing specific datasets with only the target individuals included for each analysis in order to maximise the number of loci assembled. Finally, a test run (data not included) conducted in pyRAD that included the Pitheciinae samples in the minimum coverage per locus (rather than as add-on taxa) resulted in much fewer loci overall, and no significant impact on the proportion of loci for these samples, thus it is unlikely that this assembly option is adding further bias in the number of shared loci.

Heterozygosity varies significantly across Callicebinae taxa (85% clustering threshold: 0.0012 – 0.0031, mean = 0.0022), and is much higher than for Pitheciinae (0.0007 – 0.0018, mean = 0.0012). *Plecturocebus cinerascens* (especially clade B) and *P. hoffmannsi* individuals are the most polymorphic, while *P. moloch* and *P.* cf.

moloch are the least, and among *Cheracebus* taxa, *C. lucifer* individuals show the highest heterozygosity (Tables 4.7, A3.1).

4.4.2 Phylogenetic inference: ddRADseq

Four of the concatenated ddRADseq loci data matrices assembled in pyRAD were used for phylogenetic inference: A85 and A92 (all samples); B85 (select individuals from each taxon); and PD85 (only individuals of the *P. moloch* group and *Cheracebus*). These datasets contained between 1129 and 1987 ddRADseq loci, with a total concatenated length of between 345K and 603K bp, and around 11K to 15.4K parsimony informative sites across all samples (Table 4.9). A total of nine phylogenetic trees were reconstructed using the ddRADseq data; maximum-likelihood and Bayesian (MrBayes) trees were inferred for each of the datasets, and phylogeny and diversification times were jointly estimated for the B85 dataset using BEAST. Species relationships recovered across all trees for all datasets are identical, with only minor topological conflict at some intraspecific nodes (e.g., within *P. bernhardi* and *P. caligatus*). As most nodes show perfect support (bootstrap percentage, BP = 100%; posterior probability, PP = 1.0) or strong support (e.g., BP > 95%; PP > 0.99) across the analyses (see Table 4.11), only insignificant (BP < 70%; PP < 0.95) or intermediate support values are mentioned below. Note posterior probability values for the B85 BEAST and MrBayes analyses are not included in Table 4.11 because all nodes show perfect support (PP = 1.0).

In agreement with all previous molecular phylogenies (e.g., Chapter 2 + 3, Carneiro et al., 2016), *Cheracebus* is recovered as the earliest diverging lineage within Callicebinae, and *Callicebus* (represented by *C. personatus*) and *Plecturocebus* are sister clades (Figure 4.3). Despite the extremely low sequencing coverage, the divergence between *C. personatus* and *Plecturocebus* is strongly resolved in most analyses (moderate support in A92, BP = 86%). Among taxa of the genus *Cheracebus*, *C. lucifer* and *C. purinus* are sister species, and for *C. lugens*, individuals from the left and right bank of the Rio Negro (LN and RN) form a clade to the exclusion of the left bank Rio Japurá (LJ) sample, but with low support at the LN/RN node in some ML trees (e.g., A85 + A92, BP = 53%).
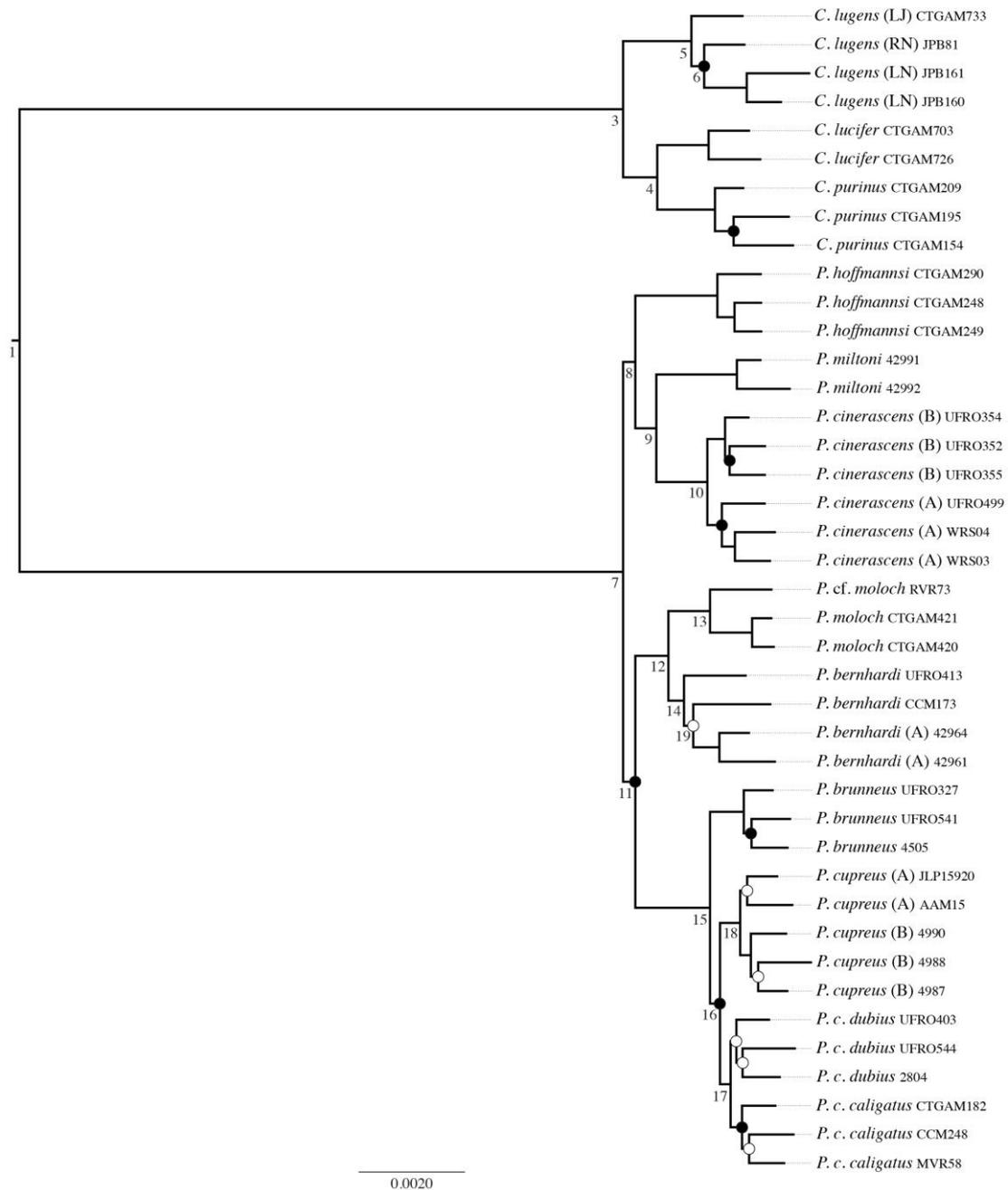
Figure 4.3 Maximum likelihood tree inferred with the ddRADseq PD85 dataset (*Plecturocebus* and *Cheracebus*). Unmarked nodes were strongly supported (BP = 100%), nodes marked with a black circle received significant support (BP = 80 – 99%), while the nodes marked with white circles were recovered without significant support (BP < 70%). Support values are listed according to node numbers in Table 4.11.

The *Plecturocebus moloch* species group is divided into three major clades, as recovered in the BEAST phylogeny based on seven loci in Chapter 3, and the informal clade names suggested are followed here; the Aripuanã-Tapajós clade containing *P. hoffmannsi, P. miltoni* and *P. cinerascens;* the Eastern clade containing

P. bernhardi, P. cf. moloch, and *P. moloch*; and the Western clade comprised of *P. brunneus, P. caligatus,* and *P. cupreus*. The Aripuanã-Tapajós clade is the earliest diverging lineage within the *P. moloch* group, while the Eastern and Western clades are sister (B85 BP = 91%; PD85 BP = 87%). Among the Aripuanã-Tapajós species, *Plecturocebus hoffmannsi* is sister to a *P. cinerascens + P. miltoni* clade, and among Eastern clade taxa, *P. bernhardi* is sister to a clade containing *P. moloch* and *P. cf. moloch*. Three relatively divergent lineages are represented by the four *P. bernhardi* specimens in the ddRADseq datasets, and most analyses recover the UFRO(413) and CCM(173) individuals as successive sister lineages to the *P. bernhardi* clade A specimens (note, the CCM173 sample wasn't included in the B85 dataset), but with low support at the CCM vs. clade A node in the ML trees (A85 BP = 63%; PD85 BP = 65%). An alternative topology, however, suggests that the UFRO and CCM *P. bernhardi* individuals are sister (A92 BP = 87%; Figure A3.1, A3.2). This is the only notable conflict between the A85 and A92 analyses, and thus, the clustering threshold used to assemble the ddRADseq datasets appears to have had minimal impact on the recovered species relationships.

Among Western clade taxa, *P. cupreus* and *P. caligatus* are recovered as sister species but with low support in the ML trees (A85 BP = 60%; A92 BP 78%; B85 BP = 61%), while *P. brunneus* is the earliest diverging lineage within the clade. In the A85 and A92 analyses including all samples, *P. c. dubius* is paraphyletic, although some of the nodes within the *P. caligatus* complex are poorly supported (see Figure A3.1 – A3.4). Subsequently, one *P. c. dubius* (UFRO427) and one *P. c. caligatus* (CTGAM181) sample were excluded from the PD85 dataset because of low sequencing coverage and, in these analyses, *P. c. dubius* and *P. c. caligatus* are recovered as monophyletic sister taxa (note, only one sample from each was included in the B85 analyses). Overall, the topology recovered across the ddRADseq phylogenetic analyses is identical to the BEAST phylogeny based on seven concatenated loci in Chapter 3, except *P. caligatus* is sister to *P. cupreus* in the former and *P. brunneus* in the latter.

Age estimates in the ddRADseq timetree (Figure 4.4, Table 4.11) are generally younger than those recovered in the combined and nuclear dataset analyses in Chapter 2, especially at species-level, but broadly concordant with the results from Chapter 3. The most recent common ancestor (MRCA) of extant Callicebinae taxa is estimated to have diverged in the late Miocene (10.7 Ma; 95% HPD = 7.9 – 13.9),

132

giving rise to the progenitor of the genus *Cheracebus* and of the *Callicebus+Plecturocebus* clade. The divergence between *Callicebus* and *Plecturocebus* also occurred in the late Miocene at *c.* 7.4 Ma (95% HPD = 5.2 – 10.0). All other diversification events among the Callicebinae taxa represented occurred in the Pleistocene. The MRCA of extant *Cheracebus* taxa is estimated to have lived in the early Pleistocene (1.9 Ma; 95% HPD = 1.3 – 2.4), while *C. purinus* and *C. lucifer* diverged at *c.* 1.4 Ma (95% HPD = 1.0 – 1.9), and the *C. lugens* lineages diverged at around 0.9 Ma [(95% HPD = 0.6 – 1.2) LJ vs. RN+LN], and 0.7 Ma [(95% HPD = 0.5 – 1.0) RN vs. LN].

The major *P. moloch* group clades diversified relatively rapidly in the early Pleistocene, with the Aripuanã-Tapajós clade diverging from the ancestor of the Eastern and Western clades at *c.* 2.0 Ma (95% HPD = 1.5 – 2.6), the Eastern and Western clades diverging at *c.* 1.8 Ma (95% HPD = 1.3 – 2.3), and *P. hoffmannsi* diverging from other species of the Aripuanã-Tapajós clade also at *c.* 1.8 Ma (95% HPD = 1.4 – 2.4). Thus, within around 200K years, four distinct *P. moloch* group lineages had emerged which likely gave rise to all known extant taxa. The sister taxa of the Aripuanã-Tapajós clade (*P. cinerascens* and *P. miltoni*) and the Eastern clade (*P. bernhardi* and *P. moloch/P.* cf. *moloch*) are estimated to have diverged at 1.5 Ma (95% HPD = 1.1 – 2.0) and 1.3 Ma (95% HPD = 1.0 – 1.7), respectively, while the *P. cinerascens* (clade A and B), *P. moloch* (vs. *P.* cf. *moloch*), and *P. bernhardi* (UFRO and clade A) lineages all diverged between *c.* 0.8 – 0.9 Ma. The earliest diverging lineage within the Western clade, *P. brunneus*, is estimated to have diverged at *c.* 1.0 Ma (95% HPD = 0.7 – 1.3), followed by the divergence between *P. cupreus* and *P. caligatus* at *c.* 0.9 Ma (95% HPD = 0.6 – 1.1), and subsequently, the *P. cupreus* lineages (clade A and B) and the *P. caligatus* subspecies (*P. c. caligatus* and *P. c. dubius*) diverged at *c.* 0.6 Ma (95% HPD = 0.4 – 0.8) and 0.7 Ma (95% HPD = 0.5 – 0.9), respectively. Individual trees including outgroups and node support values for each MrBayes and RAxML analysis are found in Appendix 3 (Figure A3.1 – A3.8) and the full timetree including outgroups is found in Appendix 3 (Figure A3.9) A summary of node support for all analyses (except the B85 BEAST and MrBayes analyses, PP = 1.00 for all nodes) and divergence date estimates and 95% HPD intervals for the B85 dataset BEAST analysis is presented in Table 4.11.

Table 4.11 Summary of node support for the ddRADseq phylogenetic analyses and age estimates from the B85 dataset BEAST analysis. Node numbers correspond to those on Figure 4.3, 4.4. Bold indicates low support (PP < 0.95; BP < 70%). Note, PP = 1.00 for all ingroup nodes in the B85 BEAST and MrBayes analyses.

Node	Divergence	B85				A85		A92		PD85	
		Mean age	95% HPD		ML BP	MB PP	ML BP	MB PP	ML BP	MB PP	ML BP
			Lower	Upper							
1	*Cheracebus* vs. *Callicebus (personatus) + Plecturocebus*	10.7	7.86	13.9	100	1.00	100	1.00	100	1.00	100
2	*Callicebus (personatus)* vs. *Plecturocebus*	7.41	5.22	9.97	99	1.00	96	1.00	86	NA	NA
3	*C. purinus + C. lucifer* vs. *C. lugens*	1.85	1.31	2.43	100	1.00	100	1.00	100	1.00	100
4	*C. purinus* vs. *C. lucifer*	1.42	0.98	1.94	100	1.00	100	1.00	100	1.00	100
5	*C. lugens*: L bank Rio Japurá vs. R bank Rio Negro + L bank Rio Negro	0.92	0.64	1.24	100	1.00	100	1.00	100	1.00	100
6	*C. lugens:* R bank Rio Negro vs. L bank Rio Negro	0.7	0.47	0.97	86	0.95	**53**	0.99	**53**	1.00	87
7	Aripuanã-Tapajós clade vs. Eastern + Western *P. moloch* clades	2.01	1.51	2.6	100	1.00	100	1.00	100	1.00	100
8	*P. hoffmannsi* vs. *P. cinerascens + P. miltoni*	1.84	1.35	2.37	97	1.00	98	1.00	97	1.00	100
9	*P. cinerascens* vs. *P. miltoni*	1.49	1.09	1.95	100	1.00	100	1.00	100	1.00	100
10	*P. cinerascens:* clade A *vs* clade B	0.85	0.56	1.18	100	1.00	100	1.00	100	1.00	100
11	Eastern vs. Western *P. moloch* clades	1.8	1.34	2.33	91	1.00	95	1.00	95	1.00	87
12	*P. bernhardi* vs. *P. cf. moloch + P. moloch*	1.32	0.97	1.73	100	1.00	100	1.00	100	1.00	100
13	*P. cf. moloch* vs. *P. moloch*	0.77	0.52	1.04	100	1.00	100	1.00	100	1.00	100
14	*P. bernhardi:* UFRO413 vs. remaining	0.91	0.61	1.25	100	1.00	97	NA	NA	1.00	100
19	*P. bernhardi:* CCM173 vs. clade A	NA [1]	NA [1]	NA [1]	NA [1]	1.00	**63**	NA	NA	1.00	**65**
--	*P. bernhardi:* UFRO413 + CCM173 vs. clade A	NA [1]	NA [1]	NA [1]	NA [1]	NA	NA	1.00	87	NA	NA
15	*P. brunneus* vs. *P. cupreus + P. caligatus*	1	0.74	1.31	100	1.00	100	1.00	100	1.00	100
16	*P. cupreus* vs. *P. caligatus*	0.86	0.63	1.14	**61**	1.00	**60**	1.00	**78**	1.00	94
17	*P. c. caligatus* vs. *P. c. dubius*	0.67	0.47	0.9	100	**0.69** [2]	**43** [2]	**1.00** [2]	**63** [2]	1.00	100
18	*P. cupreus:* clade A *vs* clade B	0.58	0.39	0.78	1.00	1.00	100	1.00	100	1.00	100

[1] CCM173 was not included in the BEAST analysis

[2] *P. c. dubius* is paraphyletic
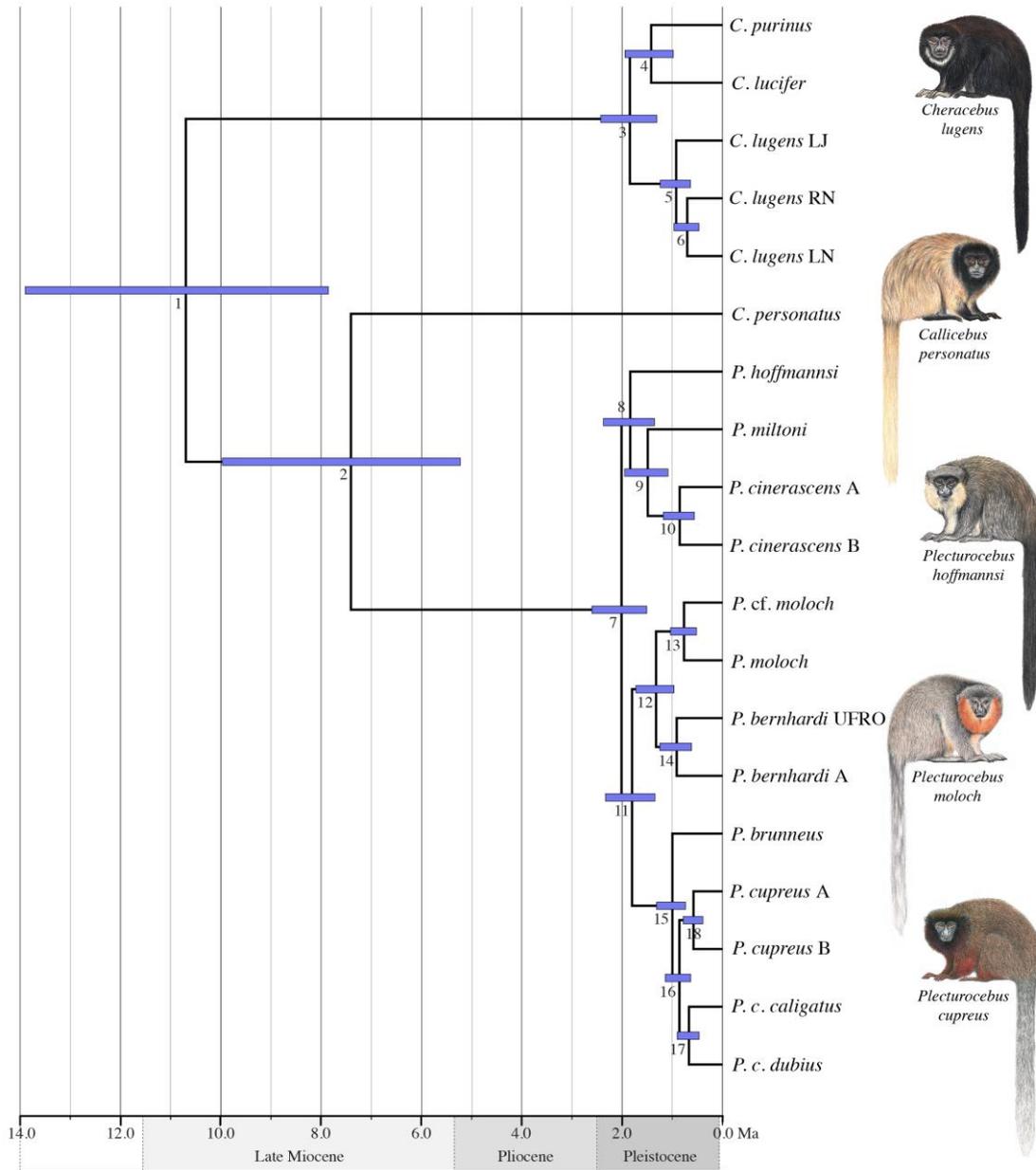
Figure 4.4 A time-calibrated phylogeny for Callicebinae inferred with the ddRADseq B85 dataset. All nodes received full support (PP = 1.00). Node bars indicate the 95% highest posterior density (HPD). Estimated divergence ages and 95% HPDs are listed according to node numbers in Table 4.11. See Figure A3.9 for the full timetree with outgroups. Illustrations by Stephen D. Nash ©Conservation International.

4.4.3 Phylogenetic inference: StarBEAST2

In comparison to the concatenated multi-locus phylogenies based on 22 loci (Chapter 2), the StarBEAST2 coalescent-based species tree (Figure 4.5) is relatively concordant with the relationships among the *P. moloch* group species recovered in the ddRADseq analyses.

135
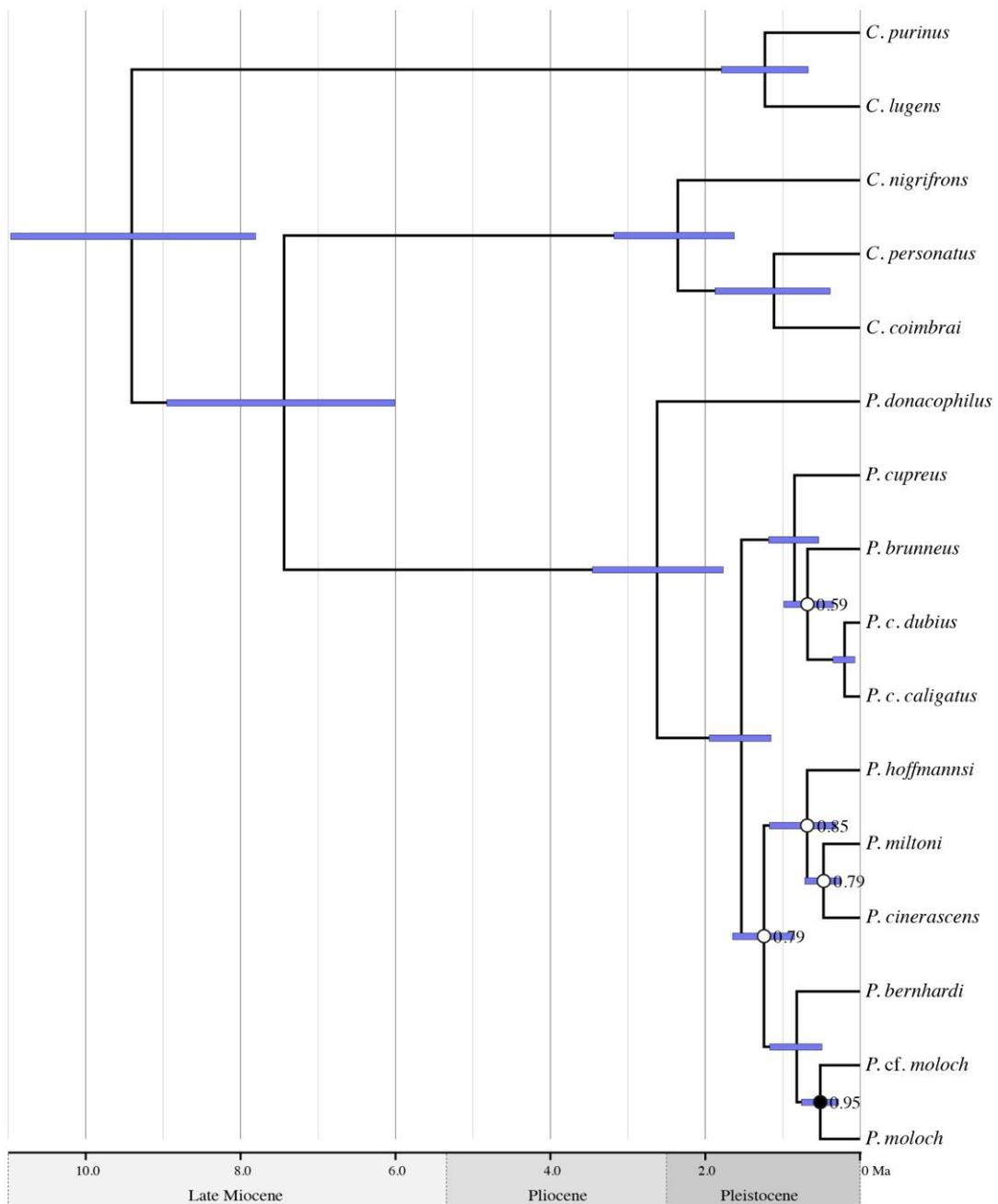
Figure 4.5 A time-calibrated coalescent-based species tree for Callicebinae inferred with multi-locus data using StarBEAST2. Unmarked nodes received full support (PP = 1.00), the node marked with a black circle received significant support (PP = 0.95), while the nodes marked with white circles were recovered without significant support (PP < 0.95). Node bars indicate the 95% highest posterior density (HPD). See Figure A3.10 for the full timetree with outgroups.

Both the StarBEAST2 and ddRADseq trees suggest that the *P. moloch* group is divided into the same three major clades, however, the Eastern and Aripuanã-Tapajós clades are sister lineages in the StarBEAST2 species tree instead of the Eastern and Western clades (ddRADseq and Chapter 3). Species relationships among Western clade taxa are identical in the coalescent-based species tree and concatenated multi-locus (combined) phylogenies, i.e., *P. brunneus* and *P. caligatus* are sister species, rather than *P. cupreus* and *P. caligatus* (ddRADseq). Thus, the ddRADseq trees and the StarBEAST2 species tree are only in disagreement as to whether the Western clade or Aripuanã-Tapajós clade is sister to the Eastern clade, and whether *P. brunneus* or *P. cupreus* is the sister taxon to *P. caligatus*. Notably, these nodes are recovered with low support in the StarBEAST2 phylogeny [*P. brunneus* vs. *P. caligatus* (PP = 0.59); Aripuanã-Tapajós vs. Eastern clade (PP = 0.79)]. Species relationships within the Aripuanã-Tapajós clade are also inferred with low support [*P. hoffmannsi* vs. *P. cinerascens* + *P. miltoni* (PP = 0.85); *P. cinerascens* vs. *P. miltoni* (PP = 0.79)], which might be expected in light of the lack of resolution in the concatenated datasets owing to the strong discordance in the phylogenetic signal in the nuclear vs. mitochondrial loci (see Chapter 2). All other nodes are recovered with full support (PP = 1.0) except the divergence between *P.* cf. *moloch* and *P. moloch* (PP = 0.95).

Uncertainty in the species tree is illustrated by the DensiTree plot (Figure 4.6) showing the four most probable topologies (out of a total of 172 topologies). The second (10.05% of trees) and third (8.75%) most probable topologies are in conflict with the consensus species tree (i.e., first, 29.37%) regarding the relationships among the Western clade species; either *P. cupreus* and *P. brunneus* (2nd; light blue in Figure 4.6) or *P. cupreus* and *P. caligatus* (3rd; purple in Figure 4.6; same as ddRADseq trees) are recovered as sister species. The fourth most probable topology (dark blue in Figure 4.6) differs from the consensus species tree in placing the Aripuanã-Tapajós clade as the earliest diverging lineage within the *P. moloch* group (8.3%), in agreement with the ddRADseq phylogeny. Thus, species relationships recovered in the StarBEAST2 consensus species tree that are in conflict with the ddRADseq phylogeny are characterised by low support, and both conflicts are resolved in one of the three most probable alternative species tree topologies.
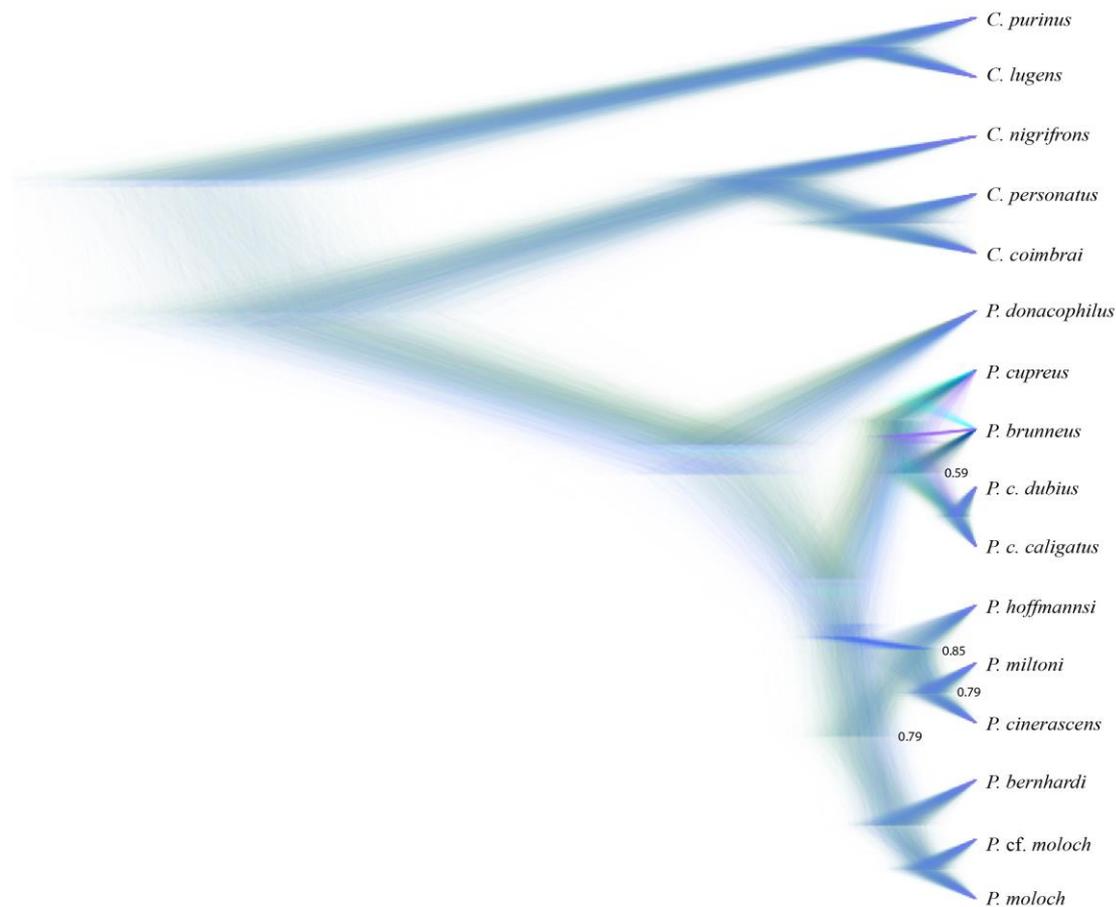
Figure 4.6 DensiTree plot illustrating uncertainty in the coalescent-based species tree. The four most probable topologies are shown: the most probable in green (29.37%); 2nd in light blue (10.05%); 3rd in purple (8.75%); and 4th in dark blue (8.3%). Nodes recovered without significant support (PP < 0.95) are indicated in the figure.

All 172 possible topologies for the species tree differ only in the relationships among the *P. moloch* group species, with all other nodes identical across 100% of the trees. Most of the remaining alternative species tree topologies are found in < 1% of the trees (topologies 16–172), or 1–2% (9–15). This high number of possible topologies for the *P. moloch* group is likely to be partially associated with the weak phylogenetic signal in the nuclear loci, most of which contain only a few informative sites that distinguish taxa within this group. This is reflected in the low support for most of the recovered species relationships within the *P. moloch* group across the individual nuclear gene trees (some examples are shown in Figure 4.7). In contrast, the mitochondrial gene tree is significantly supported (PP > 0.95) at all nodes except one, and the number of informative sites across the two mitochondrial loci is much greater (e.g., more PI sites in the two mitochondrial loci than all 20 nuclear loci in

Chapter 2, see Table 2.7). There is strong discordance among most gene trees, in particular regarding the placement of the Aripuanã-Tapajós clade taxa (e.g., see Figure 4.7). Overall, these results suggest that there may be significant gene tree heterogeneity potentially owing to stochastic processes such as incomplete lineage sorting or more recent gene flow between species of the *P. moloch* group.
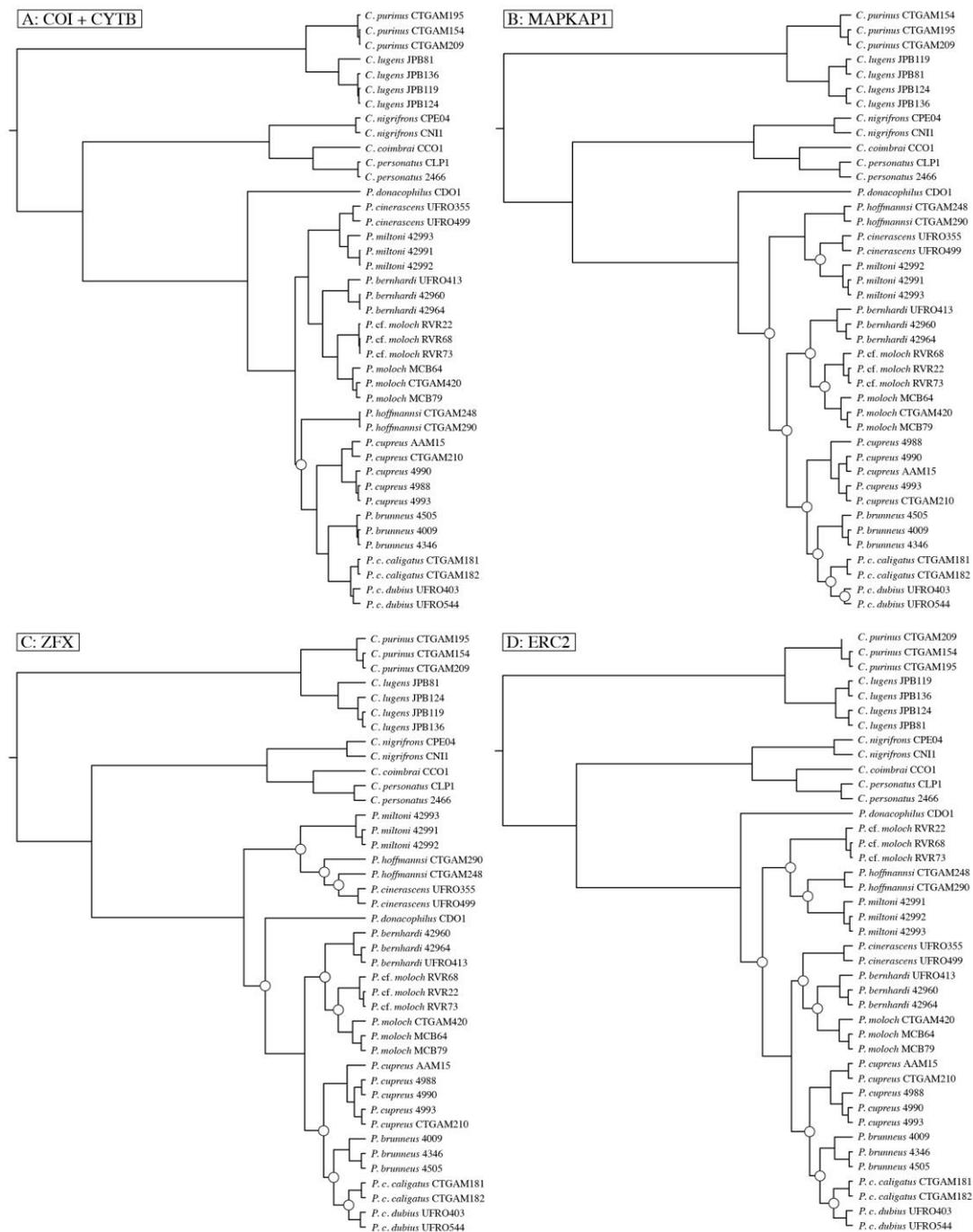
Figure 4.7 Individual gene trees for select loci inferred in the coalescent-based species tree analyses: (A) COI and CYTB, mitochondrial gene tree; (B) MAPKAP1; (C) ZFX; and (D) ERC2. Nodes marked with white circles were recovered without significant support (PP < 0.95).

Divergence dating estimates across the StarBEAST2 phylogeny (Figure 4.5) are consistently younger than those inferred in the concatenated multi-locus (Chapter 2 + 3) and generally younger than the ddRADseq analyses. It is well-documented that the multispecies coalescent model typically recovers younger estimated divergence dates than concatenation owing to implicit differences in the two approaches (e.g., McCormack et al., 2011; Meyer et al., 2016; Ruane et al., 2014); coalescent-based species tree estimation accounts for genetic divergence that arose prior to speciation, and thus, provides more realistic estimates of divergence times, while concatenation assumes that all gene trees and the species tree are identical (Edwards et al., 2016) and it is likely to overestimate speciation ages (Burbrink & Pyron, 2011). In both the StarBEAST2 and ddRADseq analyses, *Callicebus* and *Plecturocebus* are estimated to have diverged in the late Miocene at *c.* 7.4 Ma (StarBEAST2 95% HPD = 6.0 − 9.0) which may be related to the amount of missing data for *C. personatus* in the ddRADseq dataset given that all other dating estimates are older in the ddRADseq timetree.

The most notable of the StarBEAST2 age estimates is the much younger divergence between the *P. moloch* group and *P. donacophilus* (2.62 Ma; 95% HPD = 1.8 − 3.5) than the concatenated multi-locus dating analyses (*c.* 4 Ma; not represented in the ddRADseq datasets), as well as the comparatively recent divergence between *P. c. caligatus* and *P. c. dubius* (0.2 Ma; 95% HPD = 0.07 − 0.3). The major *P. moloch* group clades are estimated to have diverged in the Pleistocene at *c.* 1.5 Ma (95% HPD = 1.2 − 2.0) and 1.2 Ma (95% HPD = 0.9 − 1.6), and all other taxa of the *P. moloch* group diverged between *c.* 0.5 − 0.9 Ma. Notably, the divergence between *P. hoffmannsi* and *P. cinerascens + P. miltoni* is estimated to have occurred at 0.68 Ma (95% HPD = 0.3 − 1.2). The earliest diverging lineage among *Callicebus* taxa, *C. nigrifrons*, is estimated to have diverged at the start of the Pleistocene, *c.* 2.4 Ma (95% HPD = 1.6 − 3.2), while *C. coimbrai* and *C. personatus* diverged at *c.* 1.1 Ma (95% HPD = 0.4 − 1.9). Finally, the *Cheracebus* vs. *Callicebus + Plecturocebus* divergence occurred in the late Miocene around 9.4 Ma (95% HPD = 7.8 − 11.0), and species of the genus *Cheracebus*, *C. lugens* and *C. purinus*, are estimated to have diverged at 1.23 Ma (95% HPD = 0.67 − 1.79). The full timetree with outgroups is presented in Appendix 3, Figure A3.10.

