A phylogenomic perspective on the robust capuchin monkey (Sapajus) radiation: first evidence for extensive population admixture across South America


http://dx.doi.org/10.1016/j.ympev.2018.02.023

<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>A phylogenomic perspective on the robust capuchin monkey (Sapajus) radiation: first evidence for extensive population admixture across South America</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type</strong></td>
<td>Article</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td>This version is available at: <a href="http://usir.salford.ac.uk/46258/">http://usir.salford.ac.uk/46258/</a></td>
</tr>
<tr>
<td><strong>Published Date</strong></td>
<td>2018</td>
</tr>
</tbody>
</table>

USIR is a digital collection of the research output of the University of Salford. Where copyright permits, full text material held in the repository is made freely available online and can be read, downloaded and copied for non-commercial private study or research purposes. Please check the manuscript for any further copyright restrictions.

For more information, including our policy and submission procedure, please contact the Repository Team at: usir@salford.ac.uk.
A phylogenomic perspective on the robust capuchin monkey (Sapajus) radiation

Marcela G. M. Lima1,2, Jessica W. Lynch Alfaro1,3, Janet C. Buckner4, Alexandre Aleixo2, David Cerny4, Jimmy Zheng4, Michael E. Alfaro4, Amely Martins5,6, Jean P. Boublí7, José de Sousa e Silva-Júnior2

1. Institute for Society and Genetics, University of California, Los Angeles, CA, USA
2. Curso de Pós-Graduação em Zoologia, Universidade Federal do Pará/Museu Paraense Emílio Goeldi, Belém, PA, Brazil
3. Department of Anthropology, UCLA, Los Angeles, CA, USA
4. Department of Ecology and Evolutionary Biology, University of California, Los Angeles, USA
5. Department of Anthropology, University of Texas at Austin, Austin, TX, USA
6. Centro Nacional de Pesquisa e Conservação de Primatas Brasileiros, ICMBio, MMA, Brazil
7. School of Environment and Life Sciences, University of Salford, UK

Corresponding Author: Marcela G. M. Lima

Universidade Federal do Pará, Rua Augusto Corrêa, 01 – Guamá, Belém, PA, Brasil,
CEP 66075-110

E-mail: marcela_gml@yahoo.com.br
Highlights

● Phylogenomic analyses support *Sapajus* and *Cebo* clades within capuchin monkeys
● Molecular data support *Sapajus nigritus*, *S. robustus* and *S. xanthosternos* as species
● UCE phylogeny lumps *Sapajus* Amazonian and grassland morphospecies
● SNP data separate *S. flavius* and *S. libidinosus* as sister species
● We recommend collapsing *S. apella*, *S. macrocephalus* and *S. cay* as one species

Abstract

Phylogenetic relationships among robust capuchin monkeys (*Sapajus*) are poorly understood. Taxonomies for this group based on morphology have considered from one to twelve different species. Current IUCN classification lists eight robust capuchins: *S. xanthosternos*, *S. nigritus*, *S. robustus*, *S. flavius*, *S. libidinosus*, *S. cay*, *S. apella* and *S. macrocephalus*. Here we assembled the first phylogenomic data set for robust capuchin monkeys using ultra-conserved elements (UCEs) to construct a robust capuchin phylogeny using RAxML. We extracted SNPs from the UCE data set, and created SNP phylogenies using Bayesian and Maximum Likelihood methods. We estimated a species tree using SVDquartets analyses. All phylogenomic analyses strongly supported *Sapa-
jus and Cebus clades within capuchin monkeys, and Sapajus nigritus, S. robustus and S. xanthosternos as species. However, the UCE phylogeny lumped morphospecies S. cay, flavius, libidinosus, apella, macrocephalus, and flavius together as a single widespread evolutionary lineage. The Bayesian SNP phylogeny was better resolved, and recovered S. flavius and S. libidinosus as sister species, together as sister to an S. apella + macrocephalus + cay clade; S. apella, S. cay, and S. apella individuals were interspersed together in the topology with no evidence for monophyly for any of these three morphological species. The species tree topology differed from the UCE and SNP topologies in that it reconstructed two major clades for robust capuchin monkeys: one Atlantic Forest clade (S. robustus, S. xanthosternos, and S. nigritus) and one widely distributed clade (S. flavius, S. libidinosus, plus north and south Amazonian robust capuchins). As morphological and molecular subdivisions of the Amazonian group + southern grasslands group (currently recognized as S. cay, S. apella and S. macrocephalus) are discordant, we recommend lumping all Amazonian plus southern grassland robust capuchin taxa as S. apella without subspecies.

Keywords

Neotropical primates, phylogeny, single nucleotide polymorphisms (SNPs), species tree, Ultraconserved elements (UCEs)

1. Introduction
Robust capuchin monkeys (*Sapajus*) comprise a widespread Neotropical primate genus found across cis-Andean Latin America, from the Colombian Llanos to the Guianas and throughout the Amazon basin as well as in the Atlantic Forest, Cerrado, Caatinga and Central Grasslands of South America, as far south as northern Argentina (Rylands et al., 2013). These primates as a group are true habitat generalists, with an incredible diet breadth compared to other Neotropical primates. While fruit and insects form the bulk of their diets, their robust jaw morphology coupled with behavioral adaptations for tool use and manipulative and extractive foraging together allow for the exploitation of encased and hidden foods unavailable to most other non-human animals (Fragaszy et al., 2004; Lynch Alfaro et al., 2012b).

Taxonomists have disagreed about the proximity of the relationship of robust capuchins to gracile capuchins. Elliot (1913) created a taxonomic key that divided the genus *Cebus* into tufted and non-tufted groups on the basis of the presence or absence of hair tufts on the frontal region of the head. However, only after Hershkovitz (1949) was there a general consensus about this division, with just one species (*Cebus apella* Linnaeus, 1758) recognized among the tufted group. Hill (1960) also considered all robust capuchins as one cosmopolitan species, *Cebus apella*, placed within the gracile capuchin genus, *Cebus*. Groves (2001, 2005) considered capuchins to form two species groups: (1) *C. capucinus* group with *C. capucinus*, *C. albifrons*, *C. olivaceus*, and *C. kaapori*; and (2) *C. apella* group with *C. apella*, *C. libidinosus*, *C. nigritus*, and *C. xanthosternos* (Table 1). Silva-Júnior (2001) separated robust capuchins as a different subgenus (*Sapajus*) from gracile capuchins (*Cebus*) on the basis of distinct cranial, postcranial and pelage morphology. Subsequently, genetic research validated the separation of robust and gracile capuchins as two distinct and equally diverse clades using mitochondrial (Lynch Alfaro et al., 2012a; Lima et al., 2017) and a combination of mtDNA
and nuclear (Perelman et al., 2011) markers. Two Alu elements provide strong evidence for the monophyly of robust versus gracile capuchins: Alu element S49P is present in *Sapajus* but not *Cebus* (Viana et al., 2015) and the AluSc8 insertion is found in *Cebus* but not *Sapajus* (Martins Jr. et al., 2015). A recent review justified the splitting of robust and gracile capuchins into two genera (*Cebus* for gracile capuchins and *Sapajus* for robust capuchins) based on the distinct morphology, biogeographic history, behavior, and ecology of each type (Lynch Alfaro et al., 2012b).

