Capuchin monkey biogeography: understanding Sapajus Pleistocene range expansion and the current sympatry between Cebus and Sapajus

Marcela G. M. Lima1,2,†, Janet C. Buckner3, José de Sousa e Silva-Júnior2, Alexandre Aleixo7, Amely B. Martins4,5, Jean P. Boublí6, Andréis Link7, Izeni P. Farias9, Maria Nazareth da Silva9, Fabio Röh9,10, Helder Queiroz11, Kenneth L. Chiou12, Anthony Di Fiore4, Michael E. Alfaro1,3 and Jessica W. Lynch Alfaro1,3

ABSTRACT

Aim Our aim was to examine gracile capuchin (Cebus) and robust capuchin monkey (Sapajus) diversification, with a focus on recent Sapajus expansion within Amazonia. We wanted to reconstruct the biogeographical history of the clade using statistical methods that model lineages’ occupation of different regions over time in order to evaluate recently proposed ‘Out of Amazonia’ and ‘Reinvasion of Amazonia’ hypotheses as alternative explanations for the extensive geographical overlap between reciprocally monophyletic gracile (Cebus) and robust (Sapajus) capuchin monkeys.

Location Central and South America.

Methods We reconstructed a time-calibrated molecular phylogeny for capuchins under Bayesian inference from three mitochondrial genes. We then categorized 12 capuchin clades across four Neotropical centres of endemism and reconstructed the biogeographical history of the capuchin radiation using six models implemented in ‘BioGeoBEARS’. We performed a phylogeographical analysis for a robust capuchin clade that spans the Atlantic Forest, Cerrado, Caatinga and Amazonia.

Results We find support for a late Miocene vicariant Cebus-Sapajus divergence and a Pleistocene Sapajus invasion of Amazonia from the Atlantic Forest. Our new analyses confirm Sapajus diversified first in the Atlantic Forest, with subsequent range expansion into widespread sympatry with Cebus in Amazonia, as well as multiple expansions into drier savanna-like habitats. We do not find mitochondrial molecular congruence with morphological species distinctions for Sapajus flavius, S. cay, S. macrocephalus, S. luidinosus and S. apella; instead, these five morphological types together form a single widespread clade (Bayesian posterior probability = 1) with geographical substructure and shared ancestry during the Pleistocene.

Main conclusions Our results support vicariance dividing ancestral capuchin populations in Amazonia versus the Atlantic Forest, and a Pleistocene ‘Amazonian invasion’ by Sapajus to explain the present-day sympatry of Cebus and Sapajus.

Keywords Amazonia, Atlantic Forest, BEAST phylogeography, BioGeoBEARS, Caatinga, Cebus, Cerrado, range expansion, Sapajus
INTRODUCTION

Neotropical primates are the major diurnal mammalian radiation in the Neotropics, and include over 140 species found across Amazonia, the Andes, the Atlantic Forest, and even habitats such as the Llanos, Caatinga, Cerrado and Central Grasslands that are too hot and dry to support many types of mammals. Neotropical primate diversity is highest within the tropical rain forest habitat in Amazonia. Major questions include whether Neotropical primates diversified first in Amazonia, and how and when they spread throughout the various biomes they now inhabit.

The wide distribution of capuchin monkeys, with species ranging from Central America to northern Argentina, makes them a potential model system for addressing questions of Neotropical primate biogeography. Capuchins comprise two major forms, the more lightly built gracile capuchin, *Cebus*, and the robust capuchin, *Sapajus*. Although recent studies have confirmed the genetic distinctiveness of these two lineages (Lynch Alfaro et al., 2012a; Martins et al., 2014), key questions remain regarding their present-day distribution, including the extensive sympathy between Amazonian *Cebus* and *Sapajus* species.