4.4.4 Bayesian clustering analyses

Bayesian clustering analyses (STRUCTURE) were performed using four of the ddRADseq datasets; P85 with all *P. moloch* group species (1944 SNPs); Pi85 with taxa of the Aripuanã-Tapajós clade (2943 SNPs); Pii85 with Eastern Amazonian taxa (1569 SNPs); and Piii85 with Western Amazonian taxa (1905 SNPs).

For the P85 dataset with all *P. moloch* group taxa, likelihood was maximised at $K = 10 – 12$ (see Figure 4.8 a), although the Evanno ΔK method strongly selected $K = 2$ with small peaks at $K = 3$ and $K = 10$ (see Figure 4.8 b), likely as a result of the evident hierarchical structure.
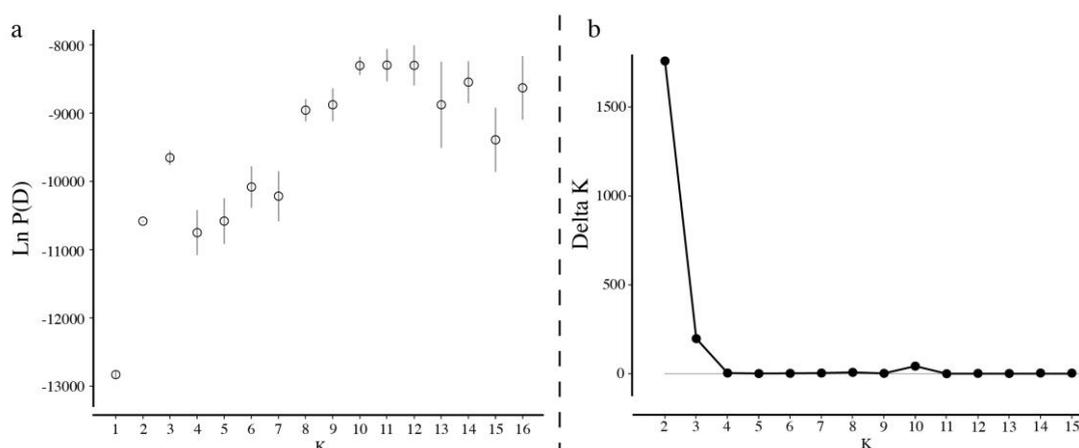
Figure 4.8 (a) Mean likelihood [LnP(D) ± SD] and (b) ΔK calculated for the P85 dataset (all *P. moloch* group taxa) from 6 independent runs for each value of K from 1 to 16.

When K = 2 is assumed, the genetic clusters correspond to the Western clade vs. the Eastern + Aripuanã-Tapajós clade (Figure 4.9) and all individuals are assigned perfectly to each group (*Q* value = 1.0). When a third cluster is added (K = 3), *P. hoffmannsi* and the Eastern clade taxa are also clearly differentiated, while *P. cinerascens* and *P. miltoni* show ancestry in both the *P. hoffmannsi* cluster (mean *Q* value = 0.82 / 0.77) and the Eastern clade cluster (mean *Q* value = 0.18 / 0.23). At K = 10, each of the five species of the Eastern and Aripuanã-Tapajós clades form near distinct clusters, and most individuals have very high membership coefficients (> 0.97) with the exception of the two *P. bernhardi* samples and one *P. miltoni* individual that share ancestry with the Western clade (*Q* value = 0.11 – 0.28) as well as one *P. cinerascens* clade B individual that shares ancestry with *P. bernhardi* (*Q* value = 0.11).
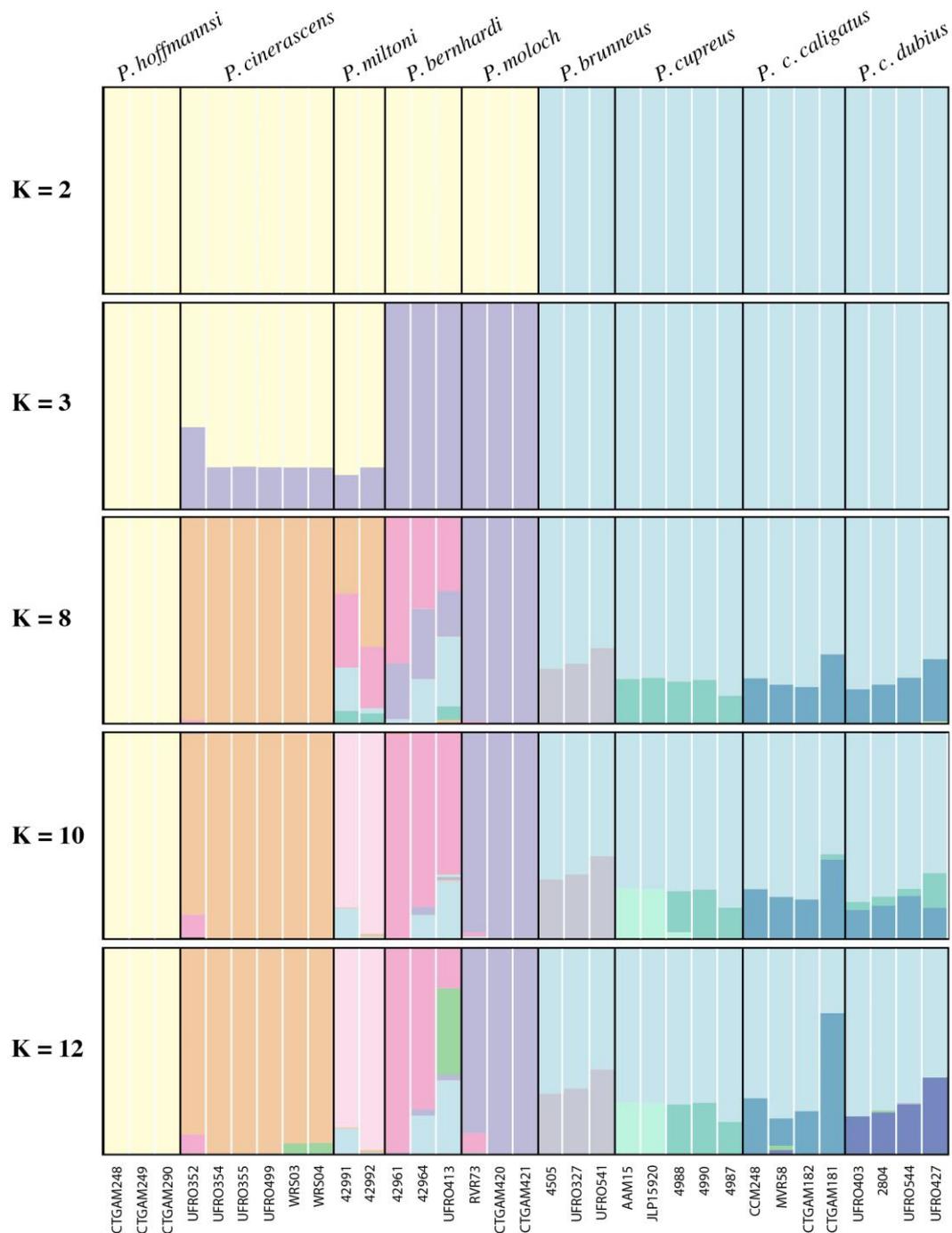
Figure 4.9 Genetic structure of the *P. moloch* group inferred from the P85 dataset (including all *P. moloch* group taxa) using Bayesian clustering analyses. Sample IDs are shown.

In contrast, all individuals of the Western clade show high proportions of ancestry in the same cluster (mean *Q* value = 0.67 – 0.78), and each taxon is only differentiated by a moderate proportion of ancestry. The remaining four clusters are assigned to the Western clade taxa as follows; (1) *P. brunneus* (mean *Q* value = 0.33); (2) *P. cupreus* clade A (mean *Q* value = 0.24); (3) *P. cupreus* clade B (mean *Q* value

= 0.19) and *P. c. dubius* (mean *Q* value = 0.07); and (4) *P. c. caligatus* (mean *Q* value = 0.25) and *P. c. dubius* (mean *Q* value = 0.16). When K = 12 is assumed, *P. c. dubius* shows a comparative proportion of ancestry in a distinct cluster (mean *Q* value = 0.25), rather than shared with *P. c. caligatus* or *P. cupreus* clade B. The *P. bernhardi* UFRO lineage is differentiated from *P. bernhardi* clade A with an intermediate level of ancestry (mean *Q* value = 0.42) in a cluster in which two *P. cinerascens* (clade A) individuals also show a small proportion of ancestry (mean *Q* value = 0.04), while *P.* cf. *moloch* shares ancestry with *P. bernhardi* clade A individuals (mean *Q* value = 0.09). Posterior probabilities for K = 10 − 12 are almost identical (mean LnP(D) ~ -8320), and although the variation between replicate runs is smallest at *K* = 10, the additional clusters assigned at *K* = 11 + 12 are informative and differentiate distinct lineages (*P. c. dubius* and *P. bernhardi* UFRO) in agreement with phylogenetic and geographic evidence. The strong hierarchical structure among *P. moloch* group taxa is evident: at low *K* values, the genetic clusters correspond largely to the major clades; at intermediate values, the deepest divergence within each clade is typically identified; and at higher values, all species and some intraspecific clades are differentiated to some extent. The Western clade taxa share a high proportion of ancestry at all *K* values, *P. cupreus* and *P. caligatus* are indistinguishable until *K* = 7 − 8, and it is not possible to differentiate *P. c. dubius* from *P. c. caligatus* until around *K* = 10.

4.4.4.1 Aripuanã-Tapajós clade (Pi85)

When only the Aripuanã-Tapajós clade taxa are included (Pi85 dataset), likelihood was maximised with relatively small variance between runs at *K* = 4 to 5 (LnP(D) = -14228 and -14237), while Δ*K* supported *K* = 2 with decreasing support for each added cluster (Figure 4.10 I.a + I.b). The difference in the most probable *K* between these methods may be a result of the comparatively deep divergence between *P. hoffmannsi* and *P. miltoni* + *P. cinerascens*, and thus, LnP(D) is considered above Δ*K*. At *K* = 4, *P. hoffmannsi* and *P. miltoni* are assigned to clusters and all *P. cinerascens* individuals share ancestry across the remaining two clusters (Figure 4.11), while at *K* = 5, the *P. cinerascens* clade (A + B) are differentiated showing a proportion of their ancestry in distinct clusters (mean Q value = 0.34). In addition to showing lower variance between runs, *K* = 5 is consistent with phylogenetic evidence and thus taken as the most likely *K*.
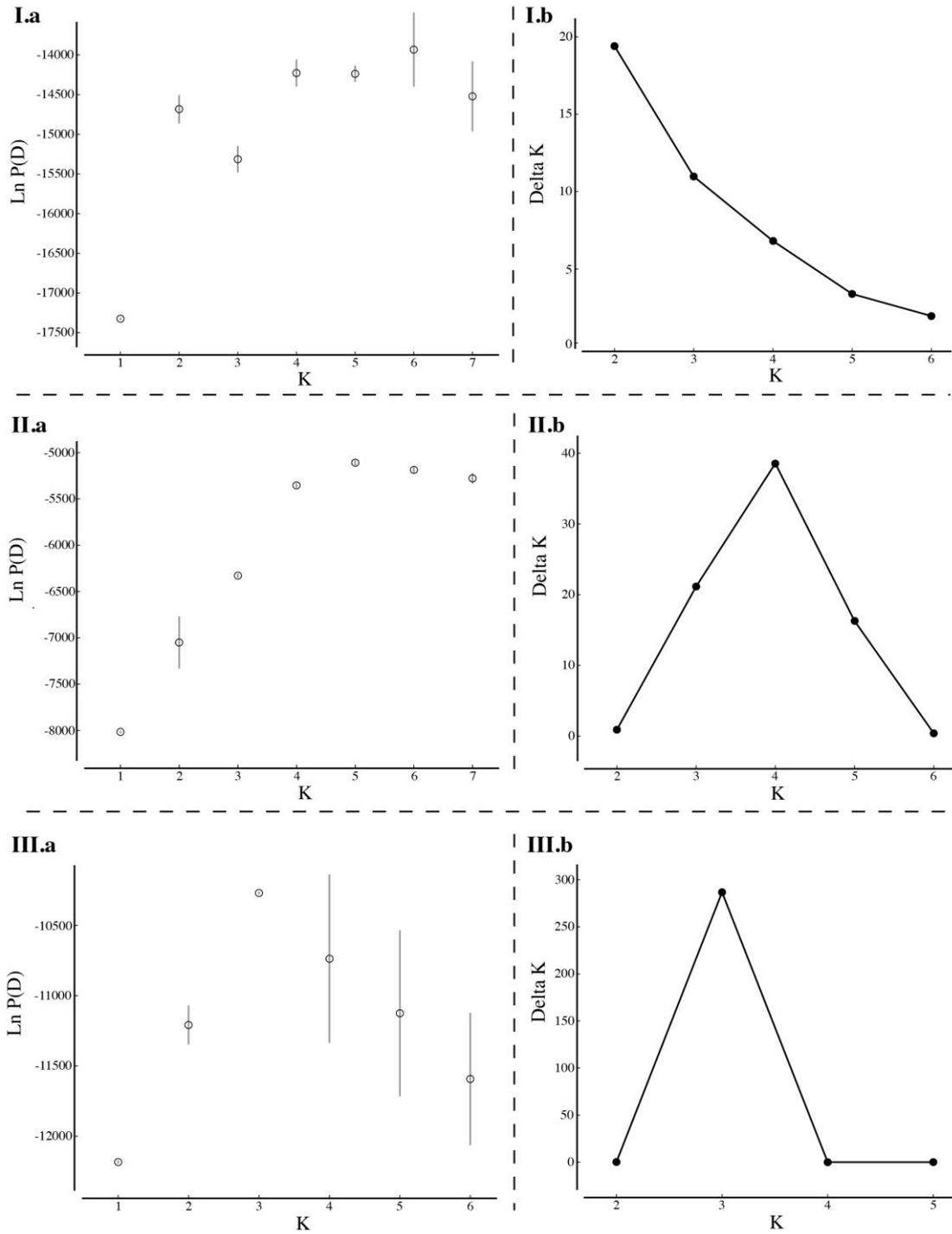
Figure 4.10 (a) Mean likelihood [LnP(D) ± SD] and (b) ΔK calculated for the subsampled datasets from 5 independent runs for each value of K from 1 to 6 or 7. Results for the Pi85 dataset are shown at the top (I: Aripuanã-Tapajós clade), Pii85 dataset are shown in the middle (II: Eastern clade), and Piii85 dataset are shown at the bottom (III: Western clade).
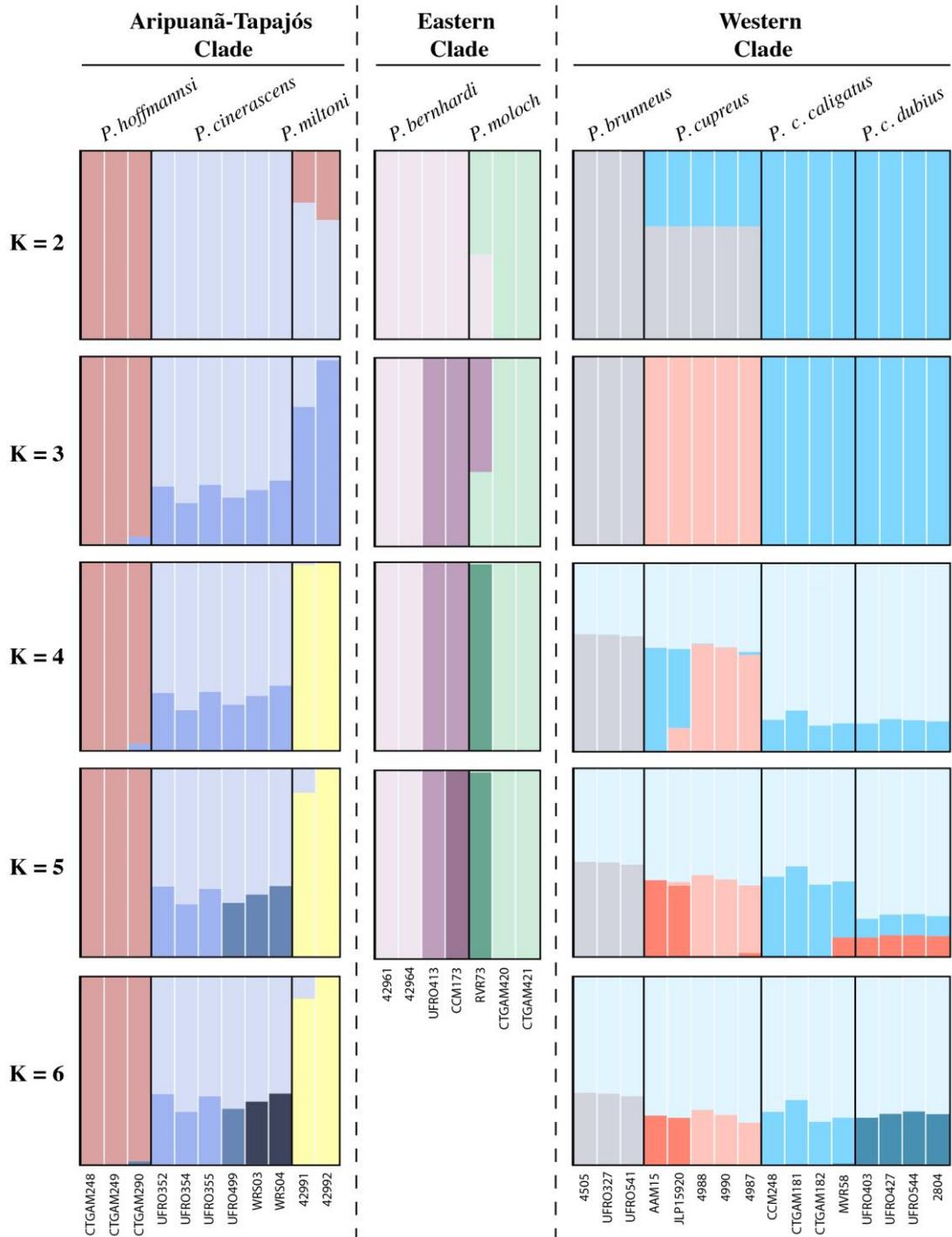
Figure 4.11 Genetic structure of the *P. moloch* group inferred from the subsampled datasets using Bayesian clustering analyses. Results for the Pi85 dataset are shown on the left (Aripuanã-Tapajós clade), Pii85 dataset are shown in the centre (Eastern clade), and Piii85 dataset are shown on the right (Western clade). Sample IDs are shown.

4.4.4.2 Eastern clade (Pii85)

For the analyses including only Eastern clade taxa (Pii85 dataset), likelihood was maximised at *K* = 5 (LnP(D) = -5110), although *K* = 4 and *K* = 6 have similar mean

LnP(D) values (-5355 and -5188) with equally low variance between replicate runs (see Figure 4.10 II.a). The Evanno ΔK method selected $K = 4$, but also showed support for $K = 3$ and $K = 5$ (see Figure 4.10 II.b). At $K = 4$, *P. moloch*, *P.* cf. *moloch*, *P. bernhardi* clade A, and *P. bernhardi* UFRO + CCM form distinct genetic clusters, while at $K = 5$, the *P. bernhardi* CCM lineage is distinguished from the *P. bernhardi* UFRO cluster (Figure 4.11). When six or seven clusters are assumed, the results follow $K = 5$ with the additional clusters assigned in tiny proportions to all individuals. For two out of three datasets used for phylogenetic inference, the UFRO and CCM individuals are recovered as successive sister lineages to the *P. bernhardi* clade A, which is more consistent with the individual assignments when five clusters are assumed, and thus, $K = 5$ is taken as the most likely. At lower values of K, *P.* cf. *moloch* (RVR73) shares almost equal ancestry with *P. moloch* and *P. bernhardi*, whereas in the analyses based on the P85 dataset, *P.* cf. *moloch* is almost indistinguishable from *P. moloch* until $K = 12$. Thus, for both the Aripuanã-Tapajós clade (Pi85) and Eastern clade (Pii85) subsampled analyses, evident substructure is recovered within some taxa that is concordant with phylogenetic evidence but not found in the overall dataset including all individuals (P85). The results from these analyses are otherwise strongly in agreement about the assignment of genetic clusters across these clades, and generally correspond well with taxonomic classification.

4.4.4.3 Western clade (Piii85)

When only the Western clade taxa are included (Piii85 dataset), both ΔK and LnP(D) suggest that $K = 3$ is the most likely (see Figure 4.10 III.a + III.b), with *P. brunneus, P. cupreus*, and *P. caligatus* each forming a distinct genetic cluster (Figure 4.11). The variation between runs increases and the likelihood decreases at higher values of K, and all individuals show some degree of shared ancestry in the same cluster, as in the P85 analyses. The assignment of individuals to clusters at $K = 5$ and $K = 6$ is generally strongly concordant with the results obtained for Western Amazonian taxa in the P85 analysis at $K = 10$ and $K = 12$, respectively, except *P. c. dubius* shares ancestry with *P. cupreus* clade A ($K = 5$, Piii85), rather than *P. cupreus* clade B ($K = 10$, P85). Although $K = 3$ is evidently the "true" K, interesting clustering patterns are uncovered at higher values of K, which may be informative. When four clusters are assumed, one cluster is shared among all individuals, while *P. caligatus* and *P. cupreus* clade A (i.e., AAM15 + JLP1590 samples) share most of their remaining

ancestry in same cluster (in differing proportions), and it is notable that this arrangement has a higher likelihood than when the *P. cupreus* lineages, *P. c. caligatus* and *P. c. dubius* show partial ancestry in distinct genetic clusters (i.e., $K = 6$). The latter is more consistent with the current taxonomic classification, while the former is defensible in a geographic context (i.e., all Western clade individuals from between the Rio Jutaí and Rio Madeira share a proportion of ancestry in one cluster).

To assess the possibility of underlying substructure shared between taxa, three additional datasets were assembled in pyRAD with one Western clade species selectively excluded as follows; *P. cupreus* + *P. brunneus* (2118 unlinked SNPs)*; P. caligatus* + *P. brunneus* (1889 unlinked SNPs)*;* and *P. caligatus* + *P. cupreus* (2136 unlinked SNPs). It is expected that *P. brunneus* and *P. cupreus* will be clearly differentiated at all values of *K* when *P. caligatus* is excluded from the dataset if their shared ancestry is owing to the relationship between both these species and *P. caligatus*, rather than a putative independent shared history between *P. brunneus* and *P. cupreus*. In this scenario, *P. caligatus* is expected to share ancestry with both *P. cupreus* and *P. brunneus* at higher *K* values, even when one of these lineages is excluded. These analyses were conducted from $K = 1 - 5$ under the same settings as the main Western clade STRUCTURE analysis (Piii85; see section 4.3.5). When only *P. cupreus* + *P. brunneus* are included, ΔK suggests that $K = 2$ is most likely, with each species assigned to a distinct cluster (Figure 4.12), while likelihood is maximised at $K = 4$ with the additional clusters assigned to the AAM15 and JLP15920 *P. cupreus* individuals.

No shared ancestry is recovered between *P. brunneus* and *P. cupreus*, regardless of the number of clusters assumed. When *P. caligatus* and either *P. brunneus* or *P. cupreus* are considered, the LnP(D) and ΔK suggest $K = 2$ is most likely, with each species forming a distinct genetic cluster, although likelihood is similar at $K = 3$. For both, when more than two clusters are assumed, all *P. caligatus* individuals show a proportion of shared ancestry with the other species (*P. brunneus* or *P. cupreus*). Further analyses such as alternatives to model-based clustering methods, for example, discriminated analysis of principle components (DAPC; Jombart et al., 2010), may shed further light on these results, as well as the inclusion of other taxa from the Western clade.
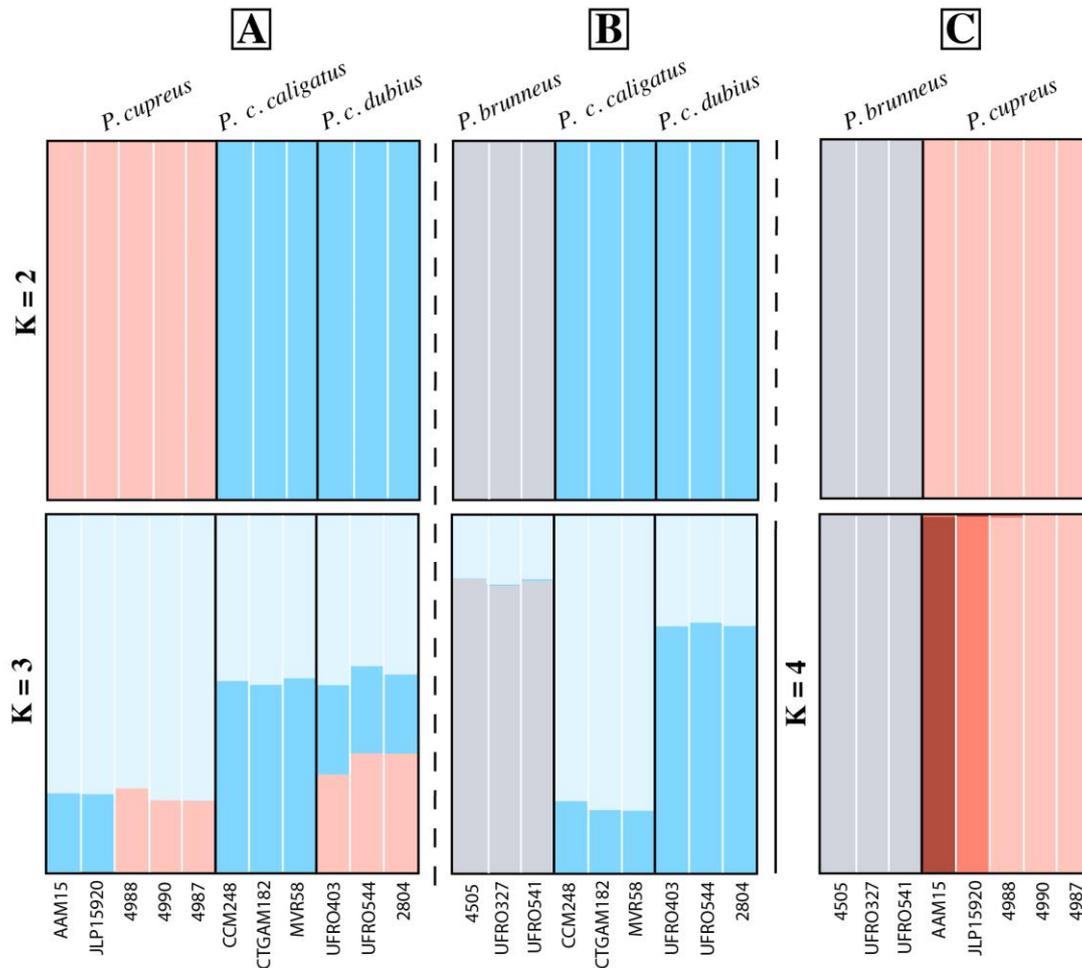
Figure 4.12 Genetic structure of the Western clade taxa of the *P. moloch* group inferred using Bayesian clustering analyses. Results for the analyses with only *P. cupreus* and *P. caligatus* are on the left (A), only *P. brunneus* and *P. caligatus* are in the middle (B), and only *P. cupreus* and *P. brunneus* are on the right (C).

Overall, these results are taken to suggest that although the three species form relatively well-defined genetic clusters, there is significant heterogeneity in the relationships among the Western clade taxa and among different lineages within and across these species, which may be explained by stochastic processes such as incomplete lineage sorting or more recent gene flow/introgression (particularly between adjacent taxa i.e., *P. caligatus* and the other two species). This is reflected in the topological conflict among the phylogenetic trees inferred based on the multi-locus vs. ddRADseq datasets regarding the species relationships among the Western clade taxa (i.e., if *P. cupreus* or *P. brunneus* is the sister taxon to *P. caligatus*), and also in the StarBEAST2 coalescent-based species tree.

148

Pairwise genetic distances between individuals of the Western clade calculated from the ddRADseq data show that *P. caligatus* shows closer genetic similarity to the other two species (see Appendix 3, Table A3.2); *P. cupreus* or *P. brunneus* individuals show the highest genetic distances (mean = 0.0022), with intermediate values between *P. caligatus* and *P. cupreus* or *P. brunneus* (mean = 0.0018 and 0.002, respectively), and members of the same species showed similar genetic distances on average for each of the species (mean = 0.0013). Pairwise distances between *P. c. dubius* individuals or *P. c. caligatus* individuals (mean = 0.0013 and 0.0012, respectively) are only slightly lower on average than those between *P. c. dubius* and *P. c. caligatus* (mean = 0.0014)*, and lower pairwise distances are often found across the taxa (e.g., 0.0011 between *P. c. dubius* UFRO403 and *P. c. caligatus* CTGAM182, and 0.0014 between the former individual and *P. c. dubius* UFRO544).

4.4.5 Introgression analyses

Four of the initial four-taxon D-statistic tests (1 – 39) returned significant Z-scores (alpha = 0.01) after correction for multiple testing, and inspection of the ABBA/BABA patterns identified *P. cinerascens* as the recipient lineage (P2) in each test. The results indicate that *P. bernhardi* shares derived alleles with *P. cinerascens* to the exclusion of *P. miltoni* and *P. hoffmannsi* (tests 3 + 8, Table 4.12), and the other member of the Eastern clade, *P. moloch*, also shares derived alleles with *P. cinerascens* to the exclusion of *P. hoffmannsi* (test 7). The final test suggests that *P. caligatus* and *P. cinerascens* are admixed (test 11), but only one combination of individuals (out of 108) is significant. Note, the P1 and P2 taxa are arranged such that P2 is the recipient lineage in the significant tests, i.e., ABBA > BABA in the four-taxon test and ABBBA > BABBA in the partitioned tests (Table 4.12). No significant four-taxon D-statistic tests detected admixture uniformly across all iterations of individuals, partially owing to the inclusion of multiple divergent lineages within many of the defined taxa.

Inspection of the results for each significant iteration in the above four tests indicated that specific *P. cinerascens* lineages may have been involved, and as such, a further set of four-taxon D-statistic tests were conducted in which the two *P. cinerascens* lineages (clade A and clade B) were differentiated (tests 40 – 44, Table 4.12). Tests 42 – 44 also served to rule out misleading D-statistics if the incorrect species tree was assumed (discussed below).

Table 4.12 Four-taxon D-statistic tests for admixture. Taxa are arranged such that ABBA>BABA. Bold indicates significance at alpha = 0.01 after correction for multiple testing.

Test no.	P1 taxon	P2 taxon	P3 taxon	Range Z	Sign./total (pre-correct)[1]	No. loci
1	*P. miltoni*	*P. cinerascens*	*P. hoffmannsi*	0.0 - 3.1	0/36 (1)	78 - 566
2	*P. miltoni*	*P. cinerascens*	*P. moloch*	0.0 - 3.4	0/36 (3)	97 - 558
3	*P. miltoni*	*P. cinerascens*	*P. bernhardi*	0.1 - **4.1**	**3/48** (11)	62 - 577
4	*P. miltoni*	*P. cinerascens*	*P. brunneus*	0.0 - 1.9	0/36 (0)	71 - 555
5	*P. miltoni*	*P. cinerascens*	*P. cupreus*	0.0 - 2.3	0/60 (0)	80 - 584
6	*P. miltoni*	*P. cinerascens*	*P. caligatus*	0.0 - 2.1	0/72 (0)	89 - 557
7	*P. hoffmannsi*	*P. cinerascens*	*P. moloch*	0.6 - **5.1**	**3/54** (23)	252 - 568
8	*P. hoffmannsi*	*P. cinerascens*	*P. bernhardi*	0.5 - **7.2**	**19/72** (35)	84 - 604
9	*P. hoffmannsi*	*P. cinerascens*	*P. cupreus*	0.1 - 3.4	0/90 (12)	117 - 617
10	*P. hoffmannsi*	*P. cinerascens*	*P. brunneus*	0.0 - 3.1	0/54 (2)	176 - 574
11	*P. hoffmannsi*	*P. cinerascens*	*P. caligatus*	0.2 - **4.3**	**1/108** (26)	189 - 582
12	*P. hoffmannsi*	*P. miltoni*	*P. moloch*	0.4 - 2.7	0/18 (2)	143 - 568
13	*P. hoffmannsi*	*P. miltoni*	*P. bernhardi*	0.1 - 2.6	0/24 (1)	73 - 585
14	*P. hoffmannsi*	*P. miltoni*	*P. brunneus*	0.6 - 3.1	0/18 (3)	78 - 570
15	*P. hoffmannsi*	*P. miltoni*	*P. cupreus*	0.2 - 2.9	0/30 (5)	106 - 584
16	*P. hoffmannsi*	*P. miltoni*	*P. caligatus*	0.3 - 3.3	0/36(7)	121 - 562
17	*P. bernhardi*	*P. moloch*	*P. hoffmannsi*	0.0 - 2.2	0/36 (0)	172 - 602
18	*P. bernhardi*	*P. moloch*	*P. miltoni*	0.0 - 1.3	0/24 (0)	92 - 593
19	*P. bernhardi*	*P. moloch*	*P. cinerascens*	0.0 - 2.7	0/72 (1)	94 - 613
20	*P. bernhardi*	*P. moloch*	*P. brunneus*	0.0 - 2.0	0/36 (0)	89 - 623
21	*P. bernhardi*	*P. moloch*	*P. cupreus*	0.1 - 2.8	0/60 (2)	123 - 659
22	*P. bernhardi*	*P. moloch*	*P. caligatus*	0.0 - 2.7	0/72 (2)	170 - 643
23	*P. caligatus*	*P. cupreus*	*P. hoffmannsi*	0.0 - 2.0	0/90 (0)	182 - 623
24	*P. caligatus*	*P. cupreus*	*P. cinerascens*	0.0 - 3.0	0/180 (3)	137 - 637
25	*P. caligatus*	*P. cupreus*	*P. miltoni*	0.0 - 3.0	0/60 (2)	132 - 578
26	*P. caligatus*	*P. cupreus*	*P. bernhardi*	0.0 - 3.1	0/120 (1)	123 - 685
27	*P. caligatus*	*P. cupreus*	*P. moloch*	0.0 - 2.4	0/90 (0)	230 - 629
28	*P. caligatus*	*P. cupreus*	*P. brunneus*	0.0 - 3.1	0/90 (2)	141 - 620
29	*P. brunneus*	*P. caligatus*	*P. hoffmannsi*	0.0 - 3.0	0/54 (1)	184 - 593
30	*P. brunneus*	*P. caligatus*	*P. cinerascens*	0.0 - 2.7	0/108 (2)	148 - 587
31	*P. brunneus*	*P. caligatus*	*P. miltoni*	0.0 - 2.3	0/36 (0)	94 - 565
32	*P. brunneus*	*P. caligatus*	*P. bernhardi*	0.0 - 3.4	0/72 (3)	78 - 636
33	*P. brunneus*	*P. caligatus*	*P. moloch*	0.0 - 2.1	0/54 (0)	218 - 568
34	*P. brunneus*	*P. caligatus*	*P. cupreus*	0.0 - 2.9	0/90 (1)	141 - 620
35	*P. brunneus*	*P. cupreus*	*P. hoffmannsi*	0.0 - 3.4	0/45 (1)	119 - 609
36	*P. brunneus*	*P. cupreus*	*P. cinerascens*	0.0 - 3.2	0/90 (7)	95 - 610
37	*P. brunneus*	*P. cupreus*	*P. miltoni*	0.2 - 2.1	0/30 (0)	88 - 578
38	*P. brunneus*	*P. cupreus*	*P. bernhardi*	0.0 - 3.4	0/60 (4)	66 - 663
39	*P. brunneus*	*P. cupreus*	*P. moloch*	0.0 - 2.0	0/45 (0)	147 - 595
40	*P. hoffmannsi*	*P. cinerascens* (A)	*P. bernhardi*	0.5 - 3.5	0/36 (5)	174 - 603
41	*P. hoffmannsi*	*P. cinerascens* (B)	*P. bernhardi*	1.8 - **7.2**	**20/36** (30)	84 - 604
42	*P. cinerascens* (A)	*P. cinerascens* (B)	*P. bernhardi*	0.5 - **7.0**	**7/36** (17)	106 - 639
43	*P. cinerascens* (A)	*P. cinerascens* (B)	*P. moloch*	0.4 - **3.6**	**1/27** (8)	283 - 591
44	*P. cinerascens* (B)	*P. cinerascens* (A)	*P. caligatus*	0.0 - 2.4	0/54 (0)	224 - 610

[1] Significant tests over possible sampled individuals

When introgression is suggested between *P. cinerascens* and an Eastern Amazonian taxon (tests 3, 7 + 8,), the iterations with *P. cinerascens* clade B individuals as the P3 lineage almost always had higher Z-scores formed nearly all of the significant iterations across these tests. This is demonstrated in tests 40 + 41 which equate to test 8 split into two separate runs whereby P3 is defined as only *P. cinerascens* clade A or clade B, respectively. Only *P. cinerascens* clade B individuals share a significant proportion of derived alleles with *P. bernhardi* to the exclusion of *P. hoffmannsi* (note, one less iteration is significant in test 8 than test 41 because more combinations are tested). When P1 and P2 are each set as the *P. cinerascens* lineages, *P. cinerascens* clade B individuals share derived alleles with *P. bernhardi* to the exclusion of *P. cinerascens* clade A (test 42), as well as with *P. moloch* to lesser extent (test 43). Among *P. cinerascens* clade B individuals, the highest Z-scores in each test were typically recovered for iterations including UFRO352.

In contrast, significant iterations in test 11 generally include *P. cinerascens* clade A as the P3 lineage, however, *P. cinerascens* clade A individuals don't share derived alleles with *P. caligatus* to the exclusion of *P. cinerascens* clade B (test 44). No significant D-statistics were recovered when intraspecific lineages among the Western clade taxa were differentiated (see Appendix 3, Table A3.3). Every possible iteration of individuals in the significant four-taxon tests (tests 7, 8, 11, 40, 41, 42. Table 4.12) showed more ABBA site patterns than BABA (positive D), and almost all iterations in test 3, thus although the degree of asymmetry varied, the pattern of shared derived alleles between the P2 and P3 lineages to the exclusion of the P1 is consistent across all individuals of each taxon in these tests. Furthermore, other tests which have many iterations of individuals with moderately high but non-significant Z-scores show consistent ABBA > BABA patterns across nearly all combinations of samples, especially for tests involving the Eastern clade as the P3 lineage and the P2 lineage defined as *P. cinerascens/P. miltoni* (e.g., tests 2 + 12).