Taxonomists have also disagreed about the number of species encompassed by extant robust capuchins based on morphology (Table 1). Elliot (1913) recognized twelve species of robust capuchins, but Cabrera (1957) and Hill (1960) placed all robust forms into one species, *Cebus apella*, while retaining 11 and 16 subspecies, respectively. For the four decades between 1960 and 2000, most researchers lumped all robust capuchins as one species irrespective of place of origin, usually without regard for subspecies designations (e.g. Cole, 1992; Daegling, 1992; Ford and Hobbs, 1996; Master- son, 1997; Wright, 2005a; 2005b, 2007), leading to obfuscation of species or population differences within the robust capuchin literature (see Lynch Alfaro et al., 2014 for discussion). However, Torres de Assumpção (1983) pointed to distinct geographical variation in morphology among robust capuchin populations within Brazil, and especially within the Atlantic Forest. More recent morphological analyses have provided evidence for multiple *Sapajus* species (Groves, 2001, 2005; Silva-Júnior, 2001, 2002, 2005; Rylands et al., 2005, 2012, 2013; Rylands and Mittermeier, 2009). The robust capuchin group is now considered by most taxonomists to be comprised of four to eight species (Silva-Júnior., 2001; Groves, 2001; Rylands and Mittermeier, 2009; Rylands et al., 2005, 2012, 2013). The IUCN (2015) currently recognizes eight species: *Sapajus flavius*, the blonde capuchin; *S. xanthosternos*, the yellow-breasted capuchin; *S. robustus*,
the robust tufted capuchin; *S. nigritus*, the black-horned capuchin; *S. apella*, the brown capuchin; *S. macrocephalus*, the large-headed capuchin; *S. cay*, Azara’s capuchin; and *S. libidinosus*, the bearded capuchin.

Recent biogeographic analyses based on mitochondrial DNA suggest that the time depth of the radiation of extant robust capuchins is about 2.5 My of diversification, with diversity accumulating first in the Atlantic Coastal Forest of Brazil, and a recent expansion of robust capuchins throughout the Amazon Basin and Cerrado, Caatinga and Central Grasslands in the last 500,000 years (Lynch Alfaro et al., 2012a; Lima et al., 2017). These analyses suggest that while the Atlantic Forest populations are relatively old and distinct, and can be separated as up to four different species, the Amazon/Grasslands radiation is better considered a highly polymorphic single species or species complex (Lima et al., 2017). If our current nuclear data set is congruent with the mtDNA data, we would expect to see evidence for four to five species: *S. nigritus*, *S. robustus*, and *S. xanthosternos* each as reciprocally monophyletic clades, with *S. flavius* either nested within or as the sister group to a single clade that extends across the Amazon and grasslands habitats in South America (and encompasses *S. apella*, *S. libidinosus*, *S. macrocephalus* and *S. cay* morphospecies) (Lima et al., 2017).

Here we use phylogenomic markers, ultraconserved elements (UCEs), to infer the phylogeny for robust capuchin monkeys, and to assess the evidence for congruence with species assignment by morphology and by mitochondrial and Alu markers. The UCE-based approach enriches DNA libraries for hundreds or thousands of UCEs and their flanking regions; then employs massively parallel sequencing for these libraries, and informatic tools to assemble, align and analyze the data (Faircloth et al., 2013). The UCE approach has been used successfully to resolve historically contentious taxonomical questions (McCormack et al., 2012; Crawford et al., 2012) including Pleistocene
radiations (McCormack et al., 2015). Previous studies using nuclear markers for capuchin phylogeny have utilized a limited number of taxa and used captive individuals from unknown provenance as species exemplars (i.e. Perelman et al., 2011, Springer et al., 2012). The present study marks the first test of robust capuchin phylogeny using phylogenomic markers to analyze genetic relationships across species-representative individuals from known provenance and assigned morphologically to each of the eight currently recognized *Sapajus* species. Based on the most comprehensive mtDNA analysis for the capuchin monkey radiation (Lima et al., 2017) we expect that much of the diversification within the *Sapajus* genus has occurred relatively recently, within the Pleistocene. We use SNP (Single Nucleotide Polymorphisms) data recovered within the UCE results in order to refine our understanding of robust capuchin diversification, as this technique was successful recently in elucidating the scrub-jay phylogeny across a similar geologic time frame (McCormack et al., 2015).

2. Material and methods

2.1. Samples, DNA extraction and sequencing

We sampled 67 individuals from 8 species of the genus *Sapajus* and 4 species of the genus *Cebus* from 62 localities distributed throughout the Atlantic Forest, Amazon, Central Grasslands habitats and Central America (Figure 1 and Table 2). The total genomic DNA was extracted from muscle and blood samples using the Qiagen DNeasy Blood & Tissue Kit, according to the manufacturer’s protocol. Library preparation, sequence capture and sequencing of ultraconserved elements were performed by RAPiD Genomics (Gainesville, FL, USA). Samples were quantified, normalized and sheared to an average fragment length of 350 base pairs (bp) for library preparation. Samples were
dual-indexed with unique i5 and i7 8bp indexes. Libraries were then pooled with
equimolar concentrations and the target sequence was captured using a custom set of
4715 probes targeting approximately 2300 UCE loci and 46 exons. Capture libraries
were then pooled with equimolar concentrations for multiplexed dual-end (2x100bp)
sequencing on an Illumina HiSeq 2500 v4 machine.

2.2. Sequence read quality control, assembly and UCE identification

We performed quality control using the trimming tool Trimmomatic 0.32.1
(Bolger et al., 2014) which trimmed sequences for adapter contamination, barcodes and
low-quality regions using the parallel wrapper script in Illumiprocessor 2.0.6 (Faircloth,
2013) (https://github.com/faircloth-lab/illumiprocessor). We assembled the contigs for
each sample using Trinity software package (vers. 2-25-2013) with default parameters
using Phyluce 1.5.0 (Faircloth, 2016). We matched our assembled contigs to 4715 UCE
loci custom-designed probe set using phyluce_assembly_match_contigs_to_probes in-
tegrating LASTZ 1.02.00 (Harris, 2007) from the Phyluce 1.5.0 (Faircloth, 2016) to
remove any contigs that did not match probes or that matched multiple probes designed
from different UCE loci. We performed in Phyluce 1.5.0 (Faircloth, 2016) the align-
ment of the contigs using the program phyluce_align_seqcap_align with MAFFT 7.271
(Katoh and Standley, 2013).

2.3. Phylogenetic analyses

For the phylogenetic analyses, we used a concatenated data set in a single
alignment constructed in Phyluce 1.5.0 (Faircloth et al., 2012; Faircloth, 2016). We
used two data sets of UCE alignments that included greater than 95% of taxa present for each UCE locus (5% missing) and greater than 75% of taxa present for each UCE locus (25% missing), totaling 1838 UCEs with five exons (RAPGEF1, NAT15, GRIA21, CLOCK e BDNF) and 1388 UCEs with two exons (NAT15, GRIA21) respectively. We performed phylogenetic tree reconstruction under maximum likelihood (ML) in RAxML 8.0.19 (Stamatakis, 2014), using a GTRCAT model of nucleotide substitution, 1000 replicate searches to identify the optimal tree and we generated non-parametric bootstrap replicates using the autoMRE option of RAxML. To find the best partitioning scheme, we used PartitionFinder (Lanfear et al., 2012). We considered each UCE as a data block and enabled hcluster (Lanfear et al., 2014) with equal weights. To evaluate the fit of each model we used the Bayesian information criterion (BIC).

