University of São Paulo
"Luiz de Queiroz" College of Agriculture
Center for Nuclear Energy in Agriculture

Phylogenomics, diversification, and biogeography of Neotropical squirrels (Sciurillinae and Sciurinae: Sciurini)

Edson Fiedler de Abreu Júnior

Thesis presented to obtain the degree of Doctor in Science. Area: Applied Ecology

Piracicaba
2020

Edson Fiedler de Abreu Júnior
Bachelor in Biological Sciences

Phylogenomics, diversification, and biogeography of Neotropical squirrels (Sciurillinae and Sciurinae: Sciurini)

Advisor:
Prof. Dr. **ALEXANDRE REIS PERCEQUILLO**

Thesis presented to obtain the degree of Doctor in Science. Area: Applied Ecology

Piracicaba
2020

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Abreu Júnior, Edson Fiedler de

Phylogenomics, diversification, and biogeography of Neotropical squirrels (Sciurillinae and Sciurinae: Sciurini) / Edson Fiedler de Abreu Júnior - - Piracicaba, 2020.

210 p.

Tese (Doutorado) - - USP / Escola Superior de Agricultura “Luiz de Queiroz”. Centro de Energia Nuclear na Agricultura.

1. Sistemática 2. Taxonomia 3. Biogeografia 4. Filogenia 5. Sciuridae I. Título

In memory of my grandma, Alba Fiedler

Acknowledgments

With no doubt I am a lucky person. Lucky for being always surrounded by so many wonderful people who have made my Ph.D. journey so much easier and pleasant. I am very grateful for all unconditional love and support from my parents, Edson and Carmen. They have always embraced my dreams and spared no efforts to give me the best conditions to pursue an academic formation. Felipe, who boarded this journey without knowing what was coming next and, despite all adversities along the way, stayed by my side encouraging me with his admiration and love. He also turned out to be a good field biologist and taxidermist! Jake, who was more than a housemate or a friend, over the past years we became like brothers and we have shared uncountable awesome moments, and also most of the anxiety and stress of the postgrad life. I am glad that he suddenly landed in Brazil in 2015.

I especially thank to my Ph.D. advisor, Alexandre R. Percequillo, for almost ten years of friendship, scientific guidance, and for providing so many great academic and life opportunities since I started as a master's student in his Lab in 2011. He has taught me all the basic and classic principles of mammalogy, and, among many other things, that taxonomy is a rigorous subject but it should not be locked down for new methodological or conceptual advances. I am also very thankful to my supervisor during the abroad period at the Smithsonian, Jesus E. Maldonado, for making this challenging project possible with his genomic expertise, and for all generous help and support during my time in Washington D.C. This project was indeed a large undertaking in which I had the honor to share the leadership with a friend and great scientist, Silvia E. Pavan. Since the beginning, when we realized that we were independently starting almost the same project, we decided to collaborate and put our efforts together in order to tackle the phylogenomics and taxonomy of South American squirrels. Our project grew up along the way and fortunately we found some amazing people to join our team. Mirian T. N. Tsuchiya trained and helped us in the genomic lab and provided crucial bioinformatics guidance. Don E. Wilson supported this project in so many ways. He is one of the kindest people that I have meet in the scientific community and it has been an incredible honor to work with one of the greatest mammalogists from all times.

Friends and lab mates provided crucial help during numerous steps of this project. Two long field expeditions were performed, one to the Brazilian Amazon (Rio Purus) and another to the Andes of Peru (Parque Nacional del Río Abiseo). In the Rio Purus, I had incredible

partnership with Jacob Charters, he embraced my crazy idea and helped chasing squirrels for almost a month. In Río Abiseo, Silvia Pavan leaded the expedition with bravery, and we had amazing help from Pamela Sanchez, Edgardo Rengifo, Felipe Câmara, Pedro Peloso, Roberta Fonseca, and Renzo Pradel. I also wish to express my gratitude to the local people who housed and/or helped us during these field expeditions. I am particularly grateful to Lillian D. Parker, Madhvi Venkatraman, Susette Castañeda Rico and Nancy R. McInerney, who provided crucial help during lab work. I am also grateful to Joyce R. Prado, Jeronymo Dallapicola and Bryan McLean, who helped with scrips, analyses, and computational problems solving. I thank Pedro Peloso, Felipe Arantes, and Samuel Oliveira for allowing to use photographs of squirrels taken by them in several places. I also thank to Gustavo Libardi, who kindly illustrated representatives of all Sciuridae subfamilies.

I am deeply indebted to Mario de Vivo for his generosity in sharing his profound knowledge on squirrels, and for kindly donating all the qualitative and quantitative data that he amassed from several museums worldwide over many years. I am also indebted to Ana Paula Carmignotto, who, together with Mario, published the last taxonomic arrangement for South American squirrels, which was the foundation of the present study. I am thankful to Louise Emmons, who shared precious insights and thoughts about Neotropical squirrels during my visits to the USNM, collected several samples analyzed in this study, and generously provided photographs and information from specimens at the USNM. I would like to acknowledge Robert Voss, who supported my application to the Collection Study Grant at the AMNH, and for sharing valuable insights on the taxonomy and evolution of Neotropical squirrels. Sharon Jansa has been an important academic mentor, and the knowledge that I have acquired as an intern at her lab are the underpinning of my genetics background.

I am grateful to the following curators and collection support staff, who loaned preserved tissue samples, allowed to analyze and obtain destructive samples from valuable historical specimens under their care, or provided important information regarding vouchers of samples included in this study: R.S. Voss, E. Hoeger, and N. Duncan (AMNH); I. Moya (CBF); M.R. Alvarez (CMARF); B.D. Patterson and A.W. Ferguson (FMNH); J. Valsecchi and I. Junqueira (IDSM); C.R. Silva and A.F. Sobrinho (IEPA); R.M. Timm and M. Eifler (KU); D.L. Dittmann and R.T. Brumfield (LSU-MZM); E. Pasa (MCN-FZB); C.G. Costa (MCN-M); A.U. Christoff, F.B. Peters and A.M. Gandini (MCNU); J.A. Oliveira and M. Weksler (MN); J. Silva Junior (MPEG); M. Campbell, J. Cook and J. Dunnum (MSB); M.T. Rodrigues and B.M.S. Batista (MTR); V. Pacheco

(MUSM); C. Conroy, J.L. Patton, and F.J.M. Pascal (MVZ); M. Vivo, L.F. Silveira and J.G. Barros (MZUSP); J.K. Braun and B.S. Coyner (OMNH); H. Garner, R.D. Bradley, and C. Phillips (TTU); L.P. Costa, Y.L.R. Leite and M.P. Nascimento (UFES-CTA); R. Rossi and T. Semedo (UFMT); A.C. Mendes Oliveira (UFPA); J. Cherem, M. Graipel and E.C. Grisard (UFSC); D.P. Lunde, S.C. Peurach, N.R. Edmison, I. Rochon, M. Krol, J. Jacobs, C. Ludwig, C. Huddleston, D. DiMichele, M. Braun, L.H. Emmons, and A.L. Gardner (USNM); and F. Catzeflis (ISEM). I also thank to A. Ravetta, G.T. Garbino and L.P. Godoy, who provided tissue samples of recently collected and uncatalogued specimens.

I am thankful to the Programa de Pós-Graduação Interunidades em Ecologia Aplicada (PPGI-EA) and the University of São Paulo for housing me as a graduate student and for providing outstanding academic opportunities that were crucial to my formation as a scientist. I am also very thankful to the Smithsonian Conservation Biology Institute's Center for Conservation Genomics (CCG) for having me as Predoctoral Fellow and for providing the best conditions for the development of this project. At PPGI-EA, I am especially grateful to Mara Casarin, Maria Victoria Ballester and Giancarlo Conde Xavier Oliveira, who are always welcome to help students regarding administrative and bureaucratic duties. At CCG, Jesus Maldonado, Robert Fleisher, Michael G. Campana and Nancy McInerney provided crucial help with administrative issues. I had valuable suggestions that certainly improved this project from my Graduate Committee, composed of Ana Carolina Pavan, Katia M. P. M. B. Ferraz and Giancarlo C. X. Oliveira, and from my Qualifying Exam Committee, composed of Ana Paula Carmignotto and Ana Carolina Loss.

Lastly, I would like to express my deepest gratitude to all friends and colleagues (I will not dare myself to name them as I would surely miss someone, not for ingratitude but because there were some many great people that crossed my way over the last years in Brazil and in the USA) that shared with me a coffee, a beer, a cheap chat or a profound discussion, that housed me during field trips or trips to visit museums, that helped in so many different ways or were extremely important just for being there. Thank you all!

This project was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) through regular doctoral fellowships (processes numbers 147145/2016-3 and 165553/2017-0) and through an abroad fellowship (SWE; 203692/2017-9); by the American Society of Mammalogists (ASM) through the Latin American Student Field Research Award, providing support for a field expedition to the Rio Purus, BR; by the American Museum

of Natural History (AMNH) through a Collection Study Grant, allowing me to collect morphological data from Neotropical squirrels housed at the AMNH; by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) through a Young Investigator Grant to A. Percequillo (09/16009-1), providing support for several field expeditions to the Amazon, BR; by the National Geographic Society (NGS) through an Explorers Grant to S. Pavan (NGS-381R-18), providing support to visit and collect samples housed at the MUSM and for a field expedition to Parque Nacional del Rio Abiseo, PE; and by the Smithsonian Institution (SI), providing research funds that covered costs associated to laboratory procedures, DNA sequencing, and publication charges.

Rio Japurá, Amazonas, BR

“É engraçado como o bem-estar não depende do conforto, da tranquilidade ou de situações favoráveis, mas simplesmente e unicamente da sensação de ir em frente.”

— Amyr Klink

Contents

RESUMO **12**
ABSTRACT **13**
1. INTRODUCTION **15**
1.1 Family Sciuridae **15**
1.2 Subfamily Sciurillinae **16**
1.3 Subfamily Sciurinae, Tribe Sciurini **17**
1.3.1 Taxonomic classifications **17**
1.3.2 Phylogenetic relationships **19**
1.3.3 Spatiotemporal diversification **21**
1.4 The sampling shortfall for Neotropical squirrels **22**
1.5 Thesis structure and chapters presentation **24**
References **26**
2. CHAPTER 1: Museomics of tree squirrels: a dense taxon sampling of mitogenomes reveals hidden diversity, phenotypic convergence, and the need of a taxonomic overhaul **29**
Abstract **30**
2.1 Background **30**
2.2 Results **33**
2.2.1 Summary of mitochondrial genome sequencing success, assembly and synteny **33**
2.2.2 Phylogenetic inferences and the effect of missing data **34**
2.2.3 A mitogenomic hypothesis: Sciurini and its species groups **35**
2.2.4 Species monophyly and species recognition **37**
2.2.5 The evolution of premolars and mammae within Sciurini **39**
2.3 Discussion **40**
2.3.1 The importance of museum specimens for the study of Neotropical tree squirrels **40**
2.3.2 Missing data versus missing taxa **42**
2.3.3 Mitochondrial phylogeny and taxonomic arrangements proposed for Sciurini **43**
2.3.4 Comments on species recognition and novelties **47**
2.3.5 Phylogenetic and biogeographic remarks **51**
2.3.6 Taxonomic consequence of the use of homoplastic traits in the study of tree squirrels **52**
2.4 Conclusions **53**
2.5 Methods **54**
2.5.1 Sampling **54**
2.5.2 Taxonomic identifications **54**
2.5.3 DNA extractions **55**
2.5.4 Library preparations and mtDNA amplification **56**
2.5.5 Quantification and sequencing **56**
2.5.6 Data processing **57**
2.5.7 Sequence alignment and dataset composition **58**
2.5.8 Mitochondrial genome recovery for historical museum samples **58**
2.5.9 Phylogenetic analyses **59**
2.5.10 Species delimitation analyses **59**
2.5.11 Morphological character evolution **60**
List of abbreviations **61**

Declarations ... **62**
Funding ... **62**
Authors' contributions ... **63**
Acknowledgements ... **63**
References ... **65**
Tables ... **70**
Figures ... **74**
Additional files ... **82**
3. CHAPTER 2: Spatiotemporal diversification of tree squirrels: is the South American invasion and speciation really that recent and fast? ... **99**
Abstract ... **100**
3.1 Introduction ... **100**
3.2 Material and Methods ... **103**
3.2.1 Dataset taxonomic composition ... **103**
3.2.2 DNA matrix, alignment and saturation test ... **103**
3.2.3 Divergence times estimate ... **104**
3.2.4 Ancestral range and diversification events estimate ... **105**
3.2.5 Lineage through time and speciation rates ... **105**
3.3 Results ... **106**
3.3.1 Phylogeny of the tribe Sciurini ... **106**
3.3.2 Model selection in ancestral range estimates ... **106**
3.3.3 Spatiotemporal diversification of Sciurini at a global scale ... **106**
3.3.4 Spatiotemporal diversification of Sciurini in the Neotropics ... **108**
3.3.5 Diversification rates of the tribe Sciurini ... **109**
3.4 Discussion ... **109**
Conflict of Interest ... **118**
Author Contributions ... **118**
Funding ... **118**
Acknowledgments ... **118**
Contribution to the Field Statement ... **119**
Data Availability Statement ... **119**
References ... **120**
Tables ... **125**
Figures ... **127**
Supplementary Material ... **131**
4. CHAPTER 3: Ultraconserved Elements resolve the squirrel tree at deep time scales but deliver lack of phylogenomic consistency in a rapid Neotropical radiation ... **142**
Abstract ... **143**
4.1 Introduction ... **143**
4.2 Material and methods ... **146**
4.2.1 Sampling ... **146**
4.2.2 DNA extraction and library preparation ... **147**
4.2.3 UCE capture and sequencing ... **147**
4.2.4 Data processing ... **148**
4.2.5 UCE sequencing success for historical museum samples ... **149**

4.2.6 UCE data filtering approaches **149**
4.2.7 Phylogenetic inferences **150**
4.2.8 Topological comparisons **151**
4.3 Results **151**
4.3.1 Sequencing summary, UCE assembly and success for historical samples **151**
4.3.2 Topological incongruences and the influence of data filtering strategies **152**
4.3.3 Phylogeny of Sciuridae: relationships among subfamilies and the genetic diversity of Sciurillinae **153**
4.3.4 Phylogeny of Sciurini: generic relationships, species monophyly and affinities **154**
4.4 Discussion **157**
4.4.1 The role of historical samples in the Neotropical squirrels UCE phylogeny **157**
4.4.2 Impact of UCE data filtering and discordance among inference methods **158**
4.4.3 UCE phylogenomic hypothesis for the relationships among subfamilies of Sciuridae **160**
4.4.4 UCE phylogenomic hypotheses for Sciurini and the mito-nuclear discordance **161**
4.5 Conclusion **165**
Acknowledgments **166**
References **168**
Tables **173**
Figures **174**
Supplementary Material **185**
5. CONCLUDING REMARKS 207

Resumo

Filogenômica, diversificação e biogeografia dos esquilos Neotropicais (Sciurillinae e Sciurinae: Sciurini)

Os esquilos são habitantes conspícuos da maioria das florestas Neotropicais, onde desempenham funções ecológicas cruciais na predação e dispersão de sementes. Duas radiações ocorrem nesta região: a subfamília Sciurillinae, representada exclusivamente pelo quatipuruzinho (*Sciurillus pusillus*); e a tribo Sciurini (subfamília Sciurinae), um grupo especioso e composto por cerca de 45 espécies de esquilos arborícolas, que ocorrem também nas regiões Neártica e Paleártica. Apesar da inquestionável importância dos esquilos nas dinâmicas dos ecossistemas e de representarem uma parte significativa da diversidade de roedores nos Neotrópicos, eles têm sido amplamente negligenciados por taxonomistas e sistematas. Como resultado, informações básicas acerca do número de gêneros e espécies ainda é controverso para o grupo e muitos aspetos da evolução e diversificação dos esquilos permanecem desconhecidos. Na presente tese, utilizei genomas mitocondriais e Elementos Ultra-conservados (UCEs) a fim de investigar a sistemática molecular e a história evolutiva das duas radiações de esquilos Neotropicais. No Capítulo 1, foram empregados mitogenomas sequenciados a partir de 232 espécimes históricos e modernos, para inferir uma hipótese filogenética robusta para os esquilos arborícolas. Os resultados filogenéticos foram contrastados com os arranjos genéricos propostos para a tribo Sciurini, foram discutidas as implicações taxonômicas e, por fim, uma nova classificação supra-específica foi proposta com o reconhecimento de 13 gêneros, empregados previamente por outros autores. Também foi constatada que a diversidade de espécies de esquilos arborícolas Neotropicais está subestimada, com o reconhecimento de ao menos seis linhagens representando espécies a serem descritas ou revalidadas. O Capítulo 2 teve como objetivo principal testar as hipóteses vigentes de tempo e modo de diversificação dos esquilos arborícolas, utilizando o banco de dados mitogenômico que incluiu 43 das 46 espécies de Sciurini. A data de origem da tribo foi estimada ao redor de 14 Ma e a sua área ancestral foi inferida com maior probabilidade para a América do Norte. A origem da radiação Neotropical ocorreu cerca de 6 Ma no noroeste da América do Sul, no domínio Pacífico. A maioria dos eventos cladogenéticos nos Neotrópicos ocorreram durante o Plioceno—logo depois da invasão Sul-Americana. Uma taxa de especiação predominantemente constante foi estimada para o grupo, o que contrastou com um pico no acúmulo de linhagens observado no Plioceno. No Capítulo 3, forneço uma perspectiva genômica nuclear da filogenia dos esquilos Neotropicais (Sciurillinae and Sciurinae: Sciurini), empregando mais de 3.700 UCEs sequenciados a partir de 184 amostras históricas e modernas. As análises filogenéticas inferiram com alto suporte as relações entre as cinco subfamílias de Sciuridae e também recuperaram de forma consistente as relações entre os ramos mais internos de Sciurini. Para a radiação Neotropical, a qual sofreu uma rápida diversificação, relações ambíguas em níveis supragenérico e interespecífico foram estimadas, dependendo da filtragem dos dados e do método de inferência. Também foram observadas inconsistências filogenéticas comparando os resultados de UCEs e mitogenomas. Por fim, nos capítulos 1 e 3 investiguei como alguns aspetos das amostras históricas podem influenciar a obtenção de dados genômicos, disponibilizando informações úteis para futuros estudos utilizando amostras antigas de museus.

Palavras-chave: Sistemática; Taxonomia; Biogeografia; Filogenia; Sciuridae

Abstract

Phylogenomics, diversification, and biogeography of Neotropical squirrels (Sciurillinae and Sciurinae: Sciurini)

Squirrels are conspicuous inhabitants of most Neotropical forests, where they play ecologically crucial roles as seed predators and dispersers. Two distinct radiations are found in this region: the subfamily Sciurillinae, represented exclusively by the Neotropical pygmy squirrel (*Sciurillus pusillus*); and the tribe Sciurini (subfamily Sciurinae), a speciose group composed of about 45 species of tree squirrels that also occur throughout the Nearctic and Palearctic regions. Despite their unquestionable importance to ecosystems dynamics and representing a substantial portion of the diversity of rodents in the Neotropics, squirrels have been largely neglected by taxonomists and systematists. As a result, basic information on number of genera and species is still ambiguous for the group, and also most aspects of their evolution and diversification remains unclear. In the present thesis, I employed mitochondrial genomes and Ultraconserved Elements to undertake the molecular systematics and the evolutionary history of the two radiations of Neotropical squirrels. In the Chapter 1, I used mitogenomic data sequenced from 232 historical and modern museum specimens to provide the first comprehensive phylogeny of tree squirrels. I contrasted the phylogenetic results with generic arrangements proposed for the tribe Sciurini, discussed the taxonomic implications, and suggested a tentative new classification at genus level employing 13 generic names used by previous authors. I also found evidence that the diversity of Neotropical tree squirrels is underestimated, with at least six lineages that represent taxa to be named or revalidated. In the Chapter 2, my main objective was to test current hypotheses on the tempo and mode of diversification of tree squirrels, employing the mitogenomic dataset including 43 of the 46 putative species of Sciurini. I estimated the date of origin of the tribe Sciurini around 14 Mya and suggested that its ancestral area was most likely in North America. The origin of the Neotropical radiation was estimated to have occurred around 6 Mya in northwestern South America, in the Pacific dominion. The majority of Neotropical cladogenetic events occurred along the Pliocene—right after the South American invasion. A fairly constant speciation rate was estimated for tree squirrels, which contrasts with the peak of lineage accumulation observed in the Pliocene. In the Chapter 3, I provided a nuclear genome-wide perspective of the Neotropical squirrels (Sciurillinae and Sciurinae: Sciurini) phylogeny, employing over 3,700 Ultraconserved Elements sequenced from 184 historical and modern samples. Phylogenetic analyzes estimated with strong support the relationship among the five subfamilies of Sciuridae, and also provided consistent and well-supported results for the relationships among the deepest branches of Sciurini. For the Neotropical radiation, which experienced a rapid diversification, conflicting relationships at both genus- and species-level were estimated upon data filtering and inference method. Inconsistences were also recovered with regards to the mitogenomic hypothesis. Finally, in both chapters 1 and 3, I took advantage of the large sampling across a diverse lineage of mammals to investigate how distinct aspects of historical samples might influence the recovery of genomic data, providing useful information for future genetic studies sampling from historical specimens.