Partitioned (five-taxon) D-statistic tests were performed to assess two putative scenarios: (I) admixture between Eastern clade taxa and *P. cinerascens*; and (II) admixture between *P. caligatus* and *P. cinerascens*. No significant results (Z-score > 2.55) were recovered for latter (II) scenario in the partitioned D-statistic tests (tests 60 – 68, Table 4.13). In fact, the highest (nearly significant) Z-score obtained indicates admixture between *P. cinerascens* clade B and *P. brunneus* rather than *P. cinerascens* clade A and *P. caligatus* (negative D, $Z_2 = 2.52$, test 65). Thus, overall, the results are

inconclusive regarding whether introgression has occurred between the Western and Aripuanã-Tapajós clades, the lineages involved, and the direction of gene flow. Future analyses with greater sampling, including more loci, individuals and Western clade species, may be required to reconstruct the history of introgression among these clades.

In contrast, partitioning shared versus uniquely derived alleles among the Eastern and Aripuanã-Tapajós taxa reveals consistent support for admixture between *P. cinerascens* clade B and *P. bernhardi*. For these tests, *P. bernhardi* clade A (Guaporé Biological Reserve) and UFRO (Machadinho D'Oeste) lineages, as well as *P. moloch* (east of the Tapajós and Teles Pires rivers) and *P* cf. *moloch* (Alta Floresta), are defined as the P3 taxa. Test $45 - 50$ (Table 4.13) found that *P. bernhardi* clade A shares uniquely derived alleles with *P. cinerascens* clade B (relative to *P. cinerascens* clade A, *P. hoffmannsi*, and *P. miltoni*) that are not shared with *P. moloch* or *P. bernhardi* UFRO (significant D_1). Derived alleles which arose in the ancestor of *P. bernhardi* and *P. moloch* are also shared with *P. cinerascens* clade B to the exclusion of the other Aripuanã-Tapajós lineages (significant D_{12}, tests $45 - 50$). However, *P. moloch* and *P. bernhardi* UFRO do not share a set of uniquely derived alleles with *P. cinerascens* clade B which are not also shared with *P. bernhardi* clade A (non-significant D_2, tests $45 - 50$).

Together, these results indicate that introgression has occurred from *P. bernhardi* clade A into *P. cinerascens* clade B, and that the set of derived alleles shared between *P. cinerascens* clade B and the other Eastern Amazonian taxa (*P. moloch* or *P. bernhardi* UFRO) most likely arose in the ancestor to the Eastern clade, rather than as a result of independent admixture between each of the lineages. In this scenario, it is also expected that *P. cf. moloch* (as the sister taxon to *P. moloch*) will share the set of derived alleles with *P. cinerascens* clade B that arose in the ancestor of the Eastern clade. The results of tests $51 - 53$ (Table 4.13) are in agreement with this expectation, such that both *P. moloch* and *P.* cf. *moloch* share the same set of derived alleles with *P. cinerascens* clade B (significant D_{12}, non-significant D_1 and D_2), although D_{12} in test 53 is slightly below significance ($Z_{12} = 2.39$).

Table 4.13 Partitioned D-statistic tests for admixture. For significant tests, taxa are arranged such that introgression of shared P3 alleles (D_{12}) is into P2 (ABBBA>BABBA). Bold indicates significant Z-scores at alpha = 0.01.

	Test No.	P1 taxon	P2 taxon	P3₁ taxon	P3₂ taxon	D_{12}	D_1	D_2	Z_{12}	Z_1	Z_2	No. loci
I	45	*P. cinerascens* (A)	*P. cinerascens* (B)	*P. bernhardi* (A)	*P. moloch*	0.51	0.70	0.28	**5.81**	**6.49**	1.49	426
	46	*P. hoffmannsi*	*P. cinerascens* (B)	*P. bernhardi* (A)	*P. moloch*	0.49	0.56	0.18	**5.51**	**3.82**	0.97	405
	47	*P. miltoni*	*P. cinerascens* (B)	*P. bernhardi* (A)	*P. moloch*	0.33	0.50	-0.08	**3.52**	**3.96**	0.43	401
	48	*P. cinerascens* (A)	*P. cinerascens* (B)	*P. bernhardi* (A)	*P. bernhardi* (UFRO)	0.46	0.67	-0.02	**4.19**	**4.54**	0.09	217
	49	*P. hoffmannsi*	*P. cinerascens* (B)	*P. bernhardi* (A)	*P. bernhardi* (UFRO)	0.57	0.59	0.13	**6.47**	**3.71**	0.53	219
	50	*P. miltoni*	*P. cinerascens* (B)	*P. bernhardi* (A)	*P. bernhardi* (UFRO)	0.49	0.59	0.02	**5.01**	**3.50**	0.10	200
	51	*P. cinerascens* (A)	*P. cinerascens* (B)	*P. moloch*	*P. cf. moloch*	0.37	-0.21	0.25	**3.58**	0.88	0.92	391
	53	*P. hoffmannsi*	*P. cinerascens* (B)	*P. moloch*	*P. cf. moloch*	0.36	-0.29	-0.07	**3.48**	1.14	0.28	354
	53	*P. miltoni*	*P. cinerascens* (B)	*P. moloch*	*P. cf. moloch*	0.24	-0.18	-0.20	2.39	0.84	0.82	386
	54	*P. hoffmannsi*	*P. cinerascens* (A)	*P. bernhardi* (A)	*P. moloch*	0.13	-0.08	-0.04	1.17	0.40	0.16	394
	55	*P. miltoni*	*P. cinerascens* (A)	*P. bernhardi* (A)	*P. moloch*	0.08	-0.12	-0.17	0.73	0.57	0.79	417
	56	*P. hoffmannsi*	*P. cinerascens* (A)	*P. bernhardi* (A)	*P. bernhardi* (UFRO)	0.18	-0.01	0.27	1.35	0.04	1.20	197
	57	*P. miltoni*	*P. cinerascens* (A)	*P. bernhardi* (A)	*P. bernhardi* (UFRO)	0.19	0.28	-0.18	1.49	1.08	0.75	199
	58	*P. hoffmannsi*	*P. cinerascens* (A)	*P. moloch*	*P. cf. moloch*	0.17	0.39	-0.14	1.43	1.28	0.50	353
	59	*P. miltoni*	*P. cinerascens* (A)	*P. moloch*	*P. cf. moloch*	0.08	-0.24	-0.29	0.72	1.24	1.11	416
II	60	*P. hoffmannsi*	*P. cinerascens* (A)	*P. c. caligatus*	*P. c. dubius*	0.25	0.08	0.12	2.04	0.22	0.32	418
	61	*P. hoffmannsi*	*P. cinerascens* (A)	*P. c. caligatus*	*P. cupreus*	0.29	0.24	-0.01	2.43	0.70	0.03	414
	62	*P. hoffmannsi*	*P. cinerascens* (A)	*P. c. caligatus*	*P. brunneus*	0.05	0.37	-0.24	0.33	0.64	0.67	236
	63	*P. cinerascens* (B)	*P. cinerascens* (A)	*P. c. caligatus*	*P. c. dubius*	-0.19	0.25	0.00	1.45	0.56	0.00	429
	64	*P. cinerascens* (B)	*P. cinerascens* (A)	*P. c. caligatus*	*P. cupreus*	-0.20	0.57	-0.83	1.67	2.03	2.47	433
	65	*P. cinerascens* (B)	*P. cinerascens* (A)	*P. c. caligatus*	*P. brunneus*	-0.34	-0.13	-0.81	2.08	0.24	2.52	245
	66	*P. miltoni*	*P. cinerascens* (A)	*P. c. caligatus*	*P. c. dubius*	0.02	-0.56	0.54	0.13	1.88	1.60	413
	67	*P. miltoni*	*P. cinerascens* (A)	*P. c. caligatus*	*P. cupreus*	0.00	-0.19	-0.50	0.01	0.67	1.39	420
	68	*P. miltoni*	*P. cinerascens* (A)	*P. c. caligatus*	*P. brunneus*	-0.22	0.35	-0.26	1.50	1.06	0.73	262

Finally, if no recent significant gene flow has occurred between the *P. cinerascens* lineages (clade A and clade B), then none of the Eastern clade taxa are expected to share derived alleles with *P. cinerascens* clade A, which is the exact result obtained (tests 54 – 59). Thus, for scenario (I), all partitioned D-statistic tests are in agreement, and together they strongly suggest that *P. cinerascens* clade B and *P. bernhardi* clade A individuals share uniquely derived alleles to the exclusion of other members of their clades. The slight negative D_1 and D_2 in some of the latter tests (51 – 59) may be an indication of a complex history of introgression between these clades involving other taxa such as *P. moloch* and *P. miltoni*.

D-statistic tests were originally applied to whole genome data (Green et al., 2010), and most examples to date using reduced representation genome-wide data typically included a much greater number of shorter loci (e.g., Chattopadhyay et al., 2016; Eaton et al., 2015). Despite the reduced statistical power in the present analyses owing to the comparatively low number of loci, significant results were obtained that are strongly concordant across the tests and with independent sources of evidence (see section 4.4.6). In combination, the conservative alpha value (0.01) and the moderate number of included loci may lead to a lack of significance when introgression has occurred, especially in the case of more ancient admixture, however, the conservative alpha value also serves to minimise spurious results if the discordant site patterns are found at few loci.

A significant D-statistic can be recovered in the absence of introgression if an incorrect species tree is assumed (i.e., if the P3 lineage forms a clade with the P2 lineage, to the exclusion of P1) as the P2 and P3 lineages will share more derived alleles owing to their closer shared history. As such, some of the D-statistics tests may be misleading if the species relationships assumed based on the ddRADseq phylogeny do not reflect the true species tree. For example, if the Aripuanã-Tapajós clade is not monophyletic (i.e., *P. hoffmannsi* is not the sister taxon to *P. cinerascens* and *P. miltoni*), then the D-statistic could be detecting shared alleles owing to a more recent shared history between *P. bernhardi* and *P. cinerascens*, rather than introgression between these species. Relatively high Z-scores are found for some iterations of individuals in all four-taxon tests with P1 defined as *P. hoffmannsi* and P2 as *P. cinerascens* or *P. miltoni* (e.g., see the number of significant iterations pre-correction, tests 7 – 16, Table 4.12), and this could be an indication that *P. hoffmannsi*, *P. cinerascens* and *P. miltoni* do not form a clade to the exclusion of the Eastern or

Western Amazonian taxa. This scenario, however, does not explain the significant D-statistics when P1 and P2 are defined as the two *P. cinerascens* lineages (clade A and B), or when P1 is *P. miltoni,* and it is unlikely that these arrangements violate the true species tree. *Plecturocebus cinerascens* and *P. miltoni* could also share a set of derived alleles with the Western clade to the exclusion of *P. hoffmannsi* owing to a complex history of gene flow between *P. cinerascens/P. miltoni* and the Eastern taxa and the closer shared history between the Eastern and Western clades.

4.4.6 Mitochondrial introgression

The new mitochondrial sequences (CYTB and COI) obtained for the *P. cinerascens* clade A UFRO354 individual confirm it has a *P. bernhardi* mitochondrial genome, which was suspected based on the classification of this sample as *P. bernhardi* by Carneiro et al. (2016). Inspection of the alignments indicates that UFRO354 shows 99.84% sequence identity (1 bp change) for COI and 99.39% sequence identity (7 bp changes) for CYTB with *P. bernhardi* clade A individuals (42960, 42961, 42964). In the maximum-likelihood tree inferred based on the mitochondrial loci, UFRO354 is sister to *P. bernhardi* clade A samples and nested within *P. bernhardi* (Figure 4.13). These results are in strong agreement with the conclusions drawn from the D-statistic tests and provide an independent source of evidence that *P. cinerascens* clade B is admixed with *P. bernhardi*, and specifically, most likely with individuals relatively closely related to the *P. bernhardi* clade A specimens in this study.

Figure 4.14 shows the collection localities for the *P. cinerascens, P. miltoni,* and *P. bernhardi* samples included in the mitochondrial and ddRADseq datasets. The putative donor lineage individuals, *P. bernhardi* clade A, were collected in the Guaporé Biological Reserve to the west of São Francisco do Guaporé, a considerable distance from all other *P. bernhardi* specimens and outside the known geographic distribution of this species. These are the closest *P. bernhardi* specimens included in this study to the *P. cinerascens* clade B individuals, followed by the *P. bernhardi* UFRO specimen. The admixed *P. cinerascens* clade B specimens (UFRO352, 354, 355) were also collected outside the known geographic distribution of the species, at the Rondon II dam between the left bank of upstream Rio Roosevelt and the two rivers that form the Rio Jiparaná (Rios Barão de Melgaço and Pimenta Bueno).
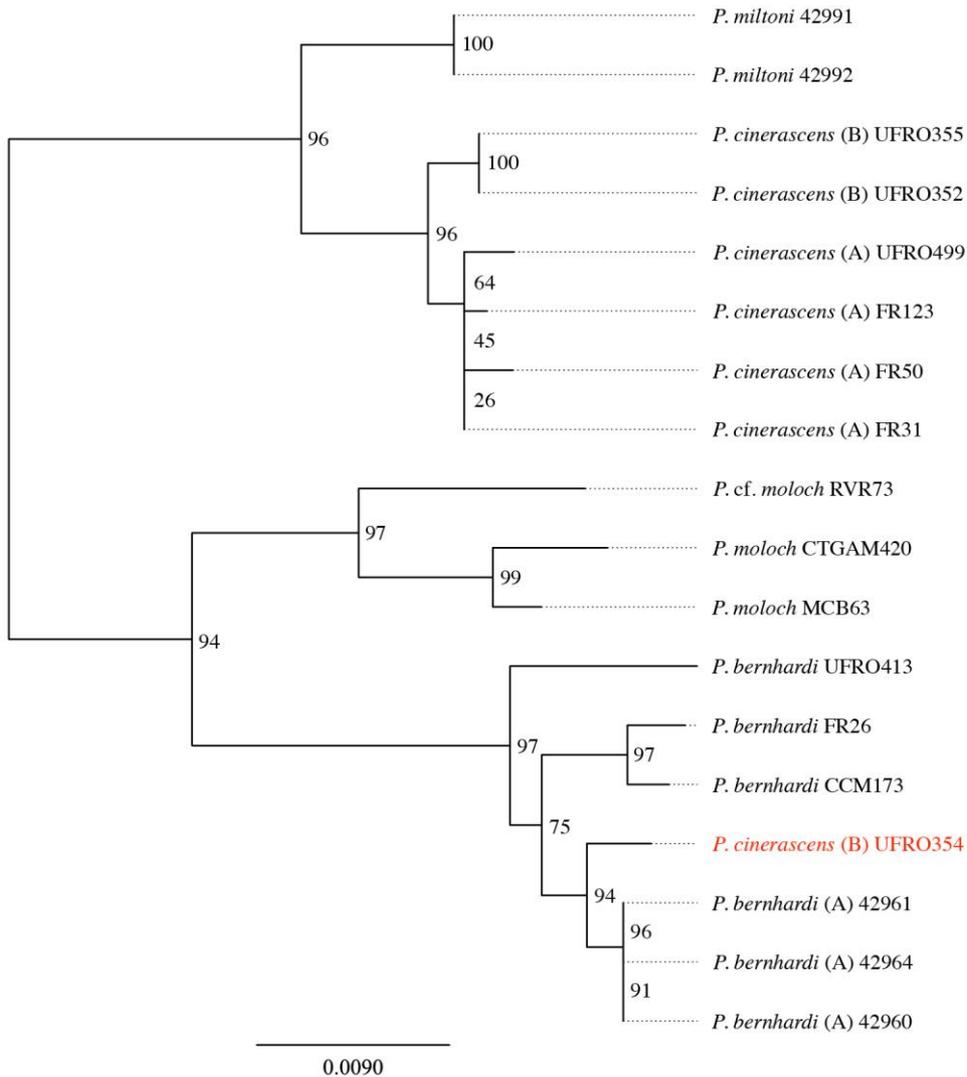
Figure 4.13 Maximum likelihood tree inferred with the mitochondrial data. The admixed *P. cinerascens* clade B individual with a *P. bernhardi* mitochondrial genome is highlighted in red.

This locality is at the edge of the southern tip of the proposed range for *P. bernhardi*, which was delineated based on these major rivers (Van Roosmalen et al., 2002), and *P. bernhardi* individuals have been recorded near to this location (around Cacoal and Pimenta Bueno, Ferrari et al., 2000). Thus, it is likely that admixture between these species occurred when *P. cinerascens* dispersed over or around the Rios Aripuanã and Roosevelt into the range of *P. bernhardi*, and in fact, may still be an ongoing process. This further suggests that the D-statistic tests may be detecting more recent introgression among the *P. moloch* group taxa included.
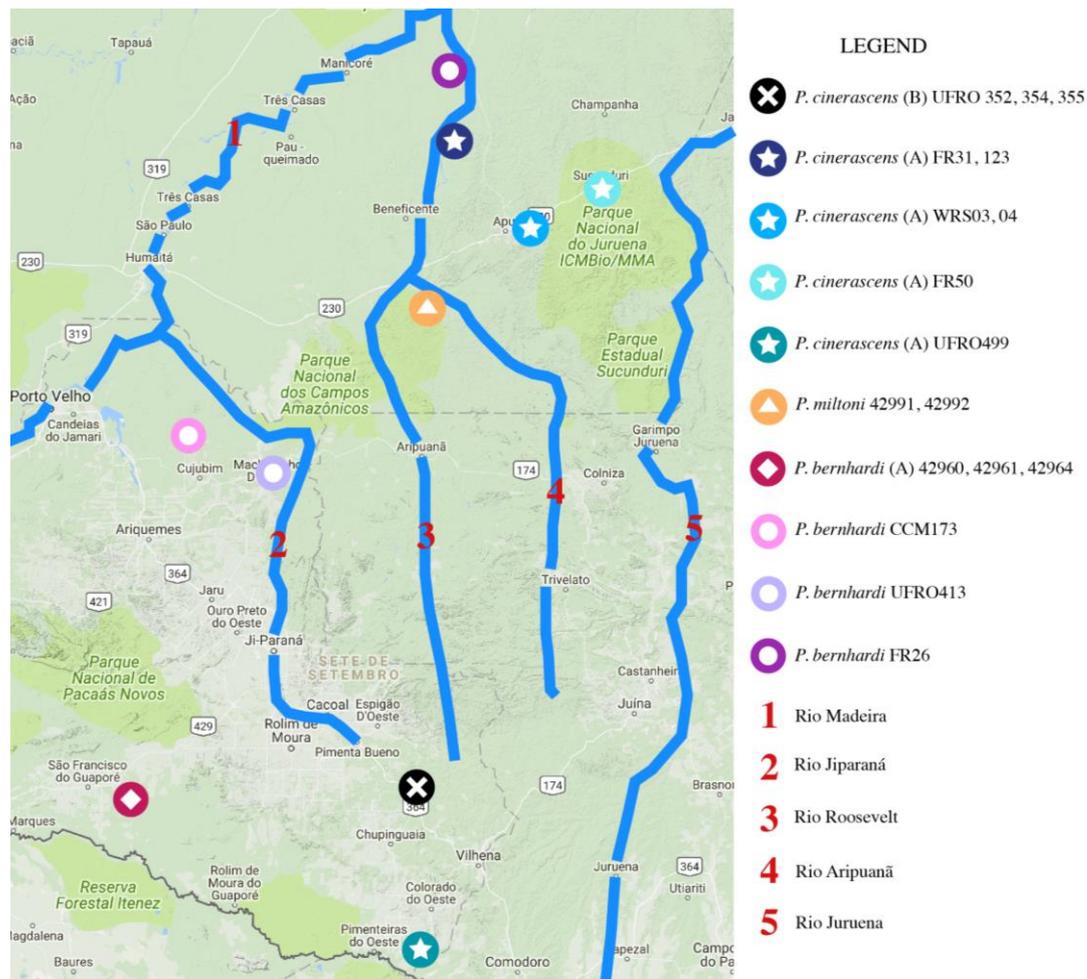
Figure 4.14 Collection localities for the *P. cinerascens, P. bernhardi* and *P. miltoni* samples included in this study. Major rivers are shown approximately.

Although UFRO354 is the specimen with the *P. bernhardi* mitochondrial genome, UFRO352 is more significantly admixed based on inspection of the Z-scores for each iteration of individuals in the four-taxon D-statistic tests. In the Bayesian clustering analyses including all *P. moloch* group taxa (P85 dataset), the same *P. cinerascens* clade B individual (UFRO352) shows shared ancestry (Q-value = 0.11) with *P. bernhardi* at $K = 10 + 12$ (see Figure 4.9). Furthermore, *P. cinerascens* clade B samples have much lower genetic distances (based on the ddRADseq data, see Appendix 3, Table A3.4) than *P. cinerascens* clade A to all *P. bernhardi* individuals (mean = 0.0028 vs. 0.0035), with UFRO352 showing the lowest genetic distances (mean = 0.0026). Two of the admixed *P. cinerascens* individuals are also the most heterozygous of all samples included in the ddRADseq dataset (UFRO352, H = 0.0030; UFRO355, H = 0.0031), while the third specimen is among the most heterozygous (UFRO354, H = 0.0026).

Images obtained of the UFRO354 specimen (Figure 4.15) indicate that it resembles typical *P. cinerascens* in the overall grey colour but with some small notable differences; creamy-white hairs on the hands, wrists and, to a lesser extent, the toes, as well as the tip of the tail; reddish-brown pelage on the dorsum; and creamy-white hairs on the chin and encircling the face. The white hands, wrists, tail tip, and toes (sometimes), as well as reddish-brown dorsal pelage, are typically found in *P. bernhardi* specimens and, therefore, some of these phenotypic differences may be a consequence of admixture. Overall, there is consistent support regarding introgressive hybridisation between *P. cinerascens* and *P. bernhardi*, and taken together, these results strongly suggest that *P. cinerascens* clade B individuals are admixed and they should be properly listed as *P. cinerascens* x *P. bernhardi*.

Figure 4.15 Dorsal (top) and ventral (bottom) images of the admixed *P. cinerascens* UFRO354 individual.

4.5 Discussion

The ddRADseq data matrices assembled in this study contained up to ~ 3K loci, with a maximum concatenated length of ~ 930K bp and up to ~ 15K parsimony informative sites, a massive increase in comparison to the data generated for the multi-locus phylogeny in Chapter 2 (22 loci, ~ 14K bp).

The assembled ddRADseq datasets contained between 33% and 56% missing data (see Table 4.9), with the outgroup samples (Pitheciinae) particularly poorly represented. Both phylogenetic distance and the mean number of raw reads significantly predicted the number of shared loci between taxa, and thus, both sequencing effort and the inclusion of disparate lineages (owing to "locus dropout") had an impact on the amount of missing data in the final matrices. Recent studies, however, have suggested that large amounts of missing data in RADseq data matrices had little impact on phylogenetic inference, with both smaller near-complete datasets and larger sparse datasets recovering similar topologies and the latter showed stronger statistical support (e.g., Eaton et al., 2015; Hou et al., 2015; Rubin et al., 2012; Wagner et al., 2013). RADseq data matrices with more loci and more missing data have been shown to provide greater statistical power and more biologically relevant information than smaller matrices with minimal missing data (Chattopadhyay et al., 2016; Huang & Knowles, 2016).

4.5.1 Phylogeny and dating estimates: concordance and conflict

Our ddRADseq phylogenetic analyses recovered a strongly supported topology for Callicebinae species that is consistent across the four datasets used for phylogenetic inference. All ddRADseq analyses support the monophyly of the three major clades within the *P. moloch* group, as well as the sister relationship between the Eastern and Western clades, in agreement with the phylogeny based on seven concatenated loci in Chapter 3. Overall, the largest conflicts among all competing phylogenetic hypotheses inferred based on the multi-locus or ddRADseq data (Chapters 2, 3 + 4) are regarding the relationship between Aripuanã-Tapajós clade taxa (*P. hoffmannsi, P. cinerascens, P. miltoni*) and the rest of the *P. moloch* group, as well as those among Western clade taxa. This uncertainty is strongly reflected in the most probable species tree topologies and the topological incongruence among the gene trees under the multispecies coalescent model (StarBEAST2; see Figure 4.6, 4.7). Despite this

uncertainty, the Aripuanã-Tapajós clade is recovered as monophyletic in the coalescent-based analysis of multi-locus data and the ddRADseq analyses, as well as based on the seven concatenated loci (nuclear and mitochondrial) included in Chapter 3. Both vastly increasing the quantity of data (ddRAD) or employing a more realistic model (multispecies coalescent model) resulted in greater resolution of species relationships among the *P. moloch* group taxa than concatenation of the 22 loci in Chapter 2 (which poorly supported *P. cinerascens* + *P. miltoni* as sister to the Eastern clade). However, the incongruence between the results obtained based on the concatenation of 22 vs. 7 of these loci (Chapter 2 vs. Chapter 3) suggests that the amount of informative sites nor the use of concatenation solely explain the conflicting results, although other factors may have also played a role such as the individuals/taxa sampled or potentially model misspecification (e.g., over-partitioning).

Although low levels of variation may also account for some of the incongruence among the StarBEAST2 gene trees, overall, these results suggest that there is significant gene tree heterogeneity owing to gene flow between species of the *P. moloch* group (e.g., admixture between *P. cinerascens* and *P. bernhardi*), or stochastic processes such as incomplete lineage sorting (Edwards, 2009; Maddison, 1997), presenting a challenge to the reconstruction of species relationships across this clade. Given the enormous difference in the quantity of data between the ddRADseq and StarBEAST2 datasets (~2000 vs. 13 loci), the ddRADseq topology is tentatively considered as more likely to represent the true species tree, i.e., the sister relationship between the Eastern and Western clades rather than the Eastern and Aripuanã-Tapajós clades. This conflict between the StarBEAST2 consensus species tree and the ddRADseq topology is resolved in one of the three most probable alternative species tree topologies. However, there is still evident uncertainty regarding the monophyly of the Aripuanã-Tapajós clade, and it is noteworthy that *P. cinerascens* and *P. miltoni* share some derived alleles with all taxa of the Eastern and Western clades to the exclusion of *P. hoffmannsi* in the D-statistic tests. D-statistic tests have previously been employed to assess competing phylogenetic hypotheses by comparing the number of shared and derived alleles among lineages under alternative species tree topologies, finding evidence of shared ancestry between taxa through testing for introgression on an incorrect species tree (Eaton et al., 2015). Given that short internal branches are problematic for the concatenation model (as discussed below), the D-statistics results may indeed indicate that *P. hoffmannsi* is sister to the remaining *P.*

moloch group taxa, rather than the earliest diverging lineage within the Aripuanã-Tapajós clade. This was also suggested in the nuclear and combined dataset multi-locus analyses in Chapter 2, and also by Carneiro et al. (2016).

Part of the difficulty in establishing the relationships between the *P. moloch* group species is that the initial divergences are estimated to have occurred over a relatively short period. Incomplete lineage sorting is particularly pronounced when lineages diverge rapidly, especially relative to effective population size (Edwards, 2009), and it is notable that the topological incongruences across these analyses are associated with the shortest internal branches in the phylogeny (e.g., in the ddRADseq timetree, see nodes numbered 7 + 11 or 15 + 16 + 17, Figure 4.4). Concatenated approaches can perform particularly poorly with even moderately short branches in the tree (Kubatko & Degnan, 2007), however, while the concatenated ddRADseq topology must be considered provisional, it is likely to be a more reliable estimate of species relationships than the smaller concatenated datasets given the enormous increase in the number of genomic regions sampled. Coalescent-based species tree analyses using the ddRADseq data (e.g., using genome wide SNPs with SNAPP; Bryant et al., 2012) will further test this phylogenetic hypothesis using a more appropriate model than concatenation. Coalescent-based methods can account for gene-tree discordance owing to incomplete lineage sorting (e.g., Linkem et al., 2016), and result in better estimates at short internodes (Edwards et al., 2016). Even these approaches, however, can be misleading in the presence of substantial gene flow between taxa (Leaché et al., 2014b), and thus, caution should be employed if putatively introgressed lineages, such as *P. cinerascens* clade B, are included (as in the StarBEAST2 tree). Alternatively such individuals should be excluded from coalescent-based analyses (e.g., Meyer et al., 2016).

Mitochondrial data (mtDNA) are particularly unsuitable for phylogenetic inference when incomplete lineage sorting is caused by short internal branches, or when contemporary introgression is suspected, owing to both the lack of intragenic recombination and the matrilineal inheritance of the mitochondrial genome. The mtDNA introgression found in the admixed *P. cinerascens* UFRO354 individual included in this study is also clear evidence that the assignment of Callicebinae specimens to species based primarily on mtDNA should be avoided [e.g., this individual was classified as *P. bernhardi* in Carneiro et al., (2016)], and that mtDNA phylogenies for Callicebinae taxa should be interpreted with caution. Overall, these

results add to the ever-growing evidence that mtDNA can mislead phylogenetic inference (e.g., Leaché, 2010; Roos et al., 2011; Yu et al., 2011). Furthermore, given the low levels of genetic differentiation between some taxa, resolution of the species relationships among all *P. moloch* group taxa, in the presence of incomplete lineage sorting and in a cost-effective manner, will likely require the generation of genome-wide molecular data.

Ages estimates recovered in the ddRADseq BEAST analysis are younger than the concatenated multi-locus analyses (Chapter 2), but generally up to 1 million years older than the StarBEAST2 tree. Across these datasets, however, age estimates were broadly comparable, with the Callicebinae genera estimated to have diverged in the late Miocene, the *Plecturocebus* species group in the Pliocene-Pleistocene, and most species divergences are dated to the Pleistocene. As discussed, younger divergence dates are expected in coalescent-based species trees owing to the implicit assumptions made by multispecies coalescent vs. concatenation models (Burbrink & Pyron, 2011; McCormack et al., 2010; Meyer et al., 2016). The timeline estimated with the multispecies coalescent model is likely to provide more realistic divergence dates across Callicebinae than previous estimates using the concatenation model in the present and other studies (Hoyos et al., 2016; Perelman et al., 2011).

4.5.2 Western clade taxa of the *P. moloch* group

In addition to the initial *P. moloch* group divergences, there is significant conflict across analyses regarding the relationships among Western clade species, with two main phylogenetic hypotheses recovered; *P. brunneus* is the earliest diverging lineage, while *P. cupreus* and *P. caligatus* are sister taxa (all ddRADseq analyses); or a sister relationship between *P. brunneus* and *P. caligatus*, with *P. cupreus* as the earliest diverging lineage (concatenated and StarBEAST2 multi-locus analyses). Statistical support varies across the analyses and all possible arrangements are recovered in the alternative most probable topologies in the StarBEAST2 species tree. Bayesian clustering results indicate that these three Western clade species form well-defined genetic clusters when only these taxa are included in the dataset (Piii85), however, all individuals are significantly admixed at all *K* values when other *P. moloch* group taxa are included (P85 dataset), as well as higher values of *K* for the subsampled datasets. When only two species are included, no admixture between *P. brunneus* and *P. cupreus* is detected, while *P. caligatus* shares ancestry with both

species at $K > 2$. No pair of taxa, however, show a significant proportion of shared derived alleles to the exclusion of the other taxon, regardless of which topology is assumed in the D-statistic test. This could be a consequence of the moderate number of ddRADseq loci and, therefore, the reduced statistical power of the D-statistic in this study, which may only be able to detect recent introgressive hybridisation when the lineages involved are included. Short internal branches are also found between Western clade species, and thus, topological incongruence across analyses may be a result of incomplete lineage sorting, as well as the low genetic variation among the taxa within this clade.

In support of the ddRADseq topology, *P. cupreus* and *P. caligatus* share more phenotypic resemblance (e.g., Hershkovitz, 1988), and geographic evidence may also suggests that these taxa are more likely to be sister than *P. caligatus* and *P. brunneus* given that the Rio Madeira (one of the largest tributaries of the Amazon) forms a more significant barrier to gene flow than the Rio Purus. *Plecturocebus brunneus* is thought to be restricted to the right bank of the Rio Madeira, while the populations of brown titis in Peru that were previously classified as *P. brunneus* are now attributed to a new taxon, *P. urubambensis* (Vermeer & Tello-Alvarado, 2015). If the *P. moloch* group originated in the Rondônia centre of endemism, as suggested in the biogeographical analyses in Chapter 3, then a sister relationship between *P. caligatus* and *P. cupreus* requires only one dispersal event across or around the Rio Madeira, whereas two are necessary for a sister relationship between *P. brunneus* and *P. caligatus*. Many Western clade species are not represented in these analyses and further studies are required with increased taxonomic sampling, as well as more individuals per lineage, in order to adequately reconstruct the evolutionary history of this clade.

4.5.3 The dubious nature of *Plecturocebus caligatus dubius*

Serrano-Villavicencio et al. (2017) recently reintroduced the subspecies rank to Callicebinae taxonomy, suggesting that *P. dubius* should be considered a subspecies rather than a synonym of *P. caligatus* (Byrne et al., 2016; see also Carneiro et al., 2016; Hoyos et al., 2016) because it represented a geographically restricted phenotype (white blaze). Although we agree with these authors that that the use of subspecies may benefit the taxonomy of Callicebinae (Serrano-Villavicencio et al., 2017), we advocate that inconclusive molecular evidence in support of monophyly, in combination with poor sampling in all molecular and taxonomic studies, extremely low genetic variation, and a lack of field data to confirm the restricted distribution of these phenotypes, should not be entirely disregarded and warrants consideration when delimiting *P. c. dubius* and *P. c. caligatus*. We also note a misinterpretation; Byrne et al. (2016) suggested that the differences between *P. c. caligatus* and *P. c. dubius* likely represented geographic variation in pelage colouration, rather than individual variation (as stated by Serrano-Villavicencio et al., 2017).

Plecturocebus c. dubius and *P. c. caligatus*, show the lowest genetic differentiation of all intraspecific lineages included in this study (e.g. lower than the *P. bernhardi, P. moloch*, and even *P. cupreus* lineages), with pairwise genetic distances between individuals often lower across the subspecies than within them. Both taxa form monophyletic clades in the ddRADseq phylogenetic analyses when the two poorly sequenced individuals are excluded, and therefore, paraphyly among *P. c. dubius* samples when all individuals are included may be a consequence of missing data. However, it is notable that the A85 and A92 MrBayes and RAxML trees suggest the same relationships among all *P. caligatus* individuals with varying statistical support (see Figure A 3.1 – 3.4), and expected species relationships are recovered for other specimens with equally poor sequencing and a closely related sister lineage, such as *P. cupreus* clade A CTGAM210. Wagner et al. (2013) recovered reciprocal monophyly among species of the Lake Victoria cichlid radiation, which originated within the last 15,000 years, using RADseq data matrices with 43% missing data. Furthermore, as discussed above, RADseq data matrices with more loci and more missing data have been shown to provide greater statistical power with little impact on phylogenetic results (e.g., Eaton et al., 2015; Huang & Knowles, 2016; Rubin et al., 2012), and thus, the putative paraphyly recovered among *P. c. dubius* individuals in some ddRADseq analyses should be investigated rather than dismissed.

Given the short internal branches within *P. caligatus*, paraphyly could also be a result of the concatenation model. Thus, coalescent-based methods will be the most appropriate for resolving the relationship between these taxa, as well as assessing alternative taxonomic hypotheses in a statistical framework using coalescent-based species delimitation approaches such as Bayes Factor Delimitation of Species (*with genomic data; BFD*; Leaché et al., 2014a). Serrano-Villavicencio et al. (2017) suggested that increasing the number of samples may resolve paraphyly recovered in mitochondrial phylogenies (see Chapter 2) as a result of potential inference error, and this is equally important in order to adequately assess monophyly of these lineages across their geographic distribution, especially given the scattered nature of the sampling in this and all previous molecular studies.

In the Bayesian clustering analyses, it is not possible to distinguish these taxa until the highest K values, and even when they are assigned partially to distinct clusters, they share most of their ancestry in the same cluster. Interestingly, all three analyses that include both *P. caligatus* and *P. cupreus* indicate that *P. c. dubius* shares a proportion of ancestry with only *P. c. caligatus* and with only *P. cupreus* (see K = 10, Figure 4.9; Western clade K = 5, Figure 4.11; *P. cupreus + P. caligatus* K = 3, Figure 4.12). Hershkovitz (1988) stated that, phenotypically, *P. c. dubius* was a mosaic of both *P. cupreus* and *P. c. caligatus* and may be a hybrid of these taxa. Although none of the D-statistic tests found a significant signal of introgression among these lineages (see Tables 4.12, A3.3), *P. c. dubius* and *P. cupreus* shared derived alleles to the exclusion of *P. c. caligatus* in the majority of iterations, with some combinations of individuals showing high but insignificant Z-scores. These results may reflect a weak signal of introgressive hybridisation between *P. c. dubius* and *P. cupreus*, or between either of the *P. caligatus* lineages and a taxon not included in this study, for example, the recently resurrected *P. toppini* (Vermeer & Tello-Alvarado, 2015).

As a further note, based on phenotypic variation, Serrano-Villavicencio et al. (2017) suggested that *P. stephennashi* is a hybrid of *P. c. caligatus* and *P. c. dubius,* and therefore the name *P. stephennashi* should be considered a homonym for these taxa. This is likely a misinterpretation of the International Code of Zoological Nomenclature (ICZN); if *P. stephennashi* is truly a hybrid, then the name should be considered as invalid rather than as a homonym. Although there is warranted doubt surrounding this taxon, there are also several curiosities not addressed by this

165

hypothesis, for example, the smaller size of the *P. stephennashi* specimens and the biogeographical context. With only four known specimens, no type locality and very little distributional data, we suggest that field studies within the proposed range of *P. stephennashi* [between the Rios Mucuim and Ipixuna, Amazonas (Van Roosmalen et al., 2002)], as well as the generation of molecular data, are required in order to adequately assess the relationship between these lineages and test this taxonomic hypothesis.

4.5.4 Introgressive hybridisation

Hybridisation among New World primates (NWP) lineages has been primarily studied in howler monkeys (*Alouatta*), with the first genetic evidence of hybridisation among NWP reported for *A. palliata* and *A. pigra* at a hybrid zone in Mexico (Cortés-Ortiz et al., 2007). There are few well-documented cases of hybridisation among all other NWP clades, and many proposed examples of interspecific hybridisation based on phenotypic evidence involve very closely related lineages, for example, as noted in the above, among the *P. moloch* group taxa of the Western clade (Hershkovitz, 1988; Serrano-Villavicencio et al., 2017). Given the lack of study systems, the evidence of introgressive hybridisation between *P. cinerascens* and *P. bernhardi* presented in this study is significant not only to Callicebinae, but also for the study of hybridisation among NWP more generally. Titi monkeys are monogamous pair-bonding primates and present a particularly interesting case to assess the dynamics of introgressive hybridisation among different primate mating systems.