Two competing hypotheses have been posited to explain current capuchin distributions (Lynch Alfaro et al., 2012a). The ‘Out of Amazonia’ (OOA) hypothesis places the ancestral range for all capuchins in Amazonia, an initial diversification of both crown robust and crown gracile capuchins within Amazonia, and subsequent independent and recent invasions of non-Amazonian regions by both gracile and robust capuchins (Lynch Alfaro et al., 2012a). The ‘Reinvasion of Amazonia’ (ROA) hypothesis proposes that gracile capuchins, *Cebus*, diverged from robust capuchins, *Sapajus*, early in capuchin evolutionary history; crown *Cebus* originated in Amazonia and crown *Sapajus* in the Atlantic Forest; and the presence of *Sapajus* in Amazonia today reflects a recent invasion from the Atlantic Forest into *Cebus*-occupied territory (Lynch Alfaro et al., 2012a).

Phylogeographical analyses using pure dispersal models strongly support the ‘ROA’ hypothesis (Lynch Alfaro et al., 2012a), and as such have transformed the interpretation of capuchin taxonomy, conservation biology, morphology and ecology (see Lynch Alfaro et al., 2012b, 2014, 2015b; Matthews, 2012; Rylands et al., 2013; Martins et al., 2014; Oliveira et al., 2014; Boublí et al., 2015; Schneider & Sampaio, 2015; Wright et al., 2015). However, limitations of these studies have left key predictions of the OOA hypothesis untested. For example, Lynch Alfaro et al. (2012a) and Nascimento et al. (2015) provided limited sampling of eastern Amazonia and Cerrado/Caatinga localities, which have the potential to strongly corroborate the OOA scenarios. In addition, capuchin monkey biogeography has yet to be interpreted in light of models that integrate processes beyond pure dispersal (Ronquist, 1997; Ree et al., 2005; Ree & Smith, 2008; Landis et al., 2013).

Here, we assemble an expanded sampling of capuchin monkeys across their range and conduct new biogeographical analyses using recently developed statistical approaches. Our overall goal is to test hypotheses explaining capuchin biogeography (Boublí et al., 2012; Lynch Alfaro et al., 2012a; Nascimento et al., 2015; Ruiz-García et al., 2016). We thus assemble a new cytochrome *b* + D-loop + cytochrome oxidase I (COI) matrix with three times the samples and more comprehensive spatial and taxonomic sampling compared to Lynch Alfaro et al. (2012a). We produce a detailed phylogeny for capuchin monkeys, with all major lineages represented, to confirm the monophyly of *Cebus* and *Sapajus* and to recover major clades within the capuchin radiation. We compare alternative biogeographical scenarios using a flexible framework for inferring spatial histories that allows for a number of different mechanisms causing range expansion or contraction. We assess whether the employment of a more complete genetic dataset and statistical methods with expanded models increase support for previous conclusions about capuchin biogeographical evolution or support new scenarios altogether. Our sampling allows us to address two new biogeographical hypotheses for this clade. The first hypothesis postulates vicariance between Amazonia and the Atlantic Forest. Divergence between *Sapajus* and *Cebus* is caused by isolation of ancestral capuchin populations in the Atlantic Forest versus Amazonian habitats, respectively, with subsequent expansion of *Sapajus* into sympatry with Amazonian *Cebus* (‘ROA’ hypothesis). The second, opposing, hypothesis postulates vicariance within Amazonia due to the formation of the Amazon River. Capuchins arise in Amazonia, and subsequently *Cebus* is isolated north of the Amazon River, with *Sapajus* isolated south of the Amazon River and expanding into the Atlantic Forest. Sympatry occurs with subsequent crossing of Amazon River in both directions (‘OOA’ hypothesis).

MATERIALS AND METHODS

DNA extraction, amplification, sequencing and alignment

We sequenced three mitochondrial genes, cytochrome *b* (*Cyt b*, 1133 bp), D-loop (1244 bp) and COI (673 bp), for 108 capuchin samples from poorly known or previously unsampled localities and species (see Table S1 in Appendix S1 in Supporting Information). For *Cyt b* and D-loop we used primers and PCR protocols from Lynch Alfaro et al. (2012a) and Boublí et al. (2012). We adapted COI primers and protocols from Ward et al. (2005) (see Appendix S2). We combined the 108 new samples with 40 samples analysed previously by Lynch Alfaro et al. (2012a) and 13 new samples with reliable species identification extracted from GenBank, totalling 161 capuchin samples. Note that samples from Ruiz-García et al. (2016) were not available for use in this study. We used only sequences over 500-bp long to improve clade support (Nascimento et al., 2015). We used the most recent comprehensive taxonomic review for capuchins (Rylands et al., 2013), in relation to specimen
morphology and collection locality, to assign species names to samples (locality data in Fig. 1 and see Table S1 in Appendix S1). Although the limitations of mitochondrial DNA studies are well understood, Cyt b, D-loop and COI have performed well in capturing phylogeny and biogeographical history in a wide range of species (Tobe et al., 2010; Zhang et al., 2011; Robins et al., 2014). In addition, mtDNA sequence can be reliably generated from museum ‘crusties’ and skins using now standard protocols (e.g. Lynch Alfaro et al., 2012a, 2015a; Mercès et al., 2015).

