Ancient DNA of the Pygmy Marmoset Type Specimen *Cebuella pygmaea* (Spix, 1823) Resolves a Taxonomic Conundrum

J. P. Boubli

M. C. Janiak

L. M. Porter

Stella de la Torre

*Universidad San Francisco de Quito*

L. Cortés-Ortiz

See next page for additional authors

Follow this and additional works at: https://digitalcommons.trinity.edu/bio_faculty

Part of the Biology Commons

Repository Citation


This Article is brought to you for free and open access by the Biology Department at Digital Commons @ Trinity. It has been accepted for inclusion in Biology Faculty Research by an authorized administrator of Digital Commons @ Trinity. For more information, please contact jcostanz@trinity.edu.
Authors
Ancient DNA of the pygmy marmoset type specimen *Cebuella pygmaea* (Spix, 1823) resolves a taxonomic conundrum

Jean P. Boublí1,2,*, Mareike C. Janiak1, Leila M. Porter3, Stella de la Torre4, Liliana Cortés-Ortiz5, Maria N. F. da Silva2, Anthony B. Rylands6, Stephen Nash7, Fabício Bertuol8, Hazel Byrne9, Felipe E. Silva9, Fabio Rohe9, Dorien de Vries1, Robin M. D. Beck1, Irene Ruiz-Gartzia10, Lukas F. K. Kuderna10, Tomas Marques-Bonet10, Tomas Hrbek7,11, Izeni P. Farias7, Anneke H. van Heteren12,13,14, Christian Roos15

1 School of Science, Engineering & Environment, University of Salford, Salford M5 4WT, UK
2 Instituto Nacional de Pesquisas da Amazônia, Manaus, Amazonas 69060-001, Brazil
3 Department of Anthropology, Northern Illinois University, DeKalb, IL 60115-2828, USA
4 College of Biological and Environmental Sciences, Universidad San Francisco de Quito, Quito 170901, Ecuador
5 Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA
6 Rewild, Austin, TX 78767, USA
7 Laboratório de Evolução e Genética Animal, Universidade Federal do Amazonas, Manaus, Amazonas 69080-900, Brazil
8 Department of Anthropology, University of Utah, Salt Lake City, UT 84112, USA
9 Research Group on Primate Biology and Conservation, Mamirauá Institute for Sustainable Development, Tefé, Amazonas 69553-225, Brazil
10 Experimental and Health Sciences Department (DCEXS), Institut de Biologia Evolutiva, Universitat Pompeu Fabra-CSIC, Barcelona 08002, Spain
11 Department of Biology, Trinity University, San Antonio, TX 78212, USA
12 Bavarian State Collection of Zoology, Zoologische Staatssammlung München, Staatliche Naturkundliche Sammlungen Bayerns, 81247 Munich, Germany
13 GeoBio-Center, Ludwig-Maximilians-Universität München, 80333 Munich, Germany
14 Department Biologie II, Ludwig-Maximilians-Universität München, 82152 Munich, Germany
15 Gene Bank of Primates and Primate Genetics Laboratory, German Primate Center, Leibniz Institute for Primate Research, 37077 Göttingen, Germany

ABSTRACT

The pygmy marmoset, the smallest of the anthropoid primates, has a broad distribution in Western Amazonia. Recent studies using molecular and morphological data have identified two distinct species separated by the Napo and Solimões-Amazonas rivers. However, reconciling this new biological evidence with current taxonomy, i.e., two subspecies, *Cebuella pygmaea* pygmaea (Spix, 1823) and *Cebuella pygmaea* niveiventris (Lönnberg, 1940), was problematic given the uncertainty as to...
whether Spix’s pygmy marmoset (Cebuella pygmaea pygmaea) was collected north or south of the Napo and Solimões-Amazonas rivers, making it unclear to which of the two newly revealed species the name pygmaea would apply. Here, we present the first molecular data from Spix’s type specimen of Cebuella pygmaea, as well as novel mitochondrial genomes from modern pygmy marmosets sampled near the type locality (Tabatinga) on both sides of the river. With these data, we can confirm the correct names of the two species identified, i.e., C. pygmaea for animals north of the Napo and Solimões-Amazonas rivers and C. niveiventris for animals south of these two rivers. Phylogenetic analyses of the novel genetic data placed into the context of cytchrome b gene sequences from across the range of pygmy marmosets further led us to re-evaluate the geographical distribution for the two Cebuella species. We dated the split of these two species to 2.54 million years ago. We discuss additional, more recent, subdivisions within each lineage, as well as potential contact zones between the two species in the headwaters of these rivers.

**Keywords:** Historic DNA; DNA taxonomy; Pygmy marmoset; Cebuella pygmaea; C. niveiventris; Amazon; Type specimen

**INTRODUCTION**

Pygmy marmosets (Cebuella) are the smallest of all living anthropoid primates. They have a wide geographic distribution across the upper Amazon region in northwestern Bolivia, Brazil, Peru, Ecuador, and southern Colombia. A recent molecular genetic study (Boubli et al., 2018) tested the validity of separating this taxon into two subspecies, namely, Cebuella pygmaea pygmaea (Spix, 1823) and Cebuella pygmaea niveiventris (Lönnberg, 1940), which have been recognized in various studies, including da Cruz Lima (1945), Napier (1976), van Roosmalen & van Roosmalen (1997), and Groves (2001, 2005). Using the mitochondrial cytchrome b (cvt b) gene and ddRAD nuclear genome sequences from geographically representative samples from both sides of the Solimões River (upper Amazon), Boubli et al. (2018) recovered two highly supported clades that diverged 2.25 million years ago (Ma), leading the authors to suggest the existence of two species of Cebuella; one comprising pygmy marmosets sampled on the northern side (left bank) of the Solimões and the other comprising samples collected on the southern side (right bank) (see also Porter et al., 2021). Reconciling the results of these molecular and morphological analyses with the current Cebuella taxonomy was confusing, however, due to the uncertainty of the provenance of Spix’s holotype of Iacchus pygmaeus (Boubli et al., 2018; Rylands et al., 2009).