Molecular (e.g., based on D-statistics, mitochondrial introgression, genetic distances, heterozygosity estimates, etc.), geographic and even phenotypic data provide independent sources of evidence, and together, strongly suggest that *P. cinerascens* clade B individuals are admixed with *P. bernhardi*. Hybridisation between these species is particularly notable given that the divergence between the progenitors to *P. bernhardi* and *P. cinerascens* may represent the deepest divergence within the *P. moloch* group i.e., Aripuanã-Tapajós clade vs. Eastern + Western clade. These species are also phenotypically distinct; typical *P. cinerascens* individuals are almost entirely grey agouti (the ashy titi), and along with *P. brunneus*, they are the only *P. moloch* group taxa that do not generally show dorsal-ventral contrast in pelage colouration, while *P. bernhardi* has contrasting orange ventral pelage and sideburns, and white hands. In future studies, "*P. cinerascens* clade B" should be properly

166

referred to as *P. cinerascens* x *P. bernhardi* in light of the above evidence. Furthermore, this information is important as this lineage should not be formally described as a valid taxon.

The admixed *P. cinerascens* specimens (UFRO352, 354, 355) were collected between the left bank of upstream Rio Roosevelt and the two rivers that form the Rio Jiparaná (Rios Barão de Melgaço and Pimenta Bueno). Field studies and molecular data for more individuals from this region will be essential to establish more information about the extent of gene flow between these species, for example, to assess; if hybridisation of the parental forms is ongoing or ancient; the extent of the hybrid zone; if the direction of introgression is biased i.e., only from *P. bernhardi* to *P. cinerascens*; and the spread of the introgressed alleles, i.e., whether gene flow is occurring between the admixed and parental population(s) outside this region or introgressed alleles are largely restricted to the hybrid/contact zone. It is clear from the sampling in the present study (see Figure 4.14), as well as several other field studies (e.g., Ferrari et al., 1996, 2000; Monção et al., 2008; Quintino & Bicca-Marques, 2013), that the geographic distributions of *P. bernhardi* and *P. cinerascens* are much broader than has been traditionally recognised (e.g., IUCN; Van Roosmalen et al., 2002).

Overall, this scenario is defensible in a biogeographic context and is concordant with the complex pattern of diversification among *P. moloch* group taxa in the Rondônia area of endemism suggested in the biogeographical reconstructions in Chapter 3. If the initial *P. moloch* group divergences were associated with Rio Roosevelt-Aripuanã such that the ancestor to the Aripuanã-Tapajós clade was restricted on the right bank to the northern part of Rondônia, the occurrence of *P. cinerascens* on left bank and upstream of the headwaters of the Rio Roosevelt-Aripuanã may represent relatively recent range expansion associated with fluvial dynamics since the Late Pleistocene (e.g., Latrubesse, 2002) In this scenario, range expansion then led to secondary contact and hybridisation between these two species when *P. cinerascens* dispersed into the geographic distribution of *P. bernhardi*. Given that these individuals are much more closely related to *P. cinerascens* than *P. bernhardi* based on nuclear data, it is evident that they are not first generation hybrids, and thus, that at least some *P. cinerascens* x *P. bernhardi* individuals are fertile. This may suggest that many *P. moloch* group species have not attained

reproductive isolation and that gene flow between species is primarily restricted by geographical barriers such as the larger rivers across Amazonia.

The identification of these individuals as admixed is also important for phylogenetic inference; some of the topological incongruence recovered across these analyses could be related to the unknowing inclusion of the admixed lineage in each study as gene flow between species isn't accounted for in concatenation or even most coalescent-based models. The effect of introgression on phylogenetic inference can be difficult to detect, however, sampling admixed individuals can have an impact on the order of species divergences and the timeline of diversification (Leaché et al., 2014b). Conducting some of the ddRADseq and multi-locus phylogenetic analyses again without the admixed *P. cinerascens* x *P. bernhardi* individuals may provide insight into their influence on phylogeny.

4.6 Conclusions

The quantity of data generated for the reduced representation ddRADseq data matrices represented a massive increase compared to the data generated for the multi-locus phylogeny, which was previously the largest molecular dataset for Callicebinae and among the largest multi-locus datasets for any group of New World primates. Phylogenomic analyses recovered a strongly supported phylogeny which, along with the coalescent-based species tree based on multi-locus data, allowed the identification of two main conflicts among the *P. moloch* group taxa that are associated with short internal branches and are still considered unresolved. Overall, our results suggest that gene flow between species of the *P. moloch* group, as well as stochastic processes such as incomplete lineage sorting, present a challenge to the reconstruction of species relationships across this species group. Phylogenetic inference with the ddRADseq data using multispecies coalescent-based models will provide the most appropriate and reliable estimation of species relationships and divergence times. Coalescent-based species tree estimation methods using large genome-wide datasets are computationally intense which has limited their application in this study, however, they will likely become more broadly accessible as multispecies coalescent models become more advanced and diverse.

This study presents the first known application of the D-statistic test to assess introgression among New World primates, one of the only known cases where genetic

evidence has been presented regarding introgressive hybridisation between any New World primates lineages other than for the howler monkeys, and also, one of the only known cases of introgression involving monogamous primates. Increasing our understanding of the nature of introgression between these species is, therefore, also of broad interest to primatologists and of particular relevance to the study of Amazonian biogeography.

Chapter 5: General discussion

A few decades of the "molecular revolution" had passed and titi monkeys had still not been the focus of any molecular phylogenetic study, in fact, the field of phylogenetics was already in the midst of a second revolution owing to next generation sequencing (McCormack et al., 2012) before this doctoral thesis had begun. Many primate species, in particular Hominidae, had assembled annotated genomes and genome-wide data for many individuals (Marques-Bonet et al., 2009; Prado-Martinez et al., 2013; Rogers & Gibbs, 2014), while the only available molecular data for Callicebinae was generated for high-level primate phylogenies and interpreting species relationships was hindered by a lack of taxonomic coverage (see Chapter 2, Perelman et al., 2011; Springer et al., 2012). When primate phylogenetic studies were already embracing the genomics era (e.g., Pecon-Slattery, 2014), the only species-level phylogeny available for titi monkeys was Kobayashi's (1995) morphological phylogenetic analysis. This doctoral thesis was borne out of the paucity of information regarding the evolutionary history of Callicebinae, as well as the availability of genetic material obtained through years of field expeditions conducted by Jean P. Boubli and many other primatologists across Brazil, without which this research would not have been possible.

5.1 Main findings

The goal of my doctoral thesis was to employ molecular data to provide insight into the evolutionary history and biogeography of Callicebinae, one of the most strikingly poorly studied and among the most species-rich groups of primates. To achieve this goal, large sequence data matrices were generated using multi-locus Sanger sequencing (20 nuclear and 2 mitochondrial loci) and reduced representation, genome-wide double-digest restriction-associated DNA sequencing (ddRADseq). In **Chapter 2**, we inferred phylogeny and diversification times using the multi-locus data and revised the taxonomy of Callicebinae based on molecular, morphological and biogeographic evidence. In **Chapter 3**, we employed a statistical biogeographical approach to perform ancestral-area estimations across the phylogeny of Callicebinae

based on time-calibrated trees inferred using the multi-locus data. In **Chapter 4**, we conducted phylogenetic analyses, assessed genetic structure and tested for interspecific gene flow using genome-wide ddRADseq data matrices, as well as performed coalescent-based species-tree estimation using the multi-locus sequences.

In **Chapter 2**, we assembled one of the largest multi-locus Sanger sequenced molecular datasets for any group of platyrrhine primates (22 loci, >14K bp in length; 15 titi species, 59 specimens), reconstructed the first comprehensive species-level molecular phylogeny for Callicebinae, and provided the first molecular review of the subfamily. Our phylogenetic analyses (based on concatenated data matrices) clarified a number of issues on the taxonomic and phylogenetic relationships among the species. We provided evidence for an early divergence (late Miocene) of three major Callicebinae lineages, reconstructed a timeline for Callicebinae diversification, and inferred a well-supported phylogeny for all species included, with the exception of *P. miltoni* and *P. cinerascens*, which required further investigation. Based on new molecular evidence and well-established differences in morphology, karyology, and biogeography, we proposed a new genus-level taxonomy for titi monkeys: *Cheracebus* (Byrne et al., 2016) in the Orinoco, Negro and upper Amazon basins, *Callicebus* Thomas, 1903, in the Atlantic Forest and neighbouring Caatinga regions, and *Plecturocebus* (Byrne et al., 2016) in the Amazon basin and Chaco region. We also reviewed the taxonomic history for Callicebinae, suggested the reintegration of *cupreus* group species (*sensu* Kobayashi, 1995) into the *moloch* group, and questioned the designation of *P. dubius* (Hershkovitz, 1988) as a valid species. This work illustrates the value of considering molecular evidence in taxonomic classification, provides a basis for future studies on the evolutionary history and taxonomy of titis, and has opened a dialogue on other taxonomic issues that had been left unattended for quite some time. The new taxonomic proposal for Callicebinae brings concordance to the classification of genera across New World primates and better describes the great diversity of this poorly studied group.

In **Chapter 3**, we provided the first known statistical biogeographical approach applied to reconstruct the biogeography of Callicebinae in an explicit phylogenetic framework. We recovered evidence for the divergence of titi monkey genera in the late Miocene via the fragmentation of a widespread ancestor distributed across the modern-day northwestern Amazon (*Cheracebus*), wet and dry savanna ecosystems (*Plecturocebus*), and Atlantic Forest (*Callicebus*). Our reconstruction

indicated that species-level diversification among the Amazonian clades initiated from a narrow area of origin (Napo, *Cheracebus*; Rondônia, *P. moloch* group), and was characterised by a sequential, long-distance dispersal model of speciation by "island-hopping" across pre-existing river barriers. These founder-events (jump dispersal) were sufficiently rare to allow diversification in isolation after dispersal, emphasising the role of major Amazonian rivers as strong barriers to gene flow among allopatric species, with notable comparisons to island biogeography. We uncovered a complex pattern of diversification among species of the *P. moloch* group, with a non-monophyletic assemblage of taxa endemic to Rondônia (area of endemism), suggesting that a complex history of river system evolution may have played an important role in driving historical distributions in this region. Our results are taken to suggest that the evolution of the Pebas system in the western Amazon may have influenced the diversification and distribution of extant Callicebinae lineages, which were absent from the western Amazon until the recession of these wetlands and the establishment of suitable forest habitat in the Pleistocene ("Young Amazon" model). This work comprises one of the first investigations of the evolutionary history of titi monkeys in the context of Amazonian and South American historical biogeography based on an explicit phylogenetic hypothesis, and sheds light on the processes that generated the great diversity found among Callicebinae taxa. Although this research provides only a large-scale reconstruction of callicebine biogeography and should be interpreted with caution, it represents a critical starting point for future research that aims to understand diversification within this subfamily.

In **Chapter 4**, we propelled research on titi monkey evolutionary history into the phylogenomics era using double digest restriction-site associated DNA sequencing (ddRADseq) to generate reduced representation genome-wide molecular data for 12 Callicebinae species (45 specimens). This work is among the first phylogenomic analyses employing genome-wide data for any New World primate group. Here, we began to address more difficult questions regarding the phylogenetic relationships among the lineages, particularly among the species of the *P. moloch* group. The ddRADseq data matrices contained up to ~ 3K loci, with a maximum concatenated length of ~ 930K bp and up to ~ 15K parsimony informative sites; a massive leap in the quantity of data in comparison to all previous molecular studies on Callicebinae. Our ddRADseq phylogenetic analyses recovered a strongly supported topology (perfect support at all nodes in some analyses) for Callicebinae

with species of the *P. moloch* group divided into three major clades (Aripuanã-Tapajós, Eastern, and Western clades). Although we advocate that the ddRADseq topology must be considered provisional until assessed using multispecies coalescent-based analyses with genome-wide data, we consider it the most reliable estimate of the *P. moloch* group species relationships currently available. We provide evidence of introgressive hybridisation using D-statistic tests that is supported by mitochondrial introgression. Collectively, our results indicate that *P. cinerascens* individuals from the left bank of upstream Rio Roosevelt, Rondônia, are admixed with *P. bernhardi*, a significant discovery of broad interest to primatologists, specifically for research on New World primates given the lack of natural hybridising study systems. This research represents the first known application of the D-statistic test among New World primates, and one of the only known cases where genetic evidence has been presented regarding admixture between any New World primates other than howler monkeys. Finally, we discuss putative sources of topological incongruence across loci and across previous studies (analytical and biological, e.g., interspecific gene flow, incomplete lineage sorting), and the complicated nature of lineage diversification across the *P. moloch* group.

Collectively, the chapters of this thesis provide a detailed picture of the evolutionary history of titi monkeys and add to our understanding of what makes the subfamily Callicebinae, the genera, and the many titi species unique. Within a few years, our knowledge of the evolutionary relationships among Callicebinae taxa has progressed rapidly, perhaps in a relatively manner unprecedented among New World primates; from species relationships based almost entirely upon Kobayashi's (1995) morphological phylogeny, to phylogenetic hypotheses derived from multi-locus or genome-wide molecular datasets that are among the largest known for any New World primate group.

5.2 Priorities and future directions

Naturally, new questions arise and inconsistencies or gaps in our knowledge are made apparent as research delves deeper into the evolutionary history of titi monkeys. Below, I highlight other avenues of scientific enquiry and areas in need of further attention, some of which emerged in this dissertation and others are longstanding issues that require addressing.

Among the most crucial of priorities for molecular phylogenetic studies, and more generally, is to responsibly collect samples to generate molecular data for Callicebinae species with no sequence data currently available in order to reconstruct a fully comprehensive species-level phylogeny. This is particularly applicable to the *Plecturocebus donacophilus* group; *P. donacophilus* is the only member of this group with data available on GenBank. Perhaps the most intriguing species lacking molecular data is *P. modestus*, which was placed in its own species group and regarded as the most primitive titi monkey species by Hershkovitz (1988, 1990) owing to its unusual elongated skull. Several species are also only represented by mitochondrial data which is unreliable for both phylogenetic inference and species assignments, as highlighted in this dissertation. Furthermore, no known population genetics studies have been conducted on any titi species, and relatively few species have habituated study populations that are the focus of long-term established field projects, which are integral to our ability to collect non-invasive samples and genetic material for enough individuals to conduct such analyses. Genetic material for many species is limited to a handful of individuals, sometimes from the same locality, which severely hinders our understanding of their genetic diversity. Given the broad distribution of titi monkeys, developing field projects and expeditions, collecting samples, and generating molecular data for these species/lineages will require an international collaborative effort. Our understanding of the diversity and the evolutionary relationships among Callicebinae will remain glaringly incomplete until this is achieved. To obtain a more comprehensive picture of the evolutionary history of Callicebinae, test the various phylogenetic hypotheses constructed to date, and address more complex phylogenetic questions, species relationships and divergence times should also be estimated using multispecies coalescent-based models with genome-wide data (e.g., using genome wide SNPs with SNAPP; Bryant et al., 2012; see also Stange et al., 2017).

Explicit phylogenetic hypotheses are a necessary component of understanding the spatial patterns of diversification, and thus, the study of the biogeography of titis is also hindered by the absence of a species-level phylogeny for all described Callicebinae taxa. This is particularly important for the species of the *P. donacophilus* group given their unusual and disjunct geographic distributions, and above all, *P. oenanthe* in the Andean foothills of Peru. Future studies with increased taxonomic sampling and geographic delineations within each region will allow a more in-depth

investigation of the biogeographic history of Callicebinae, and is essential in order to test the biogeographic scenarios and diversification patterns recovered in this work. Of particular interest are studies with a focus on the biogeography of the *P. moloch* group to provide insight into the complex diversification dynamics within the Rondônia centre of endemism, the non-monophyletic assemblage of *P. moloch* group taxa found in this region, and secondary contact between *P. cinerascens* and *P. bernhardi* in southwestern Rondônia upstream of the Rio Roosevelt-Aripuanã, which elsewhere forms a barrier to gene flow between these species. More generally, this research is important to increase our understanding of the diversification of Amazonian biota given that similar patterns of shared lineages in Rondônia are observed in other vertebrate groups such as birds (Fernandes, 2013; Thom & Aleixo, 2015) and lizards (Oliveira et al., 2016).

The identification of admixture between *P. cinerascens* and *P. bernhardi* warrants further investigation and additional molecular, phenotypic and distributional data are required to establish more information about the extent of gene flow between the two species. Field studies should initially focus on the region between the left bank of upstream Rio Roosevelt and the tributaries of the Rio Jiparaná (Rios Barão de Melgaço and Pimenta Bueno), where the *P. cinerascens* x *P. bernhardi* individuals were collected. Sampling more individuals from within this region and from the putative parental populations are required in order to assess; if hybridisation of the parental forms is ongoing; the distribution of the admixed individuals and extent of the hybrid zone; if the direction of introgression is biased i.e., only from *P. bernhardi* to *P. cinerascens*; and the spread of the introgressed alleles, i.e., whether gene flow is occurring between the admixed and parental population(s) outside this region or introgressed alleles are largely restricted to the hybrid/contact zone.

For some species, clarifying the often contradictory taxonomic history is another challenging but necessary task. The long-term blanket use of *C. torquatus* to refer to all *Cheracebus* species has hindered our understanding of the lineage that *C. torquatus* now represents (see Hershkovitz, 1988), and the identity and geographic limits of this taxon requires attention. More generally, Callicebinae taxonomy would benefit from an integrative approach where all available data are considered (Padial et al., 2010; Zapata & Jiménez, 2012). It is generally accepted that species are separately evolving lineages of populations (de Queiroz, 1998; de Queiroz, 2007; Wiley, 1978; but see also Willis, 2017), with most conflict regarding the point at which you

distinguish species when divergence is continuous (Hey, 2006; Mallet, 2008). Given that the geographic distributions of Callicebinae taxa are typically delineated by rivers, and rivers throughout Amazonia appear to form barriers to gene flow to varying degrees, many titi species are expected to show hierarchical substructure (genetic and phenotypic variation) between smaller interfluves across their distribution, and this may also be exaggerated when sampling is scattered. If very small fixed geographically restricted phenotypic or genetic differences are the sole defining criterion for delimiting lineages (even at subspecific level), without considering the extent of divergence, then the number of Callicebinae taxa may increase dramatically. The full spectrum of diversity among Callicebinae, or any group of organisms, cannot be reflected in discrete units in a hierarchical classification system. Alternative taxonomic hypotheses should also be assessed in a statistical framework, for example, using coalescent-based species delimitation approaches (Fujita et al., 2012) such as Bayes Factor Delimitation of Species (*with genomic data; BFD*; Leaché et al., 2014a).

Species-level taxonomic revisions are required for each of the genera, with updated information on species distributions to account for recent discoveries. It is evident that the limits to the distributions of many species are uncertain or unknown, and filling those gaps in our knowledge requires extensive field studies and collaborative research efforts. *Plecturocebus bernhardi* has now been recorded or collected between the Rios Jiparaná and Aripuanã-Roosevelt (original proposed distribution; Van Roosmalen et al., 2002), several localities along the left bank of the Jiparaná, and further south in the Guaporé Biological Reserve (both sampled here; see also Ferrari et al., 2000; Monção et al., 2008; Quintino & Bicca-Marques, 2013). The limits of the disparate *P. bernhardi* lineages recovered in this work are currently unknown. The sampling in this work as well as other studies (Ferrari et al., 1996, 2000) also suggests that the known distribution of *P. cinerascens* should be extended, and in a similar manner, Dalponte et al. (2014) indicated that the newly described taxon, *P. miltoni*, occurs between the Rios Roosevelt and Aripuanã rather than *P. cinerascens*.

In truth, I have already started working and collaborating on some of these topics, but that is a story for another day.

5.3 Final conclusions

This doctoral thesis provides a fascinating picture of the evolutionary history of titi monkeys and represents a major contribution to our knowledge of the phylogenetic relationships among the lineages, and the timeline, spatial patterns and mode of diversification. The revised genus-level classification for Callicebinae brings concordance to the designation of genera across New World primates and more accurately reflects titi monkey evolutionary and biogeographic history, the hierarchical relationship between the major clades, and the previously underappreciated genetic diversity represented by the "titi monkey" moniker. Callicebine researchers across many fields have started adopting the new nomenclature and it has also triggered further discussion and scientific attention towards these primates (e.g., Allgas et al., 2017; Araújo et al., 2017; Martínez & Wallace, 2016; Serrano-Villavicencio et al., 2017).

Titis are as comparatively diverse as their sister clade, the Pitheciinae, which has long included three genera, yet, this diversity is rarely recognised. This is visually exemplified on the front cover of a collection of research on the family Pitheciidae by an image including four members of this group, but only one of these is a titi monkey (Veiga et al., 2013). I believe that a lack of insight into the great diversity found among titi monkeys has had implications well-beyond how we name and classify them, however, I am reassured by the seemingly renewed interest and increasing volume of research involving these enigmatic primates over the past number of years.

Appendix 1: Supplementary material for Chapter 2: Phylogenetic relationships of the New World titi monkeys (*Callicebus*): First appraisal of taxonomy based on molecular evidence

Table A1.1 List of sequence characteristics per locus including length, variation, and sample coverage. Site information and sample coverage represent *Callicebus* taxa only.

Locus	Length (bp)	Constant sites		Variable sites		Parsimony informative sites		Sample coverage	
		bp	%	bp	%	bp	%	No. (53)	%
ABCA1	851	803	94.4	48	5.6	41	85.4	45	84.9
ADORA3	402	393	97.8	9	2.2	8	88.9	51	96.2
APP	702	691	98.4	11	1.4	4	36.4	52	98.1
COI	660	498	75.5	162	24.5	141	87	45	84.9
COI	660	488	73.9	172	26.1	142	82.6	47	88.7
CREM	414	406	98.1	8	1.9	5	62.5	46	86.7
CYTB	1140	850	74.6	290	25.4	244	84.1	47	88.7
CYTB	1140	824	72.3	316	27.7	278	88	53	100
DENND5A	705	680	96.5	25	3.5	18	72	48	90.6
DMRT1	492	481	97.8	11	2.2	8	72.7	39	73.6
ERC2	762	727	95.4	35	4.6	25	71.4	46	86.7
FAM123B	711	691	97.2	20	2.8	15	75	41	77.4
FES	456	430	94.3	26	5.7	21	80.8	52	98.1
FOXP1	552	534	96.7	18	3.2	14	77.8	52	98.1
MAPKAP1	639	630	98.6	9	1.4	9	100	51	96.2
MBD5	531	524	98.7	7	1.3	4	57.1	53	100
NEGR1	537	528	98.3	9	1.7	7	77.8	43	81.1
NPAS3.2	585	563	96.2	22	3.8	18	81.8	50	94.3
RAG1	1038	1002	96.5	36	3.5	24	66.7	44	83
RAG2	675	652	96.6	23	3.4	16	69.6	46	86.7
RPGRIP1	675	659	97.6	16	2.4	13	81.3	47	88.7
SGMS1	582	573	98.5	9	1.5	6	66.7	47	88.7
SIM1	632	619	97.9	13	2.1	6	46.2	52	98.1
ZFX	837	801	95.7	36	4.3	31	86.1	42	79.2

Table A1.2 Partitioning schemes and substitution models selected by PartitionFinder. The selected partitioning schemes were implemented in RAxML v. 8.1, MrBayes 3.2.3 or BEAST v 1.8.1. Numbers in parentheses refer to codon position for exonic mitochondrial and nuclear sequences.

Analysis	Total partitions	Partition number	Model	Loci
				Nuclear Dataset
RAxML	6	1	GTR+G	ABCA1
		2	GTR+G	ADORA3 (1), APP, DMRT1, FAM123B (1), FAM123B (3), FOXP1, MAPKAP1, MBD5, NEGR1, RAG1 (1), RAG1 (3), RAG2 (1), RAG2 (2), RPGRIP1 (1), RPGRIP1 (2), SGMS1, SIM1
		3	GTR+G	ADORA3 (2), CREM, DENND5A, FAM123B (2), FES, RPGRIP1 (3)
		4	GTR+G	ADORA3 (3), ERC2, RAG2 (3), ZFX
		5	GTR+G	NPAS3.2
		6	GTR+G	RAG1 (2)
MrBayes	5	1	K80+G	ABCA1, CREM, DENND5A, RAG1 (2)
		2	GTR+I	ADORA3 (1), ADORA3 (3), APP, DMRT1, FOXP1, MAPKAP1, MBD5, NEGR1, RAG1 (1), RAG2 (2), RPGRIP1 (2), SGMS1, SIM1
		3	HKY+G	ADORA3 (2), FAM123B (1), FAM123B (2), FES, RPGRIP1 (3)
		4	HKY+G	ERC2, NPAS3.2, RAG2 (3), ZFX
		5	K80	FAM123B (3), RAG1 (3), RAG2 (1), RPGRIP1 (1)
BEAST	6	1	K80+G	ABCA1, CREM, DENND5A, RAG1 (2)
		2	HKY+I	ADORA3 (1), ADORA3 (3), APP, MAPKAP1, RPGRIP1 (2), SGMS1, SIM1
		3	HKY+G	ADORA3 (2), FAM123B (1), FAM123B (2), FES, RPGRIP1 (3)
		4	HKY	DMRT1, FOXP1, MBD5, NEGR1, RAG1 (1), RAG2 (2)
		5	HKY+G	ERC2, NPAS3.2, RAG2 (3), ZFX
		6	K80	FAM123B (3), RAG1 (3), RAG2 (1), RPGRIP1 (1)
				Combined dataset
RAxML	7	1	GTR+G	ABCA1, ADORA3 (3), CREM, CYTB (2), DENND5A, ERC2, RAG1 (2), RAG2 (3), ZFX
		2	GTR+G	ADORA3 (1), APP, COI (2), DMRT1, FOXP1, MAPKAP1, MBD5, NEGR1, NPAS3.2, RAG1 (1), RAG2 (2), RPGRIP1 (2), SGMS1, SIM1
		3	GTR+G	ADORA3 (2), FAM123B (2), FES, RPGRIP1 (3)
		4	GTR+G	FAM123B (1), FAM123B (3), RAG1 (3), RAG2 (1), RPGRIP1 (1)
		5	GTR+G	COI (1), CYTB (1)
		6	GTR+G	COI (3)
		7	GTR+G	CYTB (3)
MrBayes	8	1	K80+G	ABCA1, CREM, DENND5A, RAG1 (2)
		2	GTR+I	ADORA3 (1), ADORA3 (3), APP, DMRT1, FOXP1, MAPKAP1, MBD5, NEGR1, RAG1 (1), RAG2 (2), RPGRIP1 (2), SGMS1, SIM1
		3	HKY+G	ADORA3 (2), FAM123B (1), FAM123B (2), FES
		4	HKY+G	ERC2, NPAS3.2, ZFX
		5	K80	FAM123B (3), RAG1 (3), RAG2 (1), RPGRIP1 (1)
		6	HKY+I	COI (2), CYTB (2), RAG2 (3), RPGRIP1 (3)
		7	K80+I	COI (1), CYTB (1)
		8	GTR+I+G	COI (3), CYTB (3)
BEAST	9	1	K80+G	ABCA1, CREM, DENND5A, RAG1 (2), RPGRIP1 (2), RPGRIP1 (3)
		2	HKY+I	ADORA3 (1), APP, COI (2)
		3	HKY+I	ADORA3 (2), FAM123B (1), FAM123B (2), FES
		4	GTR+G	ADORA3 (3), DMRT1, ERC2, FOXP1, NPAS3.2, ZFX
		5	K80+I	FAM123B (3), RAG1 (3), RAG2 (1), RPGRIP1 (1)
		6	HKY+I	MAPKAP1, MBD5, NEGR1, RAG1 (1), RAG2 (2), SGMS1, SIM1
		7	HKY+I	CYTB (2), RAG2 (3)
		8	TrNef+G	COI (1), CYTB (1)
		9	TrN+I+G	COI (3), CYTB (3)
				Mitochondrial dataset
RAxML	3	1	GTR+G	COI (1), CYTB (1)
		2	GTR+G	COI (2), CYTB (2)
		3	GTR+G	COI (3), CYTB (3)
MrBayes	5	1	K80+1	COI (1)
		2	F81	COI (2)
		3	GTR+G	COI (3), CYTB (3)
		4	HKY+G	CYTB (1)
		5	HKY+I	CYTB (2)

Table A1.3 Divergence matrix for the cytochrome *b* locus for selected *Callicebus* species. Bold indicates distance values < 0.01.

	Species	ID	1 UFRO541	2 4019	3 4346	4 4505	5 4009	6 CTGAM181	7 CTGAM182	8 MVR58	9 CCM248	10 UFRO403	11 UFRO544	12 4990	13 4984	14 4993	15 4988	16 AAM15	17 CTGAM210
1	*C. brunneus*	UFRO541	-																
2	*C. brunneus*	4019	**0.008**	-															
3	*C. brunneus*	4346	**0.005**	**0.004**	-														
4	*C. brunneus*	4505	**0.005**	**0.004**	**0.000**	-													
5	*C. brunneus*	4009	**0.005**	**0.004**	**0.000**	**0.000**	-												
6	*C. caligatus*	CTGAM181	0.030	0.029	0.025	0.025	0.025	-											
7	*C. caligatus*	CTGAM182	0.028	0.029	0.025	0.025	0.025	**0.002**	-										
8	*C. caligatus*	MVR58	0.027	0.028	0.024	0.024	0.024	**0.003**	**0.001**	-									
9	*C. caligatus*	CCM248	0.029	0.028	0.024	0.024	0.024	**0.001**	**0.001**	**0.002**	-								
10	*C. dubius*	UFRO403	0.033	0.032	0.028	0.028	0.028	**0.005**	**0.005**	**0.006**	**0.004**	-							
11	*C. dubius*	UFRO544	0.029	0.028	0.024	0.024	0.024	**0.003**	**0.003**	**0.004**	**0.002**	**0.006**	-						
12	*C. cupreus* (B)	4990	0.032	0.031	0.029	0.029	0.029	0.031	0.029	0.028	0.030	0.032	0.032	-					
13	*C. cupreus* (B)	4984	0.032	0.031	0.029	0.029	0.029	0.031	0.029	0.028	0.030	0.032	0.032	**0.004**	-				
14	*C. cupreus* (B)	4993	0.030	0.029	0.027	0.027	0.027	0.029	0.027	0.026	0.028	0.030	0.030	**0.002**	**0.002**	-			
15	*C. cupreus* (B)	4988	0.031	0.030	0.028	0.028	0.028	0.030	0.028	0.027	0.029	0.031	0.031	**0.004**	**0.004**	**0.002**	-		
16	*C. cupreus* (A)	AAM15	0.035	0.036	0.034	0.034	0.034	0.036	0.034	0.033	0.035	0.037	0.037	0.018	0.018	0.016	0.018	-	
17	*C. cupreus* (A)	CTGAM210	0.038	0.037	0.035	0.035	0.035	0.037	0.035	0.034	0.036	0.038	0.038	0.019	0.019	0.017	0.019	0.012	-
18	*C. cupreus* (A)	JLP15920	0.038	0.037	0.035	0.035	0.035	0.035	0.033	0.032	0.034	0.035	0.036	0.017	0.017	0.015	0.017	0.012	**0.007**
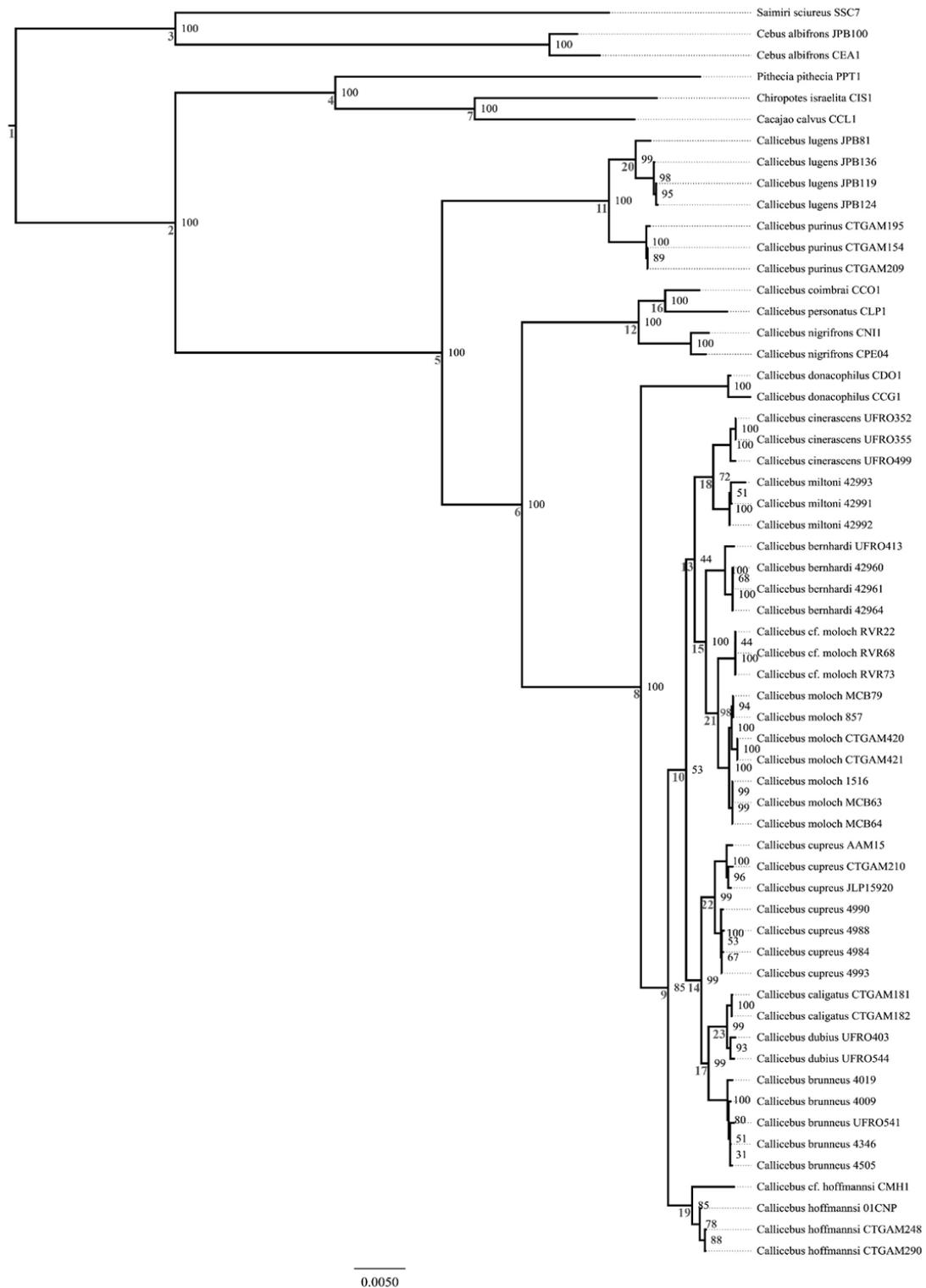
Figure A1.1 Maximum likelihood phylogeny inferred from the combined dataset. Node numbers represent nodes of interest listed in Table 2.9.