2.4. SNPs Analyses

Upon identifying the target UCE loci, we computed the coverage at each base of each contig using a python wrapper included in Phyluce (phyluce_assembly_get_trinity_coverage_for_uce_loci). We then employed a de novo SNPs calling approach by aligning all raw reads against our sample of S. robustus, the reference sample with the highest coverage across all UCE loci enriched. This method integrated BWA (v 0.7.7-1) and PICARD (v 1.106-0) to output de novo aligned alignments in BAM format, repair any formatting violations, add read group header information, and mark duplicates in each BAM. We then merged all resulting BAMs into one file, realigning the data and calling SNPs and indels using GATK (v 3.5-0-g36282e4). To ensure high-quality SNPs in downstream analyses, we hierarchically filtered the data according to stringent quality and validation parameters, excluding
SNPs with QUAL under 25, low variant confidence, and poor validation. Finally, the resulting VCF was passed through VCFTOOLS (v 0.1.14) to remove all loci that missed SNP calls for over 25% of all 67 samples.

On a parallel track, we passed our SNP data through a recently developed automatic pipeline called SNPhylo (Lee et al. 2014), designed to efficiently reconstruct trees based on genome wide SNPs. We modified our filtered VCF file by manually filling in autosomal chromosome positions for each SNP call, a necessary condition in order to run the program. We then set the Minor Allele Frequency threshold to 0.04 and negated the LD threshold to enable a more inclusive dataset for phylogenetic inference. We also bypassed the default low-quality data removal step, because the dataset had already undergone quality filtration with GATK. As a final step, the SNPhylo pipeline employs DNAML to generate a maximum likelihood hypothesis and passes the tree through PHANGORN, which generates 1000 bootstrap replicates for the final result.

Additionally, in ExaBayes 1.4.1 (Aberer et al., 2014), we performed two independent runs, each with four chains (three heated and one cold), from random starting topologies for 10 million generations with a sampling frequency of 500 generations. Posterior distributions of trees were summarized with the consensus script and combined with the postProcParam script. Convergence and stationarity of parameter estimates were verified using Tracer 1.6.0 (Rambaut et al., 2013).

We estimated a species tree using SVDquartets analyses (Singular Value Decomposition Scores for Species Quartets; Chifman and Kubatko, 2014) implemented in PAUP* v4.0a147 (Swofford, 2002). This method infers quartets based on summaries of SNPs in a concatenated sequence matrix species using a coalescent model. We randomly sampled 10 million quartets from the data matrix to infer a species tree and we meas-
ured uncertainty in relationships using nonparametric bootstrapping with 1000 replicates. For this analysis, we did not include the samples from the widely distributed clade that did not form a part of the Northern Amazon or Southern Amazon subclades in the Bayesian (Exabayes) and maximum likelihood (SNPhylo) trees.

2.5. Divergence dating analyses

For the purposes of divergence time estimation, the 75% complete dataset was re-analyzed in PartitionFinder 2 (Lanfear et al., 2017) using the k-means algorithm described by Frandsen et al. (2015) and the BIC as the model selection method. We identified the fastest-evolving partition based on the rate multipliers reported in auxiliary files generated using the “--save-phylofiles” flag. This partition, totaling 10,316 sites, was then used to conduct a time tree analysis in BEAST 1.8.2 (Drummond et al., 2012).

We used the birth-death branching process (Gernhard 2008) with default hyperpriors placed on the growth rate and relative death rate hyperparameters to generate the joint prior distribution on tree topology and node heights. The uncorrelated lognormal relaxed clock was used to model the distribution of branch rates across the tree. In order to constrain the branch rate distribution to biologically realistic values, we placed a lognormal hyperprior with a mean of 0.005 (in real space) and a standard deviation of 1 on the ucld.mean hyperparameter (initial value of 0.005), and assigned a truncated exponential distribution with support from 0 to 1 and a mean of 0.3 to the ucld.stdev hyperparameter (initial value of 0.1). GTR+I was specified as the nucleotide substitution model; all of its free parameters were assigned default priors, the base frequencies were estimated rather than fixed, and the gamma rate heterogeneity distribution was discretized into 4 categories.
We ran the analysis under the fixed topology operator mix as specified in
BEAUTi v1.8.4 (Drummond et al., 2012), with the tuning of the ucl.d.mean and
ucl.d.stdev operators set to 0.9 and their weight increased to 6.0. All remaining operators
were kept at their default values. The topological constraint we employed (Supplementary
Figure 1) was based on the species tree inferred with SVDquarters (see below),
with one callitrichid and seven catarrhine outgroups manually added to the tree based on
the generally accepted phylogeny of the Simiiformes (Perelman et al., 2011; Springer et
al., 2012). The data for outgroup species were generated from Faircloth et al. 2012.
Since most of the calibration points we used were concentrated within the catarrhine
part of the tree, we pruned the capuchin taxon sample down to 4 species, with 2 repre-
sentatives of the genus *Cebus* (*C. capucinus* and *C. olivaceus*) and 2 representatives of
*Sapajus* (*S. apella* and *S. xanthosternos*) in order to increase the ratio of calibrated to
uncalibrated nodes, as well as to achieve a more uniform placement of fossil data
throughout the tree.

To calibrate the tree, we used all of the fossil dates previously employed by
Springer et al. (2012) that were applicable to our restricted taxon sample (Table 3). To
assess the sensitivity of the posterior node age distribution to the root age prior, we also
ran an additional analysis using an older root calibration derived from the age of *Pep-
rupithecus* (Bond et al., 2015). Each calibration point was assigned an offset exponen-
tial density such that the upper bound specified by Springer et al. (2012) corresponded
to the 95th percentile of the distribution. In contrast to the uniform densities utilized by
Springer et al. (2012), exponential distributions have the advantage of concentrating
most probability mass close to the lower bound. As single-parameter distributions, ex-
ponentials are also less arbitrary than lognormal priors commonly used in BEAST time
tree analyses, which can render the posterior overly sensitive to the choice of calibration density hyperparameters (Warnock et al., 2012).

The Markov chain Monte Carlo analysis was run for 400 million generations, sampling every 1000 generations and removing the initial 10% of samples as burnin. We assessed convergence of the chain using the effective sample sizes (ESS) reported for each parameter in Tracer 1.6.0 (Rambaut et al., 2013) by ensuring that all the ESS values exceeded 200. The posterior distribution of time trees was summarized into a maximum clade credibility tree using TreeAnnotator 1.8.3 (Rambaut and Drummond, 2015a).

3. Results

3.1. Quality control

We sequenced a total of 178 million read pairs (mean = 2,661,695.4) for all samples. An average of 3309 contigs per sample (min = 1162, max = 6170) was assembled from 67 individuals (Table 2). After alignment and trimming as described above, we got an average of 1882 unique contigs matching UCE loci from each sample. We produced a 75% complete data matrix containing 1843 alignments of UCE loci, which produced a concatenated matrix of 550,515 bp (average length: 298.70 bp per alignment) and a 95% complete data matrix containing 1390 alignments of UCE loci, which produced a concatenated matrix of 439,190 bp (average length: 315.96 bp per alignment).