**Phylogenetic reconstruction**

We used MrBayes 3.2.6 (Ronquist et al., 2012) to reconstruct the phylogenetic relationships among capuchins, using Saimiri as an outgroup. We partitioned Cyt b and COI by codon position for the analysis and ran the Markov chain Monte Carlo (MCMC) for 50 million generations. We used ‘sumt’ to calculate the Bayesian consensus tree. We then used BEAST 1.8.1 (Drummond et al., 2012) to reconstruct a time tree appropriate for our biogeographical analyses by pruning the alignment to a single exemplar taxon for each recovered well-supported clade. We considered branch lengths and recovered polytomies in addition to posterior probabilities at nodes when distinguishing clades. This resulted in a new, pruned alignment with 13 tips, representing 12 capuchin clades and 1 outgroup. We specified a HKY+G model of nucleotide substitution and partitioned by codon position, using Cyt b only to avoid branch length bias due to missing data in the other genes. We used Saimiri as an outgroup and the fossil Neosaimiri (Kay, 2015) to calibrate the split at 12.5 Ma (lognormal prior, offset = 12.5, mean = 0, SD = 1) between capuchins and squirrel monkeys for the time tree. We ran the MCMC analysis for 100 million generations, sampling states every 10,000 generations with a Yule prior on the diversification rate, and an uncorrelated, lognormal relaxed molecular clock.

**Biogeographical modelling**

**Biogeographical scenarios**

In addition to testing for evidence for ‘OOA’ versus ‘ROA’ hypotheses (see Introduction), we tested for reciprocal monophyly for Sapajus in Amazonia (S. macrocephalus + S. apella); the Cerrado + Caatinga + Central Grasslands (S. cay + S. libidinosus); and the Atlantic Forest (Sapajus flavius + S. xanthosternos + S. robustus + S. nigritus). Lack of reciprocal monophyly across habitats would suggest multiple independent invasions and convergent adaptations to habitat types by independent Sapajus lineages.

**Historical biogeography**

To test alternative hypotheses of capuchin origins and radiation we used the R-package ‘BioGeoBEARS’, a recently developed flexible maximum likelihood framework for inferring spatial histories (Matzke, 2013, 2014; see Buckner et al., 2015 and Lynch Alfaro et al., 2015a for its recent application to primate taxa). ‘BioGeoBEARS’ software allows for a statistical comparison of a suite of models that include alternative biogeographical processes related to cladogenesis and that allow for species distributions across multiple regions (Matzke, 2014).

Our Bayesian consensus tree was pruned to include one representative tip for each ‘species’ recovered in the MrBayes analysis. We lumped S. flavius, S. cay, S. libidinosus, S. apella and S. macrocephalus (Sapajus clade 4, see below) as one species. We also combined C. unicolor + C. albifrons, C. olivaceus + C. kaapori and C. versicolor + C. cesarae. This resulted in a new, pruned alignment with 12 tips. We coded presence/absence in four major geographical regions for each tip based on known species distribution using Rylands et al. (2013): Atlantic Forest (AF), Cerrado/Caatinga/Central Grasslands (CC), Amazonia (AM) and Central America/Andes (CA) (Fig. 1). We used the

![Figure 1](https://example.com/figure1.png)
pruned, time-calibrated tree with regional coding in 'BioGeoBEARS' (Matzke, 2013, 2014). Our candidate pool consisted of six models: DEC, DEC-J, DIVA, DIVA-J, BAYAREA-LIKE, BAYAREALIKE-J and we used model weights calculated from corrected Akaike information criterion (AICc) to identify the best model or set of models for our data (Matzke, 2014). The 'BioGeoBEARS' script is available in Appendix S2.