182
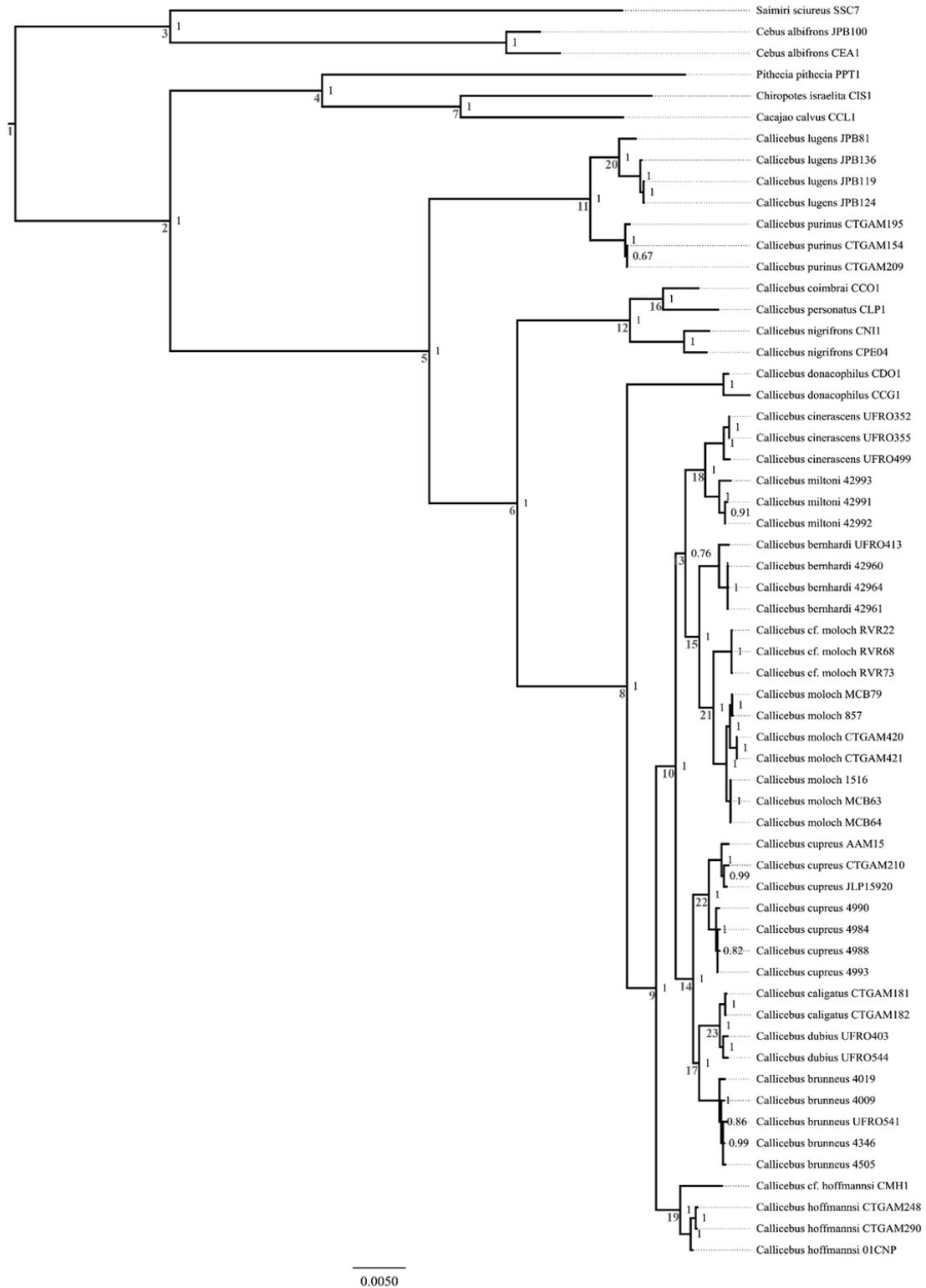
Figure A1.2 Bayesian phylogeny inferred from the combined dataset (MrBayes). Node numbers represent nodes of interest listed in Table 2.9.
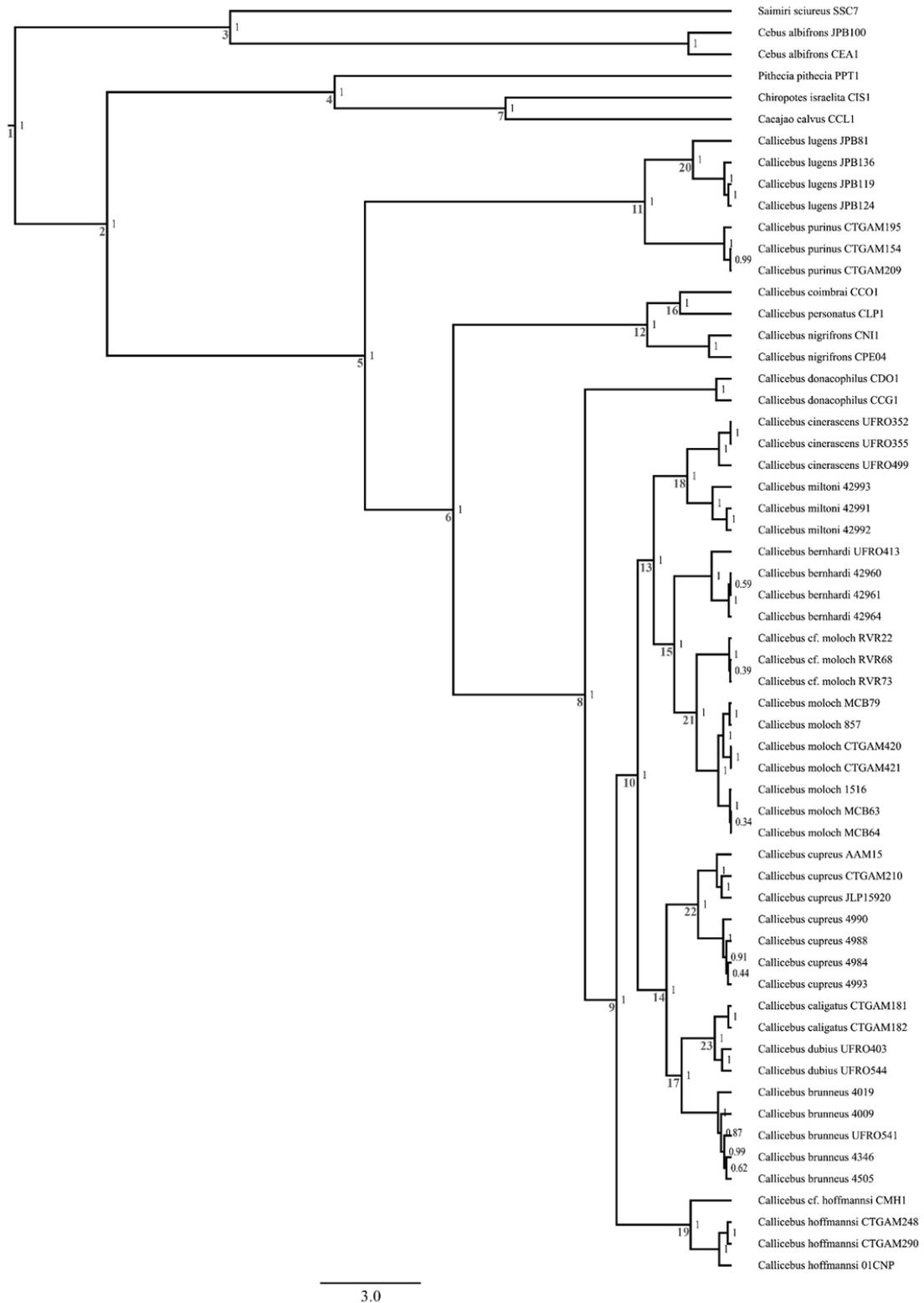
Figure A1.3 Bayesian phylogeny inferred from the combined dataset (BEAST). Node numbers represent nodes of interest listed in Table 2.9, 2.10.
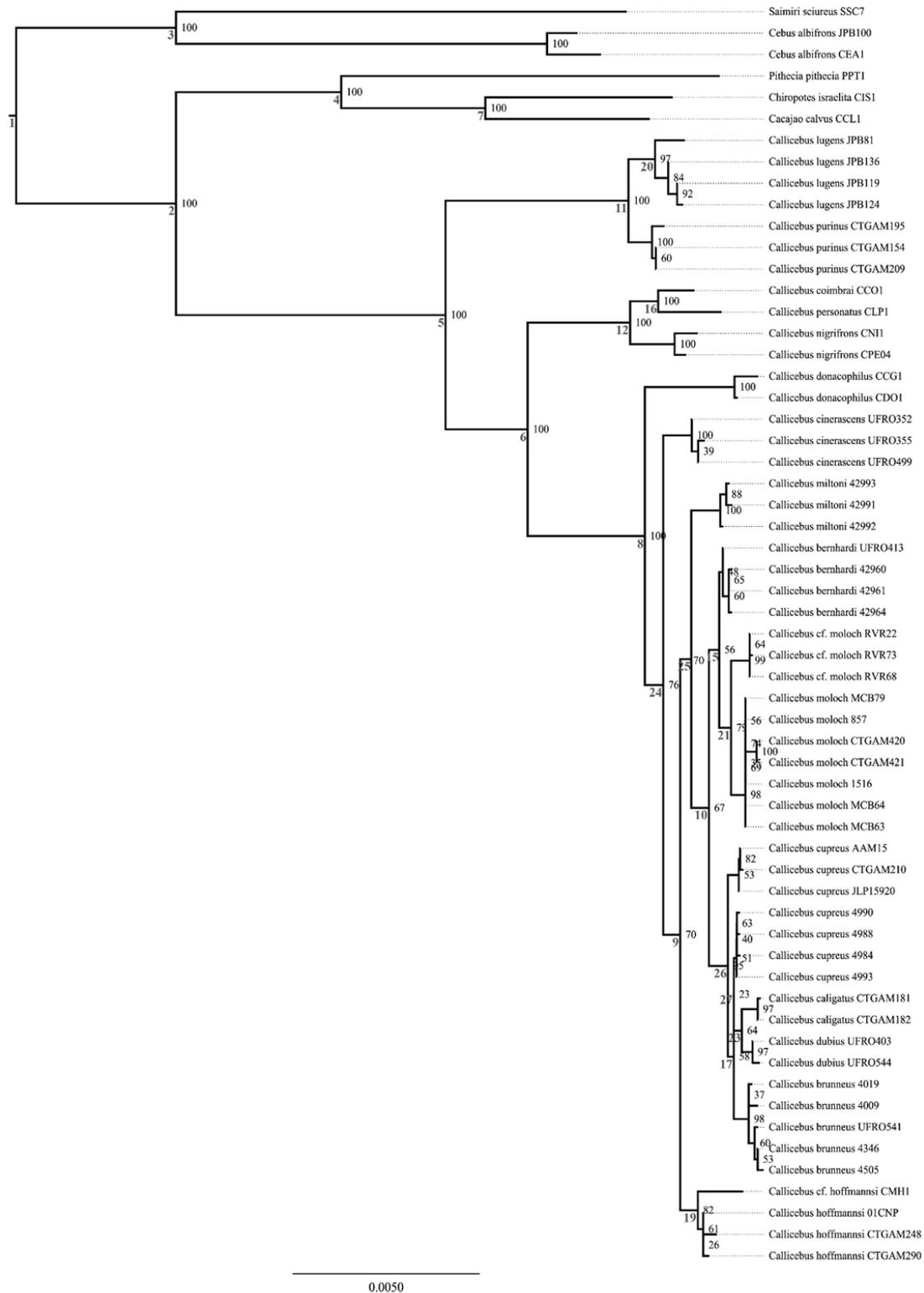
Figure A1.4 Maximum likelihood phylogeny inferred from the nuclear dataset. Node numbers represent nodes of interest listed in Table 2.9.
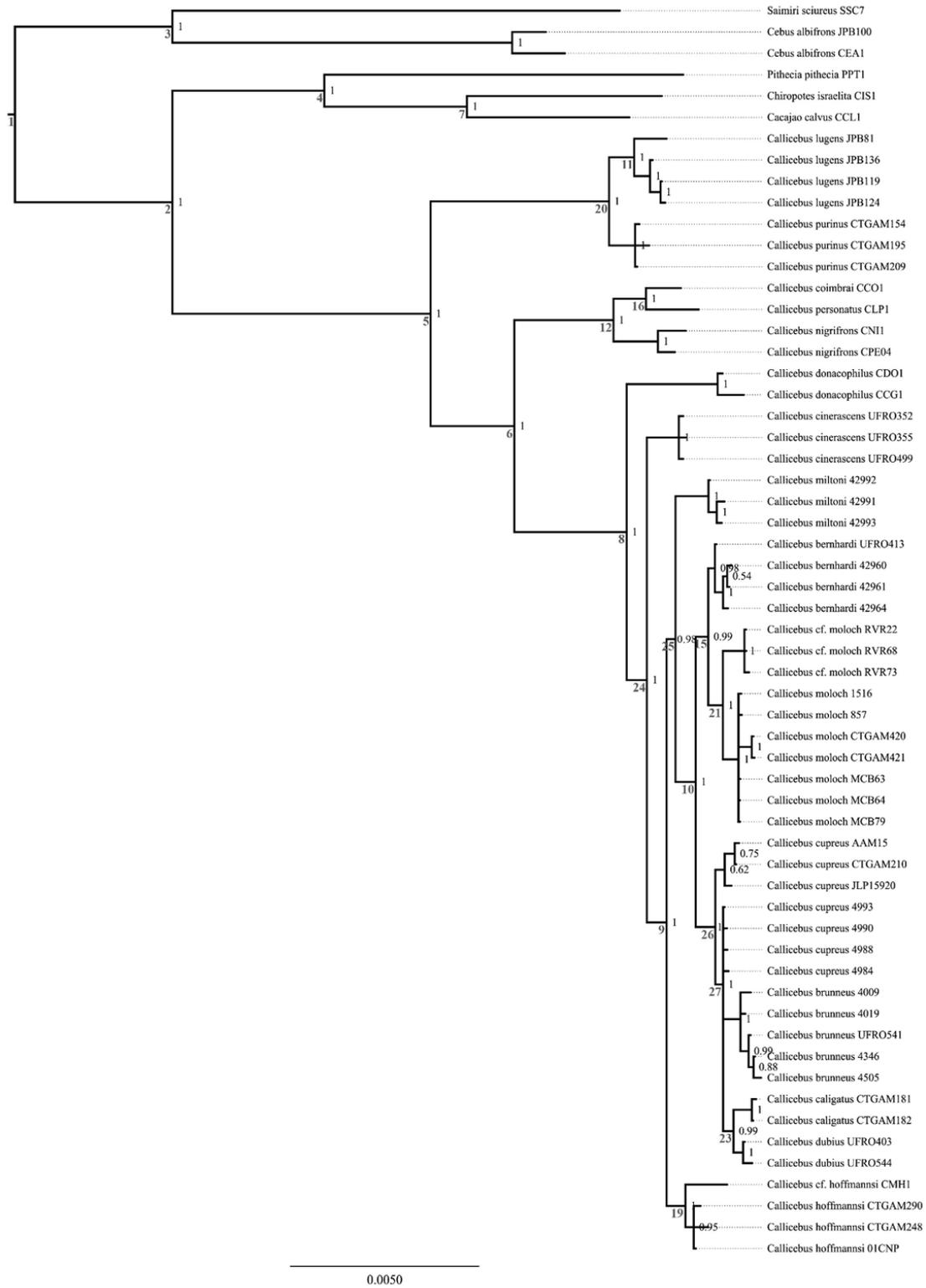
Figure A1.5 Bayesian phylogeny inferred from the nuclear dataset (MrBayes). Node numbers represent nodes of interest listed in Table 2.9.

186
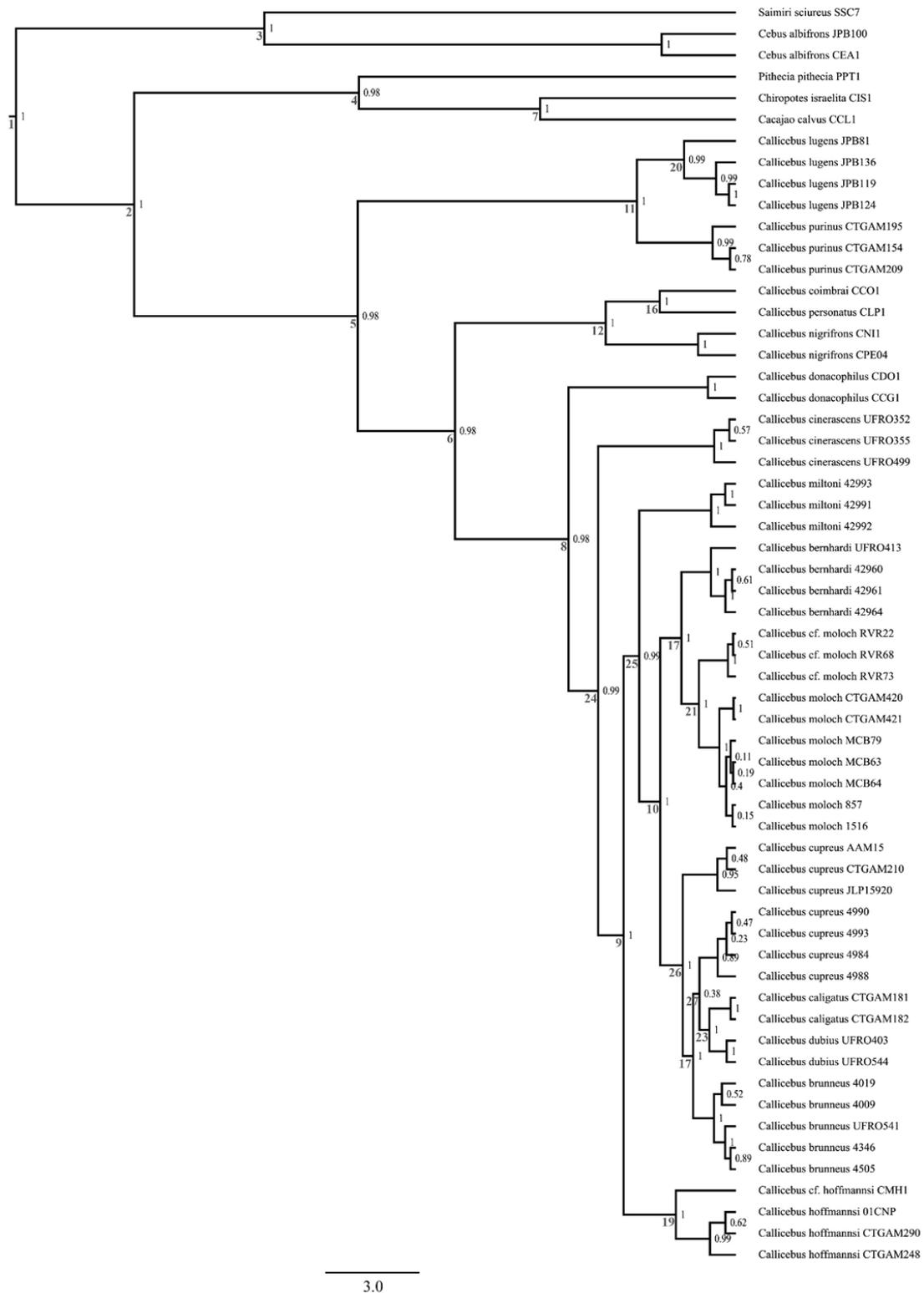
Figure A1.6 Bayesian phylogeny inferred from the combined dataset (BEAST). Node numbers represent nodes of interest listed in Table 2.9, 2.10.
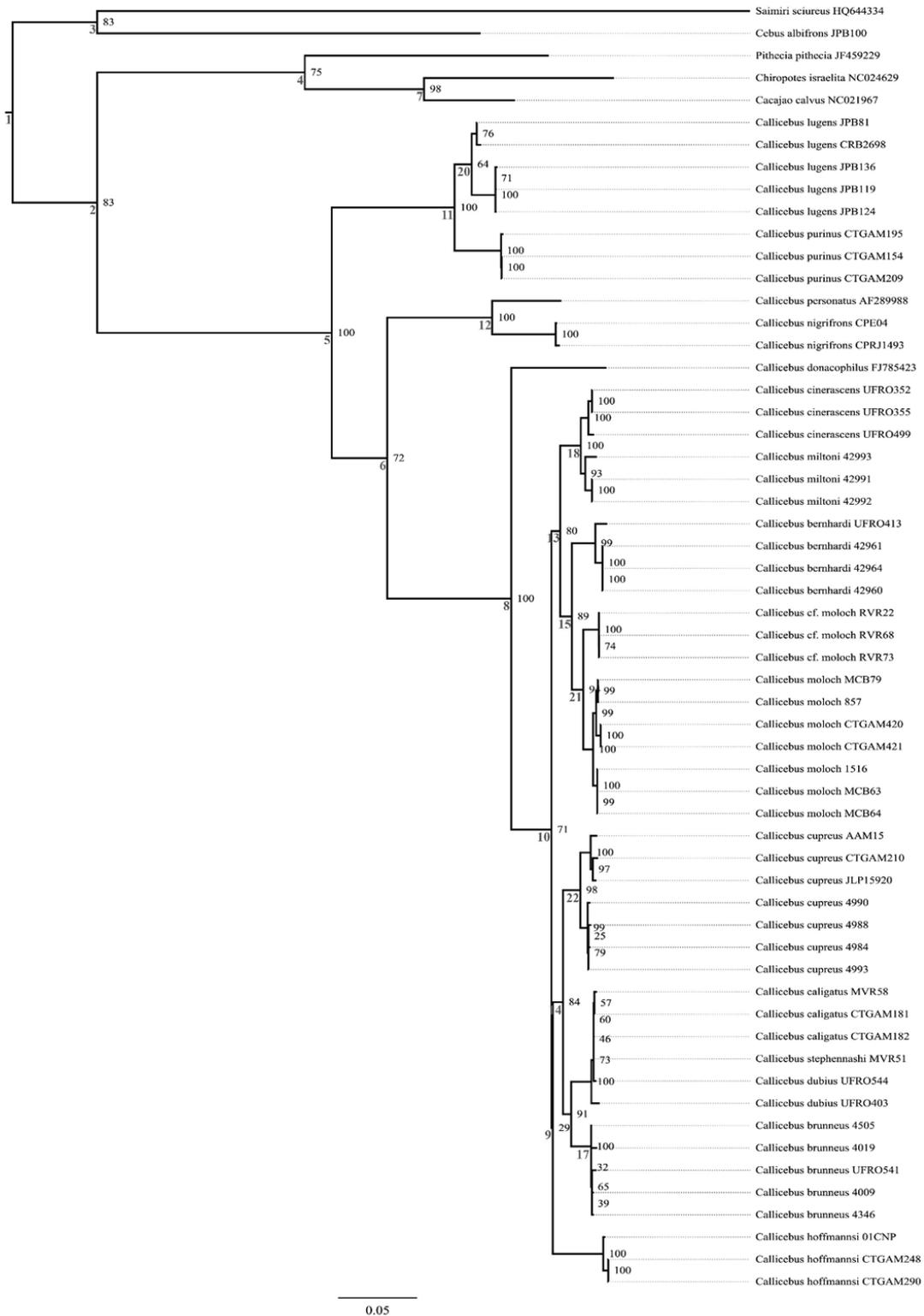
Figure A1.7 Maximum likelihood phylogeny inferred from the mitochondrial dataset. Node numbers represent nodes of interest listed in Table 2.9.

188
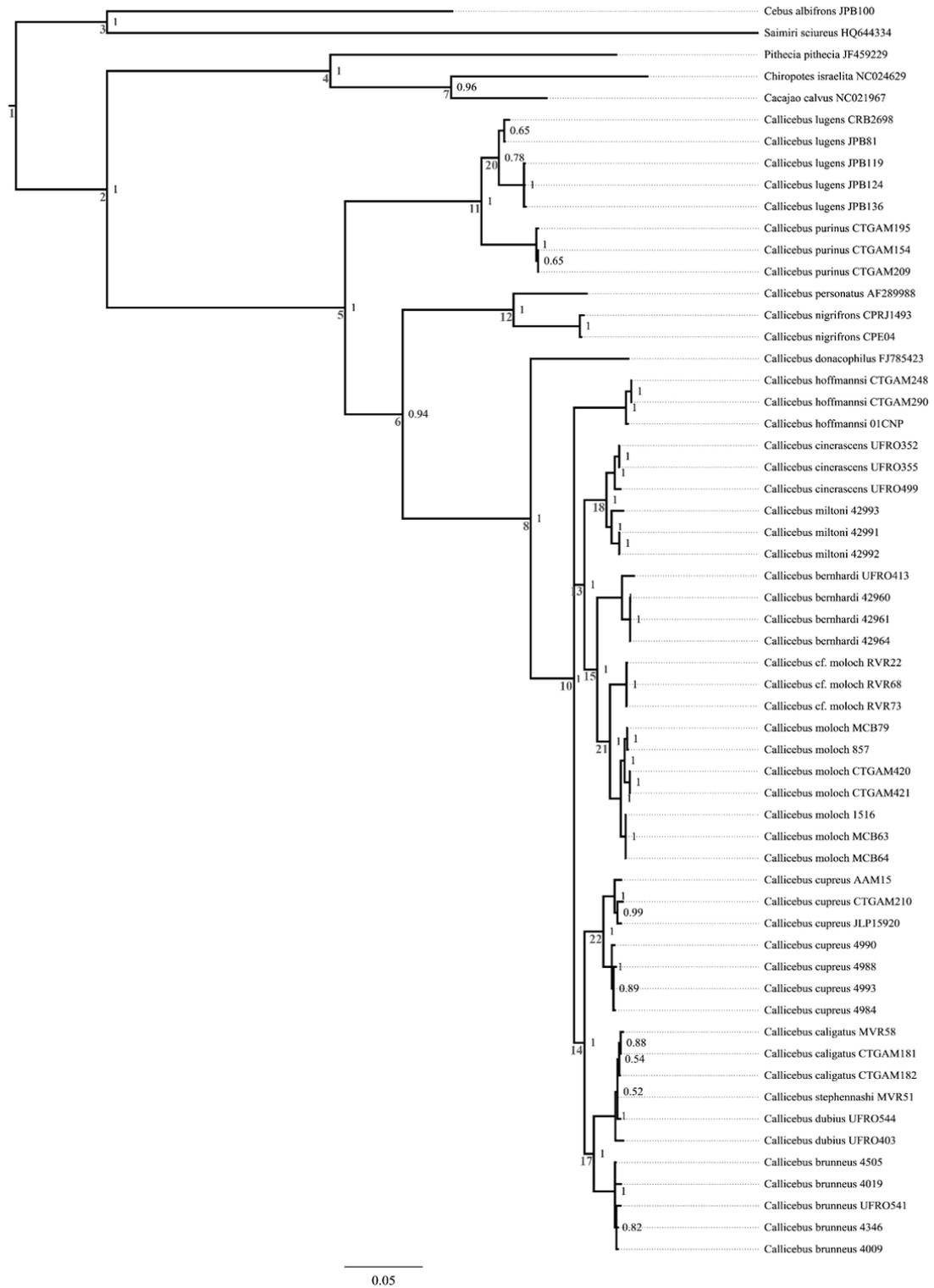
Figure A1.8 Bayesian phylogeny inferred from the mitochondrial dataset (MrBayes). Node numbers represent nodes of interest listed in Table 2.9.
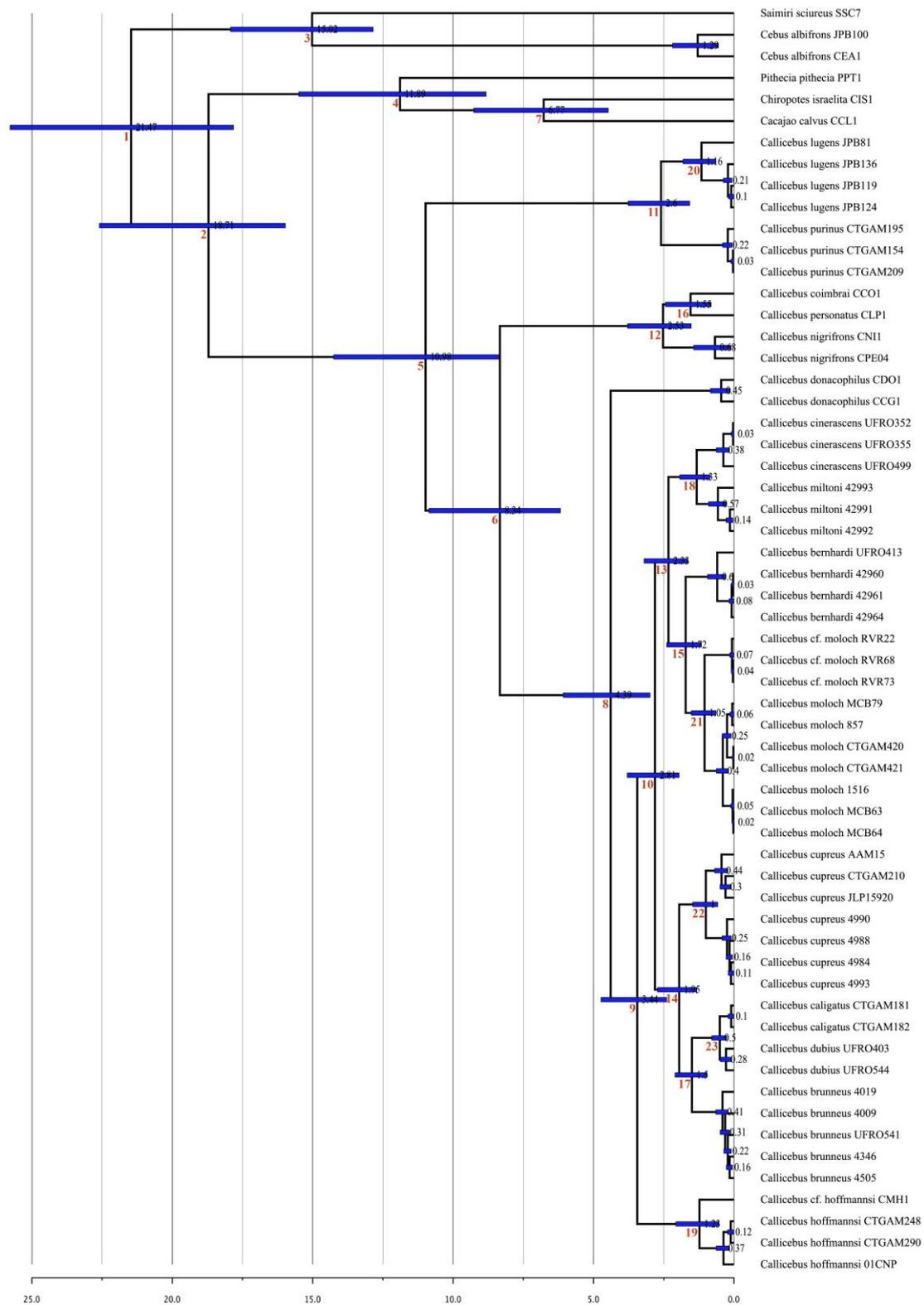
Figure A1.9 BEAST time-calibrated phylogeny inferred from the combined dataset. Node bars indicate the 95% highest posterior density. Red numbers represent nodes of interest listed in Table 2.9, 2.10.

190
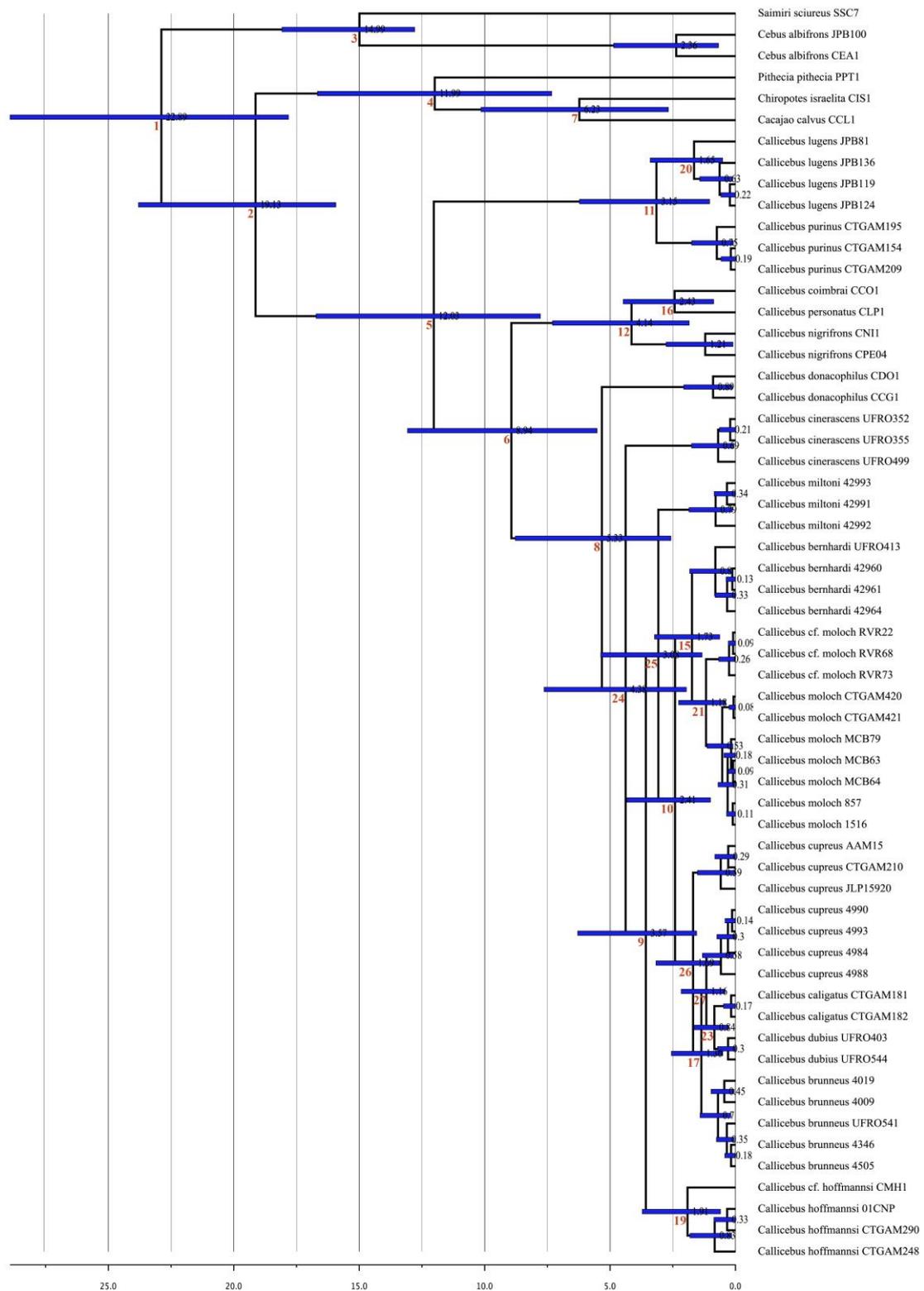
Figure A1.10 BEAST time-calibrated phylogeny inferred from the nuclear dataset. Node bars indicate the 95% highest posterior density. Red numbers represent nodes of interest listed in Table 2.9, 2.10.

191

Appendix 2: Supplementary material for Chapter 3: Biogeography of the titi monkeys (Callicebinae)
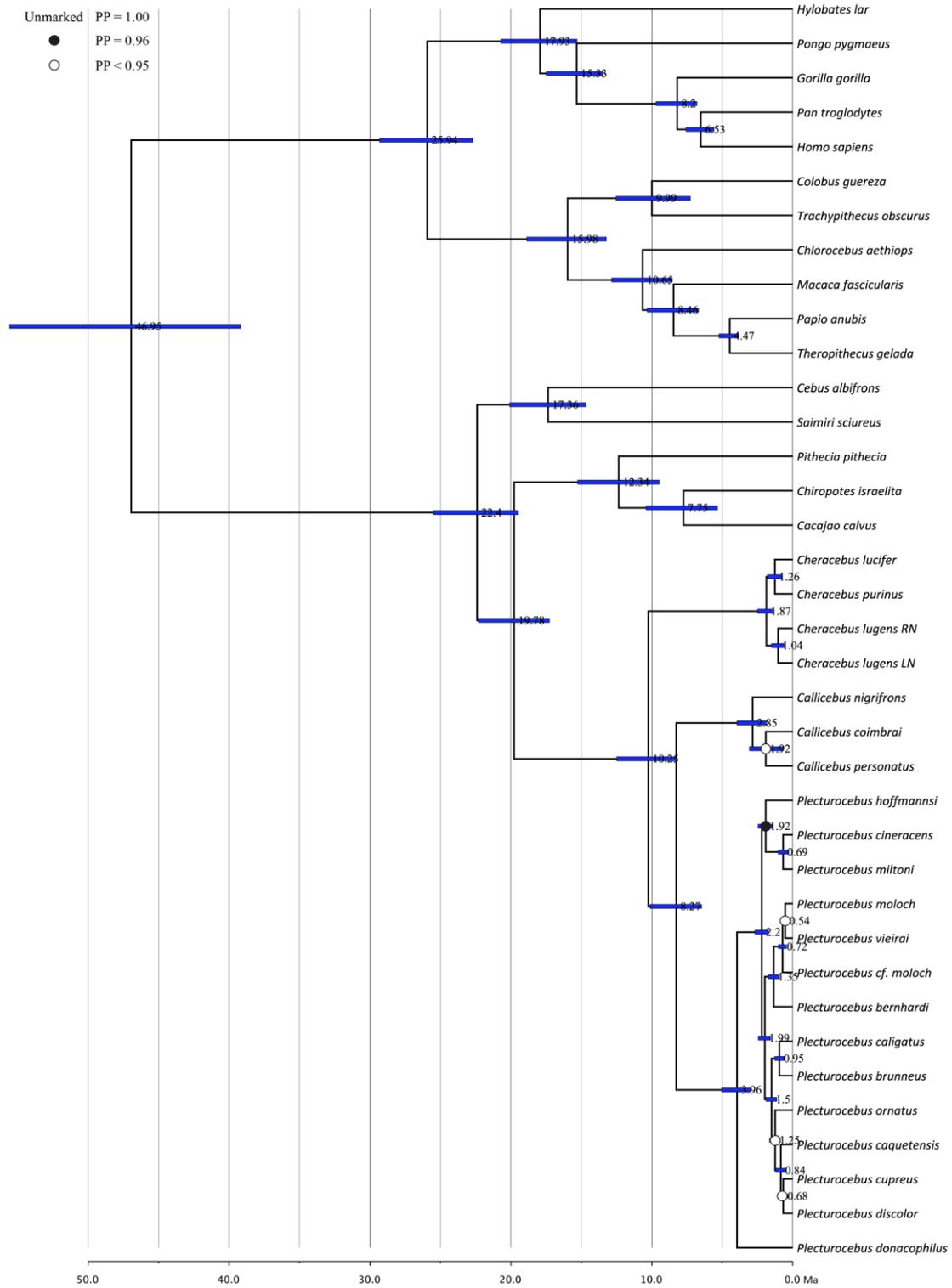
Figure A2.1 A time-calibrated phylogeny of Callicebinae with outgroups. Node bars indicate the 95% highest posterior density (HPDs).
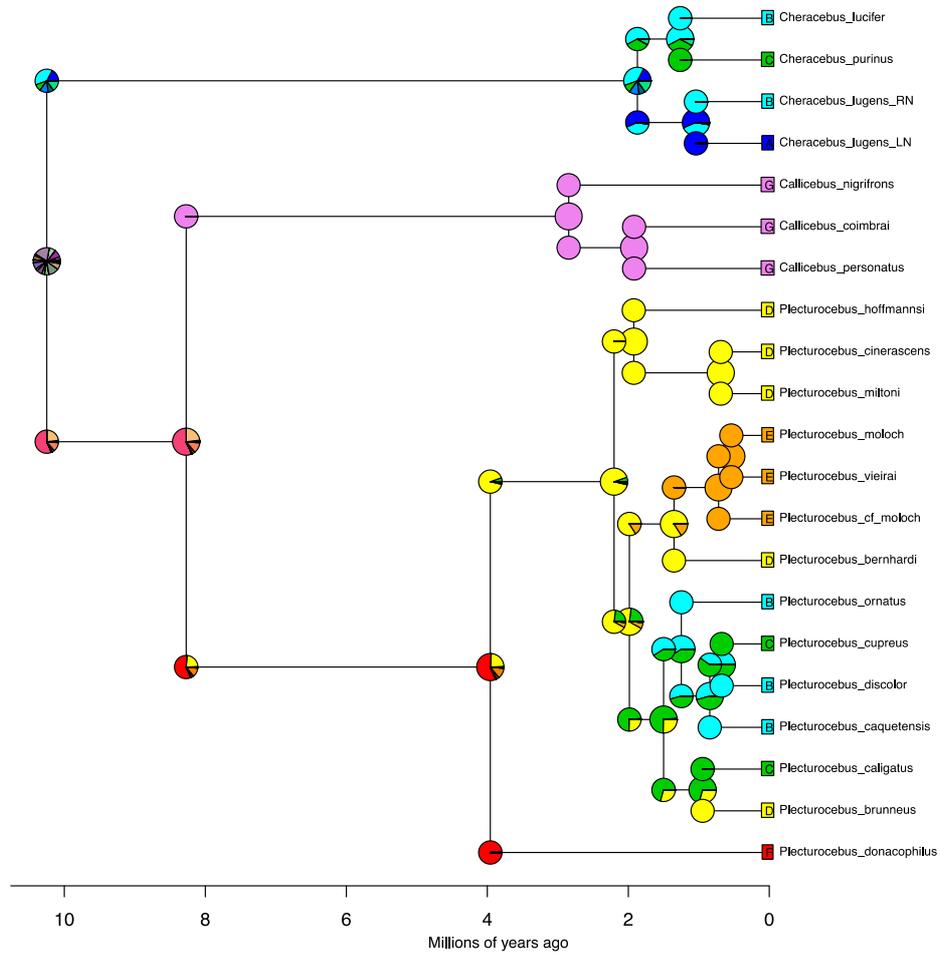
Figure A2.2 DIVALIKE+J+X ancestral area reconstruction state probabilities.