3.2. Phylogenomic analyses
We recovered strong support in the tree topology from our RAxML (75% and 95%) analyses for reciprocal monophyly between the *Sapajus* and *Cebus* clades (Figure 2 and Supplementary Figure 2). Our analyses show strong molecular support for three of the morphological species within the genus *Sapajus*: *S. robustus*, *S. xanthosternos* and *S. nigritus*, all within the Atlantic Forest of Brazil. All other morphologically defined species within the genus (*S. flavius*, *S. libidinosus*, *S. apella*, *S. cay*, and *S. macrocephalus*) group together with high support in a widely distributed clade (from the Atlantic Forest to the Amazon), but there is no support for any subclades within this group in either the 75% or 95% taxa sets. Thus, the RAxML tree suggests four species of *Sapajus*: *S. robustus*, *S. xanthosternos* and *S. nigritus* from the Atlantic Forest of Brazil, and a widespread species that encompasses morphotypes *S. flavius*, *S. libidinosus*, *S. apella*, *S. cay*, and *S. macrocephalus*.

### 3.3. SNPs Analyses

After filtering out low quality SNPs, we retained a total of 19,583 SNPs across all samples. We then filtered for missing data and included only the SNPs that were parsimony-informative sites, generating a 75% complete matrix with a total of 11,462 informative high quality SNPs.

Similar to the RAxML analyses, our Maximum Likelihood and Bayesian trees using SNPs from the UCE data recover *S. xanthosternos* and *S. nigritus* as monophyletic clades, with the single *S. robustus* sample as the sister group to *S. xanthosternos* (Figure 3). However, within the widely distributed clade in the SNP trees, there are two distinct subclades. One subclade recovers monophyly of the species *Sapajus flavius* and also contains all *S. libidinosus* samples in a clade with *S. apella* specimens from Tucu-
ruí. The other subclade contains *S. cay*, *S. apella*, and *S. macrocephalus*; clusters within this subclade are geographically coherent but do not correspond to the current morphological taxonomy of the genus *Sapajus*. There is a clear division between Amazonian *Sapajus* north and south of the Amazon River, with some exceptions. Thus, our phylogenomic SNP data provides some support for six distinct species within *Sapajus*: *S. nigritus*, *S. robustus*, *S. xanthosternos*, *S. flavius*, *S. libidinosus* and a widespread Amazonian and southern grasslands species.

While the ExaBayes and SNPhylo had similar topologies, the two trees differed in the strength of their support for particular clades. For example, the SNPhylo tree resolved *S. nigritus* as the sister group to the widespread *Sapajus* clade (98), and *S. robustus* as sister to *S. xanthosternos* (96). SNPhylo also resolved *S. flavius* + (*S. libidinosus* + Tucuruí *S. apella*) clade as the sister group to *S. apella* + *S. macrocephalus* + *S. cay* (100). On the other hand, the ExaBayes tree provided higher support for the *S. flavius* + (*S. libidinosus* + Tucuruí *S. apella*) clade (0.99) and for the *S. cay* + Rondônia *S. apella* clade (0.99). Within the widespread Amazonian *S. apella* + *S. macrocephalus* + *S. cay* clade, ExaBayes recovered a northwestern *S. macrocephalus* subclade (0.99) and a northeastern *S. apella* subclade (0.99) that were strongly supported as sister to each other (0.97). ExaBayes also supported the sister relationship (0.95) between the *S. cay* + Rondônia *S. apella* subclade and a south-central Amazonian *S. macrocephalus* clade (Atalaia, Purus, Jirau, Canutama, Cujubim, Mamiraua, Japura, Jamari; 0.91). In contrast, the internal topology for the subclades of the *S. apella* + *S. macrocephalus* + *S. cay* clade was less well-supported in SNPhylo.

In the species tree recovered using SVDquartets analyses (Figure 4), we found strong support (100) in the tree topology for reciprocal monophyly between *Sapajus* and *Cebus*. The internal topology differed in some regards for *Sapajus* when compared to
our RAxML, ML and Bayesian trees using SNPs from the UCE data. As in other analyses, *Sapajus xanthosternos* and *S. robustus* were strongly supported as sister taxa (100), but here *S. nigritus* was weakly supported (77) as sister to *S. xanthosternos + S. robustus*. While in the other trees, *S. apella, S. macrocephalus, S. cay, S. flavius*, and *S. libidinosus* formed a subclade nested within the Atlantic forest robust capuchin clade and sister to *S. nigritus*, here this widespread group forms a second and well-supported (100) clade distinct from the Atlantic forest clade, with *S. flavius* supported (90) as sister to *S. libidinosus*, and Northern Amazonian and Southern Amazonian robust capuchins together forming a clade (100).

### 3.4. Divergence time analyses

While the BEAST run with the younger root calibration (based on *Aegyptopithecus* at 28.3 Ma) reached convergence after the specified number of generations (ESS values ≥ 250 for all parameters), the analysis employing the *Perupithecus*-derived 36 Ma minimum on the age of the root failed to converge, as indicated by an effective sample size of <200 for the age of the hominoid-cercopithecid divergence (node 6). An additional run of 100 million generations was performed and combined with the first chain using LogCombiner 1.8.3 (Rambaut and Drummond, 2015b); however, the resulting ESS values were lower than those obtained from the first run alone, suggesting that the two chains had sampled from different distributions. To overcome this problem, a third chain of 500 million generations was run in BEAST under the same settings. The ESS values for both the third run alone and the total combined run of 900,000 samples exceeded 200 for all parameters.
Regardless of the choice of root prior, the 95% highest posterior density (HPD) intervals of all calibrated nodes were well within the bounds used to construct the respective calibration densities (compare Tables 3 and 4). Use of the Perupithecus calibration shifted the marginal posterior distribution of the root age from the Late to Middle Eocene but exercised comparatively little influence on the estimated ages of shallower divergences (Table 4). The intrageneric divergences within both Cebus and Sapajus (Table 4; nodes 7 and 8) were consistently older and less precise (marked by wider 95% HPD intervals) when estimated under the Perupithecus-derived root age prior. The mean estimated split between robust and gracile capuchins (Table 4; node 9) shifted from 5.4 to 6.8 Ma when Perupithecus was used to calibrate the platyrrhine-catarrhine divergence, while the width of the corresponding 95% HPD interval remained unchanged.

4. Discussion

Together our analyses provide genetic support for six distinct species within Sapajus: five morphological species (strong support for S. robustus, S. xanthosternos, S. nigritus, and more equivocal support for S. libidinosus and S. flavius) and one morphologically diverse Amazonian + Central Grasslands species that contains two major clades separated by distributions in Northern versus Southern Amazonia. Recent mitochondrial studies provide some additional support for the species status of S. robustus, S. xanthosternos and S. nigritus though the exact relationships among species varies (Lima et al., 2017; Ruiz-Garcia et al., 2012). S. flavius is recovered as a monophyletic group with mitochondrial data, but is embedded within the widespread clade, or positioned as sister to the widespread clade (Lima et al., 2017), whereas the nuclear results
here place *S. flavius* and *S. libidinosus* as sister taxa. Both the mtDNA and the nuclear DNA topologies are discordant with Groves’ (2001) taxonomic hypothesis that *S. robustus* is a subspecies of *S. nigritus*, because *S. nigritus* and *S. robustus* do not group together as sister taxa within *Sapajus*. In the previous studies employing large numbers of concatenated loci to elucidate primate relationships (Perelman et al., 2011; Springer et al., 2012), *S. robustus* and *S. xanthosternos* are recovered as sister taxa to the exclusion of *S. apella*. In Springer et al. (2012) *S. apella* is recovered as sister to *S. libidinosus*, consistent with our present phylogeny.