**Phylogeographical reconstruction of widespread Sapajus**

To examine phylogeographical patterns for a widespread robust capuchin group recovered in the MrBayes tree, we performed both symmetric and asymmetric discrete states phylogeographical analyses in **BEAST** under a constant size coalescent prior. These analyses allowed us to take into consideration the uncertainty in the relationships by simultaneously estimating the topology and the distribution history. To compare the asymmetric and symmetric analyses for their fit to our data, we performed marginal likelihood estimation (MLE) using path sampling (PS) and stepping stone (SS) sampling (Baèlé et al., 2012, 2013). We used the default priors for the MLE. We categorized the locality of each tip for the widespread Sapajus clade in the Bayesian consensus tree as occurring in one biogeographical 'region' (Fig. 2). We divided Amazonia into four regions using the Amazon River and two major tributaries, Negro and Madeira, as described by Wallace (1852) and later Cracraft (1985): Guianas (GU), Negro (NE), Inambari (IN), Rondonia (RO). We also separated the drier open ecosystems Central Grasslands, Cerrado and Caatinga (CC), from the Atlantic Forest (AF). These regions were modified from Jameson Kiesling et al. (2015) and Cracraft (1985). As mtDNA is inherited maternally, it tracks the movement of female lineages over time. Capuchin monkeys are female philopatric with male-biased dispersal (Lynch Alfaro et al., 2014), so tracing female lineage was a conservative test for the amount of movement over time in this taxon.

Within Amazonia, we tested for isolation by river barriers for widespread Sapajus. If gene flow within Amazonian capuchin populations was restricted to major interfluves, we expected a pattern of closely related subclades within, not across, four major Amazonian regions delineated by the Amazon River and its major tributaries. Evidence for the river barrier hypothesis was found for the sister group to capuchins, the Amazonian squirrel monkeys (Lynch Alfaro et al., 2015a), as well as for marmosets and tamarins (Buckner et al., 2015) uakaris, widow titi monkeys, and white-fronted capuchins (Boublí et al., 2015). Robust capuchin lineages with clades spanning multiple interfluves would provide evidence against rivers as significant isolating mechanisms.

**RESULTS**

**Capuchin phylogenetics**

We find strong support for reciprocally monophyletic clades of gracile and robust capuchins in the phylogenetic reconstruction from MrBayes (Fig. 3a). Our species tree time from **BEAST** (see Figure S1 in Appendix S3) estimates the mean divergence time between Sapajus and Cebus at 5.8 Ma [95% higher posterior density (HPD) = 2.24–10.94] (late Miocene).

For robust capuchins we recover strong support for the clades S. xanthosternos [posterior probability (pp) = 1] and S. nigritus (pp = 1), but little evidence for reciprocally monophyletic 'species' clades for the rest of Sapajus. Two S. robustus samples form a weakly supported clade (pp = 0.74). All other robust capuchins form one large widely distributed clade ('Sapajus CLADE 4,' pp = 1), comprising S. apella, S. macrocephalus, S. cay, S. libidinosus and S. flavius (Fig. 3b). Subclades within this group are not always congruent with morphological hypotheses about robust capuchin species, a result concordant with the mitochondrial results from Ruiz-Garcia et al. (2012, 2016). Most strikingly we find S. apella sensu stricto to be extensively paraphyletic with respect to other Sapajus species. Within the widespread clade, we recover several geographical subclades: (1) S. flavius (pp = 0.95); (2) S. libidinosus together with eastern S. apella (pp = 0.80); (3) S. cay from Paraguay and Rondônia/Mato Grosso plus S. apella from Guaporé and Apuí (pp = 0.81); (4) S. cay from Chapada dos Guimarães and S. apella from Alta Floresta, both in Mato Grosso, Brazil (pp = 1); (5) S. apella/macrocephalus (pp = 1) from north of the Amazon and Solimões rivers, as well as from regions just south of the Amazon River near the Xingu and Tapajos rivers, and just south of Manaus, together reaching all four Amazonian quadrants, as far northwest as Colombia; (6) S. apella/macrocephalus (pp = 0.95) from Brazil and Peru. Minimum clade ranges for each of the six subclades of the widespread Sapajus clade are geographically coherent, although several do not match the current morphological hypotheses about robust capuchin taxonomy (Fig. 3c).