Johann Baptist von Spix described pygmaea based on a pet pygmy marmoset given to him, probably by a local Tikuna Indian, during his visit to the upper Solimões River on the Brazilian-Colombian border (von Spix, 1823; von Spix & von Martius, 1824). He assigned the type locality as near Tabatinga, then a small village in Brazil, near the Colombian border on the left bank of the Solimões. In 1940, Lönnberg described niveiventris as a distinct pygmy marmoset from the Lago Ipixuna, approximately midway between the Tefé and Purus rivers, a little west of Coari on the right (south) bank of the Solimões (Lönnberg, 1940). To describe this new taxon, he compared his type series with specimens collected by the Olalla brothers near the Brazilian village of Eirunepé (then known as João Pessoa) on the right bank of the Juruá River (see Boubli et al., 2018). For him, the Juruá pygmy marmosets were typical pygmaea, even though Eirunepé is more than 200 km south of Tabatinga and on the opposite bank of the Solimões. In his publication, he affirmed that the Spix’s type came from the right bank of the Solimões, and thus the opposite bank of Tabatinga (Lönnberg, 1940).

Doubt concerning the origin of Spix’s type specimen (left or right bank of the Solimões at Tabatinga) complicated the attribution of the proper names of the two clades identified by Boubli et al. (2018). In that study, the pygmy marmosets from the south (right) bank of the Solimões formed a clade to the exclusion of those from the north bank and northwards to the right bank of the Japará River. If Spix’s type came from the right bank of the Solimões — specifically the mouth of the Javari River, as stated by Lönnberg (1940) — niveiventris could be a junior synonym of pygmaea. In this case, animals from the left bank (north) bank of the Solimões would have no available name and thus constitute a new taxon. On the other hand, if the pygmaea type originally came from near Tabatinga on the left bank of the Solimões, then niveiventris would be available for pygmy marmosets on the right bank of the Solimões and pygmaea would be available for pygmy marmosets on the left bank, as suggested by van Roosmalen & van Roosmalen (1997).

In the absence of novel historical evidence, obtaining genetic material from Spix’s type specimen, as well as material from the supposed type locality Tabatinga, presented the best means to resolve this taxonomic deadlock. Since its first use in museum collections in the 1980s, the retrieval and amplification of ancient DNA (Higuchi et al., 1984; Pääbo, 1985) has provided new insights into the evolution, biology, and taxonomy of organisms across the tree of life (Burrell et al., 2015). Perhaps the most famous example, the recovery of the first fragments (Green et al., 2006; Noonan et al., 2006) and finally complete genomes (Prüfer et al., 2014) of Neanderthals (Homo neanderthalensis), provided evidence of interbreeding between this species and our own. Ancient DNA can also be a way to improve museum collections, by assigning likely origins and species designations to museum specimens of uncertain provenance (Shepherd et al., 2013) and ascertaining the accuracy of specimen data (Verry et al., 2019). Additionally, many taxonomic puzzles have been solved by analysis of DNA from museum specimens, including equids (Orlando et al., 2009), monk seals (Scheel et al., 2014), leaf monkeys (Roos et al., 2020), and most recently the reclassification of the dire wolf from Canis dirus to Aenocyon dirus, which had been largely ignored since it was first proposed in 1918 (Perri et al., 2021).
Ancient DNA from Spix’s type specimen of *pygmaea* could help determine its relationship with the two clades of *Cebuella* recovered previously (Boubli et al., 2018) and establish the provenance of this specimen (i.e., right or left bank of the Solimões), allowing valid names to be attributed to the two species of pygmy marmosets.

In this study, we present the first molecular data from Spix’s type specimen of *Cebuella pygmaea* (housed in the Bavarian State Collection of Zoology), as well as novel mitochondrial genomes (mitogenomes) from modern pygmy marmosets sampled in the purported type locality of Spix’s specimen, near the town of Tabatinga. By placing these data in the context of available cyt *b* mitochondrial barcode gene sequences from pygmy marmosets across their distribution, we provide the first range-wide phylogenetic analysis of this taxon and confirm the correct naming of the two species of *Cebuella*, thereby resolving the present taxonomic uncertainty. We also re-evaluate the current geographical distribution hypothesis for the two *Cebuella* species (Mittermeier et al., 2013) based on our results and those of Porter et al. (2021).

**MATERIALS AND METHODS**

**Holotype sampling and sequencing**

Dried tissue skin scrapings were obtained from the stuffed type specimen of *pygmaea* from the Bavarian State Collection of Zoology, Munich, Germany. The skin scrapings were sampled from the inner side of the damaged skin together with pieces of the damaged skin. To minimize the risk of environmental (including human) and cross-sample contamination, DNA extraction and library preparation were performed in an ancient DNA laboratory, in which all standards for such laboratories were implemented (UV light decontamination, positive air pressure, separate sterile working areas, protective clothing, and negative controls during DNA extraction and library preparation for sequencing).