While all *Sapajus libidinosus* samples with light yellow pelage phenotype found across *S. libidinosus* distribution in the relatively dry biomes of Caatinga and Cerrado cluster together in one clade, that clade also includes samples that present standard *S. apella* pelage at the border of the two species distributions, near Tucuruí, Pará, on the eastern side of the lake that was formed by the damming of the Tocantins River for a Hydroelectric Plant (Figure 5b). These same individuals with *S. apella* morphotypes from Tucuruí cluster genetically with all sampled individuals with *S. libidinosus* pelage from within *S. libidinosus* distribution when using mitochondrial markers (Lima et al., 2017). Tucuruí capuchins have darker pelage and live in tropical forest habitat, while nearby *S. libidinosus* are adapted to open Cerrado and Caatinga habitats, and have light-er pelage. *S. libidinosus* has also been shown to have cranial and post-cranial adaptations to increased ground use and encased fruit extraction (Wright et al., 2015). Morphometric data are not available for the Tucuruí specimens, to determine if their cranial and post-cranial characteristics cluster with *S. libidinosus* or *S. apella*. Their external coloration should also be studied in detail to compare with other *Sapajus* specimens.

The unexpected topology leaves us with various possibilities; it may be that the *S. libidinosus* lineage has expanded from the Cerrado biome to make inroads into the Amazon,
and that *S. libidinosus* populations living in forested areas evolve darker pelage, so that they converge in appearance with *S. apella*. This could be a result of genetic adaptation, or it could be that capuchins have a developmental response with coat color adjusting to habitat conditions. Either way, this suggests ecological forces may be driving coat color and morphological characteristics. A second possibility is that *S. apella* east of the Tocantins River became isolated from other robust Amazonian capuchins, and over time gave rise to the Caatinga and Cerrado populations of *S. libidinosus*. A third possibility is that *S. apella* and *S. libidinosus* have come into secondary contact at the borders of their distribution, and that despite significant gene flow, the two populations maintain their pelage characteristics. More morphological, genetic and ecological data will need to be collected in the Cerrado-Amazon transition zone in order to better understand relationships among capuchin populations here.

Note that *Sapajus libidinosus* + Tucuruí samples formed a clade with *S. flavius*. For this study, we sampled across western Caatinga and Cerrado for *S. libidinosus*, but we do not have samples here for eastern Caatingas where *S. libidinosus* is found close to *S. flavius* in northeastern Brazil (Figure 5b). More data from the Cerrado-Amazon transition zone and the Caatinga-Atlantic Forest transition zone could resolve if *S. flavius* and *S. libidinosus* are geographical variants of the same species, two distinct species, or are best lumped within the widespread *S. apella* group described below.

The molecular distinctiveness of the other morphological species currently assigned to *Sapajus* is not supported. Within the widespread *Sapajus* clade recovered in the SNP tree, there were strong indications for shared evolutionary history among morphotypes *S. cay*, *S. apella* and *S. macrocephalus*. There was no reciprocal monophyly between any of these morphologically defined species; instead, we observed geographic coherence for recovered lineages that did not correspond to current species hypotheses
for Amazonian and grassland *Sapajus*. The pattern is more concordant with an isolation-by-distance model across the entire ‘widespread *Sapajus*’ clade, and morphological variation driven by habitat type. The samples designated as *S. cay* formed a clade with geographically proximate *S. apella* samples, indicating either a high index of gene flow between the two, or that the two types actually are within the same species and have evolved phenotypic variation related to habitat type. Another possibility is that there is more than one taxon encompassed within the current taxonomic classification of *S. cay*. Some studies have already indicated that *S. cay* from the Brazilian Pantanal and from Paraguay may not be a monophyletic group (Casado et al., 2010; Lima et al., 2017), but in this study, we do not have samples from both areas. *S. macrocephalus* as defined by Rylands et al. (2013) is also paraphyletic in our study, with two distinct lineages, one found north of the Solimões and Japurá rivers and south of the Rio Negro (recovered as sister to *S. apella* north of the Amazon River: Figure 5c) and the other in south-central Amazon south of the Amazon and Solimões rivers (recovered as the sister group to south Amazonian *S. apella* and *S. cay*: Figure 5d). Note that our study extends the *S. macrocephalus* morphotype east of the Madeira River, into the Brazilian state of Rondônia. *S. apella* appears in multiple places across the topology of both the RAxML and SNP trees, divided among various lineages which do not form a monophyletic group, but instead are interspersed with clades of *S. libidinosus, S. macrocephalus*, and *S. cay*.

It is important to note that the geographic boundaries and taxonomic affinities for *S. apella, S. cay, S. libidinosus* and *S. macrocephalus* are disputed by the two predominate morphological authorities (Groves 2001, 2005; Silva-Júnior, 2001, 2002). For example, Groves (2001) considers *S. cay* as two distinct subspecies of *S. libidinosus* (called *Cebus libidinosus paraguayanus* and *Cebus libidinosus pallidus*), and *S. macrocephalus* as a subspecies of *S. apella* (*Cebus apella macrocephalus*). Neither mitocho-
drial (Lynch Alfaro et al., 2012a; Lima et al., 2017) nor nuclear data from the present study recovered reciprocal monophyly for *S. cay*, *S. apella*, or *S. macrocephalus*. Combining genetic and morphological data, we interpret that these morphotypes are not clearly defined and discrete species, but instead form one morphologically diverse, recently evolved pan-Amazonian plus grassland clade of robust capuchins. If we collapse these three taxa into one species, the taxonomic name would be *Sapajus apella*, which has priority over the other names because it was given first by Linnaeus in 1758. We do not recommend the use of subspecies within this cosmopolitan species, because molecular and morphological subdivisions are discordant with one another suggesting a high index of morphological plasticity and convergence within the species.

We also note that while the two major *Sapajus* clades within the Amazon are divided roughly by the Amazon River (see Figures 5c and d), that some samples within the Northern clade were from individuals south of the Amazon, and vice versa. In most cases these were individuals that were very close geographically to the Amazon River itself, and may be the result of human-mediated transport across the rivers in recent or modern times. It is also possible that capuchins cross the Amazon at low frequency in areas where there are many seasonal islands. Squirrel monkeys show a similar pattern in the eastern Amazon basin, where the Amazon River forms the border for the distributions of *Saimiri sciureus* and *S. collinsi*, with some cases of limited dispersal to the opposite bank of the Amazon River for each species in the Juruti and Faro regions of Pará State, Brazil (Mercês et al., 2015).