Within Cebus, we recover eight strongly supported clades: Cebus versicolor forms a clade (pp = 1) with C. cesarae within the Magdalena river valley in the Colombian Andes, as part of a larger clade including C. capucinus in Panama and Colombia (pp = 1), and C. imitator in Central America (pp = 1). In the eastern Andes, C. brunnneus (pp = 1) is sister to C. leucocephalus (pp = 1). Within Amazonia, there is strong support for a C. unicolor + C. albifrons clade (pp = 1). C. yuracus forms a clade with one C. unicolor sample (pp = 1) in western Brazil and Ecuador, and C. o. oливaceus + C. o. castaneus + Cebus kaapori form a strongly supported clade (pp = 1). Our C. kaapori sample is from the holotype, collected by Queiroz (1992), from the type locality Chega-Tudo, in Caratapuera Municipality, Maranhão State, Brazil. Here, we provide a correction to the original Queiroz (1992) for the coordinates of that locality, 02°20' S 46°05' W (see Table S1 in Appendix S1).

**Reconstruction of capuchin biogeographical history**

Model weights based on AICc favoured the DEC model (Table 1, Fig. 4). Under this scenario the ancestral capuchin
population is inferred to have had a wide range that included both Amazonia and the Atlantic Forest. The divergence between *Sapajus* and *Cebus* is a vicariance event, where *Sapajus* is isolated in the Atlantic Forest, while the lineage leading to gracile *Cebus* is isolated in Amazonia. Crown *Sapajus* is recovered with a distribution restricted to the Atlantic Forest, with subsequent rapid expansion to the Caatinga/Cerrado/Central Grasslands and Amazonia regions at c. 500 ka. This means that *Cebus*, established in Amazonia at c. 2.4 Ma, diversified for c. 2 Myr before the appearance of any *Sapajus* clade in that region.

For crown *Cebus*, a single expansion event to the Central America/Andes region occurred at c. 2.2 Ma, with subsequent range contractions to just Central America/Andes for
the lineage leading to *C. capucinus*, *C. imitator* and *C. versicolor*, and to Amazonia only for both the lineage leading to *C. brunneus* (note that the ‘Amazonia’ region extends to the northern Venezuelan coast; see Fig. 4), and for the lineage leading to *C. unicolor*.

**Reconstruction of widespread *Sapajus* radiation**

We do not find any statistical difference in support for the symmetric versus the asymmetric analyses (K ≈ 0). Both the phylogeographical reconstructions detail the explosive Pleistocene range expansion throughout Amazonia, Cerrado, Caatinga and southern grasslands for *Sapajus* (Fig. 2 and see Figure S2 in Appendix S3). In the asymmetric reconstruction, robust capuchins enter Amazonia by a single expansion from AF to RO at around 300–500 ka, with a second unique expansion across the Amazon River to GU at 200 ka, and at least two independent expansions across the Madeira River from RO to IN. There is also evidence for expansion from GU to IN, and IN to GU, as well as back-crossing from both IN and GU to RO. No strong geographical pattern demarcates rivers as firm barriers separating the four Amazonian regions; instead, there has been significant and recurrent recent movement by *Sapajus* across the Amazon, Solimões, Madeira and Negro rivers. *Sapajus* also spread from RO three times into the Cerrado/Caatinga (CC) region, all within the last 200 ka.

The addition of a distribution partition to our beast analysis provides support for the monophyly of *Sapajus robustus* (0.93) and its position as the sister group to a widespread Amazonian-Grasslands-Atlantic Forest clade (pp = 1.0). *Sapajus robustus* is restricted to the Atlantic Forest, making it clear that the explosive radiation of robust capuchin monkeys began with a recent migration into Amazonia and grasslands.