Keywords: Systematics; Taxonomy; Biogeography; Phylogeny; Sciuridae

1. INTRODUCTION

1.1 Family Sciuridae

Sciuridae is the third most diverse family of rodents, with about 60 genera and 300 species (Thorington et al. 2012; Vivo and Carmignotto 2015; Burgin et al. 2018). Its members occur throughout all continents, excepted in Australia and Antarctica, and they have adapted to a great variety of environments, from dense tropical forests to semiarid deserts (Thorington et al. 2012). Squirrels are predominantly herbivorous, feeding on fruits, seeds and nuts, and due this habit they play important ecological roles as seed predators and dispersers (Peres 1999; Steele et al. 2005; Mendes et al. 2019). Squirrels are also often used as model organisms to address a wide range of ecological and evolutionary questions, both in their native ecosystems (Stapanian and Smith 1978; Benkman 1995; Mendes and Cândido 2014; Ketcham et al. 2017; Goldstein et al. 2018) and as invasive species (Gurnell et al. 2004; Bertolino et al. 2014).

The temporal origin of the family Sciuridae have been estimated to about 36 Mya (Millions of years ago; Mercer and Roth 2003; Fabre et al. 2012; Upham et al. 2019), which coincides with the earliest known fossil, *Douglassciurus jeffersoni* (Douglass 1901), from the late Eocene (Emry and Thorington 1984). This fossil, as well as other older fossils assigned to this family, were found in the northern hemisphere, suggesting a Holarctic origin for squirrels (Mercer and Roth 2003; Emry et al. 2005). This hypothesis was corroborated by a recent study based on genetic data, which estimated Eurasia as the ancestral range of Sciuridae (Rocha et al. 2016).

Traditional morphologic-based classifications have divided Sciuridae into two subfamilies: Sciurinae, composed of tree squirrels and ground squirrels; and Pteromyinae, composed of flying squirrels (Moore 1959; Hoffmann et al. 1993; Thorington and Hoffmann 2005). This classificatory scheme, however, was not sustained by modern molecular phylogenies, which consistently recovered five major lineages within Sciuridae and did not find Sciurinae and Pteromyinae as reciprocally monophyletic groups (Mercer and Roth 2003; Steppan et al. 2004; Fabre et al. 2012; Zelditch et al. 2015; Upham et al. 2019). It is noteworthy that these five major lineages presented identical compositions in all above mentioned studies, whereas the relationships among them vary across the inferences (Figure 1).

Therefore, Sciuridae is currently organized into five subfamilies and two of those also present a tribal arrangement (Steppan et al. 2004; Thorington et al. 2012): Ratufinae, composed of the Asian giant tree squirrel; Sciurillinae, composed of the Neotropical pygmy squirrel; Sciurinae, including Pteromyini (composed of flying squirrels) and Sciurini (composed of Holarctic and Neotropical tree squirrels); Callosciurinae, composed of Southern Asia tree squirrels; and Xerinae, including Xerini (composed of African ground squirrels), Marmotini (composed of Holartic ground squirrels), and Protoxerini (composed of African tree squirrels).

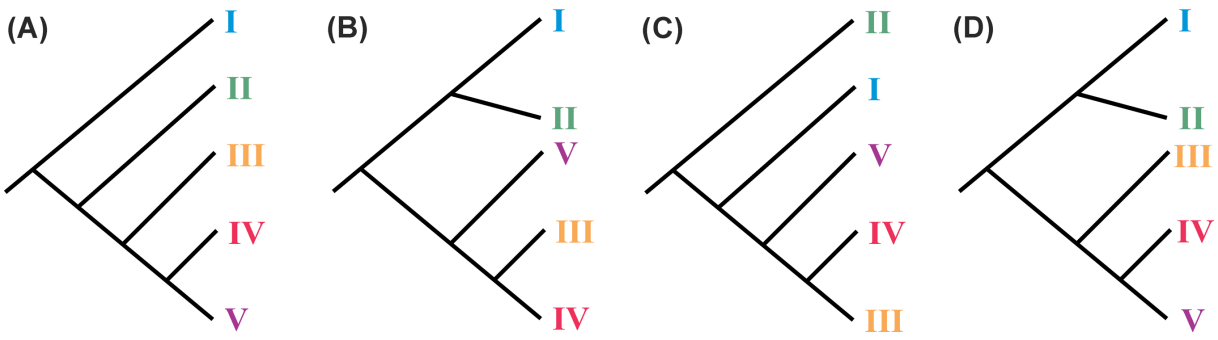


Figure 1. Phylogenetic hypotheses for the relationships among the five major lineages of Sciuridae recovered in (**A**) Mercer and Roth (2003), (**B**) Steppan et al. (2004) and Upham et al. (2019), (**C**) Fabre et al. (2012), and (**D**) Zelditch et al. (2015). Sciuridae subfamilies: (I) Sciurillinae, (II) Ratufinae, (III) Callosciurinae, (IV) Xerinae, (V) Sciurinae.

Two distinct radiations are present in the Neotropics. One of those is represented by the monotypic Neotropical pygmy squirrel, *Sciurillus pusillus*, which diverged early in the history of the family (about 35 Mya) and is included in its own subfamily, Sciurillinae (Mercer and Roth 2003; Steppan et al. 2004; Roth and Mercer 2008). The other lineage, the tribe Sciurini (subfamily Sciurinae), comprises all Neotropical tree squirrels and represents a relatively recent radiation (about 14 Mya), which also includes Nearctic and Eurasian representatives (Mercer and Roth 2003; Steppan et al. 2004; Roth and Mercer 2008).

1.2 Subfamily Sciurillinae

The subfamily Sciurillinae has a single representative, the Neotropical pygmy squirrel, *Sciurillus pusillus* (Thorington et al. 2012). This enigmatic species is one of the oldest lineages amongst all Sciuridae radiations, arising from the base of the squirrels tree, but with

controversial phylogenetic position (Figure 1): Mercer and Roth (2003) recovered it as the sister lineage to all other Sciuridae; in Fabre et al. (2012) Sciurillinae diverged after Ratufinae; and in some other inferences Sciurillinae and Ratufinae appear as sister subfamilies, this clade sister to all other lineages (Steppan et al. 2004; Zelditch et al. 2015; Upham et al. 2019). Due its antiquity and the lack of close relatives, all aspects of the evolutionary history of *Sciurillus*, save its temporal origin in the late Eocene, remain unclear.

Sciurillus pusillus is the smallest species of squirrel occurring in the New Word (head and body length averaging 113 mm and body mass averaging 44 g), and it is also characterized by a set of unique cranial features, as the lateral wall of the skull with squamosal extending dorsally more than halfway between the base of zygomatic process of squamosal and the base of postorbital process of frontal, one transbullar septum, no masseteric tubercle, maxilla contributing more than half of the lateral side of rostrum, and slender zygoma lacking superior process (Vivo and Carmignotto 2015).

This species has a wide geographical range, inhabiting Amazon rainforests of six South American countries: Brazil, Colombia, Guiana, French Guiana, Peru, and Suriname (Vivo and Carmignotto 2015). A taxonomic revision of the genus *Sciurillus* has never been tackled, however, some authors have suggested the existence of two or three subspecies (see Allen 1915; Cabrera 1957; Thorington and Hoffmann 2005) on the basis of morphological (mostly external) variation. Also, Mercer and Roth (2003) reported high genetic divergences in mitochondrial and nuclear genes comparing two specimens from the farthest limits of the species range, one from French Guiana and another from Peru. However, Vivo and Carmignotto (2015) provisionally suggested *S. pusillus* as monotypic, as they were unable to find a clear pattern of morphological variation associated with the geography in the samples available.

1.3 Subfamily Sciurinae, Tribe Sciurini

1.3.1 Taxonomic classifications

The tribe Sciurini have been largely neglected by mammalian systematists, and as a result of such disregard, basic information on the number of genera and species is still very controversial for this group. The taxonomic arrangements presented so far are greatly discrepant among them (see below). Those classifications were exclusive based on phenotypic data, which can be misleading in a group with conservative morphology (Moore

1959). In addition, inconsistencies on species recognition might attest for the extensive variation in coat color, a character that have been widely employed to diagnose taxa of the genus and species groups (Vivo and Carmignotto 2015).

In 1915, Joel A. Allen published the most comprehensive taxonomic revision of the South American squirrels, the culmination of a couple of decades of impressive research describing and organizing the diversity of this group. Allen (1915) recognized the tribe Sciurini as including eight genera and 32 species in South America, and he also recognized another 35 species in eight genera from Central America and North America. Although no other comprehensive revisionary study has been published for this tribe since then, several subsequent authors have adopted different taxonomic arrangements for Sciurini in South America (see Table 1 for the taxa of the genus-group level employed by different authors).

Table 1. Taxa of the genus-group level employed by different authors for tree squirrels from South America.

Allen (1915)	Cabrera (1957)	Moore (1959)	Thorington et al. (2012)	Vivo and Carmignotto (2015)
Guerlinguetus Gray, 1821		*Guerlinguetus* Gray, 1821		*Guerlinguetus* Gray, 1821
Hadrosciurus Allen, 1915				*Hadrosciurus* Allen, 1915
Leptosciurus Allen, 1915				
Mesosciurus Allen, 1915				
Microsciurus Allen, 1895	*Microsciurus* Allen, 1895	*Microsciurus* Allen, 1895	*Microsciurus* Allen, 1895	*Microsciurus* Allen, 1895
Notosciurus Allen, 1914				*Notosciurus* Allen, 1914
Simosciurus Allen, 1915				*Simosciurus* Allen, 1915
Urosciurus Allen, 1915				
	Sciurus Linnaeus, 1758	*Sciurus* Linnaeus, 1758	*Sciurus* Linnaeus, 1758	
				Syntheosciurus Bangs, 1902

Cabrera (1957) in the “Catálogo de los mamíferos de América del Sur” recognized only two genera and 12 species; Moore (1959) along the publication “Relationships among the living squirrels of the Sciurinae” recognized three genera and 12 species; Hoffmann et al. (1993) and Thorington and Hoffmann (2005), in two editions of the book “Mammal Species

of the World", and Thorington et al. (2012) in "The Squirrels of the World" recognized two genera and 14 species. More recently, Vivo and Carmignotto (2015) published a new taxonomic proposal, where they recognize six genera and 18 species for Sciurini in South America—a lower diversity when compared to that proposed by Allen (1915), but greater when compared to subsequent authors, especially at the generic level. The huge discrepancy between these arrangements emphasizes the need for systematic reassessment using independent sources of information, such as genetic data.

1.3.2 Phylogenetic relationships

South American tree squirrels have been suggested as one of the most rapidly diversifying branches of mammals (Roth and Mercer 2008). However, a comprehensive phylogeny is still lacking for this group. Some pioneer molecular phylogenetic studies have included sparse samples of tree squirrels in order to address broader evolutionary questions on the family Sciuridae (Mercer and Roth 2003; Steppan et al. 2004; Roth and Mercer 2008). More recently, larger phylogenetic studies for squirrels (Zelditch et al. 2015), rodents (Fabre et al. 2012), and mammals (Upham et al. 2019) have analyzed a broader sampling of sciurids, including 19 (four of which from South America), 18 (also four from South America), and 24 (seven of which from South America) samples, respectively. Other studies have focused to infer phylogenetic relationships within specific groups of tree squirrels, such as the genus *Tamiasciurus* (Arbogast et al. 2001; Herron et al. 2004; Hope et al. 2016), the Eurasian species of *Sciurus* (Oshida and Masuda 2000; Oshida et al. 2009), and the Mesoamerican tree squirrels from genera *Microsciurus*, *Sciurus*, and *Syntheosciurus* (Villalobos and Cervantes-Reza 2007; Villalobos and Gutierrez-Espeleta 2014).

Pečnerová and Martínková (2012) and Pečnerová et al. (2015) published the most representative phylogenetic hypothesis for Sciurini so far, aiming to address biogeographic and evolutionary questions, but with no regards on the systematics and taxonomy of the group. Still, the most taxonomically and geographic inclusive of those (Pečnerová et al. 2015) comprised only nine samples representing less than one third of the South American species (sensu Vivo and Carmignotto 2015) and it was based on a supermatrix of five genes with over 60% of missing data (Figure 2).

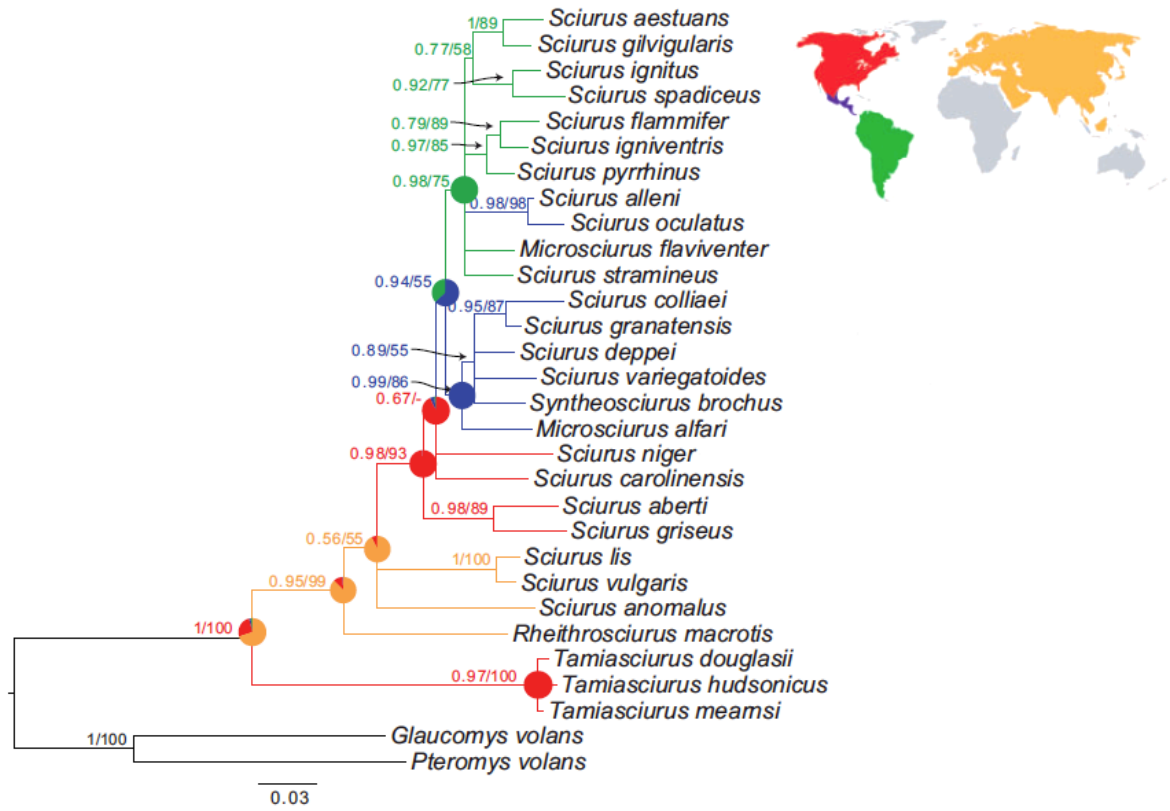


Figure 2. Most representative phylogenetic hypothesis for the tribe Sciurini, modified from Pečnerová et al. (2015). Bayesian inference topology based on supermatrix of eight loci: 12S rRNA, 16S rRNA, MT-CYB, D-loop, MYC exon 2, MYC exon 3, IRBP, and RAG1.

Results of Pečnerová et al. (2015), showed the Neotropical species of tree squirrels (including representatives from South America and Central America) as a sister-group of the Nearctic species of *Sciurus*. Both groups comprising a clade sister to the Palearctic species of *Sciurus*. The first lineage to diverge inside Sciurini was the North American genus *Tamiasciurus*, followed by the Borneo endemic genus *Rheithrosciurus*. Moreover, analyses of Pečnerová et al. (2015), as well previous phylogenetic inferences (Mercer and Roth 2003; Pečnerová and Martínková 2012; Villalobos and Gutierrez-Espeleta 2014) have shown the genera *Sciurus* and *Microsciurus*—as currently recognized (e.g. by Thorington et al. 2012 and Vivo and Carmignotto 2015)— as non-monophyletic. Clearly, both taxonomic and character sampling must be improved to enable a comprehensive phylogenetic hypothesis for tree squirrels, and a robust framework to address further questions on the taxonomic diversity and evolutionary history of this group.

1.3.3 Spatiotemporal diversification

The sister-group relationship between Sciurini and Pteromyini have been recovered in distinct phylogenetic analyzes with strong support (e.g. Mercer and Roth 2003; Steppan et al. 2004; Zelditch et al. 2015), and the origin of this major clade was estimated to about 24-25 Mya (Roth and Mercer 2008; Fabre et al. 2012; Upham et al. 2019). The timing of the origin of the tribe Sciurini is imprecise and has been estimated to have occurred from 19 to 13 million years ago, depending upon the taxonomic coverage and data sets employed (Mercer and Roth, 2003; Fabre et al., 2012; Pečnerová et al., 2015; Zelditch et al., 2015). In addition, the location of their ancestral geographic range is also uncertain, contributing to the blurred inferences of the evolutionary history of tree squirrels. Pečnerová and Martínková (2012) suggested that Sciurini originated in the northern hemisphere, not distinguishing between Eurasia and North America. A similar result was found by Rocha et al. (2016), who proposed a Holarctic distribution for the ancestor of tree squirrels. Subsequently, Pečnerová et al. (2015) estimated a more restricted ancestral range of Sciurini, confined to the Palearctic region.

Pečnerová and Martínková (2012) suggested two alternative hypotheses for the initial diversification and dispersal of tree squirrels. The first one postulates that the ancestor of the genus *Sciurus* dispersed from North America to Eurasia after diversification from *Tamiasciurus*. This ancestor diverged again in Eurasia and returned to North America via the Bering Land Bridge (Figure 3A), subsequently conquering other portions of the new world. The second hypothesis claims that ancestors of both *Tamiasciurus* and *Sciurus* originally occupied Eurasia and colonized the Americas in two distinct occasions, also crossing over Beringia (Figure 3B). However, the results obtained by Pečnerová and Martínková (2012) do not allow them to choose conclusively between these two hypotheses.

The diversification along the Neotropical region, according to Pečnerová and Martínková (2012) and Pečnerová et al. (2015), was probably a directional process following the latitudinal gradient from North America to South America giving rise to all Neotropical taxa. Besides, the entire radiation of South American tree squirrels supposedly descended from a single lineage, which arrived in the continent through the Panamanian Isthmus, about 3 Mya (Mercer and Roth 2003; Steppan et al. 2004; Pečnerová et al. 2015).