The time trees based generated from our BEAST analysis placed the mean estimated divergence time for gracile and robust capuchins at 5.4 Ma using the *Aegyptopithecus* tree root prior, or 6.8 Ma, using the *Perupithecus* tree root prior. These compare to previous mean estimates for divergence between *Cebus* and *Sapajus* at 5.8 Ma,
using mitochondrial data (Lima et al., 2017), at 6 Ma using a BEAST analysis for nuclear genes (Perelman et al. 2011), and 6.6 Ma for the MCMC tree in PAML utilizing autocorrelated rates and soft-bounded constraints for a supermatrix of both nuclear and mitochondrial genes (Springer et al., 2012). In other words, all analyses converge on a late Miocene divergence time for robust and gracile capuchin monkeys. This timing is consistent with the formation of the savanna-like Cerrado leading to vicariance of a widespread capuchin ancestor previously spanning the Amazon to the Atlantic Forest (Lynch Alfaro et al., 2015; Lima et al., 2017).

5. Conclusions

Our phylogenomic data provided strong support for *Cebus* and *Sapajus* as two reciprocally monophyletic clades. This is concordant with morphological evaluations of distinctiveness between robust and gracile capuchins (Elliott, 1913; Hershkovitz, 1949; Groves, 2001, 2005; Silva-Júnior, 2001, 2002; Lynch Alfaro et al., 2012b), and mitochondrial and Alu element data that also point to this split (Lynch Alfaro et al., 2012a; Lima et al., 2017; Martins Jr. et al., 2015; Viana et al., 2015). We recovered a late Miocene split for robust and gracile capuchins, concordant with previous molecular studies. The timetree mean estimate for the initial diversification of robust capuchins was at 2.1 Ma (using the *Aegyptopithecus* root calibration) or 2.6 Ma (using the *Perupithecus* root calibration); this early Pleistocene diversification is also consistent with previous studies using mitochondrial data (Lynch Alfaro et al. 2012a; Lima et al., 2017).

In general, our phylogenies based on ultraconserved elements were congruent with mitochondrial phylogenies for robust capuchins (Lynch Alfaro et al., 2012; Lima et al., 2017), although the placement of *S. robustus* as sister to *S. xanhosternos* was
unique to the nuclear phylogenomic data, as was the recovery of a sister relationship
between *S. flavius* and *S. libidinosus*. Our UCE tree distinguished only four *Sapajus*
species, but the ExaBayes SNP tree provided more support for six robust capuchin spe-
cies, *S. xanthosternos*, *S. robustus*, *S. nigritus*, *S. flavius*, *S. libidinosus*, and *S. apella*
(which subsumes *S. cay* and *S. macrocephalus*), although *S. apella* morphotypes from
Tucuruí were found within the *S. libidinosus* clade. The major division for Amazonian
capuchins according to molecular data is a North-South division (both in the present
work and from mitochondrial data in Lima et al., 2017), whereas the morphological
division of *S. macrocephalus* and *S. apella* has been described as more of an East-West
division, with the Madeira and Negro rivers as the suggested dividing line (Groves,
2001, 2005; Silva-Júnior, 2001, 2002). As morphological and molecular subdivisions of
the Amazonian group are discordant, we recommend lumping all Amazonian plus
southern grassland robust capuchin taxa as *S. apella* without subspecies. However, this
does not discount the importance of populational differences in behavior, morphology
and ecology in *S. apella* across the Amazon and southern grasslands; these populational
differences may serve as a model for understanding the rapid evolution of populational
differences across diverse habitats in other highly polymorphic species, such as humans.

The taxonomic relationship of *S. nigritus* to other capuchins is not well support-
ed, with the species tree placing it as the sister group to *S. xanthosternos* + *S. robustus*,
but the gene trees placing it as the sister group to the widespread clade of robust capu-
chins (*S. flavius*, *S. libidinosus*, *S. apella* as above). In contrast, mitochondrial phyloge-
netic reconstructions have placed *S. nigritus* as the sister to all other *Sapajus* (Lima et
al., 2017). More work needs to be done delineating the relationship and geographical
boundaries between *S. nigritus nigritus* from Minas Gerais to Sao Paulo, Brazil and *S.
*n. cucullatus* from southern Brazil and Argentina, and their relationships to other capu-
chins. Future work is also needed to determine the relationship of Critically Endangered
*S. apella margaritae* endemic to Margarita Island, Venezuela to the other Amazonian
and Guianan robust capuchins.

**Acknowledgments**

Special thanks to Stephen D. Nash/IUCN SSC Primate Specialist Group to Illustrations
copyright 2013. 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). Some of the
field expeditions were funded by CNPq/FAPEAM SISBIOTA Program (563348/2010-
0) to J.P.B. UCE data was generated with support from NSF-FAPESP (grant 1241066 -
and its environment: an integrated approach) to A.A. Divergence time analyses were
performed with support from the Whitcome Research Fellowship to D.C.

**References**

doi:10.1093/molbev/msu236

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: A flexible trimmer for Illu-


Springer, M.S., Meredith, R.W., Gatesy, J., Emerling, C.A., Park, J., Rabosky, D.L., Stadler, T., Steiner, C., Ryder, O.A., Janečka, J.E., Fisher, C.A., Murphy, W.J.,


### Table 1: Taxonomies of robust capuchins.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Cebus apella</em></td>
<td><em>Cebus apella</em></td>
<td><em>Cebus apella</em></td>
<td><em>Cebus apella</em></td>
<td><em>Cebus apella</em></td>
</tr>
<tr>
<td><em>Cebus fatuellus</em></td>
<td></td>
<td><em>C. a. apella</em></td>
<td><em>C. a. apella</em></td>
<td><em>C. a. apella</em></td>
</tr>
<tr>
<td><em>C. f. fatuellus</em></td>
<td></td>
<td><em>C. a. margaritae</em></td>
<td><em>C. a. margaritae</em></td>
<td><em>C. a. fatuellus</em></td>
</tr>
<tr>
<td><em>C. f. peruanus</em></td>
<td></td>
<td><em>C. a. macrocephalus</em></td>
<td><em>C. a. fatuellus</em></td>
<td><em>C. a. macrocephalus</em></td>
</tr>
<tr>
<td><em>Cebus macrocephalus</em></td>
<td></td>
<td><em>C. a. libidinosus</em></td>
<td><em>C. a. peruanus</em></td>
<td><em>C. a. peruanus</em></td>
</tr>
<tr>
<td><em>Cebus libidinosus</em></td>
<td></td>
<td><em>C. a. paraguayanus</em></td>
<td><em>C. a. tocantinus</em></td>
<td><em>C. a. tocantinus</em></td>
</tr>
<tr>
<td><em>Cebus azarae</em></td>
<td><em>C. a. pallidus</em></td>
<td><em>C. a. macrocephalus</em></td>
<td><em>C. a. margaritae</em></td>
<td><em>C. a. margaritae</em></td>
</tr>
<tr>
<td><em>C. a. azarae</em></td>
<td><em>C. a. xanthosternos</em></td>
<td><em>C. a. libidinosus</em></td>
<td>Cebus libidinosus</td>
<td></td>
</tr>
<tr>
<td><em>C. a. pallidus</em></td>
<td></td>
<td><em>C. a. versutus</em></td>
<td><em>C. a. cay</em></td>
<td>C. l. libidinosus</td>
</tr>
<tr>
<td><em>Cebus frontatus</em></td>
<td></td>
<td><em>C. a. nigritus</em></td>
<td><em>C. a. pallidus</em></td>
<td>C. l. pallidus</td>
</tr>
<tr>
<td><em>Cebus variegatus</em></td>
<td></td>
<td><em>C. a. vellerosus</em></td>
<td><em>C. a. frontatus</em></td>
<td>C. l. paraguayanus</td>
</tr>
<tr>
<td><em>Cebus versuta</em></td>
<td></td>
<td><em>C. a. robustus</em></td>
<td><em>C. a. xanthosternos</em></td>
<td>C. l. juruanus</td>
</tr>
<tr>
<td><em>Cebus cirrifer</em></td>
<td></td>
<td><em>C. a. nigritus</em></td>
<td>Cebus nigritus</td>
<td></td>
</tr>
<tr>
<td><em>Cebus crassiceps</em></td>
<td></td>
<td><em>C. a. robustus</em></td>
<td>C. n. nigritus</td>
<td></td>
</tr>
<tr>
<td><em>Cebus caliginosus</em></td>
<td></td>
<td><em>C. a. magnus</em></td>
<td>C. n. robustus</td>
<td></td>
</tr>
<tr>
<td><em>Cebus vellerosus</em></td>
<td></td>
<td><em>C. a. juruanus</em></td>
<td>C. n. cucullatus</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>C. a. maranonis</em></td>
<td>Cebus xanthosternos</td>
<td></td>
</tr>
</tbody>
</table>
Table 2: List of samples, locality data and resulting for UCE data.