DNA from the holotype was extracted with a column-based method that recovers small DNA fragments (Dabney et al., 2013; Rohland et al., 2004). The DNA concentration was measured with a Qubit 4.0 fluorometer (ThermoFisher Scientific, USA), and DNA quality and degradation status were checked on a Bioanalyzer 2100 (Agilent Technologies, USA). Approximately, 50 ng of genomic DNA was subjected to shotgun library preparation with a NEBNext Ultra II DNA Library Prep Kit (E7103, New England Biolabs, USA) following the standard protocols of the supplier. However, DNA fragmentation prior to library preparation was omitted as the DNA was already largely degraded (50–150 bp). A sequencing library was also prepared from the DNA extraction negative control. After end repair, adapter ligation, and ligation cleanup (without size selection) using the kit’s purification beads, the libraries were indexed with multiplex oligos and then cleaned again with the purification beads. Library concentration and size distribution were measured with a Qubit fluorometer and Bioanalyzer, respectively, and molarity was quantified via quantitative polymerase chain reaction (qPCR) using a NEBNext Library Quant Kit (New England Biolabs, USA). Sequencing was conducted on an Illumina HiSeq 4000 (100 bp single-end reads) at the NGS Integrative Genomics (NIG) core unit at the University Medical Center Göttingen, Germany. Raw sequencing reads were demultiplexed with Illumina software. Subsequent bioinformatics analyses were performed with the Geneious package v11.1.2. First, we trimmed and quality-filtered the reads with BBduk 37.64 in the BBTools package (https://jgi.doe.gov/data-and-tools/bbtools/) and removed duplicate reads with Dedupe 37.64 (BBTools package); both filtering steps were conducted with standard settings. For assembly, reads were mapped onto the mitogenome of the *C. pygmaea* reference mitogenome NC_021942.1 and a C. niveiventris specimen (JT95, see below) using the Geneious assembler. We applied different settings (high and custom sensitivity, 5–10 iterations), while the advanced standard settings remained unchanged. To check for potential DNA damage (C to T or G to A changes) in historical specimens (Rambaut et al., 2009), we calculated the base frequencies in our *C. pygmaea* dataset with PAUP v4.0a (Swofford, 2002) and compared the base frequencies of the holotype with those of modern specimens. The newly produced mitogenome was manually checked and then annotated with Geneious v11.1.2.

**Modern specimen sampling and sequencing**

For mitogenome analysis, we used samples of pygmy marmosets from the town of Tabatinga, Brazil (Tabatinga 1 henceforth, TAB1), and from the upper Japurá River (JAP720, JAP723) and Jutai River (JT56, JT57, JT79, JT95). These specimens were deposited in the Zoological Collection of the Instituto Nacional de Pesquisas da Amazônia (INPA), Manaus. We extracted total genomic DNA from muscle tissues preserved in alcohol using standard protocols (Sambrook et al., 1989). Complete mitogenomes for these specimens were generated via shotgun high-throughput sequencing using a NEBNext Ultra II FS DNA Library Prep Kit (New England Biolabs, USA) (TAB1, JAP720, JT95) or the Qiagen MagAttract HMW DNA Kit according to the manufacturer’s specifications (JAP723, JT56, JT57, JT79). Before library preparation, 100 ng of DNA was fragmented (200–450 bp) during the kit’s fragmentation step. All follow-up steps adhered to the kit protocols. Sequencing at NIG was conducted on an Illumina HiSeq 4000 (50 bp single-end reads) (TAB1, JAP720, JT95) or an Illumina NovaSeq6000 machine (150 bp paired-end reads (JAP723, JT56, JT57, JT79). Read mapping and mitogenome assembly were conducted as described for Spix’s holotype.

The mitochondrial cyt *b* gene was amplified for a second individual from Tabatinga (Tabatinga 02) and another individual from the Jutai River (JT32) by PCR with the primers MonkeyGluF1 (5’-CCATGACTAAATGAATGAAARCC-3’) and MonkeyProR1 (5’-AGAATSTCAGCCTTGATTGTTG-3’) (see Boubli et al., 2018). The PCR products were purified using ExoSap (Werie et al., 1994) and subjected to fluorescent dye-terminator (ddNTP) sequencing following the manufacturer’s recommended protocols for BigDye sequencing chemistry (Applied Biosystems, USA) and using the primers MonkeyCytbF2 (5’-GATGACAYAYCIRCTCAGG-3’), MonkeyCytbR1 (5’-GGBCCTCAGAAGDTATTTTGG-3’) and MonkeyCytbR2 (5’-CGTARGTATGCRATGCRRAA-3’) (Boubli et al., 2018). After the cycle sequencing reaction, the products
were precipitated with 100% ethanol/125 mmol/L EDTA solution, re-suspended in Hi-Di formamide, and resolved on an ABI 3130xl automatic sequencer (Applied Biosystems, USA). Sequences were assembled, edited, and trimmed using Geneious v8.1.8.

Additional sequences
We supplemented the newly sequenced mitochondrial data with cyt b sequencing data from previous publications (Boubli et al., 2018; Porter et al., 2021), as well as mitogenomes available in GenBank. Our final dataset consisted of cyt b sequences from 65 individuals and full mitogenomes from 15 individuals. Of the 65 individuals included in the cyt b alignment, 52 belonged to the genus Cebuella, including both contemporary samples collected from wild primates and museum samples, covering most of the geographical distribution of pygmy marmosets (Table 1, Figure 1, 2). Importantly, this included a pygmy marmoset from the town of Benjamin Constant, Brazil, which is located on the right (south) bank of the Solimões, directly across the river from Tabatinga (Figure 2). Additionally, we included 13 cyt b sequences from other platyrhine species (Aotus, Saginus, Callimico, Callithrix, Mico, and Callibella) to provide a broader phylogenetic context. Alignment of the complete mitogenomes included the newly sequenced pygmy marmosets from Tabatinga (TAB1), Japurá (JAP720, JAP723), Jutai (JT56, JT57, JT79, JT95), and Spix’s pygmaea holotype (dDNA578), as well as publicly available genomes from Cebuella pygmaea, Callithrix jacchus, Callimico goeldii, Saginus oedipus, Aotus azarai, L. lemurinus, and A. nancymaeae. Details and accession numbers for all sequences included in this study are presented in Supplementary Table S1.