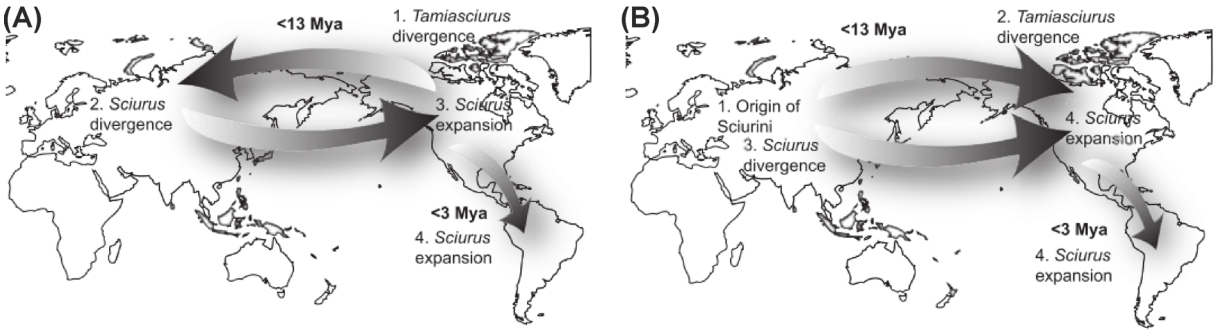


Figure 3. Alternative scenarios proposed by Pečnerová and Martínková (2012) for the initial diversification of Sciurini. (A) Ancestral area in North America and (B) origin in Eurasia.

1.4 The sampling shortfall for Neotropical squirrels

The lack of comprehensive phylogenies for Sciurini is likely a consequence of the difficulties in assessing phylogenetically informative morphological traits in a diverse group, with small sample sizes available in museums that are widely distributed throughout three continents (Thorington et al. 2012). This is also likely due to their conservative cranial morphology, as stated by Moore (1959: 201): "An enumeration and comparison of the taxonomic skull characters of the genera of Sciurinae reveal indications of great conservatism in genera occupying the typical tree squirrel niche, […], whereas genera occupying other sciurine niches appear to have had greater freedom to acquire skull character specializations". Another very plausible reason for the long-term neglect of this diverse group by molecular systematists is the relative rarity of ethanol-preserved/frozen tissues available in scientific collections. In contrast to Nearctic squirrels, Neotropical forms are elusive, trap-shy, and generally restricted to well-preserved forests that can be difficult to access (Emmons and Feer 1997). Conventional traps have proven inefficient to capture tree squirrels, while shotguns have been shown to be much more efficient (Voss and Emmons 1996); although, due to logistical difficulties and permit regulations and restrictions for carrying firearms in many Latin American countries, their use has been virtually abandoned on South American small mammal inventories in the end of the last century and in the current century.

Historical tissue samples (e.g. dried tissue snipped from skins or scraped from skeletal material), on the other hand, can be more readily obtained from museum specimens—which were collected in early expeditions when the use of shotguns for scientific specimen collection was feasible and the commonest sampling method. However, not long ago, DNA derived from historical museum specimens was considered very difficult

to obtain and inefficient to use in large scale phylogenetic analyses, due to the small quantities of degraded DNA that these samples yield, and genetic data was restricted to small fragments of a few mtDNA genes (Hofreiter et al. 2015).

More recently, the techniques for obtaining whole mitogenomes and a variety of nuclear markers from historical museum specimens (also called “Museomics”; see Mandrioli 2016) have undergone significant advances (Rowe et al. 2011; Burrell et al. 2015; McDonough et al. 2018) and historical samples have been demonstrated to be a reliable and effective source of genetic data, especially if applied to next-generation sequencing methods (Miller et al. 2009; Guschanski et al. 2013; Chang et al. 2017). These novel techniques (e.g. shotgun sequencing, targeted sequencing via hybridization-based captures or via restriction enzyme-based enrichment) are highly advantageous for obtaining large-scale genomic data from historical samples in comparison with traditional sequencing methods (e.g. Sanger sequencing; Lemmon and Lemmon 2013), as they are very efficient in sequencing short fragments of DNA, which are expectedly abundant in old museum samples (Burrell et al. 2015).

Therefore, the possibility of successful recovery of consistent genetic data from historical museum samples, reducing the dependency on fresh-tissue samples, provided a promising methodological basis to undertake the molecular systematics of a highly diverse group, widespread across most Neotropical forests, but difficult to obtain samples in current field trips and with scarce preserved tissue samples in scientific collections. In the present thesis, I used mitochondrial genome data and Ultraconserved Elements (UCEs) of the nuclear genome to elucidate several long-standing taxonomic and phylogenetic issues in both radiations of Neotropical squirrels (Sciurillinae and Sciurini). Furthermore, my objectives (see below detailed in “Thesis structure and chapters presentation”) also included investigating some aspects of their phenotypic evolution, diversification, and biogeography. One of the ultimately goals of this thesis was to provide a comprehensive estimate of the diversity of Neotropical squirrels—at genetic and taxonomic (taxa of the genus and species groups) levels—, which could be further used as a sound basis to conservation plans and strategies of this extremely important group for the preservation and maintenance of tropical forests.

1.5 Thesis structure and chapters presentation

Based on the background presented above, which exposed several conflicting and/or not properly addressed issues with regards to the taxonomy, systematics and biogeography of Neotropical squirrels, this thesis was organized in three chapters with specific objectives but together outlining a pioneer effort to tackle their molecular systematics and evolutionary history using genomic data.

CHAPTER 1: Museomics of tree squirrels: a dense taxon sampling of mitogenomes reveals hidden diversity, phenotypic convergence, and the need of a taxonomic overhaul

The main objectives of this chapter were (1) to provide the first phylogenetic hypothesis of Neotropical tree squirrels based on mitochondrial genome data obtained from a dense taxonomic sampling of modern and historical specimens; (2) to contrast the results with generic arrangements proposed for the tribe Sciurini and discuss the taxonomic implications; and (3) to investigate the evolution of number of upper premolars and number of pairs of mammae—two morphological characters traditionally employed for taxonomic classification of tree squirrels. I also took advantage of the large sampling across a diverse lineage of mammals to investigate how distinct aspects of historical samples (e.g. date of collection, museum of provenance, type of sample) might influence mitogenome recovery, and to provide useful information for future genetic studies sampling from dry museum specimens.

CHAPTER 2: Spatiotemporal diversification of tree squirrels: is the South American invasion and speciation really that recent and fast?

In this chapter, my central objective was to test current hypotheses on the tempo and mode of diversification of tree squirrels (Pečnerová and Martínková 2012; Pečnerová et al. 2015) employing a comprehensive dataset (mitogenome data) and a diverse taxonomic coverage (43 putative species of Sciurini). Specific objectives included (1) to provide a time-scaled phylogeny to infer the timing of the origin and diversification of the main lineages; (2) to perform biogeographic analyses to estimate ancestral ranges and evaluate the drivers of diversification at global and Neotropical scales; and (3) to test the hypothesis that tree

squirrels represent a rapid-diversified radiation (Roth and Mercer 2008) by estimating speciation rates and investigating diversification rate fluctuations through time.

CHAPTER 3: Ultraconserved Elements resolve the squirrel tree at deep time scales but deliver lack of phylogenomic consistency in a rapid Neotropical radiation

The last chapter of my thesis was conceived to provide a nuclear genome-wide perspective of the Sciuridae phylogeny, with emphasis on the subfamily Sciurilinnae and on the tribe Sciurini (subfamily Sciurinae), employing over 3,700 UCE loci sequenced from 184 historical and modern museum specimens. In this chapter, I also intended (1) to examine the impact of two dataset filtering approaches (taxa representativeness per UCE loci and proportion of variable sites per UCE loci) in order to verify how matrix completeness and informative content could affect the phylogenetic consistency; and (2) to assess the sensitivity of three conceptually distinct optimality criteria (a concatenated method in RAxML, a coalescent gene-tree based method in ASTRAL-III, and a coalescent site-based method in SVDquartets) to estimate the phylogeny of squirrels. Additionally, I investigated if the sample age (i.e. the year in which the specimen was collected) influenced the number of UCE loci recovered and also the average length of those loci.

References

ALLEN, J. A. 1915. Review of the south american Sciuridae. Bulletin American Museum of Natural History 34:147–309.

ARBOGAST, B. S., R. A. BROWNE, AND P. D. WEIGL. 2001. Evolutionary Genetics and Pleistocene Biogeography of North American Tree Squirrels (*Tamiasciurus*). Journal of Mammalogy 82:302–319.

BENKMAN, C. W. 1995. The impact of tree squirrels (*Tamiasciurus*) on limber pine seed dispersal adaptations. Evolution 49:585–592.

BERTOLINO, S., N. C. DI MONTEZEMOLO, D. G. PREATONI, L. A. WAUTERS, AND A. MARTINOLI. 2014. A grey future for Europe: Sciurus carolinensis is replacing native red squirrels in Italy. Biological Invasions 16:53–62.

BURGIN, C. J., J. P. COLELLA, P. L. KAHN, AND N. S. UPHAM. 2018. How many species of mammals are there? Journal of Mammalogy 99:1–14.

BURRELL, A. S., T. R. DISOTELL, AND C. M. BERGEY. 2015. The use of museum specimens with high-throughput DNA sequencers. Journal of Human Evolution 79:35–44.

CABRERA, A. 1957. Catalogo de los mamideros de America del Sur. Imprenta Y Casa Editora Coni, Buenos Aires.

CHANG, D. ET AL. 2017. The evolutionary and phylogeographic history of woolly mammoths: A comprehensive mitogenomic analysis. Scientific Reports 7:1–10.

EMMONS, L. H., AND F. FEER. 1997. Neotropical Rainforest Mammals: a field guide. Second. The University of Chicago Press, Chicago and London.

EMRY, R. J., W. W. KORTH, AND M. A. BELL. 2005. A Tree Squirrel (Rodentia, Sciuridae, Sciurini) from the Late Miocene (Clarendonian) of Nevada. Journal of Vertebrate Paleontology 25:228–235.

EMRY, R. J., AND R. W. THORINGTON. 1984. The Tree Squirrel *Sciurus* (Sciuridae, Rodentia) as a Living Fossil. Pp. 23–31 in Living Fossils (N. Eldredge & S. M. Stanley, eds.). Springer, New York.

FABRE, P.-H., L. HAUTIER, D. DIMITROV, AND E. J. DOUZERY. 2012. A glimpse on the pattern of rodent diversication: a phylogenetic approach. BMC Evolutionary Biology 12:88.

GOLDSTEIN, E. A., M. J. MERRICK, AND J. L. KOPROWSKI. 2018. Low survival, high predation pressure present conservation challenges for an endangered endemic forest mammal. Biological Conservation 221:67–77.

GURNELL, J., L. A. WAUTERS, P. W. W. LURZ, AND G. TOSI. 2004. Alien species and interspecific competition: Effects of introduced eastern grey squirrels on red squirrel population dynamics. Journal of Animal Ecology 73:26–35.

GUSCHANSKI, K. ET AL. 2013. Next-generation museomics disentangles one of the largest primate radiations. Systematic Biology 62:539–554.

HERRON, M. D., T. A. CASTOE, AND C. L. PARKINSON. 2004. Sciurid phylogeny and the paraphyly of Holarctic ground squirrels (*Spermophilus*). Molecular Phylogenetics and Evolution 31:1015–1030.

HOFFMANN, R. S., C. G. ANDERSON, R. W. J. THORINGTON, AND L. R. HEANEY. 1993. Family Sciuridae. Pp. 419–465 in Mammals Species of the World: a taxonomic and geographic reference (D. E. Wilson & D. M. Reeder, eds.). Second. Smithsonian Institute Press, Washington DC.

HOFREITER, M. ET AL. 2015. The future of ancient DNA: Technical advances and conceptual shifts. BioEssays 37:284–293.

HOPE, A. G. ET AL. 2016. Revision of widespread red squirrels (genus: *Tamiasciurus*) highlights the complexity of speciation within North American forests. Molecular Phylogenetics and Evolution 100:170–182.

KETCHAM, S. L., J. L. KOPROWSKI, AND D. A. FALK. 2017. Differential Response of Native Arizona Gray Squirrels and Introduced Abert's Squirrels to a Mosaic of Burn Severities. Mammal Study 42:1–12.

LEMMON, E. M., AND A. R. LEMMON. 2013. High-Throughput Genomic Data in Systematics and Phylogenetics. Annual Review of Ecology, Evolution, and Systematics 44:99–121.

MANDRIOLI, M. 2016. MUSEomica: quando la genomica entra in museo. Quaderni del Museo Civico di Storia Naturale di Ferrara 4:53–70.

MCDONOUGH, M. M., L. D. PARKER, N. ROTZEL MCINERNEY, M. G. CAMPANA, AND J. E. MALDONADO. 2018. Performance of commonly requested destructive museum samples for mammalian genomic studies. Journal of Mammalogy 99:789–802.

MENDES, C. P., AND J. F. CÂNDIDO. 2014. Behavior and foraging technique of the Ingram's squirrel *Guerlinguetus ingrami* (Sciuridae: Rodentia) in an Araucaria moist forest fragment. Zoologia 31:209–214.

MENDES, C. P., J. L. KOPROWSKI, AND M. GALETTI. 2019. NEOSQUIRREL: a data set of ecological knowledge on Neotropical squirrels. Mammal Review 49:210–225.

MERCER, J. M., AND V. L. ROTH. 2003. The Effects of Cenozoic Global Change on Squirrel Phylogeny. Science 299:1568–1572.

MILLER, W. ET AL. 2009. The mitochondrial genome sequence of the Tasmanian tiger (*Thylacinus cynocephalus*). Genome Research 19:213–220.

MOORE, J. C. 1959. Relationships among the living squirrels of the Sciurinae. Bulletin on the American Museum of Natural History 118:157–206.

OSHIDA, T., A. ARSLAN, AND M. NODA. 2009. Phylogenetic relationships among the old world *Sciurus* squirrels. Folia Zoologica 58:14–25.

OSHIDA, T., AND R. MASUDA. 2000. Phylogeny and Zoogeography of Six Squirrel Species of the Genus *Sciurus* (Mammalia, Rodentia), Inferred from Cytochrome b Gene Sequences. Zoological Science 17:405–409.

PEČNEROVÁ, P., AND N. MARTÍNKOVÁ. 2012. Evolutionary history of tree squirrels (Rodentia, Sciurini) based on multilocus phylogeny reconstruction. Zoologica Scripta 41:211–219.

PEČNEROVÁ, P., J. C. MORAVEC, AND N. MARTÍNKOVÁ. 2015. A skull might lie: Modeling ancestral ranges and diet from genes and shape of tree squirrels. Systematic Biology 64:1074–1088.

PERES, C. A. 1999. The structure of nonvolant mammal communities in different Amazonian forest types. Pp. 564–581 in Mammals of the Neotropics. Vol. 3. The central Neotropics: Ecuador, Peru, Bolivia, Brazil (J. F. EISENBERG & K. H. REDFORD, eds.). The University of Chicago Press, Chicago and London.

ROCHA, R. G., Y. L. R. LEITE, L. P. COSTA, AND D. ROJAS. 2016. Independent reversals to terrestriality in squirrels (Rodentia: Sciuridae) support ecologically mediated modes of adaptation. Journal of Evolutionary Biology 29:2471–2479.

ROTH, V. L., AND J. M. MERCER. 2008. Differing rates of macroevolutionary diversification in arboreal squirrels. Current Science 95:857–861.

ROWE, K. C. ET AL. 2011. Museum genomics: Low-cost and high-accuracy genetic data from historical specimens. Molecular Ecology Resources 11:1082–1092.

STAPANIAN, M. A., AND C. C. SMITH. 1978. A Model for Seed Scatterhoarding: Coevolution of Fox Squirrels and Black Walnuts. Ecology 59:884–896.

STEELE, M., L. A. WAUTERS, AND K. W. LARSEN. 2005. Selection, Predation and Dispersal of Seeds by Tree Squirrels in Temperate and Boreal Forests: are Tree Squirrels Keystone Granivores? Pp. 205–222 in Seed Fate: Predation, Dispersal and Seedling Establishment (P.-M. Forget, J. E. Lambert, P. E. Hulme & S. B. Vander Wall, eds.). CABI Publishing, Cambridge.

STEPPAN, S. J., B. L. STORZ, AND R. S. HOFFMANN. 2004. Nuclear DNA phylogeny of the squirrels (Mammalia: Rodentia) and the evolution of arboreality from c-myc and RAG1. Molecular Phylogenetics and Evolution 30:703–719.

THORINGTON, R. W. J., AND R. S. HOFFMANN. 2005. Family Sciuridae. Pp. 754–818 in Mammals Species of the World: a taxonomic and geographic reference (D. E. Wilson & D. M. Reeder, eds.). Third. Jhons Hopikins University Press, Washington DC.

THORINGTON, R. W., J. L. KOPROWSKI, M. A. STEELE, AND J. WHATTON. 2012. Squirrels of the World. The Johns Hopkins University Press, Baltimore.

UPHAM, N. S., J. A. ESSELSTYN, AND W. JETZ. 2019. Inferring the mammal tree: Species-level sets of phylogenies for questions in ecology, evolution, and conservation. PLOS Biology 17:e3000494.

VILLALOBOS, F., AND F. CERVANTES-REZA. 2007. Phylogenetic relationships of Mesoamerican species of the genus *Sciurus* (Rodentia: Sciuridae). Zootaxa 40:31–40.

VILLALOBOS, F., AND G. GUTIERREZ-ESPELETA. 2014. Mesoamerican tree squirrels evolution (Rodentia: Sciuridae): a molecular phylogenetic analysis. Revista de Biologia Tropical 62:649–657.

VIVO, M., AND A. P. CARMIGNOTTO. 2015. Family Sciuridae G. Fischer, 1817. Pp. 1–48 in Mammals of South America, Volume 2, Rodents (J. L. Patton, U. F. J. Pardiñas & G. D'Elía, eds.). The University of Chicago Press, Chicago and London.

VOSS, R. S., AND L. H. EMMONS. 1996. Mammalian Diversity in Neotropical Lowland Rainforests: a Preliminary Assessment. Bulletin of the American Museum of Natural History 230:1–115.

ZELDITCH, M. L., J. LI, L. A. P. TRAN, AND D. L. SWIDERSKI. 2015. Relationships of diversity, disparity, and their evolutionary rates in squirrels (Sciuridae). Evolution 69:1284–1300.

2. CHAPTER 1: Museomics of tree squirrels: a dense taxon sampling of mitogenomes reveals hidden diversity, phenotypic convergence, and the need of a taxonomic overhaul

Guerlinguetus brasiliensis (Credit: Pedro Peloso)

Published

Abreu-Jr, E.F., Pavan, S.E., Tsuchiya, M.T.N., Wilson, D., Percequillo, A.R., and Maldonado, J.E. 2020. Museomics of tree squirrels: a dense taxon sampling of mitogenomes reveals hidden diversity, phenotypic convergence, and the need of a taxonomic overhaul. *BMC Evolutionary Biology* 20(77): 1–25. https://doi.org/10.1186/s12862-020-01639-y

Museomics of tree squirrels: a dense taxon sampling of mitogenomes reveals hidden diversity, phenotypic convergence, and the need of a taxonomic overhaul

Abstract

Background: Tree squirrels (Sciuridae, Sciurini), in particular the highly diverse Neotropical lineages, are amongst the most rapidly diversifying branches of the mammal tree of life but also some of the least known. Negligence of this group by systematists is likely a product of the difficulties in assessing morphological informative traits and of the scarcity or unavailability of fresh tissue samples for DNA sequencing. The highly discrepant taxonomic arrangements are a consequence of the lack of phylogenies and the exclusive phenotypic-based classifications, which can be misleading in a group with conservative morphology. Here we used high-throughput sequencing and an unprecedented sampling of museum specimens to provide the first comprehensive phylogeny of tree squirrels, with a special emphasis on Neotropical taxa.
Results: We obtained complete or partial mitochondrial genomes from 232 historical and modern samples, representing 40 of the 43 currently recognized species of Sciurini. Our phylogenetic analyses—performed with datasets differing on levels of missing data and taxa under distinct analytical methods—strongly support the monophyly of Sciurini and consistently recovered 12 major clades within the tribe. We found evidence that the diversity of Neotropical tree squirrels is underestimated, with at least six lineages that represent taxa to be named or revalidated. Ancestral state reconstructions of number of upper premolars and number of mammae indicated that alternative conditions of both characters must have evolved multiple times throughout the evolutionary history of tree squirrels.
Conclusions: Complete mitogenomes were obtained from museum specimens as old as 120 years, reinforcing the potential of historical samples for phylogenetic inferences of elusive lineages of the tree of life. None of the taxonomic arrangements ever proposed for tree squirrels fully corresponded to our phylogenetic reconstruction, with only a few of the currently recognized genera recovered as monophyletic. By investigating the evolution of two morphological traits widely employed in the taxonomy of the group, we revealed that their homoplastic nature can help explain the incongruence between phylogenetic results and the classification schemes presented so far. Based on our phylogenetic results we suggest a tentative supraspecific taxonomic arrangement for Sciurini, employing 13 generic names used in previous taxonomic classifications.