<table>
<thead>
<tr>
<th>Code</th>
<th>Species</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Trimmed reads</th>
<th>Contigs Assembled</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><em>S. xanthosternos</em></td>
<td>-15.17</td>
<td>-39.07</td>
<td>2681597</td>
<td>3274</td>
</tr>
<tr>
<td>2</td>
<td><em>S. xanthosternos</em></td>
<td>-15.41</td>
<td>-39.5</td>
<td>2843593</td>
<td>3661</td>
</tr>
<tr>
<td>3A</td>
<td><em>S. xanthosternos</em></td>
<td>-14.79</td>
<td>-39.05</td>
<td>3196673</td>
<td>3802</td>
</tr>
<tr>
<td>3B</td>
<td><em>S. xanthosternos</em></td>
<td>-14.79</td>
<td>-39.05</td>
<td>3521726</td>
<td>4275</td>
</tr>
<tr>
<td>4</td>
<td><em>S. robustus</em></td>
<td>-19.95</td>
<td>-43.85</td>
<td>4538948</td>
<td>5198</td>
</tr>
<tr>
<td>5</td>
<td><em>S. nigritus</em></td>
<td>-23.86</td>
<td>-46.14</td>
<td>2762021</td>
<td>3471</td>
</tr>
<tr>
<td>6</td>
<td><em>S. nigritus</em></td>
<td>-23</td>
<td>-49.32</td>
<td>946881</td>
<td>1937</td>
</tr>
<tr>
<td>7</td>
<td><em>S. flavius</em></td>
<td>-6.56</td>
<td>-35.13</td>
<td>2713906</td>
<td>3096</td>
</tr>
<tr>
<td>8</td>
<td><em>S. flavius</em></td>
<td>-7.01</td>
<td>-34.96</td>
<td>4787966</td>
<td>5150</td>
</tr>
<tr>
<td>9</td>
<td><em>S. flavius</em></td>
<td>-7.02</td>
<td>-35.09</td>
<td>2877922</td>
<td>3601</td>
</tr>
<tr>
<td>10</td>
<td><em>S. libidinosus</em></td>
<td>-2.77</td>
<td>-41.81</td>
<td>2764451</td>
<td>3430</td>
</tr>
<tr>
<td>11</td>
<td><em>S. libidinosus</em></td>
<td>-2.8</td>
<td>-41.87</td>
<td>4348317</td>
<td>5094</td>
</tr>
<tr>
<td>12</td>
<td><em>S. libidinosus</em></td>
<td>-5.09</td>
<td>-42.43</td>
<td>2612178</td>
<td>3208</td>
</tr>
<tr>
<td>13</td>
<td><em>S. libidinosus</em></td>
<td>-7.93</td>
<td>-44.2</td>
<td>3068523</td>
<td>3551</td>
</tr>
<tr>
<td>14</td>
<td><em>S. libidinosus</em></td>
<td>-5.28</td>
<td>-48.3</td>
<td>3303530</td>
<td>3885</td>
</tr>
<tr>
<td>15</td>
<td><em>S. libidinosus</em></td>
<td>-14.14</td>
<td>-48.17</td>
<td>3381894</td>
<td>3603</td>
</tr>
<tr>
<td>16</td>
<td><em>S. libidinosus</em></td>
<td>-16.6</td>
<td>-49.26</td>
<td>3301692</td>
<td>3884</td>
</tr>
<tr>
<td>17A</td>
<td><em>S. apella</em></td>
<td>-3.83</td>
<td>-49.64</td>
<td>3541159</td>
<td>3793</td>
</tr>
<tr>
<td>17B</td>
<td><em>S. apella</em></td>
<td>-3.83</td>
<td>-49.64</td>
<td>2980533</td>
<td>3534</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td><strong>S. apella</strong></td>
<td>-6.15</td>
<td>-49.56</td>
<td>1908769</td>
<td>2828</td>
</tr>
<tr>
<td>19</td>
<td><strong>S. apella</strong></td>
<td>-3.36</td>
<td>-51.74</td>
<td>3391742</td>
<td>3723</td>
</tr>
<tr>
<td>20</td>
<td><strong>S. apella</strong></td>
<td>-2.61</td>
<td>-51.54</td>
<td>5485708</td>
<td>6170</td>
</tr>
<tr>
<td>21</td>
<td><strong>S. apella</strong></td>
<td>-0.58</td>
<td>-52.33</td>
<td>1311929</td>
<td>2137</td>
</tr>
<tr>
<td>22</td>
<td><strong>S. apella</strong></td>
<td>3.22</td>
<td>-52.03</td>
<td>1757726</td>
<td>2338</td>
</tr>
<tr>
<td>23</td>
<td><strong>S. apella</strong></td>
<td>0.83</td>
<td>-53.93</td>
<td>2781762</td>
<td>2805</td>
</tr>
<tr>
<td>24</td>
<td><strong>S. apella</strong></td>
<td>1.29</td>
<td>-58.7</td>
<td>2130450</td>
<td>2604</td>
</tr>
<tr>
<td>25</td>
<td><strong>S. apella</strong></td>
<td>-1.49</td>
<td>-56.8</td>
<td>1572934</td>
<td>2413</td>
</tr>
<tr>
<td>26</td>
<td><strong>S. apella</strong></td>
<td>-2.47</td>
<td>-58.4</td>
<td>3571090</td>
<td>3780</td>
</tr>
<tr>
<td>27</td>
<td><strong>S. apella</strong></td>
<td>-2.6</td>
<td>-56.18</td>
<td>2394355</td>
<td>3227</td>
</tr>
<tr>
<td>28</td>
<td><strong>S. apella</strong></td>
<td>-3.18</td>
<td>-55.8</td>
<td>1890413</td>
<td>2709</td>
</tr>
<tr>
<td>29</td>
<td><strong>S. apella</strong></td>
<td>-3.88</td>
<td>-56.78</td>
<td>1276241</td>
<td>2039</td>
</tr>
<tr>
<td>30</td>
<td><strong>S. apella</strong></td>
<td>-4.71</td>
<td>-56.44</td>
<td>1746336</td>
<td>2515</td>
</tr>
<tr>
<td>31</td>
<td><strong>S. apella</strong></td>
<td>-10</td>
<td>-56.04</td>
<td>1791793</td>
<td>2450</td>
</tr>
<tr>
<td>32</td>
<td><strong>S. apella</strong></td>
<td>-9.2</td>
<td>-59.06</td>
<td>2103015</td>
<td>2895</td>
</tr>
<tr>
<td>33</td>
<td><strong>S. apella</strong></td>
<td>-12.03</td>
<td>-60.67</td>
<td>2339872</td>
<td>3027</td>
</tr>
<tr>
<td>34</td>
<td><strong>S. apella</strong></td>
<td>-12.56</td>
<td>-63.44</td>
<td>3883141</td>
<td>4558</td>
</tr>
<tr>
<td>35</td>
<td><strong>S. cay</strong></td>
<td>-16.06</td>
<td>-57.72</td>
<td>1624662</td>
<td>2588</td>
</tr>
<tr>
<td>36</td>
<td><strong>S. cay</strong></td>
<td>-13.52</td>
<td>-60.43</td>
<td>2361492</td>
<td>2991</td>
</tr>
<tr>
<td>37</td>
<td><strong>S. macrocephalus</strong></td>
<td>-12.45</td>
<td>-62.92</td>
<td>2986344</td>
<td>3335</td>
</tr>
<tr>
<td>38</td>
<td><strong>S. macrocephalus</strong></td>
<td>-8.67</td>
<td>-62.37</td>
<td>2962283</td>
<td>3477</td>
</tr>
<tr>
<td>39</td>
<td><strong>S. macrocephalus</strong></td>
<td>-9.1</td>
<td>-62.88</td>
<td>2222218</td>
<td>2882</td>
</tr>
<tr>
<td>40</td>
<td><strong>S. macrocephalus</strong></td>
<td>-8.89</td>
<td>-63.24</td>
<td>3054313</td>
<td>3411</td>
</tr>
<tr>
<td>41</td>
<td><strong>S. macrocephalus</strong></td>
<td>-8.8</td>
<td>-63.95</td>
<td>1459387</td>
<td>2148</td>
</tr>
<tr>
<td>42</td>
<td><strong>S. macrocephalus</strong></td>
<td>-8.19</td>
<td>-64.02</td>
<td>2196025</td>
<td>2741</td>
</tr>
<tr>
<td>43</td>
<td><strong>S. macrocephalus</strong></td>
<td>-5.69</td>
<td>-63.24</td>
<td>3840307</td>
<td>4395</td>
</tr>
<tr>
<td>44A</td>
<td><strong>S. macrocephalus</strong></td>
<td>-4.99</td>
<td>-62.96</td>
<td>3199632</td>
<td>3780</td>
</tr>
<tr>
<td>44B</td>
<td><strong>S. macrocephalus</strong></td>
<td>-4.99</td>
<td>-62.96</td>
<td>1163783</td>
<td>2218</td>
</tr>
<tr>
<td>45</td>
<td><strong>S. macrocephalus</strong></td>
<td>-4.75</td>
<td>-61.28</td>
<td>2351064</td>
<td>3072</td>
</tr>
<tr>
<td>Calibrated node</td>
<td>Divergence</td>
<td>Fossil</td>
<td>Reference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>------------</td>
<td>-------------------------</td>
<td>--------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Hominina / Pan</td>
<td>Ardipithecus kadabba</td>
<td>Springer et al., 2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Hominini / Gorilla</td>
<td>Sivapithecus sp.</td>
<td>Springer et al., 2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Papio / Macaca</td>
<td>Macaca libyca</td>
<td>Springer et al., 2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Hominoidea / Cercopithecida</td>
<td>Afropithecus turkanensis</td>
<td>Springer et al., 2012</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4: Summary of the posterior distribution of divergence times (in Ma) estimated using BEAST (see Supplementary Figure 1 for node labels).