Phylogenetic analyses
All sequencing alignments were performed using the MAFFT v7.309 (Katoh & Standley, 2013) alignment plugin in Geneious v9.1.8 and checked by eye. Using the alignment of cyt b sequences from 65 individuals, including 52 Cebuella samples and 13 other platyrhines, we jointly estimated phylogeny and diversification times under a strict clock model implemented in BEAST v2.6.3 (Bouckaert et al., 2019). A strict clock model was chosen as it is most appropriate for single locus mitochondrial data, where rates are not expected to vary across branches. To date the tree, we used a fossil calibration dating the split between Aotinae and Callitrichidae based on the presence of the stem Callitrichidae Lagominico (de Vries & Beck, in prep.). The Aotinae and Callitrichidae split was given a hard minimum bound of 13.4 Ma following the age published for Lagominico (the following age listed by Kay, 2015), and a generous soft maximum bound of 35.0 Ma based on the age of the oldest stem catarrhine Catopithecus browni (using the maximum age of the L-41 locality in the Fayum Depression proposed by Seiffert, 2006). The calibration was given a log-normal distribution with an S parameter of 0.8 and M parameter of 1.755 to set the 95% quantile of the maximum age at 35.0 Ma. We partitioned the cyt b sequence alignment into three partitions based on codon position and used bModelTest (Bouckaert & Drummond, 2017) to assign appropriate substitution models to each partition. BEAST2 was allowed to run for 50 million generations, with sampling at every 5 000 trees and the first 10% discarded as burn in. Additionally, we compared our topology to a smaller subset of samples using full mitogenomes under a maximum-likelihood approach in RAxML v8 (Stamatakis, 2014). Robustness of the RAxML analyses was assessed via 1 000 bootstrap replicates. Alignment of the full mitoch kondial sequences, the xml files used as input for BEAST2, and the command used to run RAxML are provided in the Supplementary Materials.

RESULTS
Holotype sampling and sequencing
We successfully retrieved mitochondrial DNA from Spix’s Cebuella pygmaea type specimen and sequenced a total of 14 726 unambiguous base pairs, representing 89.1% of the complete mitogenomes, with an average sequencing depth of 3.8x. From the cyt b gene, we retrieved unambiguous base pairs for 59.6% of the sequence length. Base frequencies in the mitogenome of the C. pygmaea holotype (A=33.21%, C=26.31%, G=13.27%, and T=27.21%) were similar to those of modern specimens (mean: A=33.17%, C=26.51%, G=13.09%, and T=27.23%), indicating that DNA damage in the C. pygmaea holotype was minimal to zero.

Modern sampling and sequencing
We sequenced the mitogenomes for one of the newly collected specimens of pygmy marmosets from Tabatinga (TAB1) and six samples from other locations (JAP720, JAP723 JT56, JT57, JT79, and JT95) and retrieved full mitogenomes (100% coverage) from all modern samples (TAB1: 1717.7x; JAP720: 61.4x; JAP723: 4778.3x; JT56: 606.7x, JT57: 1007.1x, JT79: 3555.2; JT95: 90.0x). The newly sequenced mitogenomes of the holotype and modern samples were deposited in GenBank and are available under accession numbers MW733803–MW733806 and MZ747451–MZ747454 (Supplementary Table S1).

Phylogenetic analyses
BEAST2 analysis of the cyt b alignment supported a clear separation of two Cebuella clades and dated the split between clades at 2.54 Ma (95% highest posterior density (HPD) interval: 1.51–3.82 Ma) (Figure 3). The same topology was recovered by RAxML analysis of the full mitogenome alignment of the sample subset, with 100% bootstrap support (Figure 4). In both analyses, the sequence from Spix’s holotype grouped with contemporary animals sampled in Tabatinga on the north bank of the Solimões River, and other localities north of the Amazon/Solimões, including the Japurá River, while the pygmy marmoset from the south bank of the Solimões, near Benjamin Constant, grouped with the second clade. Of note, disregarding the missing data, the cyt b sequence of the holotype was identical to that of the two Tabatinga specimens.

Our results indicated a geographic distribution for C. pygmaea that is limited by the Napo and Solimões rivers in the south, the Andes in the west, and the Japurá-Caquetá in the north (however, there are some discrepancies that are discussed below). The distribution of C. niveiventris is concordantly limited by the Napo and Solimões rivers in the north and the Madeira River in the east. At this stage, our data
Table 1  List of voucher specimens and tissue samples used in this study and their localities