A significant difference between our asymmetric and symmetric BEAST analyses is related to *S. flavius*. In the

---

**Figure 3** Bayesian consensus tree for capuchin phylogeny: (a) *Cebus* and *Sapajus* phylogeny and (b) details of *Sapajus* CLADE 4; (c) map with minimum convex polygons to show geographical distribution of major subclades within the widespread *Sapajus* CLADE 4. Subclades are presented as minimum convex polygons that connect sample localities. Branches are colour coded to regions from polygons map. Posterior probabilities are shown as asterisks for above 0.95 and as circles for between 0.70 and 0.95. Sample species ID and number correspond to Table S1 in Appendix S1. [Colour figure can be viewed at wileyonlinelibrary.com]
symmetric beast phylogeography analysis, *S. flavius* is recovered as the sister group to *Sapajus* CLADE 4 (see Figure S2 in Appendix S3). On the other hand, *S. flavius* is embedded within the widespread clade in the asymmetric analysis (Fig. 2), suggesting an expansion back into the Atlantic Forest from the Amazonia/Grasslands regions.

**DISCUSSION**

Our expanded sample set with comprehensive representation of capuchin species provides further support for the monophyly of robust and gracile capuchins with a split at 5.8 Ma. A late Miocene *Sapajus-Cebus* split was previously supported using combined nuclear and mitochondrial data in Perelman *et al.* (2011) (6.0 Ma, 95% confidence interval 3.13–9.35 Ma) and Springer *et al.* (2012) (5.7 Ma average, 3.55–8.33 Ma composite 95% min–max). However, these time trees did not include all capuchin species; here, we add *C. kaapori, C. versicolor* and *S. flavius*. One exception to monophyly of *Cebus* and *Sapajus* has been recently reported in Ruiz-García *et al.* (2016), where some *Cebus albifrons sensu lato* individuals from the Colombian Eastern Llanos and Ecuadorian Amazonia displayed *Sapajus* mitochondrial haplotypes, suggesting the possibility of localized past or present introgression across these two genera. Unfortunately, neither sequences nor morphological data are publicly available for any of these individuals at this time.

Our biogeographical analyses reconstruct a widespread capuchin ancestral distribution across forested areas of South America, from Amazonia to the Atlantic Forest, supporting Amazonia versus Atlantic Forest vicariant origin for *Cebus* and *Sapajus*. The distribution of *Sapajus* exclusively in the Atlantic Forest until recently, and *Cebus* exclusively in the Amazonia and Central America/Andes regions from early on, suggests a cladogenetic event (vicariance) led to their initial diversification in isolation.

Throughout capuchin evolution, the Brazilian Shield (within AF, CC, RO Regions) and Guiana Shield (within GU) have been constant highland regions undergoing few physical changes in configuration but perhaps experiencing changes in floral assemblages related to climate (Aleixo & Rossetti, 2007). The divergence between *Sapajus* and *Cebus*, dated consistently at c. 6 Ma, coincides with floral assembly of the Cerrado (4–8 Ma), as indicated by the onset of diversification of many Cerrado plant lineages (Simon *et al.*, 2009). The coincidence in timing of the divergence between robust and gracile capuchins and of the Cerrado assemblage suggests that the development of the Cerrado may have acted as a vicariant agent, especially given the reconstructed distributions of ancestral *Sapajus* (Atlantic Forest) and *Cebus* (Amazonia). The only extant capuchins distributed in CC belong to our widespread *Sapajus* CLADE 4 – which invaded those biomes only in the last 500 kyr, and have recent cranial and post-cranial adaptations to drier habitats (Wright *et al.*, 2015).
Combining information from our biogeographical and phylogeographical analyses, we find strong evidence that all of the Atlantic Forest robust capuchin species are monophyletic clades, with the exception perhaps of *S. flavius*, which was found as part of the widespread *Sapajus* ‘clade 4’ in our MrBayes phylogeny. Major Amazonian rivers do not appear to be strong barriers for dispersal or gene flow for *Sapajus*, with up to 12 examples of dispersal across significant Amazonian rivers demonstrated through our best analysis by just the sampling included in this study.