194
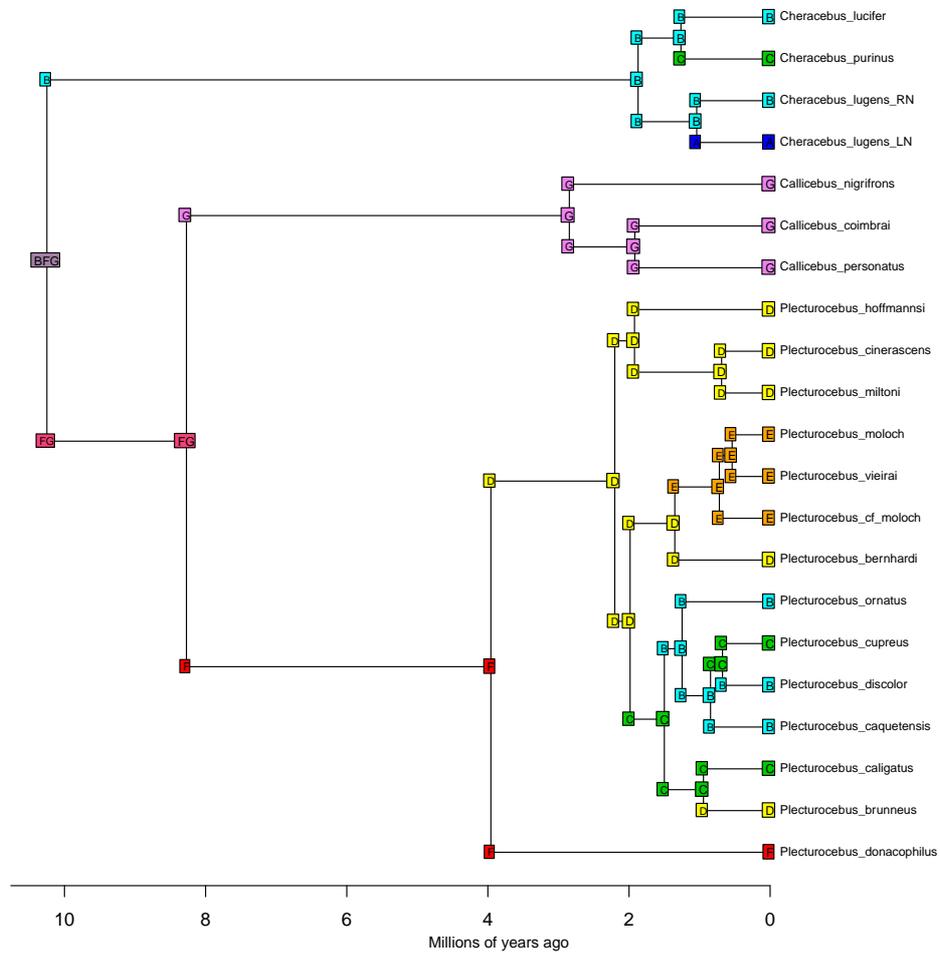
Figure A2.3 DEC+J+X reconstruction of ancestral areas for Callicebinae.
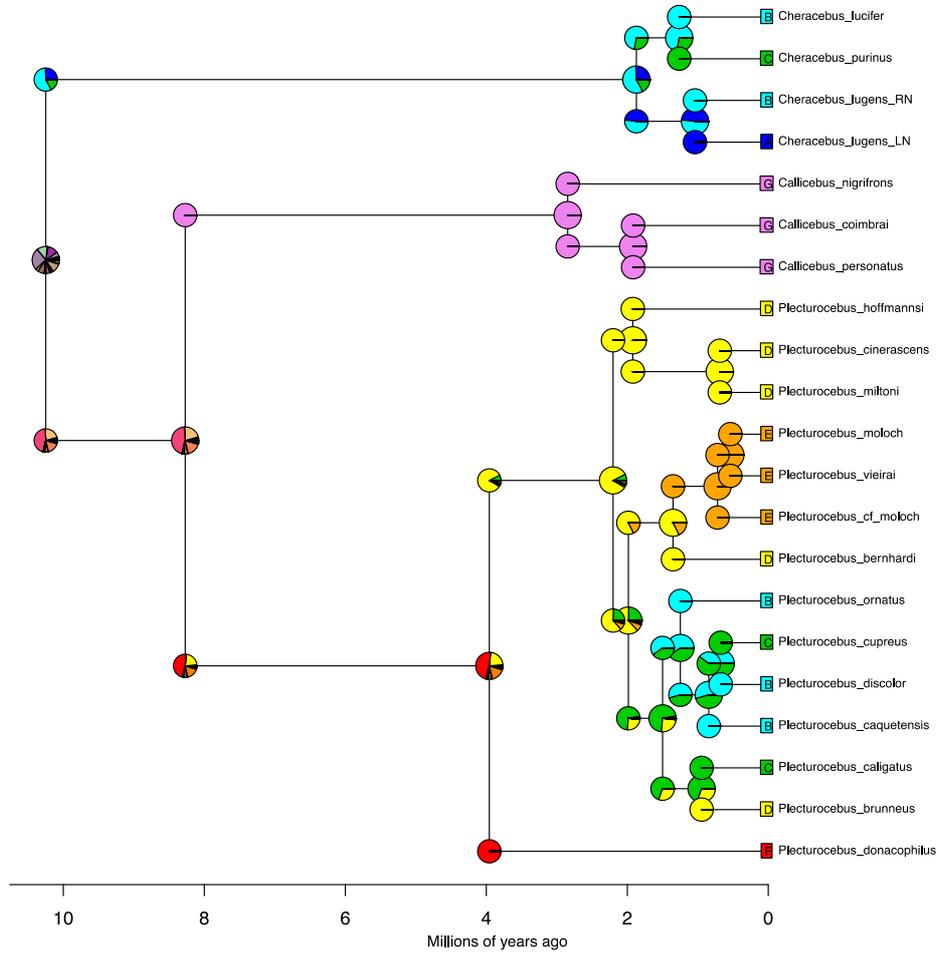
Figure A2.4 DEC+J+X ancestral area reconstruction state probabilities.
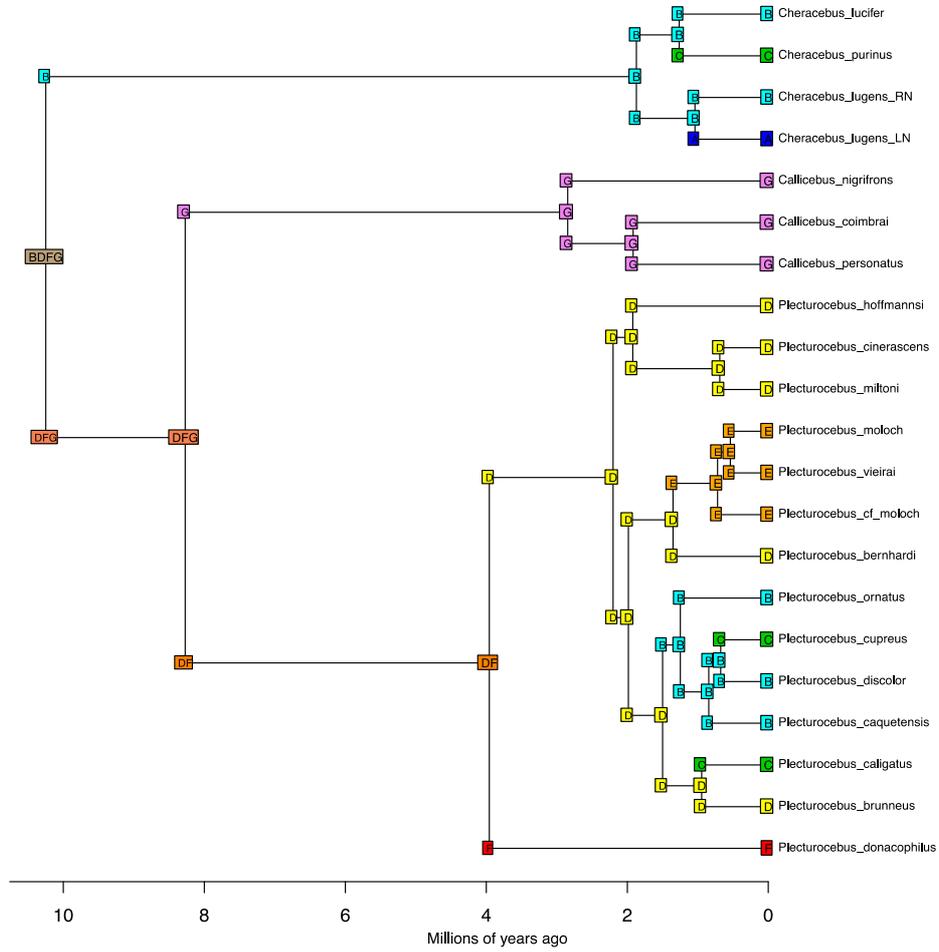
Figure A2.5 DIVALIKE+J reconstruction of ancestral areas for Callicebinae.

Appendix 3: Supplementary material for Chapter 4: Phylogenomics of titi monkeys (Callicebinae) using ddRADseq data with a focus on the *Plecturocebus moloch* group

Table A3.1 Summary of the ddRADseq data assembly (92% clustering threshold): sequencing information per sample.

Species	Sample ID	Reads	Reads passed	Clusters[1]	Avg. depth[1]	Cons. loci	No. sites	H[2]
P. hoffmannsi	CTGAM248	306337	151852	3626	16.72	3380	1001045	0.0030
P. hoffmannsi	CTGAM249	111856	55727	2701	12.49	2534	756042	0.0029
P. hoffmannsi	CTGAM290	102270	52813	2709	12.05	2515	750658	0.0029
P. miltoni	42991	103221	52457	2521	12.05	2388	711329	0.0019
P. miltoni	42992	26936	13386	774	7.43	702	210833	0.0020
P. cinerascens	UFRO352	273113	139449	3612	18.30	3385	996413	0.0030
P. cinerascens	UFRO354	340623	206741	4100	22.90	3898	1186116	0.0026
P. cinerascens	UFRO355	345291	184024	3320	15.51	3059	924006	0.0032
P. cinerascens	UFRO499	290049	175012	3905	21.36	3725	1139938	0.0023
P. cinerascens	WRS03	282987	178158	3981	21.91	3811	1164332	0.0024
P. cinerascens	WRS04	289893	174354	3801	21.68	3651	1112864	0.0024
P. bernhardi	42961	54657	27783	1648	9.18	1514	452463	0.0020
P. bernhardi	42964	162111	84364	3250	15.12	3053	907678	0.0022
P. bernhardi	UFRO413	276283	133737	2271	11.88	2088	615852	0.0023
P. bernhardi	CCM173	38910	20194	992	8.21	891	263922	0.0025
P. cf. moloch	RVR73	131902	66449	2766	13.33	2598	774822	0.0012
P. moloch	CTGAM420	150107	73122	2833	12.85	2641	794610	0.0014
P. moloch	CTGAM421	253243	120890	3426	17.61	3216	949606	0.0015
P. brunneus	4505	136376	68082	2819	13.68	2650	785664	0.0019
P. brunneus	UFRO327	403145	199843	2869	13.16	2637	770095	0.0021
P. brunneus	UFRO541	58000	29213	1624	10.20	1508	451519	0.0021
P. cupreus	AAM15	239043	119942	3759	16.38	3505	1035099	0.0017
P. cupreus	JLP15920	165045	85552	3224	15.01	3047	904992	0.0023
P. cupreus	CTGAM210	53988	25065	710	6.81	626	177410	0.0016
P. cupreus	4987	126201	60676	2569	12.93	2411	717087	0.0018
P. cupreus	4988	39680	19363	1190	8.35	1087	325986	0.0021
P. cupreus	4990	146403	75103	2902	13.28	2734	813133	0.0019
P. c. caligatus	CTGAM181	73412	33052	776	7.13	695	196204	0.0022
P. c. caligatus	CTGAM182	203884	104355	3385	16.15	3182	942583	0.0025
P. c. caligatus	CCM248	64071	30994	1858	9.65	1705	510282	0.0023
P. c. caligatus	MVR58	241597	118306	3523	16.31	3295	969515	0.0021
P. c. dubius	UFRO403	252706	124259	3383	16.34	3180	940994	0.0024
P. c. dubius	UFRO427	134817	62740	1253	8.91	1141	321840	0.0023
P. c. dubius	UFRO544	107090	47568	1944	11.83	1820	544637	0.0023
P. c. dubius	2804	127611	63963	2795	12.84	2626	779391	0.0023
C. personatus	2466	13162	6331	156	7.32	135	40451	0.0025
C. lugens	JPB160	85568	41459	1143	7.80	1048	296484	0.0018
C. lugens	JPB161	111957	51591	1065	8.42	966	270817	0.0017
C. lugens	JPB81	160013	78654	3080	14.39	2892	860491	0.0016
C. lugens	CTGAM733	101580	46833	2242	10.29	2061	630121	0.0018
C. lucifer	CTGAM703	50854	24646	1451	8.51	1326	411845	0.0024
C. lucifer	CTGAM726	47586	24161	1512	8.65	1405	421313	0.0026
C. purinus	CTGAM154	29785	13879	750	6.99	666	206696	0.0023
C. purinus	CTGAM195	82079	37058	997	7.58	902	255596	0.0020
C. purinus	CTGAM209	173405	86094	3108	15.00	2941	871620	0.0019
Pithecia mittermeieri	CTGAM215	49088	22116	1079	8.19	989	306427	0.0011
Cacajao ayresi	5667	75721	35587	1181	8.01	1069	302724	0.0009
Cacajao calvus	5241	109827	52274	1056	8.39	975	273653	0.0010
Cacajao hosomi	5698	94259	44687	1243	8.78	1146	323431	0.0009
Cacajao melanocephalus	0065	75938	35650	1025	7.83	937	264654	0.0012
Chiropotes albinasus	CTGAM213	82121	38938	1026	8.02	941	265761	0.0012
Chiropotes israelita	5713	105778	50461	1355	8.69	1263	356047	0.0014
Chiropotes sagalatus	CTGAM515	96369	45135	1188	8.41	1069	301732	0.0019

[1] After excluding loci with depth <5

[2] Heterozygosity measured as the proportion of called sites

Table A3.2 Pairwise genetic distances between individuals of the Western clade.

Species	Sample ID	4505	UFRO327	UFRO541	CCM248	CTGAM182	MVR58	UFRO403	UFRO544	2804	AAM15	JLP15920	4987	4988
P. brunneus	4505	--												
P. brunneus	UFRO327	0.0013	--											
P. brunneus	UFRO541	0.0012	0.0013	--										
P. c. caligatus	CCM248	0.0020	0.0019	0.0021	--									
P. c. caligatus	CTGAM182	0.0021	0.0018	0.0020	0.0013	--								
P. c. caligatus	MVR58	0.0021	0.0019	0.0023	0.0012	0.0011	--							
P. c. dubius	UFRO403	0.0019	0.0017	0.0020	0.0015	0.0011	0.0012	--						
P. c. dubius	UFRO544	0.0021	0.0021	0.0020	0.0016	0.0014	0.0015	0.0014	--					
P. cupreus (A)	2804	0.0019	0.0019	0.0021	0.0015	0.0013	0.0014	0.0011	0.0013	--				
P. cupreus (A)	AAM15	0.0022	0.0021	0.0025	0.0020	0.0018	0.0019	0.0016	0.0020	0.0018	--			
P. cupreus (A)	JLP15920	0.0021	0.0020	0.0022	0.0018	0.0016	0.0018	0.0014	0.0019	0.0016	0.0012	--		
P. cupreus (B)	4987	0.0021	0.0020	0.0024	0.0018	0.0017	0.0019	0.0015	0.0018	0.0017	0.0014	0.0012	--	
P. cupreus (B)	4988	0.0022	0.0021	0.0022	0.0019	0.0020	0.0021	0.0019	0.0019	0.0019	0.0017	0.0014	0.0012	--
P. cupreus (B)	4990	0.0022	0.0021	0.0022	0.0019	0.0017	0.0018	0.0014	0.0018	0.0017	0.0014	0.0012	0.0010	0.0010

Table A3.3 Additional four-taxon D-statistic tests among Western clade taxa.

P1 taxon	P2 taxon	P3 taxon	Range Z	Sign./total	No. loci
P. c. caligatus	*P. c. dubius*	*P. brunneus*	0.0 - 1.7	0/26	78 - 566
P. c. caligatus	*P. c. dubius*	*P. cupreus*	0.0 - 3.6	0/44	97 - 558
P. cupreus (A)	*P. cupreus* (B)	*P. c. dubius*	0.0 - 2.4	0/17	62 - 577
P. cupreus (A)	*P. cupreus* (B)	*P. c. caligatus*	0.1 - 1.3	0/17	71 - 555
P. cupreus (A)	*P. cupreus* (B)	*P. brunneus*	0.0 - 2.5	0/17	80 - 584

Table A3.4 Pairwise genetic distances between individuals of the Eastern and Aripuanã-Tapajós clades. Red indicates distances between *P. bernhardi* and admixed *P. cinerascens* individuals.

Species	Sample ID	CTGAM420	CTGAM421	RVR73	CCM173	UFRO413	42961	42964	UFRO352	UFRO354	UFRO355	UFRO499
P. moloch	CTGAM420	--										
P. moloch	CTGAM421	0.0008	--									
P. cf. *moloch*	RVR73	0.0021	0.0021	--								
P. bernhardi	CCM173	0.0029	0.0030	0.0027	--							
P. bernhardi	UFRO413	0.0027	0.0027	0.0027	0.0023	--						
P. bernhardi (A)	42961	0.0031	0.0033	0.0029	0.0024	0.0025	--					
P. bernhardi (A)	42964	0.0027	0.0029	0.0028	0.0021	0.0020	0.0014	--				
P. cinerascens (B)	UFRO352	0.0033	0.0032	0.0033	0.0028	0.0024	0.0027	0.0025	--			
P. cinerascens (B)	UFRO354	0.0036	0.0037	0.0036	0.0032	0.0029	0.0035	0.0031	0.0008	--		
P. cinerascens (B)	UFRO355	0.0034	0.0036	0.0036	0.0030	0.0023	0.0030	0.0027	0.0011	0.0009	--	
P. cinerascens (A)	UFRO499	0.0040	0.0039	0.0040	0.0033	0.0034	0.0037	0.0035	0.0014	0.0012	0.0014	--
P. cinerascens (A)	WRS03	0.0040	0.0040	0.0040	0.0034	0.0034	0.0036	0.0035	0.0015	0.0013	0.0016	0.0013
P. cinerascens (A)	WRS04	0.0039	0.0041	0.0040	0.0035	0.0034	0.0038	0.0035	0.0015	0.0014	0.0015	0.0014
P. miltoni	42991	0.0043	0.0043	0.0042	0.0038	0.0037	0.0039	0.0038	0.0029	0.0030	0.0033	0.0032
P. miltoni	42992	0.0046	0.0045	0.0044	0.0036	0.0038	0.0038	0.0041	0.0029	0.0032	0.0032	0.0032
P. hoffmannsi	CTGAM248	0.0042	0.0042	0.0044	0.0042	0.0038	0.0041	0.0038	0.0031	0.0033	0.0032	0.0035
P. hoffmannsi	CTGAM249	0.0042	0.0042	0.0041	0.0037	0.0037	0.0039	0.0037	0.0032	0.0033	0.0034	0.0035
P. hoffmannsi	CTGAM290	0.0040	0.0041	0.0041	0.0034	0.0035	0.0038	0.0037	0.0033	0.0033	0.0033	0.0034

Species	Sample ID	WRS03	WRS04	42991	42992	CTGAM248	CTGAM249
P. cinerascens (A)	WRS03	--					
P. cinerascens (A)	WRS04	0.0011	--				
P. miltoni	42991	0.0031	0.0032	--			
P. miltoni	42992	0.0034	0.0033	0.0011	--		
P. hoffmannsi	CTGAM248	0.0033	0.0034	0.0040	0.0044	--	
P. hoffmannsi	CTGAM249	0.0033	0.0035	0.0035	0.0038	0.0009	--
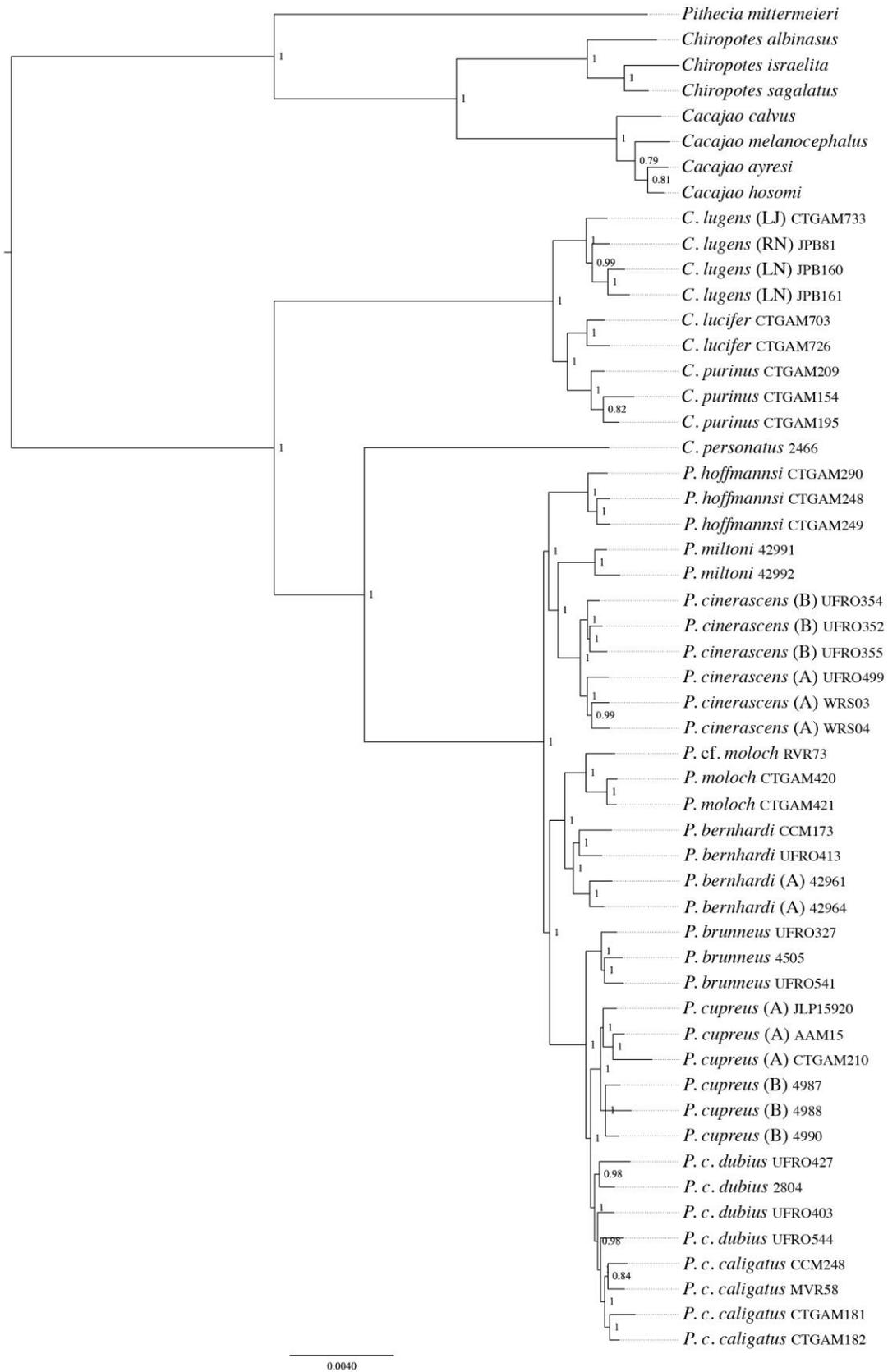
Figure A3.1 Bayesian tree inferred with the ddRADseq A92 dataset.
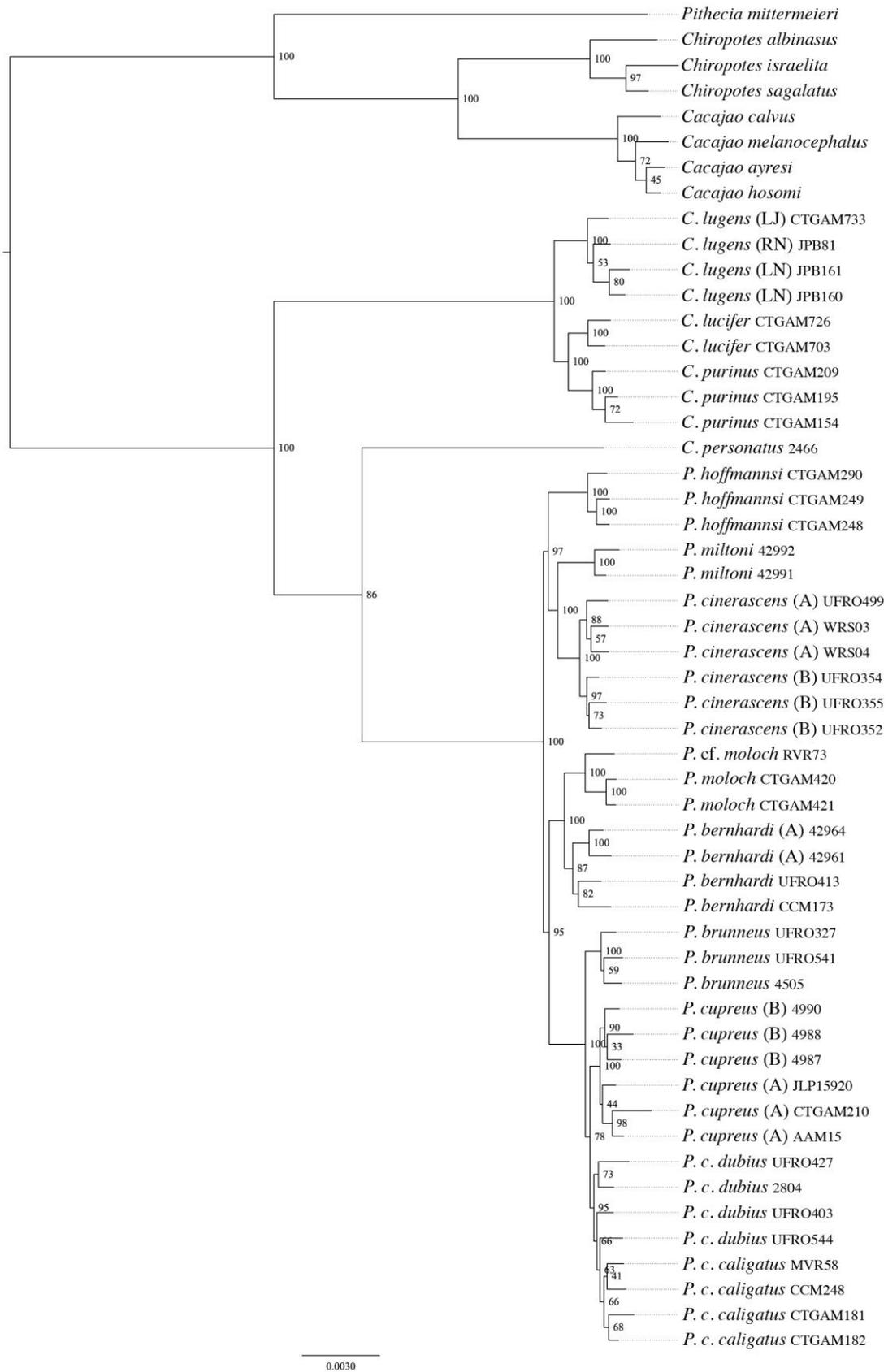
Figure A3.2 Maximum likelihood tree inferred with the ddRADseq A92 dataset.
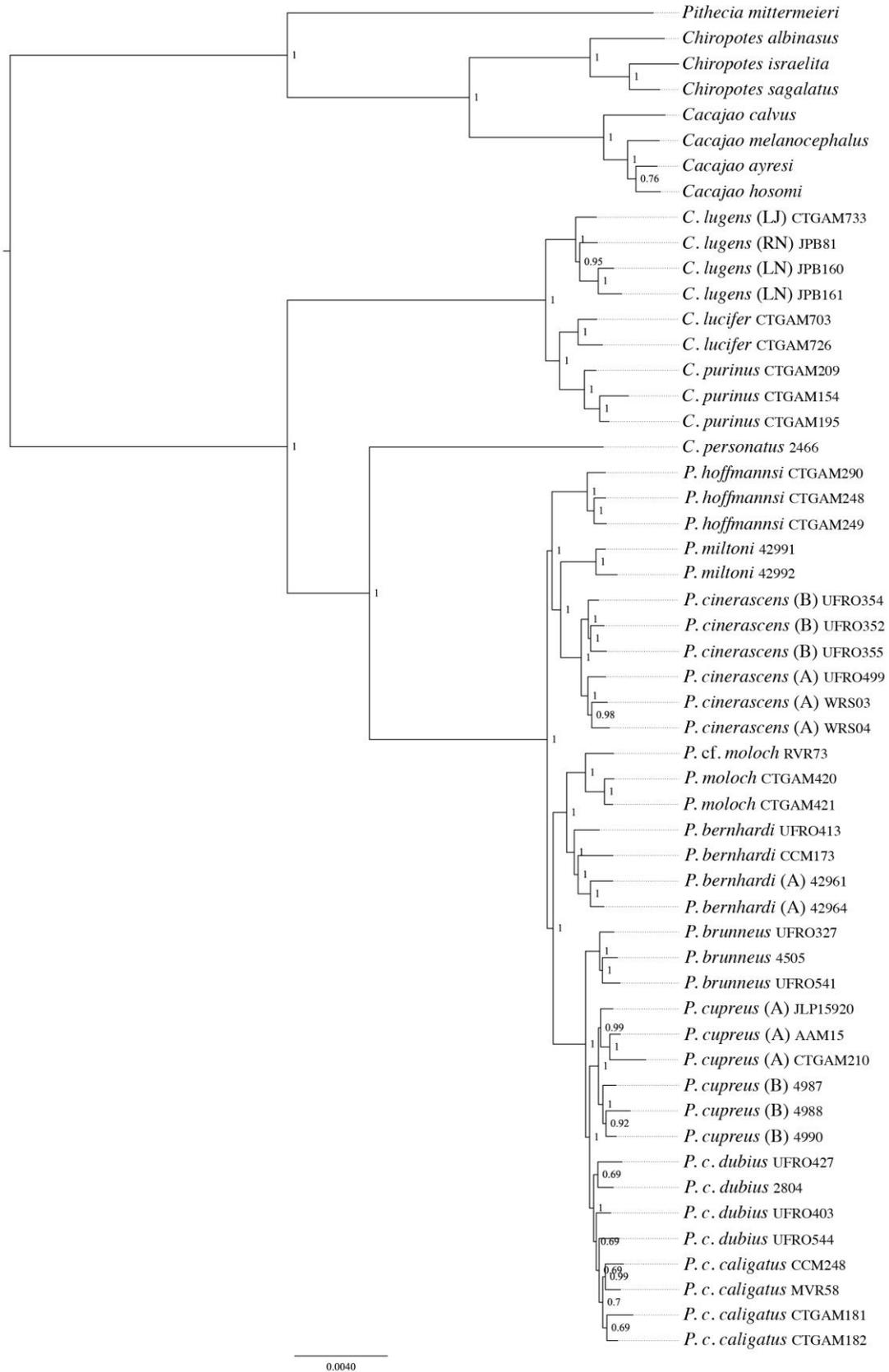
Figure A3.3 Bayesian tree inferred with the ddRADseq A85 dataset.
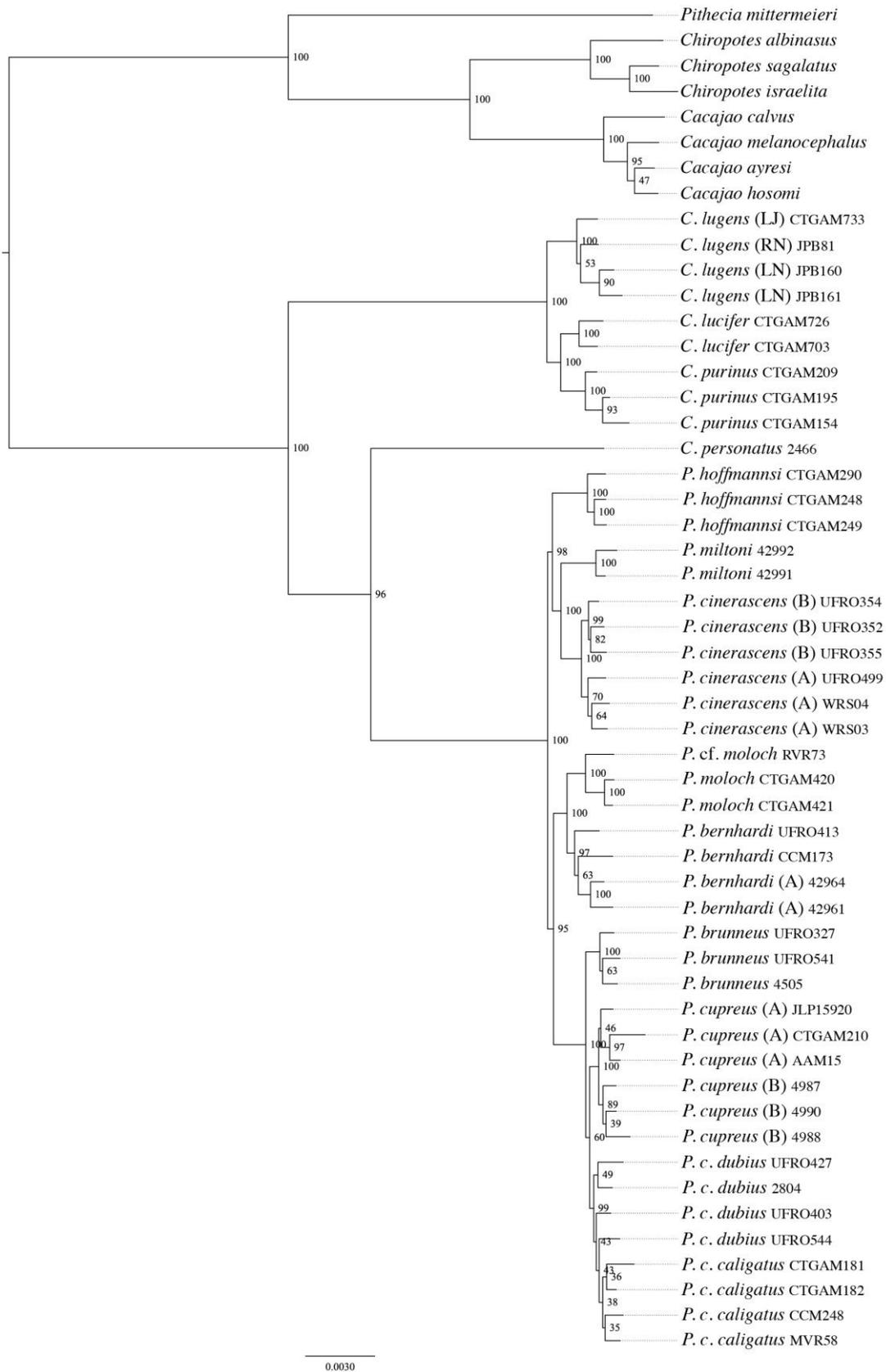
Figure A3.4 Maximum likelihood tree inferred with the ddRADseq A85 dataset.
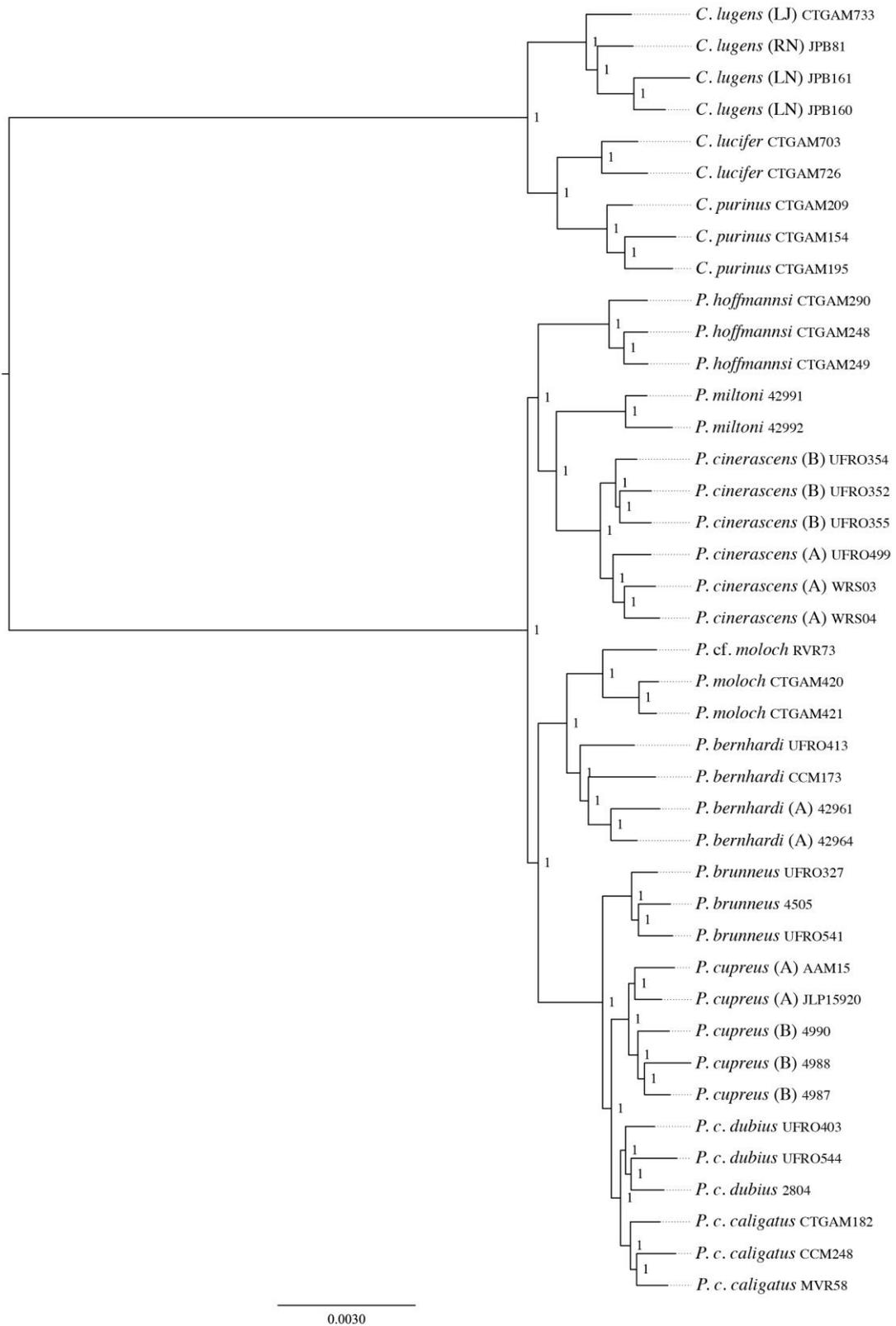
Figure A3.5 Bayesian tree inferred with the ddRADseq PD85 dataset.
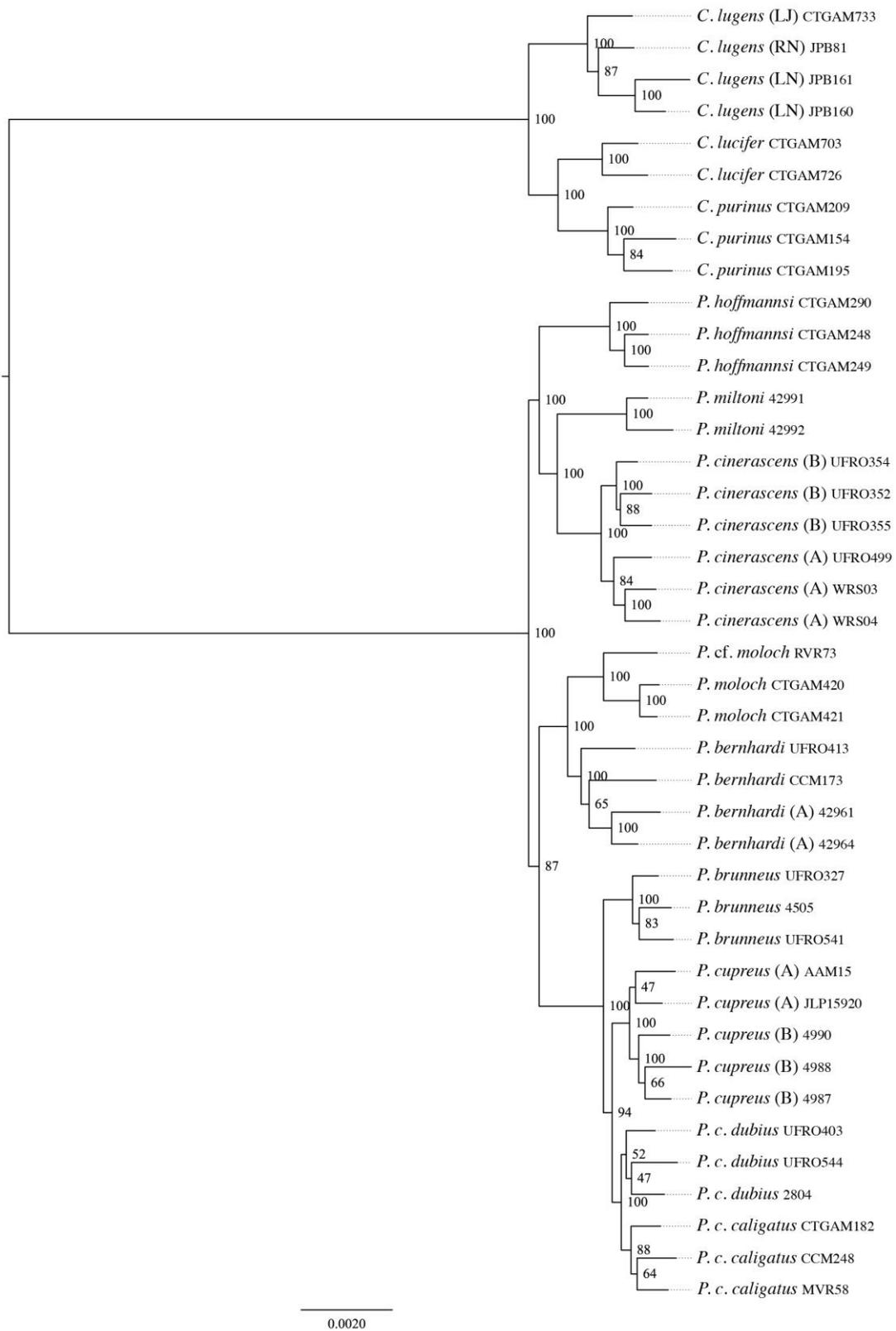
Figure A3.6 Maximum likelihood tree inferred with the ddRADseq PD85 dataset.
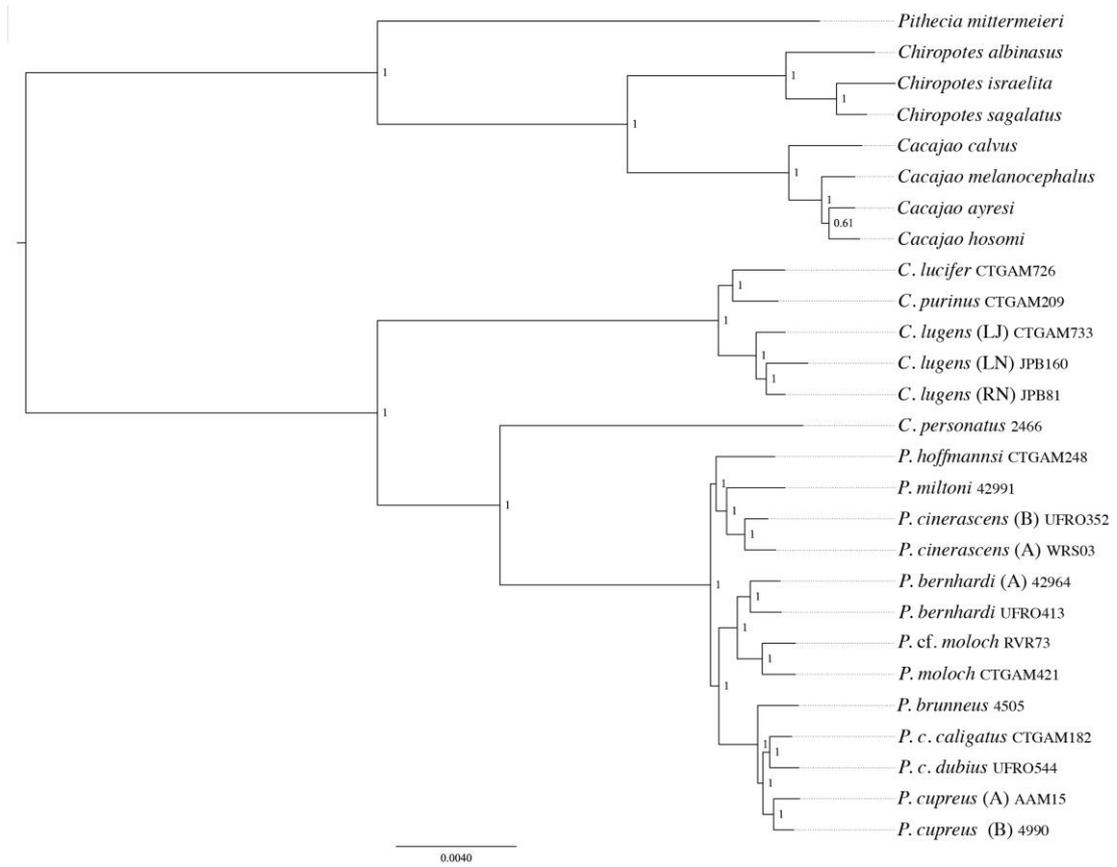
Figure A3.7 Bayesian tree inferred with the ddRADseq B85 dataset.
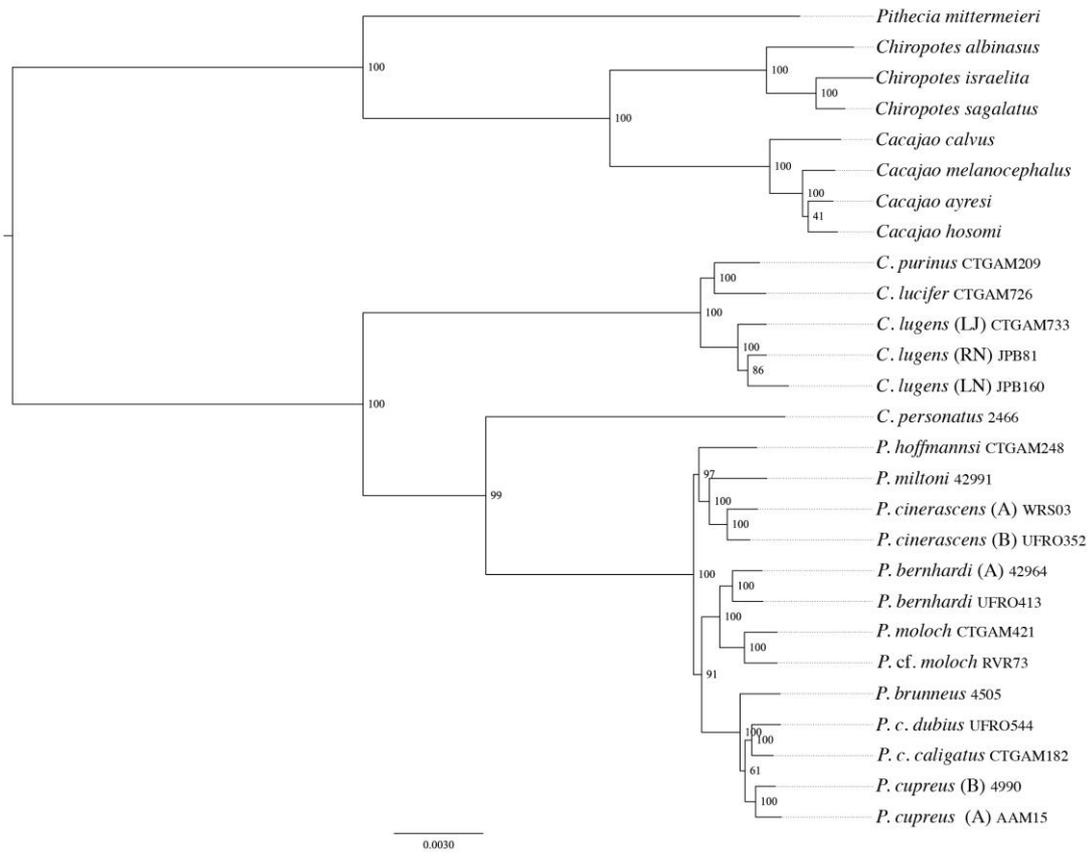
Figure A3.8 Maximum likelihood tree inferred with the ddRADseq B85 dataset.
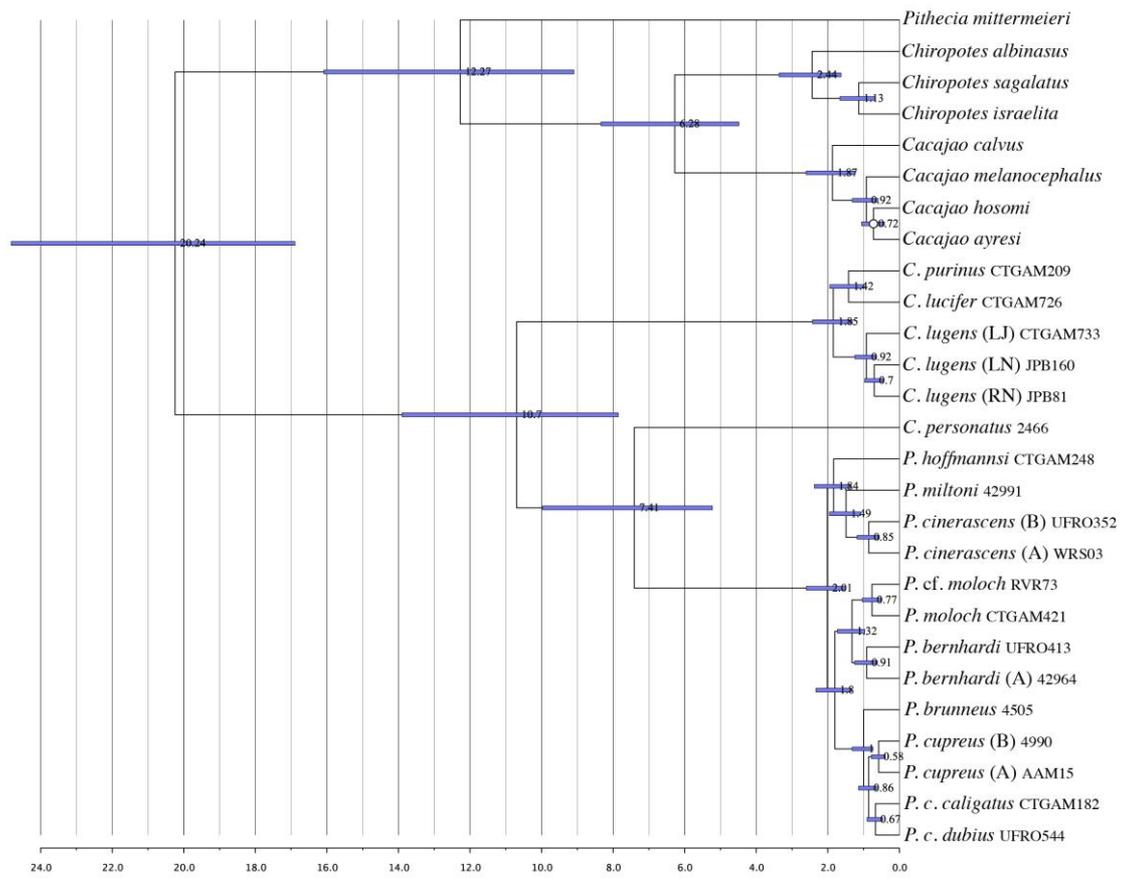
Figure A3.9 A time-calibrated phylogeny for Callicebinae with outgroups inferred with the ddRADseq B85 dataset. All nodes received full support (PP = 1.00). Node bars indicate the 95% highest posterior density (HPD).
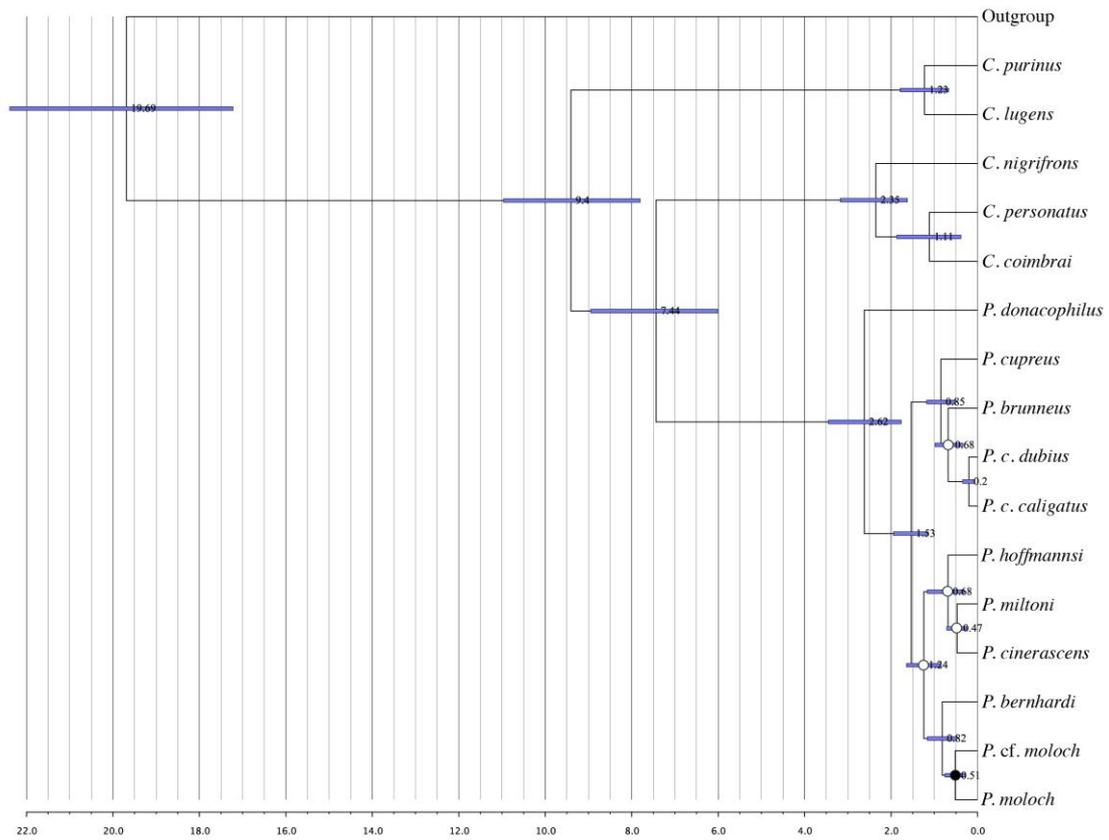
Figure A3.10 A time-calibrated coalescent-based species tree for Callicebinae with outgroups inferred with multi-locus data using StarBEAST2. Unmarked nodes received full support (PP = 1.00), the node marked with a black circle received significant support (PP = 0.95), while the nodes marked with white circles were recovered without significant support (PP < 0.95). Node bars indicate the 95% highest posterior density (HPD).