Keywords: Historical DNA, Morphology, Neotropical region, Phylogeny, Sciuridae, Systematics.

2.1 Background

Squirrels (Sciuridae) comprise the third most diverse family of rodents, with about 60 genera and 300 species organized in five subfamilies [1–3]. In the Neotropics, squirrels are

inhabitants of all forest biomes [1, 2], and crucial to ecosystem dynamics as they play a vital role in seed predation and dispersal [4, 5]. However, in contrast to other widespread Neotropical rodent groups, squirrels have been largely neglected in phylogenetic studies. South American (SA) tree squirrels (tribe Sciurini) have been suggested as one of the most rapidly diversifying branches of mammals [6]. Still, the most representative phylogenetic hypothesis [7] included only nine samples representing less than one third of the SA species (sensu [2]) and it was based on a supermatrix of five genes with over 60% of missing data. As a result, basic knowledge on phylogenetic relationships is lacking for Neotropical lineages and their evolutionary history remains unraveled.

The lack of comprehensive phylogenies is likely a consequence of the difficulties in assessing phylogenetically informative morphological traits in a diverse group, with small sample sizes available in museums that are widely distributed throughout three continents [1]. This is also likely due to their conservative cranial morphology, as stated by Moore [21: 201]: “An enumeration and comparison of the taxonomic skull characters of the genera of Sciurinae reveal indications of great conservatism in genera occupying the typical tree squirrel niche, […], whereas genera occupying other sciurine niches appear to have had greater freedom to acquire skull character specializations”. Another very plausible reason for the long-term neglect of this diverse group by molecular systematists is the relative rarity of ethanol-preserved/frozen tissues available in scientific collections. In contrast to Nearctic squirrels, Neotropical forms are elusive, trap-shy, and generally restricted to well-preserved forests that can be difficult to access [8]. Conventional traps have proven inefficient to capture tree squirrels, while shotguns have been shown to be much more efficient [9]; although, due to logistical difficulties and permit regulations and restrictions for carrying firearms in many Latin American countries, their use has been virtually abandoned on SA small mammal surveys in the current century.

Historical tissue samples (e.g. dried tissue snipped from skins or scraped from skeletal material), on the other hand, can be more readily obtained from museum specimens—which were collected in early expeditions when the use of shotguns for scientific specimen collection was feasible and the commonest sampling method. However, not long ago, DNA derived from historical museum specimens was considered very difficult to obtain and inefficient to use in large scale phylogenetic analyses, due to the small quantities of degraded DNA that these samples yield, and genetic data was restricted to

small fragments of a few mtDNA genes [10]. More recently, the techniques for obtaining whole mitogenomes from historical museum specimens (also called “Museomics”; see [11]) have undergone significant advances [12–14] and historical samples have been demonstrated to be a reliable and effective source of genetic data, especially if applied to next-generation sequencing methods (e.g. shotgun sequencing, targeted sequencing via hybridization-based captures or via restriction enzyme-based enrichment), which are very efficient in massively sequencing fragmented DNA [15–17].

Lack of knowledge of phylogenetic relationships resulted in highly discrepant taxonomic arrangements proposed for the group. In 1915, Joel A. Allen published the most comprehensive taxonomic revision of the SA squirrels. This work was the culmination of a couple of decades of impressive research describing and organizing the diversity of New World squirrels. Allen [18] recognized the tribe Sciurini as including eight genera and 32 species in SA, and he also recognized another 35 species in eight genera from Central America (CA) and North America (NA). Although no other comprehensive revisionary study has been published for this tribe since then, several subsequent authors have adopted different taxonomic arrangements that recognized two to four genera, and 12 to 14 species of Sciurini in SA (e.g. [1, 19–22]). More recently, Vivo and Carmignotto [2] published a new taxonomic proposal, where they recognize six genera and 18 species for SA Sciurini—a lower diversity when compared to that proposed by Allen [18], but greater when compared to subsequent authors, especially at the generic level. The huge discrepancy between the arrangements, all based exclusively on morphological data, attests to the extensive variation—and potential homoplastic nature—of the characters traditionally employed in taxonomy, and evidences the need for systematic reassessment using independent sources of information, such as genetic data.

Here we report the results of our analyses of mitochondrial genome data obtained from a combination of ethanol-preserved tissue samples (also referred to as “modern” samples) and tissue samples obtained from dry museum specimens (hereafter “historical” samples) representing most of the nominal taxa recognized as valid species of tree squirrels (Sciurini). We (1) provide the first phylogenetic hypothesis of Neotropical tree squirrels based on dense taxonomic sampling and on state-of-the-art methods of data generation and phylogenetic reconstruction; (2) contrast our results with generic arrangements proposed for the tribe and discuss the taxonomic implications; and (3) investigate the evolution of

number of upper premolars and number of pairs of mammae—two morphological characters traditionally employed for taxonomic classification of tree squirrels. We take advantage of our large sampling across a diverse lineage of mammals to investigate how distinct aspects of historical samples (e.g. date of collection, museum of provenance, type of sample) might influence mitogenome recovery, and to provide useful information for future genetic studies sampling from dry museum specimens.

2.2 Results

2.2.1 Summary of mitochondrial genome sequencing success, assembly and synteny

From the 271 samples that we attempted to sequence, complete mitochondrial genomes were recovered for 92 samples, partial mitogenomes with variable percentages of missing data were recovered for 172 samples, and no mtDNA sequences were recovered for seven samples (Table 1). Modern samples yield, unsurprisingly, genomes more complete than historical samples, with full mitochondrial genomes recovered for almost half of modern samples (78 out of 177), but for only about 1/6 of historical samples (14 out of 94). Even though ethanol-preserved tissues yield more complete mitogenomes, we were able to obtain partial mitogenomes from most historical samples—for which ethanol-preserved tissues were not available—producing an increment of 18 nominal taxa to our phylogenetic datasets.

When contrasting mitogenome recovery success (completeness) of historical samples with tissue type and museum location, we note that, on average, remains of muscular tissue adherent to skulls (“osteocrusts”) yielded more complete mitogenomes than skin clips, and samples from NA collections yielded more complete mitogenomes than samples from SA collections, although none of those differences were significant (X^2 = 60.114, P = 0.6146 and X^2 = 68.398, P = 0.3304, respectively; Additional file 1). We found no correlation between sample age (the year in which the specimen was collected) and completeness of mitogenomes recovered for historical samples (R^2 = 0.0028, P = 0.6762; Figure 1). We were able to obtain complete mitogenomes for specimens as old as 120 years, and partial mitogenomes with over 20% of completeness for specimens as old as 126 years.

The complete assembled mitogenomes were circular molecules with length ranging from 16,501 to 16,535 pb. For all species analyzed, including ingroup (tribe Sciurini) and outgroup (tribe Pteromyini) taxa, mitogenomes presented identical synteny, comprising 13

protein-coding genes (PCGs), two ribosomal RNA genes (rRNA), and 22 transfer RNA genes (tRNA). The GC-content ranged from 36.6 to 38.9%. The annotated circular genome of *Guerlinguetus brasiliensis* is depicted in Figure 2 to exemplify the gene organization in the group. This is the first complete mitochondrial genome published for a Neotropical squirrel.

2.2.2 Phylogenetic inferences and the effect of missing data

The resulting matrices of Datasets 1–5 included 92 to 232 specimens representing 27 to 43 Operational Taxonomic Units (OTUs). This represents the most conservative hypothesis at the species level that reconciles the initial morphological identifications and our phylogenetic results—i.e. monophyletic groups (see OTUs designation in the section "Species monophyly and species recognition"). The overall missing data ranged from 0.1–20.2%. Characteristics of each of the five matrices with different levels of missing taxa and data are summarized in Table 2. Comparisons of phylogenetic trees inferred from those datasets does not indicate that the inclusion of specimens with partial mitogenomes (missing data) produced any strong topological incongruences (Figure 3). Except for the phylogenetic position of *Neosciurus carolinensis* (see details below), the topologies recovered with Datasets 1–5 are similar. Most of the nodes were strongly supported in all of our dataset analyses, and the inclusion of specimens with up to 80% missing data had no overall impact on the nodal support of the inferred ML phylogenies, as there were no significant differences on the average Bootstrap values recovered (Wilcoxon signed-rank test $P > 0.05$).

Considering the aforementioned results, and that including as many specimens as possible would be the best option to unveil all mitochondrial lineages, the tree recovered by the Maximum Likelihood (ML) analysis of Dataset 5 (the most taxonomically comprehensive dataset, with 232 specimens) was chosen to represent the mitochondrial phylogenetic hypothesis of Sciurini. This hypothesis is shown in detail in Figure 4 with accompanying nodal support from bootstrap replicates. Bayesian inference of Dataset 5 recovered a topology similar to the one recovered by ML analysis of the same matrix. Best-fitting models of sequence evolution used on the BI are summarized in Additional file 2. Nodal support recovered as posterior probabilities by the Bayesian Inference (BI) of Dataset 5 is also shown in Figure 4.

Since our various analyses including different numbers of taxa recovered similar topologies, our results support essentially the same conclusions discussed below. Most of the differences recovered are related to the sampling difference, when species represented in some datasets were not represented in others. Any inconsistencies recovered by the different analyses are mentioned below, when appropriate.

2.2.3 A mitogenomic hypothesis: Sciurini and its species groups

The genus-level classification of Sciurini is greatly discordant among all previous taxonomic arrangements proposed for the group, and none of those classifications fully matched the phylogenies obtained by our optimality criteria. Therefore, to avoid more confusion on the use of genus-level names, we deliberately omit the generic epithet of all taxa when describing our results below, and refer the species as members of species groups (identified by capital letters), as recovered by our phylogenetic analyses. A tentative genus-level classification of Sciurini is presented in the discussion (see "Mitochondrial phylogeny and taxonomic arrangements proposed for Sciurini"), supported by our results and by other relevant taxonomic information.

Our results recovered the tribe Sciurini as a monophyletic group with full nodal support. Within Sciurini we recognize 12 major groups, A–L (Figure 4), that have been consistently recovered by all of our analyses where they were represented (Figure 3). Except for Group F, which is composed of a single specimen, the groups recognized here represent clades, all of which were recovered with high statistical support on our inferred phylogenies (ML bootstrap ≥ 75% and BI posterior probability ≥ 0.95).

Group A is represented in our analyses by the North American nominal-taxa *hudsonicus* Erxleben, 1777 and *douglasii* Bachman, 1839. Group B is monotypic, consisting of *macrotis* Gray, 1856 from Borneo. Group C includes Eurasian radiations: *anomalus* Gmelin, 1778, *lis* Temminck, 1844, and *vulgaris* Linnaeus, 1758. The three subsequent groups contain mostly North American lineages, some of which also reach Central America: Group D includes *aberti* Woodhouse, 1852 and *griseus* Ord, 1818; Group E includes *arizonensis* Coues, 1867, *nayaritensis* J. A. Allen, 1890, *niger* Linnaeus, 1758, *alleni* Nelson, 1898, and *oculatus* Peters, 1863; and Group F is composed of a single representative of *carolinensis* Gmelin, 1788. While this taxon is broadly distributed in the eastern and midwestern USA, our sample comes from the midwestern USA (Figure 4a).

Groups G–L contain all Neotropical lineages. Group G is composed of southern North American, Central American, and northern trans-Andean South American forms currently assigned to at least nine species or species complexes (Figures 4a and 5a). This geographically and taxonomically inclusive clade contains the nominal taxa *alfari* J. A. Allen, 1895 from Panama, *venustulus* Goldman, 1912 from Panama*, brochus* Bangs, 1902 from Costa Rica, *granatensis* Humboldt, 1811 from Peru, Ecuador, Colombia, Panama, Venezuela, Trinidad and Tobago, and Nicaragua, *richmondi* Nelson, 1898 from Nicaragua, *aureogaster* F. Cuvier, 1829 from Guatemala and Mexico, *colliaei* Richardson, 1839 from Mexico, *deppei* Peters, 1863 from Guatemala and Nicaragua, *yucatanensis* J. A. Allen, 1877 from Mexico, and *variegatoides* Ogilby, 1839 from Panama. Besides these taxa, Group G includes a highly divergent lineage composed of a single specimen from Chocó, Colombia, which could not be assigned to any valid species ("species 1" in Figure 4a).

Groups H–L are exclusively composed of specimens from South America and southern Panama (Figures 4b–c and 5b–d). Group H comprises specimens from mountain areas in northwestern South America allocated to six species according to Vivo and Carmignotto [2]: *mimulus* Thomas, 1898 with specimens from Colombia and Ecuador; *similis* Nelson, 1899, *otinus* Thomas, 1901 and typical *pucheranii* Fitzinger, 1867 from Colombia; *boquetensis* Nelson, 1903 from Panama; and *isthmius* Nelson, 1899 from Colombia and Panama. Group I includes two species: *nebouxii* I. Geoffroy St. Hilaire, 1855 from the coast of Ecuador, and *stramineus* Eydoux and Souleyet, 1841 from the coast of Peru.

Group J is composed of specimens distributed throughout the Atlantic Forest along the eastern coast of Brazil, and in Amazonia including Brazil, French Guiana, Guyana, and Venezuela. This clade includes two nominal taxa: *aestuans* Linnaeus, 1766 and *brasiliensis* Gmelin, 1788. Group K includes two named species: *flaviventer* Gray, 1867 from Amazonian lowlands of Brazil, Bolivia, Ecuador, and Peru; and *sabanillae* Anthony, 1922 including samples from the Amazonian lowlands of Peru to the eastern Andean mid-elevations of Ecuador. Additionally, Group K includes what seems to be another unnamed lineage, composed of specimens from the Brazilian and Peruvian Amazon, and from mid-elevation of the eastern Andes in northern Peru ("species 2" in Figure 4c).

Group L includes specimens associated with *igniventris* Wagner, 1842, *pyrrhinus* Thomas, 1898, and *spadiceus* Olfers, 1818 from lowland Amazonia in Brazil, Bolivia, Peru, Ecuador, and Colombia, and a putative unnamed lineage composed of three specimens from

Loreto, Peru ("species 3" in Figure 4c). In addition to those, Group L includes specimens assigned to *pucheranii* Fitzinger, 1867 by Vivo and Carmignotto [2] as subspecies: *pucheranii ignitus*—specimens from Brazil and Peru—, *pucheranii argentinius*—samples from Argentina— and *pucheranii boliviensis*—samples from Bolivia.

Our analyses recovered similar relationships amongst the 12 major groups within Sciurini, apart from Group F. This lineage, represented in our analyses by a single specimen of *carolinensis*, was recovered as sister to Group E by the ML analysis of Dataset 3 (Figure 3c) and the BI of Dataset 5 (not shown), or as sister to clades G–L (all Central and South American squirrels) in all remaining analyses (Figures 3a, 3b, 3d, 3e). However, the phylogenetic position of Group F was weakly supported (bootstrap < 75%, posterior probability < 0.95) in all our inferences. Amongst the intergroup relationships unanimously recovered by our analyses with high statistical support, we highlight the monophyly of the Neotropical forms (Groups G–L) and of the group that exclusively includes squirrels from South America and southern Panama (Groups H–L).

Except for comparisons between datasets with unrepresented taxa, interspecific and intraspecific relationships within the multispecies groups of Sciurini were similar in all ML analysis performed (see Figure 3). Bayesian analysis of Dataset 5 also recovered similar results, apart from relationships within Group E, where the node for *alleni*, *oculatus* and *niger* was unresolved in a polytomy sister to *nayaritensis* + *arizonensis*.

2.2.4 Species monophyly and species recognition

Our phylogenetic results corroborate most currently recognized species of Sciurini as highly supported clades (when represented by more than one specimen). However, some valid species appear nested within others, while other nominal taxa seem to include several distinct genetic lineages that do not comprise a monophyletic group. For example, a specimen originally identified as *richmondi* (USNM 48689) is nested within the clade associated with *granatensis* (Figure 4a), an individual assigned to *venustulus* (USNM 338164) is nested within the clade associated with *alfari* (Figure 4a), specimens identified as *pucheranii* (USNM 271319 and USNM 293776 in Group H, and MJ302–LHE1312 in Group L) are recovered in two phylogenetically distant lineages (Figures 4a and 4c), and samples originally identified as *aestuans* (USNM 599925–CN158 in Group J) and *brasiliensis* (AMNH

36489 and EFA41–MTR0444 in Group J) do not compose reciprocally monophyletic groups (Figure 4b).

Based on those results, we reassigned the 232 specimens used in our phylogenetic inferences to OTUs, i.e. monophyletic groups (shown along Figure 4, first column). The most relevant differences between our OTU designations and the original identifications based on Thorington et al. [1] and Vivo and Carmignotto [2] are the recognition of: a single OTU (*granatensis*) for *granatensis* and *richmondi*; a single OTU (*alfari*) for *alfari* and *venustulus*, two distinct OTUs (*pucheranii* and *ignitus*) for *pucheranii*; and three OTUs (*aestuans* "a", *aestuans* "b", *aestuans* "c" [including the specimen from Pernambuco, AMNH 36849]) for samples originally assigned to *aestuans*. Additionally, our phylogenetic analyses recovered some apparently unnamed lineages that we also consider as distinct OTUs, referred to as "species 1" (from Group G), "species 2" (from Group K), and "species 3" (from Group L; Figure 4). In total, we recognized 43 OTUs, some of which are composed of deeply divergent mitochondrial lineages that seem to merit further investigation (e.g. *granatensis*, *flaviventer* and "species 2").

As the specific status of some nominal taxa seems questionable, and some apparently unnamed species were suggested by our phylogenetic inferences, we used species delimitation analyses to provide a quantitative evaluation of the species limits within Sciurini. GMYC analyses using as input ultrametric trees generated with strict (GMYC 1) and relaxed (GMYC 2) molecular clocks resulted in highly distinct species scenarios, suggesting 66 and 39 species within Sciurini, respectively. GMYC 1 analysis failed to recognize six OTUs as distinct species, four from North America—(*douglasii*, *hudsonicus*), (*arizonensis*, *nayaritensis*)—and two from South America—(*boquetensis*, *isthmius*). This analysis also suggested an additional 26 putative species (Figure 4), aside from the original 43 OTUs.

GMYC 2 analysis provided a more conservative scenario and did not recognize 13 Eurasian, North American, and Central American OTUs—(*douglasii*, *hudsonicus*), (*lis*, *vulgaris*), (*arizonensis*, *nayaritensis*), and (*niger* (*alleni*, *oculatus*)), (*aureogaster*, *colliaei*) and (*yucatanensis*, *variegatoides*)—and six South American OTUs—(*similis* (*otinus* (*boquetensis*, *isthmius*))) and (*aestuans* "c", *brasiliensis*) as distinct species. On the other hand, this analysis suggested additional seven putative species in three species complexes associated with the Neotropical OTUs *granatensis*, "species 2", and *flaviventer* (Figure 4).

BBP analysis recovered full support (PP = 1) for the recognition of 29 species and did not recover significant support (PP > 0.95) for the recognition of 14 species. All of the non-supported species are from Eurasia, North and Central Americas and are included in five lineages: (*douglasii*, *hudsonicus*), (*anomalus*, *lis*, *vulgaris*), (*arizonensis*, *nayaritensis*), (*alleni*, *oculatus*) and (*aureogaster*, *colliaei*, *deppei*, *yucatanensis*, *variegatoides*).