<table>
<thead>
<tr>
<th>Node</th>
<th>Springer et al. root calibration</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>Mean</td>
</tr>
<tr>
<td>1</td>
<td>5.6</td>
<td>5.8</td>
</tr>
<tr>
<td>2</td>
<td>7.9</td>
<td>8.2</td>
</tr>
<tr>
<td>3</td>
<td>14.2</td>
<td>14.5</td>
</tr>
<tr>
<td>4</td>
<td>17.4</td>
<td>17.7</td>
</tr>
<tr>
<td>5</td>
<td>6.5</td>
<td>6.9</td>
</tr>
<tr>
<td>6</td>
<td>23.6</td>
<td>24.4</td>
</tr>
<tr>
<td>7</td>
<td>1.7</td>
<td>1.8</td>
</tr>
<tr>
<td>8</td>
<td>1.9</td>
<td>2.1</td>
</tr>
<tr>
<td>9</td>
<td>5.2</td>
<td>5.4</td>
</tr>
<tr>
<td>10</td>
<td>17.0</td>
<td>17.4</td>
</tr>
</tbody>
</table>

Callitrichidae / Cebidae (sensu Rylands et al., 2012)
Patasola magdalena; Lagonimico conclucatus Kay, 2015 (minimum); Springer et al., 2012 (maximum)

Catarrhini / Platyrhini
Aegyptopithecus zeuxis / Perupithecus uacayaliensis Springer et al., 2012 / Bond et al., 2015
<table>
<thead>
<tr>
<th></th>
<th>11</th>
<th>34.1</th>
<th>35.4</th>
<th>28.3–46.7</th>
<th>41.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>782</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>783</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>784</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>785</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>786</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure Captions

**Graphical Abstract.** (a) Maximum likelihood and (b) Bayesian inference for robust capuchin phylogeny based on SNP data.

**Figure 1.** Map showing the sampled localities for *Sapajus*

**Figure 2.** Maximum likelihood (RAxML) 75% phylogeny for UCE data.

**Figure 3.** (a) Maximum likelihood and (b) Bayesian inference for robust capuchin phylogeny based on SNP data.

**Figure 4.** Species tree for robust capuchins using SNP quartets.

**Figure 5.** (a) Map with minimum convex polygons to show geographic distribution of major subclades within the widespread *Sapajus* clade, (b) Minimum convex polygon for range distribution for *S. flavius* and *S. libidinosus* clades within the ExaBayes phylogeny, (c) Minimum convex polygon for range distribution for the Northern Amazonian *Sapajus* clade within the ExaBayes phylogeny and (d).Minimum convex polygon for range distribution for the Southern Amazonian *Sapajus* clade within the ExaBayes phylogeny. Larger map depicts subclades of south central Amazonian *S. macrocephalus* and southern Amazonian + grasslands *S. apella + cay*.

**Supplementary Figure 1.** Topological constraint used for divergence time estimation in BEAST.

**Supplementary Figure 2.** Maximum likelihood (RaxML) 95% phylogeny for UCE data.
Graphical Abstract.
Figure 1.
Figure 2.
Figure 4.
Supplementary Figure 1.

[Diagram showing a phylogenetic tree with species labeled as follows:
- Nomascus leucogenys
- Pongo abelii
- Gorilla gorilla
- Pan troglodytes
- Homo sapiens
- Papio hamadryas
- Macaca mulatta
- Callithrix jacchus
- Cebus olivaceus
- Cebus capucinus
- Sapajus xanthosternos
- Sapajus apella]