References

Aleixo A, Rossetti DF. 2007. Avian gene trees, landscape evolution, and geology: towards a modern synthesis of Amazonian historical biogeography? *Journal of Ornithology*, 148 (suppl. 2), S443–S453.

Alfaro ME, Zoller S, Lutzoni F. 2003. Bayes or Bootstrap? A simulation study comparing the performance of Bayesian Markov Chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. *Molecular Biology & Evolution,* 20(2), 255–266.

Allgas N, Shanee S, Shanee N, Chambers J, Tello-Alvarado JC, Keeley K, Pinasco K. 2017. Natural re-establishment of a population of a critically endangered primate in a secondary forest: the San Martin titi monkey (*Plecturocebus oenanthe*) at the Pucunucho Private Conservation Area, Peru. *Primates*, 58(2), 335–342.

Aquino R, Encarnación F. 1994. Primates of Peru. *Primate Report,* (40), 1–127.

Aquino R, Terrones W, Cornejo F, Heymann EW. 2008. Geographic distribution and possible taxonomic distinction of *Callicebus torquatus* populations (Pitheciidae: Primates) in Peruvian Amazonia. *American Journal of Primatology,* 70(12), 1181–1186.

Araújo NP, do Espírito Santo AA, do Socorro Pereira V, Stanyon R, Svartman M. 2017. Chromosome painting in *Callicebus nigrifrons* provides insights into the genome evolution of titi monkeys and the ancestral Callicebinae karyotype. *Cytogenetic & Genome Research*, DOI: 10.1159/000458748.

Auricchio P. 2010. A morphological analysis of some species of *Callicebus*, Thomas, 1903 (Pitheciidae – Callicebinae). *Neotropical Primates,* 17(2), 47–58.

Avise JC, Shapira JF, Daniel SW, Aquadro CF, Lansman, RA. 1983. Mitochondrial DNA differentiation during the speciation process in *Peromyscus*. *Molecular Biology & Evolution,* 1(1), 38–56.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLOS ONE*, 3(10), e3376.

Batalha-Filho H, Fjeldsa J, Fabre PH, Miyaki CY. 2013. Connections between the Atlantic and the Amazonian forest avifaunas represent distinct historical

events. *Journal of Ornithology*, 154, 41–50.

Baum DA, DeWitt SS, Donovan SS. 2005. The tree-thinking challenge. *Science*, 310, 979–80.

Bayzid MS, Warnow T. 2013. Naive binning improves phylogenomic analyses. *Bioinformatics,* 29(18), 2277–2284.

Benefit BR, McCrossin ML. 2002. The Victoriapithecidae, Cercopithecoidea. In: Hartwig WC, (Eds.). *The primate fossil record.* Cambridge University Press, Cambridge, pp. 241–253.

Benirschke K, Bogart MH. 1976. Chromosomes of the tan-handed titi (*Callicebus torquatus* Hoffmannsegg, 1807). *Folia Primatologica,* 25(1), 25–34.

Betancur-R R, Broughton RE, Wiley EO, Carpenter K, López JA, Li C., Holcroft NI, Arcila D, Sanciangco M, Cureton JC, Zhang F, Buser T, Campbell MA, Ballesteros JA, Roa-Varon A, Willis S, Borden WC, Rowley T, Reneau PC, Hough DJ, Lu G, Grande T, Arratia G, Ortí G. 2013. The tree of life and a new classification of bony fishes. *PLOS Currents Tree of Life,* DOI: 10.1371/currents.tol.53ba26640df0ccaee75bb1 65c8c26288.

Bicca-Marques JC, Heymann EW. 2013. Ecology and behavior of titi monkeys (*Callicebus*). In: Veiga LM, Barnett AA, Ferrari SF, Norconk MA, (Eds). *Evolutionary biology and conservation of titis, sakis and uacaris*. Cambridge University Press, Cambridge, UK, pp. 196–207.

Botero-Castro F, Tilak MK, Justy F, Catzeflis F, Delsuc F, Douzery EJ. 2013. Next-generation sequencing and phylogenetic signal of complete mitochondrial genomes for resolving the evolutionary history of leaf-nosed bats (Phyllostomidae). *Molecular Phylogenetics & Evolution,* 69(3), 728–739.

Boubli JP, Ribas C, Lynch Alfaro JW, Alfaro ME, da Silva MNF, Pinho GM, Farias IP. 2015. Spatial and temporal patterns of diversification on the Amazon: A test of the riverine hypothesis for all diurnal primates of Rio Negro and Rio Branco in Brazil. *Molecular Phylogenetics & Evolution*, 82B, 400–412.

Bouckaert RR. 2010. DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics*, 26, 1372–1373.

Bouckaert RR, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLOS Computational Biology*, 10, e1003537.

Bragg LM, Stone G, Butler MK, Hugenholtz P, Tyson GW. 2013. Shining a light on

dark sequencing: Characterising errors in Ion Torrent PGM data. *PLOS Computational Biology*, 9(4), e1003031.

Brunet M, Guy F, Pilbeam D, Mackaye HT, Likius A, Ahounta D, Beauvilain A, Blondel C, Bocherens H, Boisserie JR, De Bonis L, Coppens Y, Dejax J, Denys C, Duringer P, Eisenmann V, Fanone G, Fronty P, Geraads D, Lehmann T, Lihoreau F, Louchart A, Mahamat A, Merceron G, Mouchelin G, Otero O, Pelaez Campomanes P, Ponce De Leon M, Rage JC, Sapanet M, Schuster M, Sudre J, Tassy P, Valentin X, Vignaud P, Viriot L, Zazzo A, Zollikofer C. 2002. A new hominid from the Upper Miocene of Chad, Central Africa. *Nature*, 418, 145–151.

Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, RoyChoudhury A. 2012. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Molecular Biology & Evolution*, 29(8), 1917–1932.

Buckner JC, Lynch Alfaro JW, Rylands AB, Alfaro ME. 2015. Biogeography of the marmosets and tamarins (Callitrichidae). *Molecular Phylogenetics & Evolution*, 82B, 413–425.

Bueno ML, Defler TR. 2010. Cytogenetical approach to clarify the taxonomy of the genus *Callicebus*. *Orinoquia,* 14(1), 139–152.

Bueno ML, Torres OM, Ramírez-Orjuela C, Leibovici M. 2006. Información cariológica del género *Callicebus* en Colombia. *Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales,* 30(114), 109–115.

Burbrink FT, Pyron RA. 2011. The impact of gene-tree/species-tree discordance on diversification-rate estimation. *Evolution*, 65, 1851–1861.

Byrne H, Rylands AB, Carneiro JC, Lynch Alfaro JW, Bertuol F, da Silva MNF, Messias M, Groves CP, Mittermeier RA, Farias I, Hrbek T, Schneider H, Sampaio I, Boubli JP. 2016. Phylogenetic relationships of the New World titi monkeys (*Callicebus*): First appraisal of taxonomy based on molecular evidence. *Frontiers in Zoology*, 13(10), 1–25.

Cabrera A. 1958. Catálogo de los mamíferos de América del Sur [part 1]. *Revista del Museo Argentino de Ciencias Naturales "Bernardino Rivadavia" (Ciencias Zoológicas),* 4(1), 1–307.

Campbell Jr KE, Frailey CD, Romero-Pittman L. 2006. The Pan-Amazonian Ucayali Peneplain, late Neogene sedimentation in Amazonia, and the birth of the

modern Amazon River system. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 239, 166–219.

Canavez FC, Moreira MAM, Ladasky JJ, Pissinatti A, Parham P, Seuánez HN. 1999. Molecular phylogeny of New World primates (Platyrrhini) based on β2-microglobulin DNA sequences. *Molecular Phylogenetics & Evolution,* 12(1), 74–82.

Carneiro JC, Silva Júnior JS, Sampaio I, Pissinatti A, Hrbek T, Messias M, Röhe F, Farias I, Boubli JP, Schneider H. 2016. Phylogeny of the titi monkeys of the *Callicebus moloch* group (Pitheciidae, Primates). *American Journal of Primatology*, 78(9), 904–913.

Catchen JM, Amores A, Hohenlohe PA, Cresko WA, Postlethwait JH. 2011. Stacks: building and genotyping loci *de novo* from short-read sequences. *G3: Genes, Genomes, Genetics,* 1(3), 171–182.

Catchen JM, Hohenlohe PA, Bassham S, Amores A, Cresko WA. 2013. Stacks: an analysis tool set for population genomics. *Molecular Ecology*, 22(11), 3124–3140.

Chagas RRD, Ferrari SF. 2010. Habitat use by *Callicebus coimbrai* (Primates: Pitheciidae) and sympatric species in the fragmented landscape of the Atlantic forest of southern Sergipe, Brazil. *Zoologia (Curitiba),* 27(6), 853–860.

Chattopadhyay B, Garg KM, Kumar AV, Doss DPS, Rheindt FE, Kandula S, Ramakrishnan U. 2016. Genome-wide data reveal cryptic diversity and genetic introgression in an Oriental cynopterine fruit bat radiation. *BMC Evolutionary Biology*, 16, 41.

Chifman J, Kubatko L. 2014. Quartet inference from SNP data under the coalescent model. *Bioinformatics,* 30(23), 3317–24.

Combosch DJ, Vollmer SV. 2015. Trans-Pacific RAD-Seq population genomics confirms introgressive hybridization in Eastern Pacific *Pocillopora* corals. *Molecular Phylogenetics & Evolution*, 88, 154–162.

Cortés-Ortiz L, Duda TF, Canales-Espinosa D, García-Orduña F, Rodríguez-Luna E, Bermingham E. 2007. Hybridization in large-bodied New World primates. *Genetics*, 176(4), 2421–2425.

Cracraft J. 1985. Historical biogeography and patterns of differentiation within the South American avifauna: areas of endemism. *Ornithological Monographs*, 36, 49–84.

Cruaud A, Gautier M, Galan M, Foucaud J, Sauné L, Genson G, Dubois E, Nidelet S, Deuve T, Rasplus JY. 2014. Empirical assessment of RAD sequencing for interspecific phylogeny. *Molecular Biology & Evolution*, 31(5), 1272–1274.

Dalponte JC, Silva FE, Silva-Júnior JS. 2014. New species of titi monkey, genus *Callicebus* Thomas, 1903 (Primates, Pitheciidae), from Southern Amazonia, Brazil. *Papéis Avulsos de Zoologia (São Paulo),* 54(32), 457–472.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods,* 9(8), 772.

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics,* 12(7), 499–510.

de Boer LEM. 1974. Cytotaxonomy of the Platyrrhini (Primates). *Genen en Phaenen*, 17(1–2), 1−115.

de Queiroz K. 1998. The general lineage concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations. In: Howard DJ, Berlocher SH, (Eds). *Endless Forms: Species and Speciation.* Oxford University Press, Oxford, pp. 57–75.

de Queiroz K. 2007. Species concepts and species delimitation. *Systematic Biology*, 56(6), 879–86.

de Queiroz A, Gatesy J. 2007. The supermatrix approach to systematics. *Trends in Ecology & Evolution,* 22(1), 34–41.

Defler TR. 1994. *Callicebus torquatus* is not a white-sand specialist. *American Journal of Primatology*, 33, 149–154.

Defler TR. 2004. *Primates of Colombia*. Conservation International, Washington.

Defler TR. 2012. Studying primates in eastern Colombia: thirty five years of a primatological life. *Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales,* 36(140), 421–434.

Defler TR, Bueno ML, García J. 2010. *Callicebus caquetensis*: a new and critically endangered titi monkey from Southern Caquetá, Colombia. *Primate Conservation,* (25), 1–9.

Degnan JH, Rosenberg NA. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution,* 24(6), 332–340.

Delsuc F, Brinkmann H, Philippe H. 2005. Phylogenomics and the reconstruction of the tree of life. *Nature Reviews Genetics,* 6(5), 361–375.

Díaz-Arce N, Arrizabalaga H, Murua H, Irigoien X, Rodríguez-Ezpeleta N. 2016. RAD-seq derived genome-wide nuclear markers resolve the phylogeny of tunas. *Molecular Phylogenetics & Evolution*, 102, 202–207.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology & Evolution,* 29(8), 1969–1973.

Durand EY, Patterson N, Reich D, Slatkin M. 2011. Testing for ancient admixture between closely related populations. *Molecular Biology & Evolution*, 28, 2239–2252.

Earl DA, vonHoldt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2), 359–361.

Eaton DAR. 2014. PyRAD: assembly of *de novo* RADseq loci for phylogenetic analyses. *Bioinformatics,* 30(13), 1844–1849.

Eaton DAR, Hipp A, Gonzalez-Rodriguez A, Cavender-Bares J. 2015. Historical introgression among the American live oaks and the comparative nature of tests for introgression. *Evolution,* 69(10), 2587–2601.

Eaton DAR, Ree RH. 2013. Inferring phylogeny and introgression using RADseq data: an example from flowering plants (*Pedicularis*: Orobanchaceae). *Systematic Biology,* 62(5), 689–706.

Eaton DAR, Spriggs EL, Park B, Donoghue MJ. 2017. Misconceptions on missing data in RAD-seq phylogenetics with a deep-scale example from flowering plants. *Systematic Biology*, 66(3), 399–412.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), 1792–1797.

Edwards SV. 2009. Is a new and general theory of molecular systematics emerging? *Evolution,* 63(1), 1–19.

Edwards SV, Xi Z, Janke A, Faircloth BC, McCormack JE, Glenn TC, Zhong B, Wu S, Lemmon EM, Lemmon AR, Leaché AD, Liu L, Davis CC. 2016. Implementing and testing the multispecies coalescent model: a valuable paradigm for phylogenomics. *Molecular Phylogenetics & Evolution*, 94, 447–462.

Elliot DG. 1913. *A review of primates*. Monograph Series. American Museum of Natural History, New York.

Escudero M, Eaton DA, Hahn M, Hipp AL. 2014. Genotyping-by-sequencing as a tool to infer phylogeny and ancestral hybridization: A case study in *Carex* (Cyperaceae). *Molecular Phylogenetics & Evolution,* 79, 359–367.

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

Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, 164(4), 1567–1587.

Felton A, Felton AM, Wallace RB, Gómez H. 2006. Identification, behavioral observations, and notes on the distribution of the titi monkeys *Callicebus modestus* Lönnberg, 1939 and *Callicebus olallae*, Lönnberg 1939. *Primate Conservation*, (20), 40−46.

Fernandes AM. 2013. Fine-scale endemism of Amazonian birds in a threatened landscape. *Biodiversity & Conservation*, 22, 2683–2694.

Fernandes AM, Wink M, Aleixo A. 2012. Phylogeography of the chestnut-tailed antbird (*Myrmeciza hemimelaena*) clarifies the role of rivers in Amazonian biogeography. *Journal of Biogeography*, 39(8), 1524–1535.

Ferrari SF, Iwanaga S, Messias MR, Ramos EM, Ramos PCS, da Cruz Neto EH, Coutinho PEG. 2000. Titi monkeys (*Callicebus* spp., Atelidae: Platyrrhini) in the Brazilian state of Rondônia. *Primates*, 41(2), 229–234.

Ferrari SF, Iwanaga S, Silva JL. 1996. Platyrrhines in Pimenta Bueno, Rondônia, Brazil. *Neotropical Primates,* 4, 151–153.

Ferrari SF, Veiga LM, Pinto LP, Marsh LK, Mittermeier RA, Rylands AB. 2013. Family Pitheciidae (titis, sakis and uacaris). In: Mittermeier RA, Rylands AB, Wilson DE, (Eds). *Handbook of the Mammals of the World, Vol 3, Primates*. Lynx Edicions, Barcelona, pp. 432−483.

Fragaszy DM, Schwarz S, Shimosaka D. 1982. Longitudinal observations of care and development of infant titi monkeys (*Callicebus moloch*). *American Journal of Primatology,* 2(2), 191−200.

Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C., 2012. Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution*, 27(9), 480–488.

Garzione CN, Hoke GD, Libarkin JC, Withers S, MacFadden B, Eiler J, Ghosh P,

Mulch A. 2008. Rise of the Andes. *Science*, 320(5881), 1304–1307.

Garzione CN, Molnar P, Libarkin JC, MacFadden BJ. 2006. Rapid late Miocene rise of the Bolivian Altiplano: evidence for removal of mantle lithosphere. *Earth & Planetary Science Letters*, 241(3–4), 543–556.

Gatesy J, Springer MS. 2014. Phylogenetic analysis at deep timescales: unreliable gene trees, bypassed hidden support, and the coalescence/concatalescence conundrum. *Molecular Phylogenetics & Evolution,* 80, 231–266.

Ghosh P, Garzione CN, Eiler JM. 2006. Rapid uplift of the Altiplano revealed through 13C–18O bonds in paleosol carbonates. *Science*, 311(5760), 511–515.

Goodman M, Porter CA, Czelusniak J, Page SL, Schneider H, Shoshani J, Gunnell G, Groves CP. 1998. Toward a phylogenetic classification of Primates based on DNA evidence complemented by fossil evidence. *Molecular Phylogenetics & Evolution,* 9(3), 585–598.

Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MHY, Hansen NF, Durand EY, Malaspinas AS, Jensen JD, Marques-Bonet T, Alkan C, Prüfer K, Meyer M, Burbano HA, Good JM, Schultz R, Aximu-Petri A, Butthof A, Höber B, Höffner B, Siegemund M, Weihmann A, Nusbaum C, Lander ES, Russ C, Novod N, Affourtit J, Egholm M, Verna C, Rudan P, Brajkovic D, Kucan V, Gu I, Doronichev VB, Golovanova LV, Lalueza-Fox C, Rasilla M, Fortea J, Rosas A, Schmitz RW, Johnson PLF, Eichler EE, Falush D, Birney E, Mullikin JC, Slatkin M, Nielsen R, Kelso J, Lachmann M, Reich D, Pääbo S. 2010. A draft sequence of the Neandertal genome. *Science*, 328, 710–722.

Gregory TR. 2008. Understanding evolutionary trees. *Evolution: Education and Outreach,* 1(2), 121–137.

Gregory-Wodzicki KM. 2000. Uplift history of the central and northern Andes: a review. *Geological Society of America Bulletin*, 112(7), 1091–1105.

Groves CP. 2001. *Primate taxonomy*. Smithsonian Institution Press, Washington, DC.

Groves CP. 2005. Order Primates. In: Wilson DE, Reeder DM, (Eds). *Mammal species of the world: a taxonomic and geographic reference, Volume 1.* Johns Hopkins University Press, Baltimore, MD, pp. 111–184.

Gualda-Barros J, Nascimento FO, Amaral MK. 2012. A new species of *Callicebus* Thomas, 1903 (Primates, Pitheciidae) from the states of Mato Grosso and Pará, Brazil. *Papéis Avulsos de Zoologia (São Paulo),* 52(23), 261−279.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, 52(5), 696−704.

Haile-Selassie Y. 2001. Late Miocene hominids from the middle Awash, Ethiopia. *Nature*, 412, 178–181.

Hartwig WC, Meldrum DJ. 2002. Miocene platyrrhines of the northern Neotropics. In: Hartwig WC, (Eds). *The primate fossil record.* Cambridge University Press, Cambridge, UK, pp. 175–188.

Haugaasenn T, Peres CA. 2005. Primate assemblage structure in Amazonian flooded and unflooded forests. *American Journal of Primatology*, 67, 243–258.

Heled J, Drummond AJ. 2010. Bayesian inference of species trees from multilocus data. *Molecular Biology & Evolution*, 27(3), 570–580.

Hennig W. 1966. *Phylogenetic Systematics*. University of Illinois Press, Urbana.

Herrara S, Shank TM. 2016. RAD sequencing enables unprecedented phylogenetic resolution and objective species delimitation in recalcitrant divergent taxa. *Molecular Phylogenetics & Evolution*, 100, 70–79.

Hershkovitz P. 1963. A systematic and zoogeographic account of the monkeys of the genus *Callicebus* (Cebidae) of the Amazonas and Orinoco River basins. *Mammalia*, 27(1), 1–80.

Hershkovitz P. 1988. Origin, speciation, and distribution of South American titi monkeys, genus *Callicebus* (Family Cebidae, Platyrrhini). *Proceedings of the Academy of Natural Sciences of Philadelphia*, 140(1), 240–272.

Hershkovitz P. 1990. Titis, New World monkeys of the genus *Callicebus* (Cebidae, Platyrrhini): a preliminary taxonomic review. *Fieldiana Zoology, New Series* (55), 109pp.

Hey J. 2006. On the failure of modern species concepts. *Trends in Ecology and Evolution*, 21, 447–450.

Heymann EW, Encarnación CF, Soini P. 2002. On the diagnostic characters and geographic distribution of the "yellow-handed" titi monkey, *Callicebus lucifer*, Peru. *Neotropical Primates,* 10(3), 124–126.

Heymann EW, Nadjafzadeh MN. 2013. Insectivory and prey foraging techniques in *Callicebus*—a case study of *Callicebus cupreus* and a comparison to other pitheciids. In: Veiga, LM, Barnett, AA, Ferrari, SF, Norconk, MA, (Eds). *Evolutionary biology and conservation of titis, sakis and uacaris*. Cambridge University Press, Cambridge, UK, pp. 215–224.

Hill WCO. 1960. *Primates, comparative anatomy and taxonomy IV. Cebidae Part A*. Edinburgh University Press, Edinburgh, UK.

Hillis DM, Bull JJ. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology*, 42(2), 182−192.