Despite the differences observed in the results of species delimitation analyses, the most noteworthy result was the consistent support of all South American OTUs as distinct species by at least two analyses. The only exception is *boquetensis* and *isthmius*, which were not supported as distinct species by the GMYC analyses using both strict and relaxed clock generated chronograms, but they were recovered as distinct species with full support by BPP analysis. On the other hand, some Eurasian, North American and Central American OTUs were consistently not supported as distinct species (see Figure 4). For the taxa of non-South American lineages, we have a limited number of specimens in our analysis (one specimen per OTU in most cases), and this might be blurring the species delimitation analyses.

Based on our phylogenetic inferences and species delimitation analyses, as well as available information in the literature (e.g. phenotypic and karyotypic data, analyses of geographic variation and previous phylogenetic and species delimitation analyses; [1, 2, 21, 23–27]), we recognized the initial set of 43 OTUs as putative species of Sciurini (Figure 6), all of which are treated as distinct terminal taxa in the following analyses. From those 43 OTUs, 37 represent described species recognized as valid by the latest taxonomic hypothesis comprising those taxa ([1], [2], [43]; but *granatensis* includes *richmondi* and *alfari* includes *venustulus*), while six are putative additional species to be described or revalidated ("species 1–3", *aestuans* "a–b", and *ignitus*—currently considered as a subspecies of *pucherani* by [2], but meriting specific status as per our results). For detailed justification for taxonomic decisions and name usage, please see the discussion section "Comments on species recognition and novelties".

2.2.5 The evolution of premolars and mammae within Sciurini

The best-fitting model to explain the evolution of number of premolars within Sciurini was the Mk with equal rates (AICw = 0.686; Additional file 3), which suggests that transitions between states occurred at the same rate and with equal probabilities. For the number of pairs of mammae, the most supported model was Mk with symmetric rates (AICw = 0.763;

Additional file 4), suggesting that changes between states had equal probabilities regardless of direction, but differed in rates. Ancestral state reconstructions suggest that the most recent common ancestor (MRCA) of Sciurini had, most likely, two upper premolars (P = 0.53) and four pairs of mammae (P = 0.56; Figure 5). However, alternative states for both characters were recovered with nearly equal probabilities for this node, indicating that the MRCA of Sciurini might also have had one upper premolar (P = 0.47) and three pairs of mammae (P = 0.44). Despite the uncertainty regarding the deepest nodes of Sciurini, several of the major clades recognized within the tribe exhibited unambiguous optimizations, indicating that conditions of both characters must have evolved multiple times during the evolutionary history of tree squirrels.

Two upper premolars were likely (P > 0.70) for the MRCA of six major Groups (A, C, D, G, H, and K), while one upper premolar was likely (P > 0.70) for the MRCAs of other four Groups (E, I, J, and L). Considering the estimates within the major Groups of Sciurini, the loss of one upper premolar seem to have happened at least three independent times (*anomalus* in C, *granatensis* in G, and *pucheranii* in H). On the other hand, the MRCA of the clade formed by Groups I, J, K, and L likely had one premolar (P > 0.70), and the condition of Group K, two premolars, could be interpreted as a new gain or a return to an ancestral state. Regarding the number of mammae, three pairs were likely (P > 0.70) present on the MRCA of four Groups (C, G, H, and K), while four pairs were likely (P > 0.70) present on the MRCA of six Groups (A, D, E, I, J, and L). Considering the estimates within the major Groups of Sciurini, several changes in the number of pairs of mammae are evident, with at least three independent changes from three to four/five pairs (*anomalus* and *vulgaris* in C, and MRCA of *aureogaster*, *colliaei*, *deppei*, *yucatanensis*, and *variegatoides* in G) and two independent transitions from four to three pairs (*deppei* in G and *ignitus* in L).

2.3 Discussion

2.3.1 The importance of museum specimens for the study of Neotropical tree squirrels

All samples used on this study were gathered from specimens deposited in scientific collections. The inclusion of historical samples was crucial in the detection of several taxonomic issues reported here. Almost a third of our samples were obtained from dry museum specimens, collected between 1893 and 2010, and housed mostly in two North American museums (AMNH and USNM). We were successful in obtaining at least 20% of the

mitogenome for about 70% of the historical samples, which allowed us to include 18 nominal taxa in the Sciurini phylogeny for which ethanol-preserved tissues were not available. Moreover, two of the main groups recognized within Sciurini (B and H) were exclusively represented by historical samples.

Our success in obtaining mtDNA data from historical samples is twofold: i) it could be partially attributed to the sequencing method employed. Next-generation sequencing techniques (e.g. shotgun sequencing, targeted sequencing via hybridization-based captures or via restriction enzyme-based enrichment [28]) are highly advantageous for obtaining large-scale genomic data from historical samples in comparison with traditional sequencing methods (e.g. Sanger sequencing), as they are very efficient in sequencing short fragments of DNA [14], which are expectedly abundant in old museum samples; ii) it could also be a consequence of our sampling strategy. We prioritized obtaining fragments of muscular tissue adhered to skulls, which had been shown by previous studies [12] and confirmed here, to yield higher concentrations and longer fragments of DNA than samples obtained from skins clips.

The use of historical samples in molecular studies has increased substantially in the last two decades (see review in [14, 29]) and is helping to reveal hidden diversity, to unveil puzzling phylogenetic relationships, and to place rare and elusive mammal species/lineages in a phylogenetic context [16, 30–32], including in squirrels [33]. For Neotropical mammal groups, the study of historical museum samples using high-throughput sequencing to address phylogenetic and taxonomic questions is growing, becoming more feasible, and holding lots of promise [34, 35].

When investigating how different aspects of historical samples of Sciurini might influence the success of mitogenome recovery, our most robust result was that sample age (the year in which the specimen was collected) does not affect the completeness of mitogenomes obtained (Figure 1). Previous studies that compared age of source sample with other metrics of mitochondrial DNA recovery (e.g. copy number) also found no relationship between those two factors [12, 15, 36]. Together, those findings highlight the potential of old museum specimens—including holotypes and other taxonomically important material—for phylogenetic and evolutionary inferences. Future studies on Neotropical mammals could benefit from the large series of specimens collected in South America during the first decades of the 20th century (e.g. by A. Garbe, see [37]; and by the Olalla family, see [38]).

These valuable and irreplaceable specimens—many of those from localities that were long-ago transformed into human modified landscapes—will allow the reconstruction of comprehensive phylogenetic hypotheses for taxonomic groups for which samples of fresh, frozen, or ethanol-preserved tissue are absent, scarce, and/or difficult to obtain.

Despite the immense value of historical material, as aforementioned, we advocate that this type of sample should be used as a complement to traditional ethanol-preserved samples. As documented by our and previous studies [12], modern samples result in higher sequencing success, and they have the great advantage of not being destructive, in any sense, to the morphological vouchers. Therefore, whenever possible, they should be preferred as a primary source of genetic data. In this sense, contemporary field sampling using diverse collection techniques is crucial to increasing the representativeness and value of our repositories of biodiversity, as defended by Voss and Emmons [9]. The continuity of field expeditions, especially to remote and unsampled regions to collect new specimens, will certainly result in the discovery of many unknown species, which is imperative to uncover the still hidden biodiversity of the richest areas of the globe, including the Neotropics [39]. Ultimately, while historical specimens are a critical and essential resource, they should not obscure the potential value of obtaining additional specimens for science in the wild, which are paramount not only for taxonomic and phylogenetic refinement, but also for documenting ongoing ecological and evolutionary changes and promoting biodiversity conservation.

2.3.2 Missing data versus missing taxa

The discussion regarding how much missing data (and their effects) should be allowed in phylogenetic inferences has gained considerable attention, especially after the dissemination of next-generation sequencing methods (e.g. [40–43]). It was previously shown that missing data could obscure phylogenetic relationships and promote negative impacts on the phylogenies (e.g. [44]). Subsequent authors (e.g. [41]) have shown that including as many loci as possible in the phylogenetic analyses is benefical even with large amounts of missing data because this would increase the sampling of distinct regions across the genome. Streicher et al. [42] emphasized that the optimal approach in terms of amount of missing data incorporated without losing the accuracy of the inference depends on the dataset and the phylogenetic method employed. After exploring our data by comparing the

performance of matrices with alternative sampling strategies, we observed that the addition of samples with a limited amount of missing data did not impact the estimated relationships nor cause significant change to the nodal support of the inferred phylogenies.

2.3.3 Mitochondrial phylogeny and taxonomic arrangements proposed for Sciurini

The comparison of our results with proposed taxonomic arrangements for Sciurini ([2, 18, 21] and [1] which is identical to [20, 22]) illustrates that none of the generic arrangements fully corresponds to the phylogenetic structure recovered (Figure 5). Most currently recognized genera (by [1, 2]) are not recovered as monophyletic by our analyses of genetic data. Allen [18] suggested the greatest diversity of genera, and his hypothesis seems to be the one that best fits our results, especially regarding the Nearctic taxa.

The delimitation of the taxa at the genus-group level has not received as much attention as species delimitation [45, 46]. Recently published generic arrangements for South American rodents, which included the description of new genera [47–49], have provided a solid diagnosis for the new taxa by consistently testing phylogenetic hypotheses and employing reciprocal monophyly as a primary criteria. Their phylogenetic analyses were complemented with robust and consistent morphologic analyses and they then applied total evidence analysis or conducted *a posteriori* comparison to support taxon diagnosis.

Here we use the phylogenetic information provided by a taxonomically robust mitogenome dataset to suggest a tentative classification at the genus level for Sciurini (Figure 6). In our arrangement, we recognize reciprocal monophyletic entities as taxa at the genus-group level, and most of those entities correspond to groups A to L as recovered by our analyses; the only exception is the recognition of three reciprocally monophyletic genera within Group G, based on previous classificatory arrangements [1, 2, 18, 21]. We attribute to them the appropriate available names following the criteria established by the International Commission on Zoological Nomenclature (ICZN). Therefore, we favor the principles of priority and stability in such tentative nomenclatural acts, employing whenever possible the generic names proposed by Allen in 1915 [18], the first reviewer and author of several names of the genus-group valid and available for this radiation of squirrels.

For the sake of consistency with current generic nomenclatural acts (see above), our limited morphologic dataset (number of pairs of mammae and of upper premolars) precludes us from providing formal diagnosis and description for a presumptive taxon, a goal

that is beyond the scope of this contribution. Thus, for the lineage of the genus-group level with no available name, we apply the genus name that was historically employed for it, presenting this name between quotation marks.

Given the limitations and shortfalls of our data, which are solely based on mitochondrial DNA, we do not presume this updated generic arrangement the definitive scheme, but we intend to offer a working hypothesis that can be tested and formalized by further studies, as additional data become available. Careful taxonomic assessments with the inclusion of several lines of evidence, such as phenotypic information from sequenced and type material, are indispensable for this and other taxonomic issues of Sciurini to be properly addressed.

The first two major groups within Sciurini (A and B) compose the genera *Tamiasciurus* Trouessart, 1880 (including *T. douglasii* and *T. hudsonicus*) and *Rheithrosciurus* Gray, 1867 (including *R. macrotis*) as monophyletic groups, and we suggest the application of these names for the Groups A and B, respectively. The genus *Sciurus*, as broadly recognized in the past century—including Eurasian, Nearctic and Neotropical species (e.g. [1, 20–22])—is not monophyletic. This result has been also recovered in previous phylogenetic inferences with fewer species [7, 50–52] and is strongly supported by our analyses with denser taxon sampling. *Sciurus* can be restricted to the Eurasian clade, Group C, since it includes *vulgaris* Linnaeus, 1758, the type species of *Sciurus* Linnaeus, 1758. Other species included in the restricted concept of this genus are *S. anomalus* and *S. lis*. The remaining North American species are arranged in Groups D, E and F, for which four generic names are available, *Hesperosciurus*, *Otosciurus*, *Neosciurus* and *Parasciurus*. For Group D, two generic names were coined as subgenera by Nelson in 1899, *Otosciurus* for *aberti* and *Hesperosciurus* for *griseus*. Here we conservatively assign the oldest available genus-group name, *Hesperosciurus*, as a genus including *H. aberti* and *H. griseus*. Regarding Group E, the genus name *Parasciurus* Trouessart, 1880 is the only available one, and its type species is *niger* Linnaeus, 1758; therefore, we suggest the application of this name for this clade, composed of *P. nayaritensis*, *P. arizonensis*, *P. alleni* and *P. oculatus*, along with the type species. Finally, for Group F the only available name is *Neosciurus*, described by Trouessart, 1880 for *carolinensis* Gmelin, 1788, and this is the name that we tentatively apply to this lineage.

Group G, which comprises most Central American taxa, might be the most taxonomically conflicting group as it contains the type species of many genera, including

alfari (type of *Microsciurus* J. A. Allen, 1895), *brochus* (type of *Syntheosciurus* Bangs, 1902) *deppei* (type of *Baiosciurus* Nelson, 1899) and *aureogaster* (the senior synonym of *Sciurus hypopyrrhus* Wagler, 1831, type species of *Echinosciurus* Trouessart, 1880) (Figure 5). Allen [18] recognized five genera for the eight species or species-complex in this group. Subsequent authors recognized fewer genera, but none suggested a unique genus to contain those species. It is noteworthy that Moore [21] was the only author to anticipate a close relationship between *brochus* and *granatensis*, suggesting these taxa be placed under the genus *Syntheosciurus*. We partially follow the arrangement proposed by Allen [18] and Moore [21], and we suggest that the genus name *Microsciurus* should be applied for the clade formed by *alfari* and "species 1"; the name *Syntheosciurus* must be attributed to the group formed by *brochus* and *granatensis* (if Vivo and Carmignotto are correct [see below, on the discussion of the name of clade H] and *granatensis* is the type species of *Notosciurus*, this genus name is a junior synonym of *Syntheosciurus*); and we advocate the adoption of the name *Echinosciurus*, as the oldest available one, for the group formed by *aureogaster*, *colliaei*, *deppei*, *yucatanensis* and *variegatoides*.

Group H, composed of northwestern South American forms, includes five species traditionally allocated to the genus *Microsciurus* J. A. Allen, 1895 (*mimulus*, *similis*, *otinus*, *boquetensis*, and *isthmius*). However, the type species of the genus *Microsciurus* (*alfari* J. A. Allen, 1895), as demonstrated above, is nested within Group G and, therefore, the name *Microsciurus* cannot be applied to Group H. The other species recovered in this clade, *pucheranii* Fitzinger, 1867, is a controversial taxon. Allen [18] described the genus *Leptosciurus* and considered *pucheranii* as its type species. Moore [21] placed *pucheranii* in *Microsciurus*, and he was the only author to suggest a close relationship between this taxon and the small-sized species from the highlands of northwestern South America. Most other authors have allocated *pucheranii* to *Sciurus* (e.g. [1, 19, 20]), but Vivo and Carmignotto [2] included *pucheranii* and *granatensis* under the genus *Notosciurus* Allen, 1914. Their decision was based on the fact that: i) their concept of *N. granatensis* included *chysuros* Pucheran, 1845 as a subspecies (*N. g. chrysuros*) and *soederstroemi* Stone, 1914 as a junior-synonym of this subspecies; ii) Vivo and Carmignotto [2] followed Hershkovitz [53] who identified *N. rhoadsi* Allen, 1914 (the type species of *Notosciurus*) as a young specimen of *soederstroemi*. Therefore, for these authors, the name *Notosciurus* would be applied to *granatensis* (via its synonymy with *soederstromi*) and *pucheranii* (for their morphological similarity), and would

have priority over the name *Leptosciurus*. However, our results did not recover *granatensis* and *pucheranii* as closely related taxa. Instead, *granatensis* was recovered in Group G, sister to *Syntheosciurus brochus*. Therefore, *Leptosciurus* seems to be the only available name for Group H, and we suggest the application of this name to the six species there nested.

Group I includes *stramineus* (type species of *Simosciurus* Allen, 1915) along with *nebouxii*. Vivo and Carmignotto [2] followed Allen [18] considering *Simosciurus* a valid genus, and we recover it as monophyletic based on mitogenomic data. Since *Simosciurus* is the only available name for this clade, which includes the type species of this genus, we believe it is the appropriate name for Group I. Our analyses also support the monophyly of the genus *Guerlinguetus* Gray, 1821, as recognized by both [2, 18], represented in our analyses by Group J. *Guerlinguetus* has been consistently employed for parts of this particular group of species, as full genus or subgenus, by several authors.

Described species recovered within Group K have been assigned to the genus *Microsciurus* by all authors. The unnamed lineage recovered within this Group ("species 2") was also referred to the genus *Microsciurus* by [54, 55], but it has been referred to the genus *Syntheosciurus* by Vivo and Carmignotto [2]. Our data do not recover taxa of this group as closely related to the type species of *Microsciurus* or *Syntheosciurus*. Moreover, as the valid species in this group (*flaviventer* and *sabanillae*) were both described in genera currently occupied (*Macroxus* and *Microsciurus*, respectively), there seems to be no generic name available for Group K. Until more consistent morphologic dataset is available to allow a formal nomenclatural designation, we provisionally use the name "*Microsciurus*" for this clade (see Patton et al., 2015, for "*Handleyomys*"), as this was the name historically assigned to these species. An alternative measure would be to apply the genus name of the sister group (L) to this lineage, but we do not recommend this option as this would introduce more taxonomic confusion and instability.

Finally, Group L clustered species allocated in distinct genera according to [2, 18], or from a single but not monophyletic genus of Moore [21] and Thorington et al. [1]. At least five generic names have been applied to those species: *Notosciurus* Allen, 1914, *Leptosciurus* Allen, 1915, *Mesosciurus* Allen, 1915, *Hadrosciurus* Allen, 1915, and *Urosciurus* Allen, 1915. However, only *Hadrosciurus* and *Urosciurus* are possibly vacant here, and the correct assignment must be carefully evaluated in a comprehensive taxonomic study that includes a meticulous nomenclatural investigation for this group. However, in order to propose a

tentative nomenclatural definition, as we have done for previous clades, we tentatively apply the name *Hadrosciurus* Allen, 1915, whose type species is *flammifer* Thomas, 1904, considered by Vivo and Carmignotto [2] as a junior-synonym of *igniventris* Wagner, 1842. This name was also advocated by Vivo and Carmignotto [2].

2.3.4 Comments on species recognition and novelties

In this study, we sampled across the geographic ranges of several widespread taxa and, therefore, we were able to test the genetic integrity of currently recognized species of tree squirrels, especially those from South America. In contrast to our generic level analyses, most recognized species are highly supported as monophyletic groups in our analyses of mitochondrial genome data. Regarding Palearctic and Nearctic taxa, our sampling was remarkably inferior to the sampling for Neotropical taxa, with many species represented by as few as one or two individuals. As expected, all those species with more than one individual exhibit reciprocal monophyly in our phylogenomic analyses, following the species concepts presented by Thorington et al. [1].

Among the Central American taxa, the two cases of non-reciprocal monophyly were (i) the recovery of a sample identified as *richmondi* Nelson, 1898, from Nicaragua, nested within the clade associated with *Syntheosciurus granatensis*; and (ii) a specimen assigned to *venustulus* Goldman, 1912, from Panama, nested within the clade of *Microsciurus alfari*. Samples from *Syntheosciurus granatensis* compose two well-structured subclades, one of which includes specimens from the Ecuadorean and Peruvian Andes, Venezuela, and Trinidad and Tobago, and the other includes samples from the coast of Ecuador, Colombia, Nicaragua (referred to as *richmondi*), and Panama (Group G, Figure 4a). Without the inclusion of additional specimens referred to *richmondi* and the careful examination of voucher material, we are unable to unveil, at this point, if this is a simple case of misidentification or if this taxon needs taxonomic re-evaluation. Our molecular species delimitation analyses provide distinct resolutions for the samples assigned to *granatensis* and *richmondi*, but none of them suggested the sample assigned to *richmondi* as a distinct species from the specimens of *granatensis*. Based on our phylogenetic results corroborated by BPP analysis, we recognize a single putative species, *Syntheosciurus granatensis*, for those samples. Regarding the second case, all samples of *Microsciurus alfari*, as well as the sample initially identified as *venustulus*, are from Panama and were suggested as a single species by

all species delimitation analyses. Thus, we provisionally do not treat *venustulus* as a valid taxon until further evaluation with additional specimens.