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Genus</th>
<th>Species</th>
<th>Collection site</th>
<th>Latitude</th>
<th>Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>EC_H1</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Flor del Pantano (Group 1), Orellana, Ecuador</td>
<td>–0.451</td>
<td>–76.864</td>
</tr>
<tr>
<td>EC_H2</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Flor del Pantano (Group 3), Orellana, Ecuador</td>
<td>–0.451</td>
<td>–76.864</td>
</tr>
<tr>
<td>EC_H3</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Flor del Pantano (Group 4), Orellana, Ecuador</td>
<td>–0.451</td>
<td>–76.864</td>
</tr>
<tr>
<td>EC_H4</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>San Pablo, Sucumbios, Ecuador</td>
<td>–0.273</td>
<td>–76.421</td>
</tr>
<tr>
<td>EC_H5</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>San Pablo, Sucumbios, Ecuador</td>
<td>–0.273</td>
<td>–76.421</td>
</tr>
<tr>
<td>EC_H6</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Tiputini Biological Station, Orellana, Ecuador</td>
<td>–0.638</td>
<td>–76.149</td>
</tr>
<tr>
<td>AMNH_72033</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Curaray, Maynas, Loreto, Peru</td>
<td>–2.366</td>
<td>–74.083</td>
</tr>
<tr>
<td>AMNH_72035</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Curaray, Maynas, Loreto, Peru</td>
<td>–2.366</td>
<td>–74.083</td>
</tr>
<tr>
<td>AMNH_72036</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Curaray, Maynas, Loreto, Peru</td>
<td>–2.366</td>
<td>–74.083</td>
</tr>
<tr>
<td>AMNH_72037</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Curaray, Maynas, Loreto, Peru</td>
<td>–2.366</td>
<td>–74.083</td>
</tr>
<tr>
<td>AMNH_73751</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Orosa, Mariscal Ramon Castilla, Loreto, Peru</td>
<td>–3.533</td>
<td>–72.183</td>
</tr>
<tr>
<td>AMNH_74054</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Orosa, Mariscal Ramon Castilla, Loreto, Peru</td>
<td>–3.533</td>
<td>–72.183</td>
</tr>
<tr>
<td>AMNH_74055</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Orosa, Mariscal Ramon Castilla, Loreto, Peru</td>
<td>–3.533</td>
<td>–72.183</td>
</tr>
<tr>
<td>AMNH_74056</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Orosa, Mariscal Ramon Castilla, Loreto, Peru</td>
<td>–3.533</td>
<td>–72.183</td>
</tr>
<tr>
<td>AMNH_74366</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Apayacu, Maynas, Loreto, Peru</td>
<td>–3.483</td>
<td>–72.183</td>
</tr>
<tr>
<td>AMNH_74367</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Apayacu, Maynas, Loreto, Peru</td>
<td>–3.483</td>
<td>–72.183</td>
</tr>
<tr>
<td>AMNH_74368</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Apayacu, Maynas, Loreto, Peru</td>
<td>–3.483</td>
<td>–72.183</td>
</tr>
<tr>
<td>AMNH_74369</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Apayacu, Maynas, Loreto, Peru</td>
<td>–3.483</td>
<td>–72.183</td>
</tr>
<tr>
<td>AMNH_75280</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Sarayacu, Ucayali, Loreto, Peru</td>
<td>–6.783</td>
<td>–75.116</td>
</tr>
<tr>
<td>AMNH_76327</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Sarayacu, Ucayali, Loreto, Peru</td>
<td>–6.783</td>
<td>–75.116</td>
</tr>
<tr>
<td>AMNH_76328</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Sarayacu, Ucayali, Loreto, Peru</td>
<td>–6.783</td>
<td>–75.116</td>
</tr>
<tr>
<td>AMNH_98312</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Iquitos, Maynas, Loreto, Peru</td>
<td>–3.766</td>
<td>–73.25</td>
</tr>
<tr>
<td>FMNH_54290</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Río Copataza, Pastaza, Ecuador</td>
<td>–2.116</td>
<td>–77.449</td>
</tr>
<tr>
<td>FMNH_71003</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Leticia, Amazonas, Colombia</td>
<td>–4.15</td>
<td>–69.950</td>
</tr>
<tr>
<td>FMNH_87136</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Río Mantí, Santa Cecilia, Maynas, Peru</td>
<td>–3.433</td>
<td>–72.766</td>
</tr>
<tr>
<td>FMNH_87137</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Río Mantí, Santa Cecilia, Maynas, Peru</td>
<td>–3.433</td>
<td>–72.766</td>
</tr>
<tr>
<td>FMNH_88997</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Alto Yavari Mirim, boca Yaque, Mariscal Ramon, Peru</td>
<td>–4.449</td>
<td>–71.783</td>
</tr>
<tr>
<td>FMNH_88998</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Alto Yavari Mirim, boca Yaque, Mariscal Ramon, Peru</td>
<td>–4.449</td>
<td>–71.783</td>
</tr>
<tr>
<td>FMNH_122750</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Quistococha, Maynas, Loreto, Peru</td>
<td>–3.833</td>
<td>–73.266</td>
</tr>
<tr>
<td>FMNH_122752</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Quistococha, Maynas, Loreto, Peru</td>
<td>–3.833</td>
<td>–73.266</td>
</tr>
<tr>
<td>UMMZ_82856</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Río Napo, Intillama, Napo, Ecuador</td>
<td>–0.982</td>
<td>–77.817</td>
</tr>
<tr>
<td>UMMZ_82857</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Río Napo, Intillama, Napo, Ecuador</td>
<td>–0.982</td>
<td>–77.817</td>
</tr>
<tr>
<td>JAP270</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Río Japurú, Amazonas, Brazil</td>
<td>–1.842</td>
<td>–69.022</td>
</tr>
<tr>
<td>JAP273</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Río Japurú, Amazonas, Brazil</td>
<td>–1.842</td>
<td>–69.022</td>
</tr>
<tr>
<td>JAP274</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Río Japurú, Amazonas, Brazil</td>
<td>–1.842</td>
<td>–69.022</td>
</tr>
<tr>
<td>Tabatinga_01</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Tabatinga, Amazonas, Brasil</td>
<td>–4.241</td>
<td>–69.944</td>
</tr>
<tr>
<td>Tabatinga_02</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Tabatinga, Amazonas, Brasil</td>
<td>–4.238</td>
<td>–69.944</td>
</tr>
<tr>
<td>CTGA-M170</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Igarapé do Jacinto, Tapauá, Amazonas, Brazil</td>
<td>–5.7</td>
<td>–63.2</td>
</tr>
<tr>
<td>FR20</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Lago Xadá, Amazonas, Brazil</td>
<td>–5.262</td>
<td>–60.722</td>
</tr>
<tr>
<td>CCM19</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Benjamín Constant, Amazonas, Brazil</td>
<td>–4.382</td>
<td>–70.008</td>
</tr>
<tr>
<td>MNFS1019</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Ocidente, Acre, Brazil</td>
<td>–8.572</td>
<td>–72.8</td>
</tr>
<tr>
<td>MNFS1020</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Ocidente, Acre, Brazil</td>
<td>–8.572</td>
<td>–72.8</td>
</tr>
<tr>
<td>MNFS1361</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Ocidente, Acre, Brazil</td>
<td>–8.572</td>
<td>–72.8</td>
</tr>
<tr>
<td>CCM23</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Codajas, Amazonas, Brazil</td>
<td>–3.894</td>
<td>–62.071</td>
</tr>
<tr>
<td>CCM251</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Lago Matupiri, Rio Madeira</td>
<td>–5.598</td>
<td>–61.006</td>
</tr>
<tr>
<td>JT79</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Río Jutai, Brazil</td>
<td>–3.311</td>
<td>–67.532</td>
</tr>
<tr>
<td>JT95</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Río Jutai, Brazil</td>
<td>–3.735</td>
<td>–67.469</td>
</tr>
<tr>
<td>JT57</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Río Jutai, Brazil</td>
<td>–3.218</td>
<td>–67.334</td>
</tr>
<tr>
<td>JT56</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Río Jutai, Brazil</td>
<td>–3.218</td>
<td>–67.334</td>
</tr>
<tr>
<td>JT32</td>
<td>Cebuella</td>
<td>niveiventris</td>
<td>Río Jutai, Brazil</td>
<td>–3.218</td>
<td>–67.334</td>
</tr>
<tr>
<td>Holotype</td>
<td>Cebuella</td>
<td>pygmaea</td>
<td>Adjacent to the town of Tabatinga, Brazil</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
do not offer unambiguous evidence to propose a southern limit for *C. niveiventris* (Figure 5).