Hipp A, Eaton DAR, Cavender-Bares J, Fitzek E, Nipper R, Manos PS. 2014. A framework phylogeny of the American oak clade based on sequenced RAD data. *PLOS ONE*, 9(4), e93975.

Ho SYW, Phillips MJ. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology,* 58(3), 367–380.

Hodgson JA, Sterner KN, Matthews LJ, Burrell AS, Jani RA, Raaum RL, Stewart CB, Disotell TR. 2009. Successive radiations, not stasis, in the South American primate fauna. *Proceedings of the National Academy of Sciences,* 106(14), 5534–5539.

Hoorn C, Wesselingh FP, ter-Steege H, Bermudez MA, Mora A, Sevink J, Sanmartin I, Sanchez-Meseguer A, Anderson CL, Figueiredo JP, Jaramillo C, Riff D, Negri FR, Hooghiemstra H, Lundberg J, Stadler T, Särkinen T, Antonelli A. 2010. Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science*, 330(6006), 927–931.

d'Horta FM, Cuervo AM, Ribas CC, Brumfield RT, Miyaki CY. 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(1), 37–49.

Horvath JE, Weisrock DW, Embry SL, Fiorentino I, Balhoff JP, Kappeler P, Wray GA, Willard HF, Yoder AD. 2008. Development and application of a phylogenomic toolkit: resolving the evolutionary history of Madagascar's lemurs. *Genome Research,* 18(3), 489–499.

Hou Y, Nowak MD, Mirré V, Bjorå CS, Brochmann C, Popp M. 2015. Thousands of RAD-seq loci fully resolve the phylogeny of the highly disjunct arctic-alpine genus *Diapensia* (Diapensiaceae). *PLOS ONE*, 10(10), e0140175.

Hoyos M, Bloor P, Defler T, Vermeer J, Röhe F, Farias, I., 2016. Phylogenetic relationships within the *Callicebus cupreus* species group (Pitheciidae: Primates): Biogeographic and taxonomic implications. *Molecular Phylogenetics & Evolution*, 102, 208–219.

Huang H, Knowles LL. 2016. Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. *Systematic Biology*, 65(3), 357–365.

Humboldt A von. 1811. Sur les singes qui habitent les Rives de l'Orénoque du Cassiquiare et du Rio Negro. La Viudita. In: Humboldt A von, Bonpland AJ, (Eds). *Recueil d'observations de zoologie et d'anatomie comparée faites dans l'Océan Atlantique, dans l'interieur du nouveau continent et dans la Mer du Sud pendant les année 1799, 1800, 1801, 1802 et 1803.* Premier Volume, F. Schoell, Paris, France, pp. 319−321.

Humboldt A von. 1852. *Personal Narrative of Travels to the Equinoctial Regions of America during the years 1799–1804.* Henry Bohn, London, UK.

Jacobs SC, Larson A, Cheverud JM. 1995. Phylogenetic relationships and orthogenetic evolution of coat color among tamarins (genus *Saguinus*). *Systematic Biology*, 44, 515–532.

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, 23, 1801–1806.

Jameson Kiesling NM, Yi SV, Xu K, Sperone FG, Wildman DE. 2015. The tempo and mode of New World monkey evolution and biogeography in the context of phylogenomic analysis. *Molecular Phylogenetics & Evolution*, 82B, 386−399.

Johns AD. 1991. Forest disturbance and Amazonian primates. In: Box HO, (Eds). *Primate responses to environmental change*. Chapman & Hall, London, pp. 115−135.

Jombart T. 2008. Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403–1405.

Jombart T, Devillard S, Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, 11, 94.

Jones C, Anderson S. 1978. *Callicebus moloch*. *Mammalian Species,* (112), 1−5.

Kay RF, Fleagle JG. 2010. Stem taxa, homoplasy, long lineages, and the phylogenetic position of *Dolichocebus*. *Journal of Human Evolution,* 59(2), 218–222.

Kay RF, Fleagle JG, Mitchell TRT, Colbert M, Bown T, Powers DW. 2008. The anatomy of *Dolichocebus gaimanensis*, a stem platyrrhine monkey from Argentina. *Journal of Human Evolution,* 54(3), 323–382.

Kay RF, Johnson DJ, Meldrum DJ. 1998. A new pitheciine primate from the middle Miocene of Argentina. *American Journal of Primatology,* 45(4), 317–336.

Kay RF, Johnson DJ, Meldrum DJ. 1999. *Proteropithecia*, new name for *Propithecia* Kay, Johnson and Meldrum, 1998 non Vojnits 1985. *American Journal of Primatology*, 47, 347.

Kay RF, Meldrum DJ, Takai M. 2013. Pitheciidae and other platyrrhine seed predators. In: Veiga LM, Barnett AA, Ferrari SF, Norconk MA, (Eds). *Evolutionary biology and conservation of titis, sakis and uacaris.* Cambridge University Press, Cambridge, pp. 3–12.

Kelley J. 2002. The hominoid radiation in Asia. In: Hartwig WC, (Eds.). *The primate fossil record.* Cambridge University Press, Cambridge, pp. 369–384.

Kinzey WG. 1981. The titi monkeys, genus *Callicebus*. In: Coimbra-Filho AF, Mittermeier RA, (Eds). *Ecology and Behavior of Neotropical Primates. Volume 1*. Academia Brasileira de Ciências, Rio de Janeiro, Brazil, pp. 241−276.

Kinzey WG. 1982. Distribution of primates and forest refuges. In: Prance GT, (Eds). *Biological diversification in the tropics*. Columbia University Press, New York, pp. 445–482.

Kinzey WG, Gentry GH. 1979. Habitat utilization in two species of *Callicebus*. In: Sussman RW, (Eds). *Primate Ecology: Problem Oriented Field Studies*. John Wiley & Sons, New York, pp. 89–100.

Kluge AG. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Systematic Zoology,* 38(1), 7–25.

Kobayashi S. 1990. A morphological study of upper first and second molars in the genus *Callicebus. Journal of the Anthropological Society of Nippon,* 98(2), 121–135.

Kobayashi S. 1995. A phylogenetic study of titi monkeys, genus *Callicebus*, based on cranial measurements: I. Phyletic groups of *Callicebus*. *Primates*, 36(1), 101−120.

Kobayashi S, Langguth AL. 1999. A new species of titi monkey, *Callicebus* Thomas, from north-eastern Brazil (Primates, Cebidae). *Revista Brasileira de Zoologia,* 16(2), 531–551.

Kubatko LS, Degnan JH. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology*, 56(1), 17–24.

Laehnemann D, Borkhardt A, McHardy AC. Denoising DNA deep sequencing data: high-throughput sequencing errors and their correction. *Briefings in Bioinformatics*, 17(1), 154–179.

Landis MJ. 2017. Biogeographic dating of speciation times using paleogeographically informed processes. *Systematic Biology*, 66(2), 128–144.

Lanfear R, Calcott B, Ho SYW, Guindon S. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology & Evolution*, 29(6), 1695–1701.

Larget BR, Kotha SK, Dewey CN, Ané C. 2010. BUCKy: gene tree/species tree reconciliation with Bayesian concordance analysis. *Bioinformatics,* 26(22), 2910–2911.

Latrubesse EM. 2002. Evidence of quaternary palaeohydrological changes in middle Amazonia: the Aripuanã-Roosevelt and Jiparaná fans. *Zeitschrift für Geomorphologie*, 129, 61–72.

Latrubesse EM, Cozzuol M, da Silva-Caminha SAF, Rigsby CA, Absy ML, Jaramillo C. 2010. The late Miocene paleogeography of the Amazon Basin and the evolution of the Amazon River system. *Earth-Science Reviews*, 99, 99–124.

Leaché AD. 2010. Species trees for spiny lizards (Genus *Sceloporus*): identifying points of concordance and conflict between nuclear and mitochondrial data. *Molecular Phylogenetics & Evolution*, 54, 162–171.

Leaché AD, Fujita MK, Minin VN, Bouckaert RR. 2014a. Species delimitation using genome-wide SNP Data. *Systematic Biology*, 63, 534–542.

Leaché AD, Harris RB, Rannala B, Yang Z. 2014b. The influence of gene flow on species tree estimation: a simulation study. *Systematic Biology,* 63, 17–30.

Leaché AD, Rannala B. 2011. The accuracy of species tree estimation under simulation: a comparison of methods. *Systematic Biology,* 60(2), 126–137.

Leakey M. 1993. Evolution of *Theropithecus* in the Turkana Basin. In: Jablonski NG, (Eds.). *Theropithecus: the rise and fall of a primate genus.* Cambridge University Press, Cambridge, pp. 85–123.

Lima MGM, Buckner JC, Silva-Junior JS, Aleixo A, Martins A, Boubli JP, Link A, Farias IP, da Silva MN, Rohe F, Queiroz H, Chiou KL, Di Fiore A, Alfaro ME, Lynch Alfaro JW. 2017. Capuchin monkey biogeography: understanding *Sapajus* Pleistocene range expansion and the current sympatry between *Cebus* and *Sapajus*. *Journal of Biogeography*, 44(4), 810–820.

Linares OJ. 1998. *Mamíferos de Venezuela*. Sociedad Conservacionista Audubon de Venezuela, Caracas, Venezuela.

Linkem CW, Minin VN, Leaché AD. 2016. Detecting the anomaly zone in species trees and evidence for a misleading signal in higher-level skink phylogeny (Squamata: Scincidae). *Systematic Biology*, 65, 465–477.

Liu L. 2008. BEST: Bayesian estimation of species trees under the coalescent model. *Bioinformatics,* 24(21), 2542–2543.

Liu L, Pearl DK. 2007. Species trees from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. *Systematic Biology,* 56(3), 504–514.

Liu L, Xi Z, Wu S, Davis CC, Edwards SV. 2015. Estimating phylogenetic trees from genome-scale data. *Annals of the New York Academy of Sciences,* 1360(1), 36–53.

Liu L, Yu L, Kubatko L, Pearl DK, Edwards SV. 2009. Coalescent methods for estimating phylogenetic trees. *Molecular Phylogenetics & Evolution,* 53(1), 320–328.

Lönnberg E. 1939. Notes on some members of the genus *Callicebus*. *Arkiv för Zoologi,* 31(13), 1–18.

Lynch Alfaro JW, Boubli JP, Paim FP, Ribas CC, da Silva MNF, Messias M, Röhe F, Mercês MP, Silva Júnior JS, Silva CR, Pinho GM, Koshkarian G, Nguyen MTT, Harada ML, Rabelo RM, Queiroz HL, Alfaro ME, Farias IP. 2015a. Biogeography of squirrel monkeys (genus *Saimiri*): South-central Amazon origin and rapid pan-Amazonian diversification of a lowland primate. *Molecular Phylogenetics & Evolution*, 82B, 436–454.

Lynch Alfaro JW, Cortés-Ortiz L, Di Fiore A, Boubli JP. 2015b. Special issue: Comparative biogeography of Neotropical primates. *Molecular Phylogenetics & Evolution*, 82B, 518–529.

Lynch Alfaro JW, Silva-Júnior JS, Rylands AB. 2012. How different are robust and

gracile capuchin monkeys? An argument for the use of *Sapajus* and *Cebus*. *American Journal of Primatology,* 74, 273–286.

MacFadden BJ. 1990. Chronology of Cenozoic primate localities in South America. *Journal of Human Evolution,* 19(1), 7–21.

Maddison WP. 1997. Gene trees in species trees. *Systematic Biology*, 46(3), 523–536.

Maddison WP, Knowles LL. 2006. Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology,* 55(1), 21–30.

Mallet J. 2008. Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Philosophical Transactions of the Royal Society B*, 363, 2971–2986.

Manthey JD, Campillo LC, Burns KJ, Moyle RG. 2016. Comparison of target-capture and restriction-site associated DNA sequencing for phylogenomics: a test in cardinalid tanagers (Aves, Genus: *Piranga*). *Systematic Biology*, 65(4), 640–650.

Marques-Bonet T, Ryder OA, Eichler EE. 2009. Sequencing primate genomes: what have we learned? *Annual Review of Genomics & Human Genetics*, 10, 355−386.

Martínez J, Wallace RB. 2007. Further notes on the distribution of the Bolivian endemic titi monkeys, *Callicebus modestus* and *Callicebus olallae*. *Neotropical Primates*, 14(2), 47–54.

Martínez J, Wallace RB. 2010. Pitheciidae. In: Wallace RB, Gómez H, Porcel ZR, Rumiz DI, (Eds). *Distribución, ecología y conservación de los mamíferos medianos y grandes de Bolivia.* Centro de Ecología Difusión, Fundación Simón I. Patiño, Santa Cruz de la Sierra, pp. 305−330.

Martínez J, Wallace RB. 2013. New information about the distribution of *Callicebus* (Pitheciidae, Primates) in northern Beni Department, Bolivia. *Ecología en Bolivia*, 48(1), 57–62.

Martínez J, Wallace, RB. 2016. Ecological and behavioural factors influencing territorial call rates for the Bolivian titi monkeys, *Plecturocebus modestus* and *Plecturocebus olallae*. *Folia Primatologica*, 87(5), 279–290.

Matzke NJ. 2013a. BioGeoBEARS: BioGeography with Bayesian (and Likelihood) Evolutionary Analysis in R Scripts. R package, version 0.2.1, published July 27, 2013. <http://CRAN.R-project.org/package=BioGeoBEARS>

Matzke NJ. 2013b. Probabilistic historical biogeography: New models for founder-

event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography*, 5(4), 242–248.

Matzke NJ. 2014. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Systematic Biology*, 63(6), 951–970.

Matzke NJ. 2016. Stochastic mapping under biogeographical models. PhyloWiki BioGeoBEARS website (27/11/2016):

http://phylo.wikidot.com/biogeobears#stochastic_mapping

Mayr E. 1954. Change of genetic environment and evolution. In: Huxley J, Hardy AC, Ford EB, (Eds). *Evolution as a process*. Allen & Unwin, London, pp. 157–180.

Mayr E, Bock WJ. 2002. Classifications and other ordering systems. *Journal of Zoological Systematics & Evolutionary Research,* 40(4), 169–194.

Mayr E, Provine WB. 1998. *The evolutionary synthesis: perspectives on the unification of biology.* Harvard University Press, Cambridge, UK.

McCormack JE, Heled J, Delaney KS, Peterson AT, Knowles LL. 2011. Calibrating divergence times on species trees versus gene trees: implications for speciation history of *Aphelocoma* jays. *Evolution,* 65, 184–202.

McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT. 2012. Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics & Evolution*, 66(2), 526–538.

Meyer BS, Matschiner M, Salzburger W. 2016. Disentangling incomplete lineage sorting and introgression to refine species-tree estimates for Lake Tanganyika cichlid fishes. *Systematic Biology*, syw069.

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: *Proceedings of the Gateway Computing Environments Workshop (GCE),* New Orleans, LA.

Mirarab S, Bayzid MS, Boussau B, Warnow T. 2014. Statistical binning enables an accurate coalescent-based estimation of the avian tree. *Science,* 346(6215), 1250463.

Monção GR, Selhorst V, Soares-Filho JAR. 2008. Expansão da distribuição geográfica de *Callicebus bernhardi* a oeste do rio Ji-Paraná, Estado de Rondônia, Brasil. *Neotropical Primates*, 15, 67–68.

Moynihan M. 1966. Communication in the titi monkey, *Callicebus*. *Journal of Zoology,* 150(1), 77−127.

Muniz FL, Campos Z, Farias IP, Hrbek T. 2017. What drives diversification along the Amazon–Paraguay basin transition? A biogeographic model for semi-aquatic and aquatic taxa. *Journal of Biogeography*, in review.

Murphy WJ, Eizirik E, Johnson WE, Zhang YP, Ryder OA, O'Brien SJ. 2001. Molecular phylogenetics and the origins of placental mammals. *Nature,* 409(6820), 614–618.

Nei M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nei M, Kumar S. 2000. *Molecular Evolution and Phylogenetics.* Oxford University Press, Oxford.

Neigel JE, Avise JC. 1986. Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. In: Nevo E, Karlin S, (Eds). *Evolutionary Processes and Theory*. Academic Press, New York, pp. 515–534.

Nishihara H, Okada N, Hasegawa M. 2007. Rooting the eutherian tree: the power and pitfalls of phylogenomics. *Genome Biology,* 8(9), R199.

Norconk MA. 2011. Sakis, uakaris, and titi monkeys. In: Campbell CJ, Fuentes A, MacKinnon KC, Bearder SK, Stumpf RM, (Eds). *Primates in Perspective*. Oxford University Press, Oxford, pp. 122−139.

Noronha M de A, Spironello WR, Ferreira DC. 2007. New occurrence records and eastern extension to the range of *Callicebus cinerascens* (Primates, Pitheciidae). *Neotropical Primates,* 14(3), 137–139.

Nylander JA, Ronquist F, Huelsenbeck JP, Nieves-Aldrey J. 2004. Bayesian phylogenetic analysis of combined data. *Systematic Biology,* 53(1), 47–67.

Ogilvie HA, Bouckaert RR, Drummond AJ. 2016a. StarBEAST2 brings faster species tree inference and accurate estimates of substitution rates. *bioRxiv* preprint, DOI: 10.1101/070169.

Ogilvie HA, Heled J, Xie D, Drummond AJ. 2016b. Computational performance and statistical accuracy of *BEAST and comparisons with other methods. *Systematic Biology*, 65 (3), 381–396.

Oliveira DP de, Carvalho VT de, Hrbek T. 2016. Cryptic diversity in the lizard genus *Plica* (Squamata): Phylogenetic diversity and Amazonian biogeography.

Zoologica Scripta, 45, 630–641.

Padial JM, Miralles A, De la Riva I, Vences M. 2010. The integrative future of taxonomy. *Frontiers in Zoology*, 7(16), 1–14.

Paithankar KR, Prasad KS. 1991. Precipitation of DNA by polyethylene glycol and ethanol. *Nucleic Acids Research*, 19(6), 1346.

Palacios E, Rodríguez A. 2013. Seed eating by *Callicebus lugens* at Caparú Biological Station, on the lower Apaporis River, Colombian Amazonia. In: Veiga LM, Barnett AA, Ferrari SF, Norconk MA, (Eds). *Evolutionary biology and conservation of titis, sakis and uacaris*. Cambridge University Press, Cambridge, UK, pp. 225−231.

Palacios E, Rodríguez A, Defler TR. 1997. Diet of a group of *Callicebus torquatus lugens* (Humboldt, 1812) during the annual resource bottleneck in Amazonian Colombia. *International Journal of Primatology,* 18(4), 503–522.

Pamilo P, Nei M. 1988. Relationships between gene trees and species trees. *Molecular Biology & Evolution,* 5(5), 568–583.

Pante E, Abdelkrim J, Viricel A, Gey D, France SC, Boisselier MC, Samadi S. 2015. Use of RAD sequencing for delimiting species. *Heredity*, 114(5), 450–459.

Pecon-Slattery J. 2014. Recent advances in primate phylogenomics. *Annual Review of Animal Biosciences*, 2, 41–63.

Pelzeln A von. 1883. *Brasilische Säugethiere: Resultate von Johann Natterer's Reisen in den Jahren 1817 bis 1835. Verhandlungen 33.* Kaiserlich-Königlichen Zoologisch-botanischen Gesellschaft, Vienna, Austria.

Percequillo AR, Weksler M, Costa LP. 2011. A new genus and species of rodent from the Brazilian Atlantic Forest (Rodentia: Cricetidae: Sigmodontinae: Oryzomyini), with comments on oryzomyine biogeography. *Zoological Journal of the Linnean Society*, 161, 357–390.

Perelman P, Johnson WE, Roos C, Seuánez HN, Horvath JE, Moreira MA, Kessing B, Pontius J, Roelke M, Rumpler Y, Schneider MPC, Silva A, O'Brien SJ, Pecon-Slattery J. 2011. A molecular phylogeny of living primates. *PLOS Genetics*, 7(3), e1001342.

Peres CA. 1997. Primate community structure at twenty western Amazonian flooded and unflooded forests. *Journal of Tropical Ecology*, 13, 381–405.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012. Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLOS ONE*, 7(5), e37135.

Pilbeam DR, Walker AC. 1968. Fossil monkeys from the Miocene of Napak, Northeast Uganda. *Nature*, 220, 657–660.

Prado-Martinez J, Sudmant PH, Kidd JM, Li H, Kelley JL, Lorente-Galdos B, Veeramah KR, Woerner AE, O'Connor TD, Santpere G, Cagan A, Theunert C, Casals F, Laayouni H, Munch K, Hobolth A, Halager AE, Malig M, Hernandez-Rodriguez J, Hernando-Herraez I, Prüfer K, Pybus M, Johnstone L, Lachmann M, Alkan C, Twigg D, Petit N, Baker C, Hormozdiari F, Fernandez-Callejo M, Dabad M, Wilson ML, Stevison L, Camprubí C, Carvalho T, Ruiz-Herrera A, Vives L, Mele M, Abello T, Kondova I, Bontrop RE, Pusey A, Lankester F, Kiyang JA, Bergl RA, Lonsdorf E, Myers S, Ventura M, Gagneux P, Comas D, Siegismund H, Blanc J, Agueda-Calpena L, Gut M, Fulton L, Tishkoff SA, Mullikin JC, Wilson RK, Gut IG, Gonder MK, Ryder OA, Hahn BH, Navarro A, Akey JM, Bertranpetit J, Reich D, Mailund T, Schierup MH, Hvilsom C, Andrés AM, Wall JD, Bustamante CD, Hammer MF, Eichler EE, Marques-Bonet T. 2013. Great ape genetic diversity and population history. *Nature*, 499(7459), 471–5.

Printes RC, Jerusalinsky L, Sousa MC, Rodrigues LRR, Hirsch A. 2013. Zoogeography, genetic variation and conservation of the *Callicebus personatus* group. *Cambridge Studies in Biological & Evolutionary Anthropology*, 1(65), 43−54.

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

Pritchard JK, Wen W. 2004. *Documentation for STRUCTURE software: Version 2*. Department of Human Genetics, University of Chicago, Chicago. <http://pritch.bsd.uchicago.edu>

Pukk L, Ahmad F, Hasan S, Kisand V, Gross R, Vasemägi A. 2015 Less is more: extreme genome complexity reduction with ddRAD using Ion Torrent semiconductor technology. *Molecular Ecology Resources*, 15(5), 1145–1152.

Puritz JB, Matz MV, Toonen RJ, Weber JN, Bolnick DI, Bird CE. 2014. Demystifying the RAD fad. *Molecular Ecology*, 23(24), 5937–5942.

Quintino EP, Bicca-Marques JC. 2013. Occurrence of *Callicebus bernhardi* in Rolim de Moura, Rondônia, Brazil. *Neotropical Primates*, 20(1), 62–63.

Rannala B, Yang Z. 2008. Phylogenetic inference using whole genomes. *Annual Review of Genomics and Human Genetics,* 9, 217–231.

Razkin O, Sonet G, Breugelmans K, Madeira MJ, Gómez-Moliner BJ, Backeljau T. 2016. Species limits, interspecific hybridization and phylogeny in the cryptic land snail complex *Pyramidula*: the power of RADseq data. *Molecular Phylogenetics & Evolution*, 101, 267–278.

Rheindt FE, Fujita MK, Wilton PR, Edwards SV. 2014. Introgression and phenotypic assimilation in *Zimmerius* flycatchers (Tyrannidae): population genetic and phylogenetic inferences from genome-wide SNPs. *Systematic Biology*, 63(2), 134–152.

Ribas CC, Aleixo A, Nogueira ACR, Miyaki CY, Cracraft J. 2012. A palaeobiogeographic model for biotic diversification within Amazonia over the past three million years. *Proceedings of the Royal Society B: Biological Sciences*, 279(1729), 681–689.

Robinson JG. 1979. An analysis of the organization of vocal communication in the titi monkey *Callicebus moloch. Zeitschrift für Tierpsychologie,* 49(4), 381−405.

Roch S, Steel M. 2015. Likelihood-based tree reconstruction on a concatenation of aligned sequence data sets can be statistically inconsistent. *Theoretical Population Biology,* 100, 56−62.

Rodrigues LRR, Barros RMS, Pissinatti A, Pieczarka JC, Nagamachi CY. 2001. Cytogenetic study of *Callicebus hoffmannsi* (Cebidae, Primates) and comparison with *C. m. moloch. Cytobios*, 105, 137−145.

Rodrigues LRR, Barros RMS, Pissinatti A, Pieczarka JC, Nagamachi CY. 2004. A new karyotype of an endangered primate species (*Callicebus personatus*) from the Brazilian Atlantic forests. *Hereditas*,140(2), 87−91.

Rodríguez-Ezpeleta N, Brinkmann H, Roure B, Lartillot N, Lang BF, Philippe H. 2007. Detecting and overcoming systematic errors in genome-scale phylogenies. *Systematic Biology,* 56(3), 389–399.

Rogers J, Gibbs RA. 2014. Comparative primate genomics: emerging patterns of genome content and dynamics. *Nature Reviews Genetics*, 15(5), 347−359.

Rognes T, Flouri T, Nichols B, Quince C, Mahé F. 2016. VSEARCH: a versatile open source tool for metagenomics. *PeerJ*, 4, e2584.

Röhe, F, Silva-Júnior J de S. 2009. Confirmation of *Callicebus dubius* (Pitheciidae) distribution and evidence of invasion into the geographic range of *Callicebus stephennashi*. *Neotropical Primates,* 16(2), 71–73.

Rokas A, Williams BL, King N, Carroll SB. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature,* 425(6960), 798–804.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61(3), 539–542.

Roos C, Zinner D, Kubatko L, Schwarz C, Yang M, Meyer D, Nash S, Xing J, Batzer M, Brameier M, Leendertz F, Ziegler T, Perwitasari-Farajallah D, Nadler T, Walter L, Osterholz M. 2011. Nuclear versus mitochondrial DNA: evidence for hybridization in colobine monkeys. *BMC Evolutionary Biology,* 11, 77.

Rosenberg NA. 2004. DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes,* 4, 137–138.

Rosenberger AL, Hartwig WC, Takai M, Setoguchi T, Shigehara N. 1991. Dental variability in *Saimiri* and the taxonomic status of *Neosaimiri fieldsi*, an early squirrel monkey from La Venta, Colombia. *International Journal of Primatology,* 12(3), 291–301.

Rossetti DF, Toledo PM, Góes AM. 2005. New geological framework for Western Amazonia (Brazil) and implications for biogeography and evolution. *Quaternary Research*, 63, 78–89.

Ruane S, Bryson RW, Pyron RA, Burbrink FT. 2014. Coalescent species delimitation in milksnakes (genus *Lampropeltis*) and impacts on phylogenetic comparative analyses. *Systematic Biology*, 63(2) 231–250.

Rubin BER, Ree RH, Moreau CS. 2012. Inferring phylogenies from RAD sequence data. *PLOS ONE*, 7, e33394.

Rumiz DI. 2012. Distribution, habitat and status of the white-coated titi monkey (*Callicebus pallescens*) in the Chaco-Chiquitano forests of Santa Cruz, Bolivia. *Neotropical Primates*, 19(1), 8–15.

Rylands AB, Heymann EW, Lynch Alfaro JW, Buckner JC, Roos C, Matauschek C, Boubli JP, Sampaio R, Mittermeier RA. 2016. Taxonomic review of the New World tamarins (Callitrichidae, Primates). *Zoological Journal of the Linnean*

Society, 177(4), 1003–1028.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

San Mauro D, Agorreta A. 2010. Molecular systematics: a synthesis of the common methods and the state of knowledge. *Cellular & Molecular Biology Letters,* 15(2), 311–341.

Santana SE, Lynch Alfaro JW, Alfaro ME. 2012. Adaptive evolution of facial colour patterns in Neotropical primates. *Proceedings of the Royal Society B: Biological Sciences*, 279(1736), 2204–2211.

Schlegel H. 1876. *Muséum d'Histoire Naturelle des Pays-Bas. Revue méthodique et critique des collections déposées dans cet établissement. Tome 7. Monographie 40: Simiae.* E. J. Brill, Leiden, Germany.

Schrago CG, Mello B, Soares AER. 2013. Combining fossil and molecular data to date the diversification of New World Primates. *Journal of Evolutionary Biology,* 26(11), 2438–2446.

Schrago CG, Menezes AN, Furtado C, Bonvicino CR, Seuánez HN. 2014. Multispecies coalescent analysis of the early diversification of Neotropical primates: Phylogenetic inference under strong gene trees/species tree conflict. *Genome Biology & Evolution,* 6(11), 3105–3114.

Senut B, Pickford M, Gommery D, Mein P, Cheboi K, Coppens Y. 2001. First hominid from the Miocene (Lukeino Formation, Kenya). *Comptes Rendus de l'Académie des Sciences, Series IIA - Earth & Planetary Sciences*, 332, 137–144.

Serrano-Villavicencio JE, Vendramel RL, Garbino GST. 2017. Species, subspecies, or color morphs? Reconsidering the taxonomy of *Callicebus* Thomas, 1903 in the Purus–Madeira interfluvium. *Primates*, 58(1), 159–167.

Shendure J, Ji H. 2008. Next-generation DNA sequencing. *Nature Biotechnology*, 26(10), 1135–1145.

Silva JMC, Novaes FC, Oren DC. 2002. Differentiation of *Xiphocolaptes* (Dendrocolaptidae) across the river Xingu, Brazilian Amazonia: recognition of a new phylogenetic species and biogeographical implications. *Bulletin of the British Ornithologists' Club,* 122, 185–194.

Silva-Júnior JS, Figueiredo-Ready WMB, Ferrari SF. 2013. Taxonomy and geographic distribution of the Pitheciidae. In: Veiga LM, Barnett AA, Ferrari

SF, Norconk MA, (Eds). *Evolutionary biology and conservation of titis, sakis and uacaris*. Cambridge University Press, Cambridge, UK, pp. 31−42.

Simmons MP, Gatesy J. 2015. Coalescence vs. concatenation: Sophisticated analyses vs. first principles applied to rooting the angiosperms. *Molecular Phylogenetics & Evolution,* 91, 98–122.

Simon MF, Grether R, Queiroz LP, Skema C, Pennington RT, Hughes CE. 2009. Recent assembly of the Cerrado, a Neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *Proceedings of the National Academy of Sciences*, 106, 20359–20364.

Spix JB von. 1823. *Simiarum et vespertiliarum brasiliensis species novae.* F. S. Hübschmann, Munich, Germany.

Springer MS, Gatesy J. 2016. The gene tree delusion. *Molecular Phylogenetics & Evolution*, 94, 1–33.

Springer MS, Meredith RW, Gatesy J, Emerling C, Park J, Rabosky DL, Stadler T, Steiner C, Ryder O, Janečka JE, Fisher C, Murphy WJ. 2012. Macroevolutionary dynamics and historical biogeography of primate diversification inferred from a species supermatrix. *PLOS ONE*, 7(11), e49521.

Stallings JR. 1985. Distribution and status of primates in Paraguay. *Primate Conservation*, (6), 51−58.

Stallings JR, West L, Hahn W, Gamarra, I. 1989. Primates and their relation to habitat in Paraguayan Chaco. In: Redford KH, Esienberg JF, (Eds). *Advances in Neotropical mammalogy*. The Sandhill Crane Press Inc., Gainesville, FL, pp. 425−442.

Stamatakis A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22, 2688–2690.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), 1312−1313.

Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML Web servers. *Systematic Biology,* 57(5), 758–771.

Stange M, Sánchez-Villagra MR, Salzburger W, Matschiner M. 2017. Bayesian divergence-time estimation with genome-wide SNP data of sea catfishes

(Ariidae) supports Miocene closure of the Panamanian Isthmus. bioRxiv preprint, DOI: 10.1101/102129.

Stanyon R, Bonvicino CR, Svartman M, Seuánez HN. 2003. Chromosome painting in *Callicebus lugens*, the species with the lowest diploid number (2n = 16) known in primates. *Chromosoma,* 112(4), 201–206.

Suh A, Smeds L, Ellegren H. 2015. The dynamics of incomplete lineage sorting across the ancient adaptive radiation of neoavian birds. *PLOS Biology*, 13(8), p.e1002224.

Swofford DL. 2002. *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Sinauer, Sunderland, MA.

Takahata N. 1989. Gene genealogy in three related populations: consistency probability between gene and population trees. *Genetics*, 122(4), 957–966.

Takai M. 1994. New specimens of *Neosaimiri fieldsi* from La Venta, Colombia: a middle Miocene ancestor of the living squirrel monkeys. *Journal of Human Evolution,* 27(4), 329–360.

Takai M, Anaya F, Suzuki H, Shigehara N, Setoguchi T. 2001. A new platyrrhine from the Middle Miocene of La Venta, Colombia, and the phyletic position of Callicebinae. *Anthropological Science*, 109(4), 289–307.

Tate GHH. 1939. The mammals of the Guiana region. *Bulletin of the American Museum of Natural History,* 76, 151–229.

Tateno Y, Nei M, Tajima F. 1982. Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution,* 18(6), 387–404.

Teeling EC, Scally M, Kao DJ, Romagnoli ML, Springer MS, Stanhope MJ. 2000. Molecular evidence regarding the origin of echolocation and flight in bats. *Nature,* 403(6766), 188–192.

Thom G, Aleixo A. 2015. Cryptic speciation in the white-shouldered antshrike (*Thamnophilus aethiops*, Aves – Thamnophilidae): The tale of a transcontinental radiation across rivers in lowland Amazonia and the northeastern Atlantic Forest. *Molecular Phylogenetics & Evolution*, 82, 95–110.

Thomas O. 1903. Notes on South-American monkeys, bats, carnivores, and rodents, with descriptions of new species. *Annals and Magazine of Natural History 7th series,* (12), 455–464.

Thomas O. 1908. Four new Amazonian monkeys. *Annals and Magazine of Natural History 8th series,* (2), 88–91.

Tirira D. 2007. *Guía de Campo de los Mamíferos del Ecuador. Publicación Especial sobre los Mamíferos del Ecuador, Vol. 6*. Ediciones Murciélago Blanco, Quito, Ecuador.

Tonini J, Moore A, Stern D, Shcheglovitova M, Ortí G. 2015. Concatenation and species tree methods exhibit statistically indistinguishable accuracy under a range of simulated conditions. *PLoS Currents Tree of Life,*
DOI: 10.1371/currents.tol.34260cc27551a527b124ec5f6334b6be.

Townsend TM, Mulcahy DG, Noonan BP, Sites JW. 2011. Phylogeny of iguanian lizards inferred from 29 nuclear loci, and a comparison of concatenated and species-tree approaches for an ancient, rapid radiation. *Molecular Phylogenetics & Evolution,* 61, 363–380.

Van Dam MH, Matzke NJ. 2016. Evaluating the influence of connectivity and distance on biogeographical patterns in the south-western deserts of North America. *Journal of Biogeography*, 43(8), 1514–1532.

Van Roosmalen MGM, Van Roosmalen T, Mittermeier RA. 2002. A taxonomic review of the titi monkeys, genus *Callicebus* Thomas, 1903, with the description of two new species, *Callicebus bernhardi* and *Callicebus stephennashi*, from Brazilian Amazonia. *Neotropical Primates,* 10(suppl.), 1–52.

Veiga LM, Barnett AA, Ferrari SF, Norconk MA. 2013. *Evolutionary biology and conservation of titis, sakis and uacaris.* Cambridge University Press, Cambridge.

Venta PJ, Brouillette JA, Yuzbasiyan-Gurkan V, Brewer GJ. 1996. Gene-specific universal mammalian sequence-tagged sites: application to the canine genome. *Biochemical Genetics*, 34(7-8), 321−341.

Vermeer J, Tello-Alvarado JC. The distribution and taxonomy of titi monkeys (*Callicebus*) in central and southern Peru, with the description of a new species. *Primate Conservation*, 29, 9–29.

Vignaud P, Duringer P, Mackaye HT, Likius A, Blondel C, Boisserie JR, De Bonis L, Eisenmann V, Etienne ME, Geraads D, Guy F, Lehmann T, Lihoreau F, Lopez-Martinez N, Mourer-Chauviré C, Otero O, Rage JC, Schuster M, Viriot L, Zazzo A, Brunet M. 2002. Geology and palaeontology of the Upper

Miocene Toros-Menalla hominid locality, Chad. *Nature*, 418, 152–155.

Wagner CE, Keller I, Wittwer S, Selz OM, Mwaiko S, Greuter L, Sivasundar A, Seehausen O. 2013. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology*, 22(3), 787–798.

Wallace AR. 1852. On the monkeys of the Amazon. *Proceedings of the Zoological Society of London*, 20, 107–110.

Wallace RB, Gómez H, Felton A, Felton AM. 2006. On a new species of titi monkey, genus *Callicebus* Thomas (Primates, Pitheciidae), from western Bolivia with preliminary notes on distribution and abundance. *Primate Conservation,* (20), 29–39.

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN. 2005. DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B,* 360(1462), 1847−1857.

Wesselingh FP, Salo JA. 2006. Miocene perspective on the evolution of the Amazonian biota. *Scripta Geologica*, 133, 439–458.

Wiley EO. 1978. The evolutionary species concept reconsidered. *Systematic Zoology*, 27, 17–26.

Wiley EO, Lieberman BS. 2011. *Phylogenetics: theory and practice of phylogenetic systematics*. Wiley-Blackwell, New York.

Willis SC. 2017. One species or four? Yes!... and, no. Or, arbitrary assignment of lineages to species obscures the diversification processes of Neotropical fishes. *PLOS ONE*, 12(2), e0172349.

Wood B, Collard M. 1999. The human genus. *Science*, 284(5411), 65−71.

Wu CI. 1991. Inferences of species phylogeny in relation to segregation of ancient polymorphisms. *Genetics,* 127(2), 429–435.

Yi Z, Strüder-Kypke M, Hu X, Lin X, Song W. 2014. Sampling strategies for improving tree accuracy and phylogenetic analyses: A case study in ciliate protists, with notes on the genus *Paramecium. Molecular Phylogenetics & Evolution,* 71, 142–148.

Young NM, MacLatchy L. 2004. The phylogenetic position of *Morotopithecus*. *Journal of Human Evolution*, 46, 163–184.

Yu L, Peng D, Liu J, Luan P, Liang L, Lee H, Lee M, Ryder O, Zhang Y. 2011. On the phylogeny of Mustelidae subfamilies: analysis of seventeen nuclear non-

coding loci and mitochondrial complete genomes. *BMC Evolutionary Biology*, 11, 92.

Yu Y, Harris AJ, Blair C, He XJ. 2015. RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. *Molecular Phylogenetics & Evolution*, 87, 46–49.

Zapata F, Jiménez I. 2012. Species delimitation: Inferring gaps in morphology across geography. *Systematic Biology*, 61 (2), 179–194.