Across South American lineages, our results indicate that *pucheranii* sensu [2] forms a non-monophyletic assemblage composed of two phylogenetically distant lineages included in Groups H and L. The concept of *pucheranii* adopted by Vivo and Carmignotto [2] includes specimens with a disjunct distribution, from the Central Andes of Colombia (assigned to *pucheranii pucheranii* Fitzinger, 1867) and from Peru and Brazil, Bolivia, and Argentina [assigned to three other subspecies named *pucheranii ignitus* Gray, 1867, *pucheranii boliviensis* Osgood, 1921, and *pucheranii argentinius* Thomas, 1921, respectively]. In our analyses, specimens of *pucheranii pucheranii* are recovered as part of Group H (Figure 4b)—an Andean Trans-Andean clade composed of taxa from high elevation areas of northwestern South America. For this clade we provisionally apply the name *pucheranii* Fitzinger, 1867 to the species level, with the combination *Leptosciurus pucheranii*. Specimens associated with the remaining three subspecies were recovered as a clade nested within Group L (Figure 4c), which includes Cis-Andean lowland taxa. For this lineage we suggest the application of the name *ignitus* Gray, 1867, as it has priority over *argentinius* and *boliviensis*, with the status of a full species. We did not intend to revalidate or describe new species in this contribution, however, as we were unable to use the current species concepts for the taxa mentioned above, we tentatively suggest this alternative arrangement, which is in accordance with the classification of Thorington et al. [1] at the species-group level; the name we propose is, thus, *Hadrosciurus ignitus*.

The concepts of *Guerlinguetus aestuans* and *G. brasiliensis* adopted by Vivo and Carmignotto [2] are also not monophyletic according to our analyses. Based on the geographic distribution of the samples, the subclades *G. aestuans* "a" and *G. aestuans* "b" include specimens associated with *Guerlinguetus aestuans*. The first subclade is composed of samples from Guyana and Venezuela, and the second of Brazilian samples from the southern bank of the Amazon river, west of the Tapajós river (Figure 5d). The subclade *G. aestuans* "c" seems to encompass representatives of both *aestuans* and *brasiliensis*, since it includes specimens from north of the Amazon river (assigned to *aestuans* by those authors) and one specimen from Pernambuco, northern Atlantic Forest (assigned to *brasiliensis* by those authors). We referred to this last subclade as *G. aestuans* "c" as the great majority of samples within this lineage were previously assigned to *G. aestuans* and not to *G.*

brasiliensis. The subclade *G. brasiliensis* is apparently composed of samples assigned exclusively to this nominal taxon, from southeastern Amazonia, eastern and southern Brazil. Therefore, we recognized specimens previously identified as *Guerlinguetus aestuans* and *G. brasiliensis* as composing four distinct lineages (see Figure 4b), suggesting hidden diversity along the Amazon basin and implying an independently evolving lineage from the Gran Sabana and Mount Roraima, on the border of Brazil, Venezuela, and Guyana. This result was corroborated by most species delimitation analyses, except for one analysis (GMYC 2) in which *Guerlinguetus aestuans* "c" and *G. brasiliensis* were suggested as a unique putative species.

Our phylogenetic results also indicate the existence of three apparently unnamed lineages that might represent species to be described or revalidated, all of which were supported by molecular species delimitation methods. "Species 1" is represented by a specimen from Chocó, Colombia, which was previously identified as *Microsciurus mimulus*; however, this specimen was recovered as phylogenetically distant from other specimens of *M. mimulus* from Colombia and Ecuador (all of which clustered within Group H), and exhibited deep genetic divergence from its sister-taxa, *M. alfari* (see branch lengths on Figure 4a). "Species 2" is represented in our analyses by five specimens from Peru (San Martin, Madre de Dios) and Brazil (Acre). Voucher material of this species, from San Martin, was analyzed by [54, 55]—who referred to it as *Microsciurus* sp.—and by [2]—who referred to it as *Syntheosciurus* sp. "Species 3" is represented by three specimens from two Amazonian lowland localities in Loreto (Peru), and is apparently sympatric with *Hadrosciurus spadiceus* at Rio Galvez, Nuevo San Juan. We did not find previous mention of this putative species in the literature.

Therefore, monophyletic groups representing currently recognized species in addition to the lineages representing putative unnamed taxa composed a set of 43 OTUs that we hypothesize as distinct species of tree squirrels. All South American OTUs were corroborated as unique species by at least two out of the three species delimitation analyses performed, except for one OTU, *Leptosciurus boquetensis*, which was only supported as a distinct species by BPP. Regarding non-South American taxa, our species delimitation analyses did not fully corroborate our working hypothesis. Discrepant results in the recognition of those species are likely a product of our sampling strategy, densely focused on South American taxa. At least BPP analyses are potentially affected by the number of

samples per each presumed species, especially if using a dataset with few loci [56]. GMYC estimates might not be as affected by poorly represented species as BPP [57, 58], but can be strongly influenced by the way that the ultrametric tree is generated, which underprints the analysis [59]. Our results corroborate this assumption, as we found discrepant results suggesting 66 or 39 putative species using ultrametric trees generated with strict and relaxed molecular clocks, respectively.

Several studies have employed molecular species delimitation methods either as a standalone tool or as part of an integrative approach to delimit species [60–62]. Here, we advocate for the use of molecular species delimitation methods along with other sources of evidence, to avoid misleading species delimitation due to theoretical and/or methodological shortfalls (as exemplified above; see also [57–59]). Moreover, in many cases, when delimiting species based on a single-locus dataset, the estimates could be biased by the genealogical history of this locus which may or may not reflect the evolution of the group. As we used an exclusively mitogenomic dataset, we acknowledge that the evidence for pervasive natural selection, uniparental inheritance and the lack of recombination on the mitochondrial genome make it susceptible to evolutionary processes distinct to the nuclear genome [63, 64].

Considering the possible methodological weaknesses mentioned above and the shortfalls of our sampling of taxa and data, we evaluate the results of our molecular species delimitation analyses with special caution in some situations. For example, the genus *Tamiasciurus* was recently extensively revised through molecular analyses (including mitochondrial and nuclear genes) and ecological niche modeling, with over 250 specimens examined from throughout the distribution of the genus [23]. These authors found evidence for the recognition of *T. douglassi* and *T. hudsonicus* as valid species. Our analyses consistently failed to suggest these taxa, represented by a single sample each, as distinct species (see Figure 4a). Another example is that some species delimitation analyses did not recognized *Sciurus lis* and *S. vulgaris* (GMYC 2) or *Sciurus lis*, *S. vulgaris* and *S. anomalus* (BPP) as distinct species. These taxa, which are represented in our dataset by only one terminal each, have been consistently recognized as distinct species based on molecular [24, 65] and karyotypical [26] data. They also exhibit consistent morphological differences in the number of pairs of mammae (a trait that seems not to be variable within species of tree squirrels [2]), which is three in *S. lis*, four *in S. vulgaris*, and five in *S. anomalus* (see Figure 6).

These controversial results, especially for Eurasian and North American taxa, lead us to adopt a conservative posture that does not totally reject the hypotheses provided by the species delimitation analyses, but it is also in consonance with current taxonomic proposals based on wider sampling approaches (see examples above). For those cases of inconsistency regarding Eurasian, North American, and Central American taxa, and also for South American lineages where species complexes were suggested by one or two of the species delimitation analyses, subsequent investigations are certainly necessary. A thorough delimitation of species of Sciurini demands additional sampling for several taxa and, possibly, the inclusion of other lines of evidence such as phenotypic data and genetic data from independently evolving loci as in nuclear DNA.

2.3.5 Phylogenetic and biogeographic remarks

Despite the discordances between our mitochondrial phylogenomic hypothesis and the taxonomic arrangements previously proposed for Nearctic and Neotropical tree squirrels, our results are biogeographically coherent, and consistent with most of the results obtained by the few molecular phylogenetic studies published for Sciurini, especially regarding the deepest nodes (major clades) within the tribe. Like Pečnerová and Martínková [66] and Pečnerová et al. [7], we recovered the genus *Tamiasciurus* as the first lineage to diverge within Sciurini, followed by *Rheithrosciurus* and *Sciurus*, although our study is the first to recover strong support for these relationships. Our results also corroborate the sister-taxa relationship between *Hesperosciurus griseus* and *H. aberti* found in those previous studies. The Central American clade obtained by [7] is similar in composition to our Group G, despite the different relationships within this group, recovered by us with strong support. In previous studies, the representativeness of South American taxa was very limited, and the relationships among the very few specimens were mostly discordant from our results. One relevant difference is that we recovered the Mexican endemics *Parasciurus alleni* and *P. oculatus* clustering with North American species, instead of within a South American clade as in [7].

Concerning the biogeographic pattern, we recovered two Palearctic clades (A and B), four Nearctic (C–F), and six Neotropical—one (G) predominantly composed of Central American with a few South American specimens included (all from Andean or Trans-Andean areas) and five (H–L) composed exclusively of South American taxa and Southern Panama

specimens (Figure 4). The distribution of those five predominantly South American clades seems to be defined by the Andean Cordillera. We found two clades occupying Andean and Trans-Andean areas (H and I) and three clades distributed on the Cis-Andean portion of the Continent (Groups J–L). Group H seems largely associated with montane habitats, while Group I is restricted to low elevation coastal areas near the sea-level. Regarding the Cis-Andean groups, Group J is the most widespread, occurring from the extreme east of South America, in the Atlantic Forest, to the Guiana Shield, and throughout the Amazon basin. The sister Groups K and L are largely sympatric and composed mostly by Amazonian lowland dwellers. In Group K, however, two lineages ("*Microsciurus" sabanillae* and "species 2") reach mid-elevations on the east side of the Andean cordillera in Ecuador and Peru; and in Group L, one lineage (*Hadrosciurus ignitus*) is also found in high-altitude localities in Bolivia.

2.3.6 Taxonomic consequence of the use of homoplastic traits in the study of tree squirrels

Historically, all genera proposed for Neotropical species of Sciurini were delimited based exclusively on morphological traits. For example, species of *Notosciurus* sensu [2] were diagnosed by the presence of three pairs of mammae and one upper premolar; and the genus *Microsciurus* sensu [1, 2, 18, 21] was defined, among other traits (e.g. small size), by the presence of three pairs of mammae and two upper premolars. Our results, however, indicate that these features are homoplastic, with similar conditions of both characters having evolved multiple times during the evolutionary history of tree squirrels. Morphologic convergence has been detected among several lineages of Sciuridae [67, 68] and, according to our data, seems to be common in both cranial and external traits of Sciurini. Grouping species based primarily on homoplastic characteristics might have led to some of the incongruences that we observe between the taxonomic arrangements and the molecular phylogeny recovered for tree squirrels. For instance, the genus *Microsciurus* sensu [1, 2, 18, 21] comprises a polyphyletic assemblage that clusters species sharing the same number of premolars and mammae. Phenotypic convergence has been previously detected for cranial traits in species formerly assigned to *Microsciurus* [7], and the use of homoplastic characters to diagnose this genus (e.g. by [2, 18]) can be claimed to explain the polyphyly of this taxon.

2.4 Conclusions

The inclusion of historical samples was crucial to provide a comprehensive phylogenetic hypothesis for tree squirrels and to detect several taxonomic issues reported here. We investigated the different aspects that might have influenced the success of mitogenome recovery from historical samples of Sciurini and showed that the age of the specimen does not affect mitogenome completeness. This finding highlights the potential of old museum specimens—including holotypes and other taxonomically important material— for phylogenetic and evolutionary inferences. Our extensive sampling of museum specimens, allied with a modern next-generation sequencing approach, allowed us to recover the entire mitochondrial genome of several species of squirrels. After exploring our data by comparing the performance of matrices with alternative sampling strategies, we observed that the addition of a limited amount of missing data did not impact the estimated relationships nor caused significant change to the nodal support of the inferred phylogenies. The comparison of our results with proposed classification schemes illustrates that none of the taxonomic arrangements ever proposed fully corresponds to the phylogenetic structure recovered for Sciurini, with only a few of the currently recognized genera recovered as monophyletic. Therefore, we advance a preliminary and tentative nomenclatural designation for the taxa at the genus-group level, employing 13 names used in previous taxonomic classifications. Our phylogenetic reconstruction revealed that most recognized species are highly supported as monophyletic groups. Nevertheless, we found evidence supported by species delimitation analyses that the diversity of Neotropical tree squirrels is currently underestimated, with at least six lineages that might represent taxa to be named or revalidated. In summary, we hypothesize that the tribe Sciurini comprises 14 genera and 46 species (see Table 3)—of which 43 species were sampled here and three were not included in the present study, but we provisionally treated them as valid—, a more diverse estimate than recent catalogues [1, 2]. *Sciurus*, formerly the most diverse genus in the tribe, harbors only three species, while the genera *Leptosciurus* (with six species), *Hadrosciurus*, *Parasciurus* and *Echinosciurus* (all with five species each), are the most diverse within this radiation; the only monotypic genus is *Rheithrosciurus*. The Neotropical region harbors eight genera and 29 species. However, a detailed taxonomic investigation is necessary to carefully evaluate the applicability of the genus-level names, to provide diagnoses and or descriptions to them, as well as to evaluate the species-level taxonomy for those genera. Finally, by investigating the evolution of two

morphological traits widely employed in the taxonomy of the group we revealed their homoplastic nature, helping to explain the incongruence between phylogenetic results and classificatory schemes presented so far.

2.5 Methods

2.5.1 Sampling

In order to obtain a thorough sampling of Sciurini, we gathered a total of 271 samples from 27 scientific collections (Additional file 5), including 177 modern samples (ethanol-preserved tissue) and 94 historical samples (obtained from dry museum specimens). Historical samples were collected with the specific purpose of complementing missing taxa or important geographic variants, with special effort on Neotropical taxa. When collecting tissues from dry museum specimens, we prioritized sampling remains of muscular tissue adherent to skulls ("osteocrusts") or, if those were not available, we obtained skin clips. Sampling from dry museum specimens followed strict procedures, including changing gloves and cleaning all instruments and working surfaces with 15% bleach followed by sterilized water between each sample (see detailed protocol in [12]).

The sampled material includes 40 out of the 43 currently recognized species of Sciurini (sensu [1, 2]). The unsampled taxa include *Microsciurus santanderensis* (known from few specimens collected between the Río Magdalena and the western slopes of the Cordillera Oriental in Colombia; [69]), *Microsciurus simonsi* (known from few localities west of the Andes, in the Ecuadorian provinces of Bolívar and Pichincha; [2]), and *Tamiasciurus fremonti*, revalidated from the synonymy of *T. hudsonicus* by Hope et al. [23] (known from the southwestern United States in the southern Rockies, Sacramento Mountains in New Mexico, and the southwestern Sky Islands [23, 70]). Additionally, we sampled three species of the tribe Pteromyini (sister to Sciurini; [50]) to be used as outgroups. A complete list of the 232 specimens used in our analyses (for which we recovered at least 20% of the mitogenome) indicating the GenBank accession numbers and accompanied by geographic data and other relevant information is provided as Additional file 6.

2.5.2 Taxonomic identifications

Specimens of Sciurini were identified at the species level following the latest taxonomic hypotheses available for each taxon. Samples of the North American genus *Tamiasciurus*

were identified following Hope at al. [23]. South American material was identified following Vivo and Carmignotto [2], as well as the Central American taxa included on the taxonomic hypothesis of those authors (assigned by them to the genus *Microsciurus*: *alfari*, *boquetensis* and *venustulus*). For the remaining Central American taxa (not included on the taxonomic hypothesis of [2]), North American (except by *Tamiasciurus*), and Eurasian taxa, we have identified specimens following Thorington et al. [1].

For several specimens, especially those housed at the American Museum of Natural History (AMNH) and at the Smithsonian National Museum of Natural History (USNM), we kept the museum identifications, which had been made by some of the main authorities on tree squirrel taxonomy (e.g. R. W. Thorington and M. de Vivo). For material not previously identified, we were able to perform the identifications by examining the morphology of the vouchers, consulting original descriptions and other relevant literature. In cases for which we were not able to examine vouchers, we accepted original museum identifications if i) those identifications correspond to the known geographic distribution of the taxon in question, and ii) phylogenetic analyses of their DNA sequences were consistent with the museum identification.

2.5.3 DNA extractions

DNA of historical samples was extracted in an isolated ancient DNA facility at the Smithsonian's Center for Conservation Genomics (CCG), using a standard phenol-chloroform protocol (see detailed protocol in [12]). The ancient DNA lab at CCG is physically separated from the main laboratory, and no fresh tissue/DNA samples or PCR amplifications are allowed, to minimize and control sample contamination. Extractions included a long lysis step, between three to five days. Each batch of historical sample extraction included from seven to 11 specimens and a negative control to monitor for contamination. DNA extractions of modern tissues were performed in the main laboratory at CCG using the DNeasy® Blood & Tissue kit, following manufacture's protocol (Qiagen Inc.), with an overnight lysis step. Total DNA concentrations were measured using a Qubit 2.0 fluorometer (Thermo Fisher Scientific).

2.5.4 Library preparations and mtDNA amplification

For historical samples, an initial amount of 33 µl of DNA (regardless of concentration) was purified and concentrated using 5x SPRI magnetic beads [71]. DNA extracted from preserved tissues was sonicated to randomly shear with QSonica Q800R, using 25 % of amplitude and 5 min of on/off pulse. Sheared DNA was visualized on agarose gel to confirm the resulting fragment size around 300 bp. Approximately 500 ng of sheared modern DNA was then purified using 5x SPRI magnetic beads [71].

Library preparations were performed using the KAPA LTP Library Preparation Kit (Roche Sequencing) following the manufacturer's protocol. Subsequently, Nextera-style indices and KAPA HiFi Hotstart ReadyMix (Roche Sequencing) were used for indexing PCRs (iPCR). The iPCR profile included an initial denaturation at 98 °C for 45 sec, a final extension at 72 °C for 7 min, and 14 (for modern samples) or 16 to 18 (for historical samples) cycles of amplification, with denaturation at 98 °C for 15 sec, annealing at 60 °C for 30 sec and extension at 72 °C for 60 sec. The iPCR products were purified using 1.8x SPRI magnetic beads, quantified with a Qubit 2.0 fluorometer (Thermo Fisher Scientific) and visualized on a 1.5% agarose gel.

Libraries were multiplexed in equimolar ratios for target capture and enrichment of Ultraconserved Elements (UCEs) using similar procedures as described in [32]. We did not perform capture or enrichment of the mtDNA; the mitogenomes were obtained as a byproduct of the UCE enrichment without the need of an extra step for mitochondrial-specific enrichment or amplifications. For historical samples we pooled up to four libraries and for modern samples up to eight libraries. No historical samples were pooled with modern samples to avoid biased enrichment. Post-capture amplifications were performed using KAPA HiFi Hotstart ReadyMix (Roche Sequencing), with the following profile: initial denaturation at 98 °C for 2 min, a final extension at 72 °C for 7 min, and 15 (for modern samples) or 16 (for historical samples) cycles of amplification, with denaturation at 98 °C for 20 sec, annealing at 60 °C for 30 sec, and extension at 72 °C for 30 sec. A 1.8x SPRI magnetic bead cleanup was performed subsequently.

2.5.5 Quantification and sequencing

Cleaned amplifications were quantified using a Qubit 2.0 fluorometer (Thermo Fisher Scientific) and visualized on a Bioanalyzer (Agilent) with high sensitivity kits. Equimolar

pooling of samples for sequencing was based on the concentration (ng/µl) and on the average size (bp) of DNA fragments. High concentration of dimmers was common, especially for historical samples. This problem was solved by size-selecting the fragments of DNA between 200 and 550 bp using a Pippin Prep (Sage Science). Both size-selection and sequencing were performed at the DNA Sequencing Center at the Brigham Young University, Utah, and at the Vincent J. Coates Genomics Sequencing Laboratory at the University of California, Berkeley. Illumina sequencing was done on a Hi-Seq 2500 125 PE and on a Hi-Seq 4000 150 PE using the Illumina Free Adapter Blocking Reagent to prevent index hopping.