Our BEAST2 tree indicated further sub-structuring within both *C. pygmaea* and *C. niveiventris*. Within the *C. pygmaea* clade, animals from the region between the Putumayo and Japurá rivers clustered separately from the animals collected south of the Putumayo. In the *C. niveiventris* clade, animals from the eastern extent of the species range, in the area between the Purus and Madeira rivers (CTGA-M170, FR20, CCM251), formed a cluster mostly separate from animals west of the Purus. While *C. niveiventris* west of the Ucayali and south of the Napo formed a subcluster in the phylogeny (Figure 3).

Four of the museum specimens labeled as having been collected south of the Napo and Solimões rivers fell within the *C. pygmaea* clade (Figure 3, marked with asterisks, Figure 1): FMNH (Field Museum of Natural History) 54290, AMNH (American Museum of Natural History) 76327, AMNH 76328, and AMNH 98312.

**DISCUSSION**

Based on their phylogenetic analysis using newly generated genomic data for a relatively large sample of individuals, Boubli et al. (2018) proposed the division of pygmy marmosets into two species of *Cebuella* (see Boubli et al., 2018; Garbino et al., 2019; see also Porter et al., 2021) separated by the
Solimões-Amazonas River. However, as pointed out by these authors, confirming the names to be attributed to these two newly identified clades was hindered by the uncertainty of the type locality of Spix’s *Cebuella pygmaea* (Lönnberg, 1940; see Boubli et al., 2018; Garbino et al., 2019). Garbino et al. (2019) provided an interpretation of the historic literature and concluded that the type specimen was, in fact, obtained north of the Solimões River, contra Lönnberg (1940), who stated that its origin was the mouth of the Javari River, south of the Solimões. Our results resolve this issue definitively, as we show that Spix’s type specimen is more closely related to contemporary pygmy marmosets from Tabatinga, on the north bank of the Solimões River, than to animals from the south bank. We also reaffirm that museum collections are a valuable source for genetic and taxonomic investigations of primates, particularly of name-bearing types, and highly damaged DNA, as is typically extracted from such old material, can be analyzed with modern high-throughput sequencing technologies.

The holotype and modern Tabatinga pygmy marmosets formed a clade with samples from the right bank of the Japurá River and other locations north of the Solimões River that split from the southern clade, south of the Solimões approximately 2.54 Ma, confirming the separation of the genus *Cebuella* into two species: i.e., *Cebuella pygmaea* (Spix, 1823), north of the Solimões River, and *Cebuella niveiventris* Lönnberg, 1940, south of the Solimões in that region (in agreement with Boubli et al., 2018).

The morphological analyses of Garbino et al. (2019) and Porter et al. (2021) identified the Napo River as the southern range limit of *C. pygmaea* in Peru and Ecuador. Previously, it had been thought that the divide was marked by the Amazonas-Marañón rivers, extending west to the left bank of the Pastaza River (Mittermeier et al., 2013; van Roosmalen & van Roosmalen, 1997). Pygmy marmosets sampled south of the Solimões-Amazonas and Napo rivers largely grouped with the *C. niveiventris* clade. The sample from Tiputini, Ecuador (with haplotype EC_H6, Porter et al., 2021) collected on the right bank of the upper Napo was clearly nested in the *C. niveiventris* clade.

Interestingly, as stated by Porter et al. (2021), four of the museum specimens are noteworthy exceptions and fall within the *C. pygmaea* clade, despite sampling locations south of the Napo River. We have highlighted these specimens with asterisks in Figure 3. Sample FMNH 54290 was collected in Ecuador, near the headwaters of the Copataza River on 8

Figure 3 BEAST2 cytochrome b time-tree for 65 primate samples, including 52 pygmy marmosets and 13 other taxa as outgroups

Numbers in nodes correspond to posterior support and error bars represent 95% HPD intervals. Inset is picture of *Cebuella pygmaea* type and original drawing by Spix (von Spix, 1823). See Figures 1 and 2 for map showing localities for all specimens used in this phylogenetic analysis.