**Gracile capuchin biogeography and phylogenetics**

In our analysis, *Cebus* shows strong geographical divisions, suggesting a widespread ancestor diversified in the Andes and Central America, and was split by range expansion and subsequent range contraction in the Andes and Central America. The Amazonian group (*C. yuracu*, *C. unicolor*, *C. albifrons*, *C. olivaceus* and *C. kaapori*) is about the same age as the CA group (*C. capucinus*, *C. imitator*, *C. versicolor*, *C. cesarae*, *C. leucocephalus* and *C. brunnneus*). The recently discovered *C. kaapori* was initially considered to be a subspecies of *C. olivaceus* (Harada & Ferrari, 1996), but other morphological taxonomists have argued that it is a unique species (Groses, 2001, 2005; Silva-Júnior, 2001), with some morphological characteristics more similar to *C. albifrons sensu lato* (Masterson, 1995). In our MrBayes analysis, *C. kaapori* is recovered within the strongly supported *C. olivaceus* clade. A close relationship between *C. kaapori* and *C. olivaceus* makes sense, as *C. o. castaneus* is the gracile capuchin geographically most proximate to *C. kaapori*. *Cebus versicolor* is recovered as sister to *C. cesarae*. This is concordant with geographical distribution as both species are found in isolated pockets of the Magdalena Valley in Colombia. MtDNA COI analysis of *Cebus albifrons sensu lato* by Ruiz-García et al. (2010) found *C. cesarae* and *C. pleei* as sister to *C. versicolor.

**Robust capuchin biogeography and phylogenetics**

Our phylogeographical analysis suggests robust capuchins were able to colonize an array of divergent habitat types all within a very short time period, in the last 1 Myr. Robust capuchins show evidence of multiple moves within and between all regions, except that they are absent from the Central American and Andean region. This may be explained because their arrival at the foot of the eastern Andes was so recent that the mountain range was already fully formed and impassable (Lynch Alfaro et al., 2012a).

Consistent with this recent explosive radiation, we find little support for the internal topology of a morphologically diverse and geographically expansive clade of robust capuchins. Morphological species contained in this group include *S. libidinosus*, *S. cay*, *S. flavius*, *S. macrocephalus* and *S. apella*. While there is significant individual variation within *Sapajus* populations, these five morphological species each display distinct morphological patterns with clear geographical correspondence (Silva-Júnior, 2001), possibly as a consequence of repeated founder effects. While the recent divergence times among all the Amazonian and open region grasslands capuchins (all under 400 ka) suggest there probably has not been time enough for speciation among these morphological types, our data do point to geographical clades that deserve further study for their morphological and behavioural population characteristics. For example, morphological differences between *S. macrocephalus* north and south of the Amazon River have already been described (Rylands et al., 2013); our analysis here provides new hypotheses for geographical boundaries between major population centres within Amazonian *Sapajus*, and suggests the need for new morphological and morphometric analyses for undersampled regions. We suggest nuclear genomic methods will be necessary to provide a complementary perspective to help resolve the relationships within this clade.

Rapid Pleistocene expansion throughout Amazonia has been recovered for the sister taxon to capuchin monkeys, the squirrel monkeys (Chiou et al., 2011; Lynch Alfaro et al., 2015a). *Sapajus* presence today in western Amazonia is explained in the best phylogeography analysis as resulting from two distinct pathways: one was expanding from RO west across the Madeira River, into IN, and then spreading north into NE. The second was a move first from RO across the Amazon River into GU, then expansion west across the Negro River into NE, and from there south to IN. The exact distribution of *S. macrocephalus* in western Amazonia is contested (Groses, 2001, 2005; Silva-Júnior, 2001). Rylands et al. (2013) depict *S. macrocephalus* morphological diversity with Brazilian, Peruvian, Bolivian and Colombian forms; one explanation for this variation congruent with our results is that western Amazonia has been colonized multiple times from different *Sapajus* seed populations from different parts of Amazonia. In contrast, most of *Sapajus* diversity in eastern Amazonia can be explained by a single founder event from AF into RO, and another single founder event from RO into GU, with some later expansions back to RO from GU.