2.5.6 Data processing

Raw FASTQ files were provided by the sequencing cores. The raw data was processed to extract mtDNA as "off-target sequences" of the UCE capture [32]. Raw reads were cleaned for removal of adapter contamination and low-quality bases using Illumiprocessor 2.0 [72, 73]. Partial and complete mitochondrial genomes were recovered using Geneious R11 [74]. Clean reads (paired P1 and P2, plus singletons) were incorporated in Geneious and mapped to a reference mitochondrial genome (*Sciurus vulgaris* available in GenBank with accession number AJ238588) using the following mapping parameters: a minimum map quality of 30—which means that with 99.9% confidence the mapping is correct; a minimum overlap of 25 base-pairs for a read to be assembled into a contig; a minimum overlap identity of 85% (i.e. the minimum percentage of bases that must be identical in the overlapping region for a read to be assembled) with maximum of 15% of mismatches per read; a maximum of 10% of gap per read, with maximum gap size of 10 base-pairs. Up to five iterative mapping cycles were performed to find the greatest number of matching reads. Consensus sequences were generated with a minimum coverage of 3x. The mitochondrial genomes assembled were visually inspected and the coding genes were translated. We submitted all complete mitogenomes recovered to be annotated by MITOS [75] and the remaining partial genomes obtained were manually annotated based on the annotations provided by MITOS. All annotations were manually added to the sequences using Geneious R11 [74], where we performed visual inspection to certify that the beginning and end of the annotated coding sequences (CDS) matched with the translations of start and stop codons. We converted the mitogenome annotation of one species of tree squirrel (*Guerlinguetus brasiliensis*) into a

graphical map using OGDRAW 1.3.1[76], to exemplify the genome synteny in the group (Figure 2).

2.5.7 Sequence alignment and dataset composition

The consensus mitochondrial genomes were aligned using MUSCLE [77] with up to eight interactions. In order to examine the possible effects of including and excluding characters and taxa with missing data on our phylogenetic inferences, we generated five datasets considering distinct percentages of mitogenome completeness per sample: Dataset 1 included only specimens for which we obtained full mitochondrial genomes (92 specimens with no missing data); Dataset 2 included samples for which at least 80% of the mitogenome was recovered (162 specimens with <20% of missing data per sample); Dataset 3 included samples for which at least 60% of the mitogenome was recovered (186 specimens with <40% of missing data per sample); Dataset 4 included samples for which a minimum of 40% of the mitogenome was recovered (210 specimens with <60% of missing data per sample); and Dataset 5 included samples for which a minimum of 20% of the mitogenome was recovered (232 specimens with <80% of missing data per sample). We did not include the 39 samples for which we recovered less than 20% of the mitogenome in our datasets (Table 1). Therefore, from the 271 samples that we attempted to sequence, we only used a total of 232 samples in our analyses.

2.5.8 Mitochondrial genome recovery for historical museum samples

To investigate how different factors might influence the success of mitochondrial genome recovery for historical samples, we compared the completeness of mitogenome obtained with tissue type (osteocrusts versus skin clips), and museum location (samples from scientific collections in NA versus SA) using Pearson's Chi-squared test. We also investigated the relationship between collection year and mitogenome recovery using a linear regression model. For the last, we only included osteocursts (which compose the great majority of our historical samples) from NA museums (our main source of historical samples). This was done to avoid bias related to tissue type and storage conditions (NA museums have similar storage conditions and standardized procedures to preserve specimens, while storage conditions and procedures are highly heterogeneous in SA collections). All analyses were performed in RStudio 1.1.463 (RStudio, Inc.).

2.5.9 Phylogenetic analyses

Phylogenetic analyses were performed for each one of our five datasets using Maximum Likelihood (ML) in RaxML 8.2.7 [78]. Ten independent searches were performed under the GTR+G nucleotide substitution model (RAxML only implements GTR-based models of nucleotide substitution). The best-scoring ML trees were selected to draw the bootstrap support values obtained by running 1,000 replicates using the "thorough standard bootstrap" optimization option. In addition to the ML analyses performed on all datasets, Bayesian inference (BI) in MrBayes 3.2.6 [79] was performed exclusively on Dataset 5 (see below "Phylogenetic inferences and the effect of missing data" for details). For the BI, the best-fit partitioning scheme and models of nucleotide substitution were specified as determined by PartitionFinder [80] under the corrected Akaike Information Criterion (AIC). We defined 39 separate data blocks in our alignment: 22 transfer RNAs (tRNAs), 13 protein-coding genes (PCGs), two ribosomal RNAs (rRNAs), one origin of replication, and a control region (D-loop). The PCGs were also separated by codon positions. Therefore, our search for partitioning schemes and models occurred independently in 65 data blocks of the mitochondrial genome. The BI was performed in parallel [81] with two independent runs and with four chains each. The MCMC (Markov Chain Monte Carlo) simulations proceeded for 4 X 10^7 generations, sampling every 4,000 generations. Nodal support was obtained as posterior probability. The quality and convergence of MCMC runs were verified in Tracer 1.7 [82]. Topologies from both ML and BI analyses were edited in FigTree 1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/). Phylogenetic analyses were conducted on the Smithsonian Institution High Performance Cluster (SI/HPC; https://doi.org/10.25572/SIHPC).

2.5.10 Species delimitation analyses

In order to provide a quantitative test regarding species limits in Sciurini, we applied two species delimitation methods that are amongst the most widely used in the recent literature: the generalized mixed Yule-coalescent model (GMYC; [83]), and the Bayesian multispecies coalescent approach in the software Bayesian Phylogenetics and Phylogeography (BPP; [84]).

GMYC was initially developed for analyses of single-locus data, however, it has been frequently applied to concatenated matrix of multilocus data by assuming a common genealogical history [56]. This method aims to identify the limit between a Yule speciation

process and the intraspecific coalescence using a likelihood approach and a provided ultrametric tree [83]. We used BEAST 2.6.1 [85] to obtain ultrametric trees. We generated two trees using a concatenated matrix of the 13 PCGs, one applying a strict molecular clock and another with a relaxed log-normal clock, both with a Yule tree prior. BEAST runs were conducted by 100 million generations of Markov chain Monte Carlo (MCMC), sampling every 20,000 generations. The results of MCMC runs were inspected in Tracer 1.7 [82] to confirm a minimum of 200 effective sample size for all parameters. Ultrametric trees were summarized with TreeAnnotator v2.6.0 [85], considering 10% of burnin and selecting the maximum clade credibility as the target tree. We performed GMYC analyses using the strict molecular clock generated tree (GMYC 1) and using the relaxed log-normal clock generated tree (GMYC 2), to verify how this change may affect species delimitation. Both GMYC analyses were performed using the R package splits v1.0-19 [86] and we selected the single-threshold version of GMYC, as it has been shown to outperform the multiple-threshold version [83, 87].

Different from GMYC, BBP was designed to test models of evolution on multilocus datasets. BPP is a Bayesian MCMC based program for delimiting species under the multispecies coalescent model [84]. We used BPP v4.1.4 [88] to analyze a multilocus dataset composed of the 13 mitochondrial PCGs. We provided a guided species-tree based on our ML analysis of Dataset 5, considering as terminals (species to be tested) the 43 OTUs recognized through the morphological identifications and confirmed as monophyletic groups by our phylogenetic inferences. We performed multiple BPP analyses with varying priors of ancestral population size (θ) and root age (τ). Our exploratory analyses showed that only assuming large ancestral population sizes (θ = 0.2) and deep divergences among species (τ = 0.2) we were able to reach convergence between MCMC runs and to obtain satisfactory effective sample sizes (near or above 200) for the analytical parameters. Each BPP analysis was conducted by 300,000 generations of MCMC and 10% of the samples were discarded as burn-in. Convergence on the estimates were verified in Tracer 1.7 [82].

2.5.11 Morphological character evolution

We performed ancestral state reconstructions (ASR) of two discrete traits: “number of upper premolars” and “number of pairs of mammae” using a ML approach. These morphological characters were selected considering their broad use in taxonomic studies of Neotropical

squirrels [2, 18, 21], and they are representatives of both cranial and external features that have been traditionally employed to assign species to genera and to establish genera limits. To reconstruct the evolution of number of upper premolars, we recognized the states *one premolar*, and *two premolars*. To reconstruct the evolution of the number of pairs of mammae, we assigned species to either *three pairs*, *four pairs*, or *five pairs* of mammae. To score species, we compiled information from literature and examined specimens housed at the American Museum of Natural History (AMNH) and National Museum of Natural History (USNM), including all species included in our phylogenetic hypothesis. State coding for all species and relevant scoring details are provided as Additional file 7.

Reconstructions were inferred in RStudio 1.1.463 (RStudio, Inc.), on an ML tree derived from our Dataset 5 trimmed to the species level. We tested three models of evolution using the function "fitDiscrete" from package geiger [89]: Markov model with equal rates (Mk-ER); Markov model with symmetric rates (Mk-SYM); and Markov model with all rates different (Mk-ARD). The best models, selected with base on their support from a vector of AIC scores, were then used to perform the ancestral character estimation using the function "ace" from package ape [90]. The results were displayed on the ML tree using the function "plot.phylo" from ape [90].

List of abbreviations

AIC: Akaike Information Criterion; AICw: Akaike Information Criterion weight; AMNH: American Museum of Natural History; ASR: Ancestral state reconstructions; BI: Bayesian Inference; BPP: Bayesian Phylogenetics and Phylogeography; CA: Central America; CCG: Smithsonian's Center for Conservation Genomics; GMYC: generalized mixed Yule-coalescent model; MCMC: Markov Chain Monte Carlo; Mk-ARD: Markov model with all rates different; Mk-ER: Markov model with equal rates; Mk-SYM: Markov model with symmetric rates; ML: Maximum Likelihood; MRCA: Most recent common ancestor; NA: North America; OUT: Operational Taxonomic Units; PCGs: Protein-coding genes; SA: South America; USNM: National Museum of Natural History.

Declarations

Ethics approval and consent to participate

No animals were used in this study. All samples analyzed here were obtained from scientific collections.

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and analyzed during the current study are available in the Dryad Digital Repository (https://doi.org/10.5061/dryad.9w0vt4bc9). GenBank accession numbers are provided for all complete mitochondrial genomes in the Additional file 6.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico through doctoral fellowships to EFA (147145/2016-3, 203692/2017-9, and 165553/2017-0) and through a productivity scholarship to ARP (304156/2019-1); by the American Society of Mammalogists through the Latin American Student Field Research Award to EFA, providing support for a field expedition to the Rio Purus, BR; by the American Museum of Natural History through a Collection Study Grant, allowing EFA to collect morphological data from tree squirrels housed at the AMNH; by the Smithsonian Institution through postdoctoral fellowships to SEP and MTNT (Smithsonian Women's Committee), and through research funds provided to SEP, MTNT, DEW and JEM, covering costs associated to laboratory procedures and DNA sequencing; by the National Geographic Society through an Explorers Grant to SEP (NGS-381R-18), providing support to visit and collect samples housed at the MUSM and for a field expedition to Parque Nacional del Rio Abiseo, PE, and through Support for Women + Dependent Care to SEP (NGS 3758), allowing her to visit and collect additional data in North American museums; and by the Fundação de Amparo à Pesquisa do Estado de São Paulo to ARP (09/16009-1), providing support for several filed expeditions to

the Amazon, BR. The funding body provided resources for sampling and data generation. Funding agencies had no role in the study design, data interpretation, decision to publish, or in the manuscript preparation.

Authors' contributions

EFA, SEP conceived the project with support from ARP, JEM; EFA, SEP, MTNT generated and analyzed the data; EFA, SEP, ARP, JEM interpret and discussed the results; EFA, SEP, ARP wrote the paper with contribution from MTNT, DEW, JEM. All authors have read and approved the manuscript.

Acknowledgements

We are grateful to the following curators and collection support staff, who loaned us preserved tissue samples, allowed us to analyze and obtain destructive samples from valuable historical specimens under their care, or provided important information regarding vouchers: R.S. Voss, E. Hoeger, and N. Duncan (AMNH); I. Moya (CBF); M.R. Alvarez (CMARF); B.D. Patterson and A.W. Ferguson (FMNH); J. Valsecchi and I. Junqueira (IDSM); C.R. Silva and A.F. Sobrinho (IEPA); R.M. Timm and M. Eifler (KU); D.L. Dittmann and R.T. Brumfield (LSU-MZM); E. Pasa (MCN-FZB); C.G. Costa (MCN-M); A.U. Christoff, F.B. Peters and A.M. Gandini (MCNU); J.A. Oliveira and M. Weksler (MN); J. Silva Junior (MPEG); M. Campbell, J. Cook and J. Dunnum (MSB); M.T. Rodrigues and B.M.S. Batista (MTR); V. Pacheco (MUSM); C. Conroy, J.L. Patton, and F.J.M. Pascal (MVZ); M. Vivo, L.F. Silveira and J.G. Barros (MZUSP); J.K. Braun and B.S. Coyner (OMNH); H. Garner, R.D. Bradley, and C. Phillips (TTU); L.P. Costa, Y.L.R. Leite and M.P. Nascimento (UFES-CTA); R. Rossi and T. Semedo (UFMT); A.C. Mendes Oliveira (UFPA); J. Cherem, M. Graipel and E.C. Grisard (UFSC); D.P. Lunde, S.C. Peurach, N.R. Edmison, I. Rochon, M. Krol, J. Jacobs, C. Ludwig, C. Huddleston, D. DiMichele, M. Braun, L.H. Emmons, and A.L. Gardner (USNM); F. Catzeflis (ISEM); A. Ravetta, G.T. Garbino and L.P. Godoy provided tissue samples or allowed us to sample recently collected or uncatalogued specimens.

We especially thank R.S. Voss and S.A. Jansa, who provided helpful suggestions and ideas for an early version of this project, and K.M. Helgen, who provided crucial initial support which allowed us to apply for funding at the Smithsonian Institution. We are also thankful to M. Vivo, for his generosity in sharing his deep knowledge on squirrels with us,

and for kindly donating to us all the qualitative and quantitative data that he amassed from several museums worldwide. We are deeply indebted to L. Emmons, who shared with us precious insights and thoughts about Neotropical squirrels, collected several samples analyzed in this study during her field trips to South America, and photographed and examined specimens for us at the USNM. J.L. Patton is a continuous source of inspiration, and as such, we are indebted to him. We are particularly grateful to L. Parker, M. Venkatraman, S. Castañeda and N. McInerney, who provided crucial help during lab work. We also had amazing help during field expeditions conducted at several sites in the Amazon, Brazil, namely: Rio Purus, by J. Charters; Rio Içá, by J. Dalapicolla, P.R.O. Roth, L.P. Godoy, and L.S. Correa; Rio Japurá, by E.A. Chiquito, P.R.O. Roth, I. Junqueira, and L.S. Correa; Maturacá, by L.F. Silveira and M.T.U. Rodrigues; and at Parque Nacional del Río Abiseo, Peru, by P. Sanchez, E. Rengifo, F. Câmara, P. Peloso, R. Fonseca, and R. Pradel. We also wish to thank the local people who housed and/or helped us during fieldwork in Brazil and Peru. We thank P. Peloso for allowing us to use a photography of *Guerlinguetus brasiliensis*. Lastly, we would like to thank Dr. J.A. Baeza and two anonymous reviewers for insightful comments and suggestions that greatly improved the quality of this manuscript.

References

1. Thorington RW, Koprowski JL, Steele MA, Whatton J. Squirrels of the World. Baltimore: The Johns Hopkins University Press; 2012.
2. Vivo M, Carmignotto AP. Family Sciuridae G. Fischer, 1817. In: Patton JL, Pardiñas UFJ, D'Elía G, editors. Mammals of South America, Volume 2, Rodents. Chicago and London: The University of Chicago Press; 2015. p. 1–48.
3. Burgin CJ, Colella JP, Kahn PL, Upham NS. How many species of mammals are there? J Mammal. 2018;99:1–14.
4. Peres CA. The structure of nonvolant mammal communities in different Amazonian forest types. In: Eisenberg JF, Redford KH, editors. Mammals of the Neotropics. Vol. 3. The central Neotropics: Ecuador, Peru, Bolivia, Brazil. Chicago and London: The University of Chicago Press; 1999. p. 564–81.
5. Mendes CP, Koprowski JL, Galetti M. NEOSQUIRREL: a data set of ecological knowledge on Neotropical squirrels. Mamm Rev. 2019;49:210–25.
6. Roth VL, Mercer JM. Differing rates of macroevolutionary diversification in arboreal squirrels. Curr Sci. 2008;95:857–61.
7. Pečnerová P, Moravec JC, Martínková N. A skull might lie: Modeling ancestral ranges and diet from genes and shape of tree squirrels. Syst Biol. 2015;64:1074–88.
8. Emmons LH, Feer F. Neotropical Rainforest Mammals: a field guide. Second. Chicago and London: The University of Chicago Press; 1997.
9. Voss RS, Emmons LH. Mammalian Diversity in Neotropical Lowland Rainforests: a Preliminary Assessment. Bull Am Museum Nat Hist. 1996;230:1–115. http://orton.catie.ac.cr/cgi-bin/wxis.exe/?IsisScript=OET.xis&method=post&formato=2&cantidad=1&expresion=mfn=013548.
10. Hofreiter M, Paijmans JLA, Goodchild H, Speller CF, Barlow A, Fortes GG, et al. The future of ancient DNA: Technical advances and conceptual shifts. BioEssays. 2015;37:284–93.
11. Mandrioli M. MUSEomica: quando la genomica entra in museo. Quad del Mus Civ di Stor Nat di Ferrara. 2016;4:53–70.
12. McDonough MM, Parker LD, Rotzel McInerney N, Campana MG, Maldonado JE. Performance of commonly requested destructive museum samples for mammalian genomic studies. J Mammal. 2018;99:789–802. doi:10.1093/jmammal/gyy080.
13. Rowe KC, Singhal S, Macmanes MD, Ayroles JF, Morelli TL, Rubidge EM, et al. Museum genomics: Low-cost and high-accuracy genetic data from historical specimens. Mol Ecol Resour. 2011;11:1082–92.
14. Burrell AS, Disotell TR, Bergey CM. The use of museum specimens with high-throughput DNA sequencers. J Hum Evol. 2015;79:35–44. doi:10.1016/j.jhevol.2014.10.015.
15. Guschanski K, Krause J, Sawyer S, Valente LM, Bailey S, Finstermeier K, et al. Next-generation museomics disentangles one of the largest primate radiations. Syst Biol. 2013;62:539–54.
16. Miller W, Drautz DI, Janecka JE, Lesk AM, Ratan A, Tomsho LP, et al. The mitochondrial genome sequence of the Tasmanian tiger (*Thylacinus cynocephalus*). Genome Res. 2009;19:213–20.
17. Chang D, Knapp M, Enk J, Lippold S, Kircher M, Lister A, et al. The evolutionary and phylogeographic history of woolly mammoths: A comprehensive mitogenomic analysis. Sci Rep. 2017;7 September 2016:1–10.