Zoological Research 42(6): 761–771, 2021

767
April 1939 by R. Olalla, according to Chicago’s FMNH records. The collecting locality may be inaccurate, or the individual may have been moved by people who use them as pets (something that continues today, de la Torre pers. obs.), or changes in rivers due to meandering caused individuals from one bank to be isolated on islands that later connected to the other bank, allowing admixture between populations (see Boubli et al., 2018; Porter et al., 2021). FMNH 54290 was classified by Garbino et al. (2019) as C. niveiventris (Type ‘1’) based on its pelage pattern (Figure 6). Considering that it carries a C. pygmaea mitochondrial haplotype, this could be a case of admixture causing mitochondrial introgression. Such admixture events are not uncommon in the headwaters of Amazonian rivers (Naka et al., 2012; Weir et al., 2015), but further sampling in this area is needed to test this hypothesis.

The case for samples AMNH 76327, AMNH 76328, and AMNH 98312 is harder to reconcile as they were collected at the southern edge of the distribution of C. niveiventris, yet all three grouped with the C. pygmaea clade in our analysis. Samples AMNH 76327 and AMNH 76328 were collected for the AMNH by “Olalla & Hijos” on 16 March and 4 April 1927, respectively. The collection sites for these specimens are close to that of a third specimen, AMNH 75280, which was collected by the same group later that year (Haven Wiley, 2010). Sample AMNH 98312 was collected on 28 January 1929 and is labeled as coming from near Iquitos, a known wildlife trading hub.

It is possible that the location information available for these specimens is inaccurate. In fact, the provenance of other specimens collected by the Olalla brothers has previously been called into question (Haven Wiley, 2010; Marsh, 2014). The most problematic specimens appear to be those first purchased by Harvey Bassler (e.g., AMNH 98312, included here) before coming to the AMNH, as the Olalla brothers did not reliably record which riverbank was sampled until late 1926 (Haven Wiley, 2010). Several of the samples included here were collected during an expedition when the Olalla brothers appear to have realized the importance of river boundaries and began purposefully collecting on opposite sides of a river, detailing the bank clearly in their sample notes (Haven Wiley, 2010). Samples AMNH 74366–74369 and AMNH 73751, 74054–74056 were collected during this expedition and clearly separated into pygmaea and niveiventris based on our analysis, in accordance with provenance (Figure 3). For samples collected prior to this, it is
Likewise, Boubli et al. (2018) proposed a new morph of pygmy C. pygmaea as typical C. niveiventris under character by Garbino et al. (2019) as upper Juruá in Brazil have dark underparts and were classified "1" (see above). Which was classified by Garbino et al. (2019) as morphotype 290, varying degrees of white fur present and accounted for the "2" consisted of animals with much darker underparts with Lönnberg (1940). Morphotypes "1" and C. niveiventris examined vouchers and consisted of the typical, white-bellied vouchers into three somewhat discrete morphotypes.

Of Cebuella pelage coloration to the extent that variability in this region. Separating the two species where their ranges meet around an admixed individual as there are no natural barriers to accommodate our suggestion that FMNH 54290 is potentially C. pygmaea geographical distribution of intriguing hypothesis of Porter et al. (2021) regarding the broad distribution of C. pygmaea and thus the distribution of C. niveiventris in western Amazonia would be nested within the broad distribution of C. pygmaea (see Figure 4 of Porter et al. 2021 and question marks in Figure 5 here). We agree with the intriguing hypothesis of Porter et al. (2021) regarding the geographical distribution of C. pygmaea, which would accommodate our suggestion that FMNH 54290 is potentially an admixed individual as there are no natural barriers separating the two species where their ranges meet around the foothills of the Andes and upper reaches of the rivers in this region.

Hershkovitz (1977) and Garbino et al. (2019) reported great variability in Cebuella pelage coloration to the extent that Hershkovitz (1977) concluded that there was only one pygmy marmoset species, and all variation was intraspecific. On the other hand, Garbino et al. (2019) restricted this variability to C. niveiventris. After examining 44 museum voucher specimens of Cebuella, they proposed classification of the C. niveiventris vouchers into three somewhat discrete morphotypes. Predominant morphotype "3" accounted for 71% of the examined vouchers and consisted of the typical, white-bellied C. niveiventris sensu Lönnberg (1940). Morphotypes "1" and "2" consisted of animals with much darker underparts with varying degrees of white fur present and accounted for the minority of the samples. This is the case for FMNH 54290, which was classified by Garbino et al. (2019) as morphotype "1" (see above).

Animals collected by the Olalla brothers in Eirunepé on the upper Juruá in Brazil have dark underparts and were classified by Garbino et al. (2019) as C. niveiventris under character states "1" and "2". This is what led Lönnberg (1940) to classify them as typical C. pygmaea (see Boubli et al., 2018). Likewise, Boubli et al. (2018) proposed a new morph of pygmy marmosets from the upper Juruá in Acre, Brazil, and identified them as C. cf. pygmaea. These specimens (MNFS_1019-20, MNFS_1361) are morphologically similar to those collected by the Olalla brothers in Eirunepé, in that their underparts are darker than the typical C. niveiventris (see Boubli et al., 2018). Such variability in underpart coloration in C. niveiventris led Porter et al. (2021) to disregard such variation as meaningful for taxonomic classification. In fact, Soini (1988) and de la Torre (pers. obs.) report seeing great within-population pelage color variation in Ecuador. If we consider the distribution hypothesis for C. pygmaea proposed by Porter et al. (2021) (see above), then we should expect possible contact between the two species in the southern and western edges of the distribution of C. niveiventris and potential gene flow between the two species in the upper reaches of the Napo, Pastaza, Ucayali, and possibly Juruá, which could account for the color pattern variation observed in pygmy marmosets in these regions. Previous studies have reported on gene flow and hybridization among populations of different primate species in the New World, e.g., between Saginus midas and S. bicolor (Farias et al., 2015), Plecturocebus cinerascens, P. parecis, and P. bernhardi (Byrne et al., 2021), and P. moloch and P. vierai (Boubli et al., 2019), as well as for other organisms on other continents (e.g., frogs in Southeast Asia – Chan et al., 2020; Darwin’s finches in the Galapagos – Lamichhaney et al., 2020; Gazelles in Africa – Garcia-Erríl et al., 2021). As such, this hypothesis deserves further investigation using nuclear DNA data.