*Sapajus cay* has been the subject of controversy within robust capuchin taxonomy, and two major authorities (Groses, 2001, 2005; Silva-Júnior, 2001) disagree significantly about its distribution, in part due to differences in the localities of samples available in each of their studies. *Sapajus cay* (sensu Rylands et al., 2013) is recovered here as paraphyletic, with evidence for at least two distinct *Sapajus* populations moving independently into gallery forests in more open habitat types. *Sapajus cay* samples from Paraguay, Rondônia and Mato Grosso cluster with *S. apella* from Rondônia; while *S. cay* from northern central Mato Grosso forms a clade with nearby *S. apella* from Alta Floresta. These two populations have apparently converged morphologically to a lighter pelage phenotype in the drier, open regions. *Sapajus libidinosus* (sensu Rylands et al., 2013) marks a third expansion into CC from Amazonian *Sapajus*, with a resultant lighter pelage (Fig. 2). All expansions into drier habitats appear to
have occurred in the last 200–100 kyr, similar to findings for other Neotropical primates (Lynch Alfaro et al., 2015b).

We expected to recover *S. flavius* as sister to *S. xanthosternos* because of geographical proximity or sister to *S. libidinosus* based on a gradation of morphology between the two morphotypes (Silva, 2010); or possibly sister to all other *Sapajus* species, based on its unusual morphology (Lynch Alfaro et al., 2012a). In our *BEAST* phylogeography and MrBayes analyses, *S. flavius* is recovered as part of the widespread *Sapajus* clade 4 (Figs 2 & 3b). However, in the symmetric *BEAST* phylogeography analysis, *S. flavius* is recovered as the sister group to *Sapajus* clade 4 (see Figure S2 in Appendix S3), suggesting all four Atlantic Forest *Sapajus* taxa may have diverged from each other first prior to *Sapajus* expansion into other regions. Given the difference in topology and the range of support values among analyses, the phylogenetic placement of *S. flavius* remains uncertain, although close phylogenetic affinity to *Sapajus* clade 4 is clear.

**Sympathy between robust and gracile capuchins**

Overall, our data and analyses support the ‘ROA’ hypothesis (Lynch Alfaro et al., 2012a) by *Sapajus* as indicated by the ancestral reconstructions of the crown *Cebus* and crown *Sapajus* clades. The two capuchin genera diversify in isolation, with *Sapajus* species only later invading a *Cebus*-occupied Amazonia at c. 500 ka. The *BEAST* phylogeography analysis shows *Sapajus* invasion of Amazonia occurring first in Rondonia (south-western Amazon) and quickly spreading to all four quadrants of Amazonia. The recovered timing of entry into sympathy for gracile and robust capuchins indicate an event in Amazonia < 1 Ma, as a result of *Sapajus* counter-invasion into a region already occupied by *Cebus*.

**ACKNOWLEDGEMENTS**

Thanks to Robert Whittaker for inviting us to contribute this paper, and to Alexandre Antonelli, Marcelo Tejedor and anonymous referees for their excellent suggestions that have improved this article substantially. Special thanks to L.E. Olson and K. Fales, University of Alaska Museum. Support to M.G.M.L.’s PhD research was provided by a CNPq PhD fellowship (142141/2012-7) and CNPq SWE fellowship (201172/2014-3). Funding for this research was provided in part by NSF BCS 0833375 to J.W.L.A. Some of the molecular analyses and field expeditions were funded by CNPq/FAPEAM SISBIOTA Program (563348/2010-0) and CAPES (AXU no. 3261/2013) to I.P.F. and CNPq (grants ‘INCT em Biodiversidade e Uso da Terra da Amazônia’ 574008/2008-0; and 310880/2012-2) and NSF–FAPESP 2012/50260–6 (grant 1241066 – Dimensions US–BIOTA–São Paulo: Assembly and evolution of the Amazonian biota and its environment: an integrated approach) to A.A. This is article number 24 of the Amazonian Mammals Research Group.

**REFERENCES**


**SUPPORTING INFORMATION**

Additional Supporting Information may be found in the online version of this article:

**Appendix S1** Details of individual and GenBank accession numbers.

**Appendix S2** Supplementary details of methods.

**Appendix S3** Supporting results.

**BIOSKETCH**

Marcela G. M. Lima is a biologist interested in biogeography, phylogenetics and population genetics, with a focus on Neotropical vertebrates.


Editor: Alexandre Antonelli