Our study revealed greater lineage diversity in pygmy marmosets than ever before. In addition to the clear separation of pygmy marmosets in two distinct species, Porter et al. (2021) and our analysis also identified further structuring in both species, thus revealing four reciprocally monophyletic lineages in Cebuella: i.e., two C. pygmaea lineages separated by the Putumayo River and two C. niveiventris lineages separated by the Purus River. However, such structuring should be considered with caution as only one locus was used in some cases (cyt b). Thus, more robust nuclear data are needed to better understand the phylogenetic diversity of pygmy marmosets.

CONCLUSIONS
In this study, we resolved a long-standing taxonomical conundrum surrounding the origin of Spix’s pygmaea type by sequencing its mitogenomes from historical DNA. Unambiguously, our results showed that the type is closely related to the pygmy marmosets that currently live around the town of Tabatinga in Brazil. Thus, our data support the classification of pygmy marmosets as C. pygmaea and C. niveiventris. This is the first study to successfully use historic DNA from a type specimen to address an important taxonomical question in New World primates, thus paving the way for future studies addressing similar issues in other platyrhines.

Our results and those of Porter et al. (2021) considerably expand the range of C. niveiventris to the west and raise the possibility that the headwaters of these western Amazon tributaries may not, in their uppermost reaches, be barriers, and thus the distribution of C. pygmaea may be much larger.

Zoological Research 42(6): 761−771, 2021 769
than previously thought.

Once considered a single and widespread species, we show that Cebuella is a diverse taxon, with two full species and further cryptic diversification within them. As such, we now have four or even five (see Porter et al., 2021) evident lineages of significance as units for conservation management (sensu Moritz, 1994).

SCIENTIFIC FIELD SURVEY PERMISSION INFORMATION

Permission to conduct fieldwork and to collect tissue samples was granted by IBAMA (License No 005/2005–CGFAU/LIC) and ICMBio in Brazil, and by the Ecuadorian Ministry of the Environment (Contrato Marco de Acceso a los Recursos Genéticos Nro. MAE-DNBCM-2015-0019 to Stella de la Torre —Universidad San Francisco de Quito and Liliana Cortés-Ortiz—University of Michigan) in Ecuador.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS’ CONTRIBUTIONS

J.P.B. conceived and designed the study, acquired funding and organized sample acquisition, interpreted the findings, and led writing of the manuscript. M.C.J. led genomic analyses, interpreted findings, participated in writing of the manuscript, and created figures with input from J.P.B. and the other authors. L.M.P., S.dT.T., and L.C.-O. collected, extracted, and sequenced the samples, acquired funding, and edited the manuscript. M.N.F.d.S. conceived of the study, interpreted findings, and collected samples. A.B.R. conceived of the study, interpreted findings, and participated in writing of the manuscript. S.N. participated in interpreting the findings and editing of the manuscript. F.E.S. collected samples and edited the manuscript. D.dV. and R.M.D.B. collected data and participated in writing and editing of the manuscript. T.H. and I.P.F. acquired funding and samples, extracted and sequenced samples, and edited the manuscript. A.H.vH. collected samples from the Spix type specimen. I.R.-G., L.F.K.K., and T.M.-B. extracted and sequenced modern mitogenomes. C.R. led the extraction and sequencing of mitogenomes, acquired funding and samples, interpreted the findings, and participated in writing of the manuscript. All authors read and approved the final version of the manuscript.

ACKNOWLEDGEMENTS

Our most sincere thanks to the Bavarian State Collection of Zoology, Munich, Germany, for granting access to the Spix’s Cebuella pygmaea type specimen for our DNA work. We are indebted to Christiane Schwarz for her dedication and excellent laboratory work at DPZ. Our sincere thanks to G. Garbinho for letting us examine his photo archives of pygmy marmoset voucher specimens from museums across the world. L.P., L.C.O., and S.T. thank the staff at the Field Museum of Natural History (FMNH), American Museum of Natural History (AMNH), and University of Michigan Museum of Zoology (UMMZ) for granting them permission to access their collections, and to the FMNH staff for collecting osteocrust samples and to Priscilla and Paige Watson and Merelyn Dolins for their generous hospitality in New York City while L.P. worked at the AMNH. We thank Delaney Cargo, Anna del Valle Perts, Dayanna López, Erika Alzamora, Amalia de la Torre, Ramiro SanMiguel, and Pablo Yépez for their help in the collection of Ecuadorian samples. The Tiputini Biodiversity Station, the Secoya community of San Pablo, and the SanMiguel family greatly facilitated the field work in Ecuador. We thank Nina Q. Espinosa de los Monteros Silva, Cassandra Huerta, Patricia Billete, and Stavi Tennenbaum for their help producing the molecular data from Ecuador and Peru. Finally, L.P., L.C.O., and S.T. thank the Chicago Zoological Society’s Chicago Board of Trade Endangered Species and USFQ Chancellor and Collaboration Grants for their help funding research on samples from Peru and Ecuador. J.P.B. thanks M van Roosmalen for lively discussions on the ecology and biogeography of Amazonian primates that inspired this study.

REFERENCES