18. Allen JA. Review of the south american Sciuridae. Bull Am Museum Nat Hist. 1915;34:147–309.
19. Cabrera A. Catalogo de los mamideros de America del Sur. Buenos Aires: Imprenta Y Casa Editora Coni; 1957.
20. Hoffmann RS, Anderson CG, Thorington RWJ, Heaney LR. Family Sciuridae. In: Wilson DE, Reeder DM, editors. Mammals Species of the World: a taxonomic and geographic reference. Second. Washington DC: Smithsonian Institute Press; 1993. p. 419–65.
21. Moore JC. Relationships among the living squirrels of the Sciurinae. Bull Am Museum Nat Hist. 1959;118:157–206. http://ezproxy.stanford.edu:2048/login?url=http://digitallibrary.amnh.org/dspace/handle/2246/1265.
22. Thorington RWJ, Hoffmann RS. Family Sciuridae. In: Wilson DE, Reeder DM, editors. Mammals Species of the World: a taxonomic and geographic reference. Third. Washington DC: Jhons Hopikins University Press; 2005. p. 754–818.
23. Hope AG, Malaney JL, Bell KC, Salazar-Miralles F, Chavez AS, Barber BR, et al. Revision of widespread red squirrels (genus: *Tamiasciurus*) highlights the complexity of speciation within North American forests. Mol Phylogenet Evol. 2016;100:170–82.
24. Oshida T, Arslan A, Noda M. Phylogenetic relationships among the old world *Sciurus* squirrels. Folia Zool. 2009;58:14–25.
25. Oshida T, Masuda R. Phylogeny and Zoogeography of Six Squirrel Species of the Genus *Sciurus* (Mammalia, Rodentia), Inferred from Cytochrome b Gene Sequences. Zoolog Sci. 2000;17:405–9.
26. Özkurt Ş, Sözen M, Yiğit N, Çolak E, Verimli R. On the karyology and morphology of *Sciurus anomalus* (Mammalia: Rodentia) in Turkey. Zool Middle East. 1999;18:9–15.
27. Musser GG. A Systematic Study of the Mexican and Guatemalan Gray Squirrel, *Sciurus aureogaster* F. Cuvier (Rodentia: Sciuridae). Misc Publ Museum Zool Univ Michigan. 1968;137:1–112.
28. Lemmon EM, Lemmon AR. High-Throughput Genomic Data in Systematics and Phylogenetics. Annu Rev Ecol Evol Syst. 2013;44:99–121. doi:10.1146/annurev-ecolsys-110512-135822.
29. Wandeler P, Hoeck PEA, Keller LF. Back to the future: museum specimens in population genetics. Trends Ecol Evol. 2007;22:634–42.
30. Meyer M, Stiller M, Nagel S, Nickel B, Palkopoulou E, Mallick S, et al. Palaeogenomes of Eurasian straight-tusked elephants challenge the current view of elephant evolution. Elife. 2017;6:1–14.
31. Poinar HN, Schwarz C, Qi J, Shapiro B, MacPhee RDE, Buigues B, et al. Metagenomics to Paleogenomics: Large-Scale Sequencing of Mammoth DNA. Science. 2006;311:392–4. doi:10.1126/science.1123360.
32. Castañeda-Rico S, León-Paniagua L, Edwards CW, Maldonado JE. Ancient DNA From Museum Specimens and Next Generation Sequencing Help Resolve the Controversial Evolutionary History of the Critically Endangered Puebla Deer Mouse. Front Ecol Evol. 2020;8:1–18.
33. Hawkins MTR, Leonard JA, Helgen KM, McDonough MM, Rockwood LL, Maldonado JE. Evolutionary history of endemic Sulawesi squirrels constructed from UCEs and mitogenomes sequenced from museum specimens. BMC Evol Biol. 2016;16:1–16. doi:10.1186/s12862-016-0650-z.

34. Fabre PH, Upham NS, Emmons LH, Justy F, Leite YLR, Loss AC, et al. Mitogenomic phylogeny, diversification, and biogeography of South American spiny rats. Mol Biol Evol. 2016;34:613–33.
35. Emmons LH, Fabre P. A Review of the *Pattonomys*/*Toromys* Clade (Rodentia: Echimyidae), with Descriptions of a New *Toromys* Species and a New Genus . Am Museum Novit. 2018;3894:1–52.
36. Mason VC, Li G, Helgen KM, Murphy WJ. Efficient cross-species capture hybridization and next-generation sequencing of mitochondrial genomes from noninvasively sampled museum specimens. Genome Res. 2011;21:1695–704.
37. Pinto OM de O. Cinqüenta anos de investigação ornitológica: história das origens e do desenvolvimento da coleção ornitológica do Museu Paulista e de seu subseqüente progresso no Departamento de Zoologia da Secretaria da Agricultura. Arq Zool do Estado São Paulo. 1945;4:261–340.
38. Wiley RH. Alfonso Olalla and His Family: The Ornithological Exploration of Amazonian Peru. Bull Am Museum Nat Hist. 2010;343:1–68.
39. Gutiérrez EE, Helgen KM, McDonough MM, Bauer F, Hawkins MTR, Escobedo-Morales LA, et al. A gene-tree test of the traditional taxonomy of american deer: The importance of voucher specimens, geographic data, and dense sampling. Zookeys. 2017;697 September:87–131.
40. Roure B, Baurain D, Philippe H. Impact of missing data on phylogenies inferred from empirical phylogenomic data sets. Mol Biol Evol. 2013;30:197–214.
41. Huang H, Knowles LL. Unforeseen consequences of excluding missing data from next-generation sequences: Simulation study of rad sequences. Syst Biol. 2016;65:357–65.
42. Streicher JW, Schulte JA, Wiens JJ. How Should Genes and Taxa be Sampled for Phylogenomic Analyses with Missing Data? An Empirical Study in Iguanian Lizards. Syst Biol. 2016;65:128–45.
43. Lemmon AR, Brown JM, Stanger-Hall K, Lemmon EM. The effect of ambiguous data on phylogenetic estimates obtained by maximum likelihood and bayesian inference. Syst Biol. 2009;58:130–45.
44. Wilkinson M. Coping With Abundant Missing Entries in Phylogenetic Inference using Parsimony. Syst Biol. 1995;44:501–14.
45. Sites JW, Marshall JC. Operational criteria for delimiting species. Annu Rev Ecol Evol Syst. 2004;35 Wiens 1999:199–227.
46. Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C. Coalescent-based species delimitation in an integrative taxonomy. Trends Ecol Evol. 2012;27:480–8.
47. Weksler M, Percequillo AR, Voss RS. Ten New Genera of Oryzomyine Rodents (Cricetidae: Sigmodontinae). Am Museum Novit. 2006;3537:1. doi:10.1206/0003-0082(2006)3537[1:TNGOOR]2.0.CO;2.
48. Percequillo AR, Weksler M, Costa LP. A new genus and species of rodent from the Brazilian Atlantic Forest (Rodentia: Cricetidae: Sigmodontinae: Oryzomyini), with comments on oryzomyine biogeography. Zool J Linn Soc. 2011;161:357–90.
49. Pardiñas UFJ, Geise L, Ventura K, Lessa G. A new genus for *Habrothrix angustidens* and *Akodon serrensis* (Rodentia, Cricetidae): again paleontology meets neontology in the legacy of Lund. Mastozoología Neotrop. 2016;23:93–115.
50. Mercer JM, Roth VL. The Effects of Cenozoic Global Change on Squirrel Phylogeny. Science (80-). 2003;299:1568–72.

51. Steppan SJ, Storz BL, Hoffmann RS. Nuclear DNA phylogeny of the squirrels (Mammalia: Rodentia) and the evolution of arboreality from c-myc and RAG1. Mol Phylogenet Evol. 2004;30:703–19.
52. Villalobos F, Gutierrez-Espeleta G. Mesoamerican tree squirrels evolution (Rodentia: Sciuridae): a molecular phylogenetic analysis. Rev Biol Trop. 2014;62:649–57.
53. Hershkovitz P. Mammals of Northern Colombia, Preliminary Report No. 1: Squirrels (Sciuridae). Proc United States Natl Museum. 1947;97:1–46.
54. Pacheco V, Macedo H De, Vivar E, Ascorra C, Arana-Cardo R, Solari S. Lista Anotada de los Mamiferos Peruanos. Occas Pap Conserv Biol. 1995;2:1–35.
55. Pacheco V, Cadenillas R, Salas E, Tello C, Zeballo H. Diversidad y endemismo de los mamíferos del Perú. Rev Peru Biol. 2009;16:5–32.
56. Luo A, Ling C, Ho SYW, Zhu CD. Comparison of methods for molecular species delimitation across a range of speciation scenarios. Syst Biol. 2018;67:830–46.
57. Ahrens D, Fujisawa T, Krammer HJ, Eberle J, Fabrizi S, Vogler AP. Rarity and incomplete sampling in DNA-based species delimitation. Syst Biol. 2016;65:478–94.
58. Dellicour S, Flot JF. The hitchhiker's guide to single-locus species delimitation. Mol Ecol Resour. 2018;18:1234–46.
59. Kekkonen M, Hebert PDN. DNA barcode-based delineation of putative species: Efficient start for taxonomic workflows. Mol Ecol Resour. 2014;14:706–15.
60. Vitecek S, Kučinić M, Previšić A, Živić I, Stojanović K, Keresztes L, et al. Integrative taxonomy by molecular species delimitation: multi-locus data corroborate a new species of Balkan Drusinae micro-endemics. BMC Evol Biol. 2017;17:1–18.
61. Suárez-Villota EY, Carmignotto AP, Brandão MV, Percequillo AR, Silva MJDJ. Systematics of the genus *Oecomys* (Sigmodontinae: Oryzomyini): molecular phylogenetic, cytogenetic and morphological approaches reveal cryptic species. Zool J Linn Soc. 2018;184:182–210.
62. García-Melo JE, Oliveira C, Da Costa Silva GJ, Ochoa-Orrego LE, Garcia Pereira LH, Maldonado-Ocampo JA. Species delimitation of neotropical Characins (Stevardiinae): Implications for taxonomy of complex groups. PLoS One. 2019;14:1–22.
63. Edwards S, Bensch S. Looking forwards or looking backwards in avian phylogeography? A comment on Zink and Barrowclough 2008. Mol Ecol. 2009;18:2930–3. doi:10.1111/j.1365-294X.2009.04270.x.
64. Edwards S V., Kingan SB, Calkins JD, Balakrishnan CN, Bryan Jennings W, Swanson WJ, et al. Speciation in birds: Genes, geography, and sexual selection. Syst Orig Species Ernst Mayr's 100th Anniv. 2005;102:95–119.
65. Oshida T, Masuda R. Phylogeny and Zoogeography of Six Squirrel Species of the Genus *Sciurus* (Mammalia, Rodentia), Inferred from Cytochrome b Gene Sequences. Zoolog Sci. 2000;17:405–9.
66. Pečnerová P, Martínková N. Evolutionary history of tree squirrels (Rodentia, Sciurini) based on multilocus phylogeny reconstruction. Zool Scr. 2012;41:211–9.
67. Roth VL. Cranial Integration in the Sciuridae. Am Zool. 1996;36:14–23.
68. Harrison RG, Bogdanowicz SM, Hoffmann RS, Yensen E, Sherman PW. Phylogeny and evolutionary history of the ground squirrels (Rodentia: Marmotinae). J Mamm Evol. 2003;10:249–76.
69. Hale SL, Greer VL, Koprowski JL, Ramos-Lara N. Microsciurus santanderensis (Rodentia: Sciuridae). Mamm Species. 2018;50:166–9.
70. Arbogast BS, Browne RA, Weigl PD. Evolutionary Genetics and Pleistocene Biogeography of North American Tree Squirrels (*Tamiasciurus*). J Mammal. 2001;82:302–19.

71. Rohland N, Reich D. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Res. 2012;22:939–46.
72. Faircloth BC. Illumiprocessor: a trimmomatic wrapper for parallel adapter and quality trimming. 2013.
73. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.
74. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012;28:1647–9.
75. Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, et al. MITOS: Improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol. 2013;69:313–9. doi:10.1016/j.ympev.2012.08.023.
76. Greiner S, Lehwark P, Bock R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res. 2019;47:W59–64. doi:10.1093/nar/gkz238.
77. Edgar RC. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32:1792–7.
78. Stamatakis A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30:1312–3.
79. Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, et al. Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012;61:539–42.
80. Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. Partitionfinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol Biol Evol. 2017;34:772–3.
81. Altekar G, Dwarkadas S, Huelsenbeck JP, Ronquist F. Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. Bioinformatics. 2004;20:407–15.
82. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst Biol. 2018;67:901–4.
83. Fujisawa T, Barraclough TG. Delimiting species using single-locus data and the generalized mixed yule coalescent approach: A revised method and evaluation on simulated data sets. Syst Biol. 2013;62:707–24.
84. Yang Z. The BPP program for species tree estimation and species delimitation. Curr Zool. 2015;61:854–65.
85. Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, et al. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLOS Comput Biol. 2019;15:e1006650. doi:10.1371/journal.pcbi.1006650.
86. Ezard T, Fujisawa T, Barraclough TG. Splits: species limits by threshold statistics (version 1.0-19). 2009.
87. Talavera G, Dincă V, Vila R. Factors affecting species delimitations with the GMYC model: Insights from a butterfly survey. Methods Ecol Evol. 2013;4:1101–10.
88. Flouri T, Jiao X, Rannala B, Yang Z. Species tree inference with BPP using genomic sequences and the multispecies coalescent. Mol Biol Evol. 2018;35:2585–93.
89. Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W. GEIGER: Investigating evolutionary radiations. Bioinformatics. 2008;24:129–31.
90. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics. 2019;35:526–8.

Tables

Table 1. Details on mitogenome completeness obtained in this study, including success for modern and historical samples.

Percentage of mitogenome completeness	**Number of samples**	**Number of modern samples and percentage of success**	**Number of historical samples and percentage of success**
100%	92	78 (44.1%)	14 (14.9%)
100-80%	70	51 (28.8%)	19 (20.2%)
80-60%	24	11 (6.2%)	13 (13.8%)
60-40%	24	14 (7.9%)	10 (10.6%)
40-20%	22	10 (5.6%)	12 (12.8%)
<20%	32	11 (6.2%)	21 (22.3%)
No coverage	7	2 (1.1%)	5 (5.3%)
Total	**271**	**177**	**94**

Table 2. Summary features of the mitogenome datasets analyzed in this study.

	Dataset 1	**Dataset 2**	**Dataset 3**	**Dataset 4**	**Dataset 5**
Mitogenome completeness per specimen	100%	≥80%	≥60%	≥40%	≥20%
Minimum mitogenome length per specimen	16,376 bp	12,099 pb	7,144 bp	3,884 pb	2,118 pb
Number of specimens included	92	162	186	210	232
Alignment overall missing data	0.1%	3.7%	8.3%	14.2%	20.2%
Number of OTUs included	27	37	41	41	43
Average nodal support (bootstrap) of the inferred ML trees	97.4	98.2	97.0	96.1	95.8

Table 3. (Continued) New taxonomic arrangement proposed for the tribe Sciurini based on the species analyzed and recognized as valid in this study. We did not analyze material from *Tamiasciurus fremonti* sensu [23], *Microsciurus santanderensis* and *M. simonsi* sensu [2], which are treated as valid but not listed below.

Family Sciuridae G. Fischer, 1817
- **Subfamily Sciurinae** G. Fischer, 1817
  - **Tribe Sciurini** G. Fischer, 1817
    - **Genus *Tamiasiurus*** Trouessart, 1880
      - *Tamiasciurus douglasii* (Bachman, 1839)
      - *Tamiasciurus hudsonicus* (Erxleben, 1777)
    - **Genus *Rheithrosciurus*** Gray, 1867
      - *Rheithrosciurus macrotis* (Gray, 1856)
    - **Genus *Sciurus*** Linnaeus, 1758
      - *Sciurus anomalus* Gmelin, 1778
      - *Sciurus lis* Temminck, 1844
      - *Sciurus vulgaris* Linnaeus, 1758
    - **Genus *Hesperosciurus*** Nelson, 1899
      - *Hesperosciurus griseus* (Ord, 1818)
      - *Hesperosciurus aberti* (Woodhouse, 1852)
    - **Genus *Parasciurus*** Trouessart, 1880
      - *Parasciurus arizonensis* (Coues, 1867)
      - *Parasciurus nayaritensis* (J. A. Allen, 1890)
      - *Parasciurus niger* (Linnaeus, 1758)
      - *Parasciurus alleni* (Nelson, 1898)
      - *Parasciurus oculatus* (Peters, 1863)
    - **Genus *Neosciurus*** Trouessart, 1880
      - *Neosciurus carolinensis* (Gmelin, 1788)
    - **Genus *Microsciurus*** J. A. Allen, 1895
      - *Microsciurus alfari* J. A. Allen, 1895
      - *Microsciurus* "species 1"
    - **Genus *Syntheosciurus*** Bangs, 1902
      - *Syntheosciurus brochus* Bangs, 1902
      - *Syntheosciurus granatensis* (Humboldt, 1811)
    - **Genus *Echinosciurus*** Trouessart, 1880
      - *Echinosciurus aureogaster* (F. Cuvier, 1829)
      - *Echinosciurus colliaei* (Richardson, 1839)
      - *Echinosciurus deppei* (Peters, 1863)
      - *Echinosciurus yucatanensis* (J. A. Allen, 1877)
      - *Echinosciurus variegatoides* (Ogilby, 1839)
    - **Genus *Leptosciurus*** Allen, 1915
      - *Leptosciurus mimulus* (Thomas, 1898)
      - *Leptosciurus pucheranii* (Fitzinger, 1867)
      - *Leptosciurus similis* (Nelson, 1899)
      - *Leptosciurus otinus* (Thomas, 1901)

Table 3. (Continuation)

Leptosciurus boquetensis (Nelson, 1903)
Leptosciurus isthmius (Nelson, 1899)
Genus *Simosciurus* J. A. Allen, 1915
Simosciurus nebouxii (I. Geoffroy St.-Hilaire, 1855)
Simosciurus stramineus (Gervais, 1841)
Genus *Guerlinguetus* Gray, 1821
Guerlinguetus aestuans "a"
Guerlinguetus aestuans "b"
Guerlinguetus aestuans "c"
Guerlinguetus brasiliensis (Gmelin, 1788)
Genus *"Microsciurus"*
"Microsciurus" sabanillae Anthony, 1922
"Microsciurus" "species 2"
"Microsciurus" flaviventer (Gray, 1867)
Genus *Hadrosciurus* J. A. Allen, 1915
Hadrosciurus "species 3"
Hadrosciurus igniventris (Wagner, 1842)
Hadrosciurus pyrrhinus (Thomas, 1898)
Hadrosciurus ignitus (Gray, 1867)
Hadrosciurus spadiceus (Olfers, 1818)

Figures

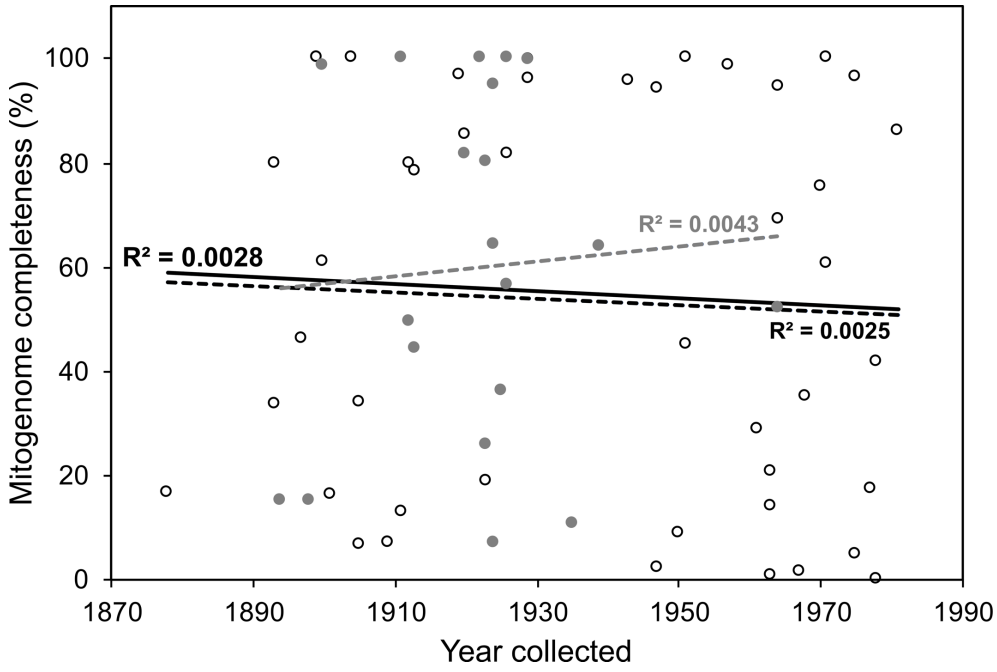


Figure 1. Relationship between specimen age and mitochondrial genome completeness recovered from historical samples of Sciurini. Open circles represent osteocrusts from the National Museum of Natural History (USNM; N = 44), while gray circles represent osteocrusts from the American Museum of Natural History (AMNH; N = 20). Dashed black line represents linear regression based on samples from USNM and dashed gray line from ANMH, while the solid line represents the linear regression based on the samples from both collections. None of the linear regressions performed were significant ($P > 0.05$).

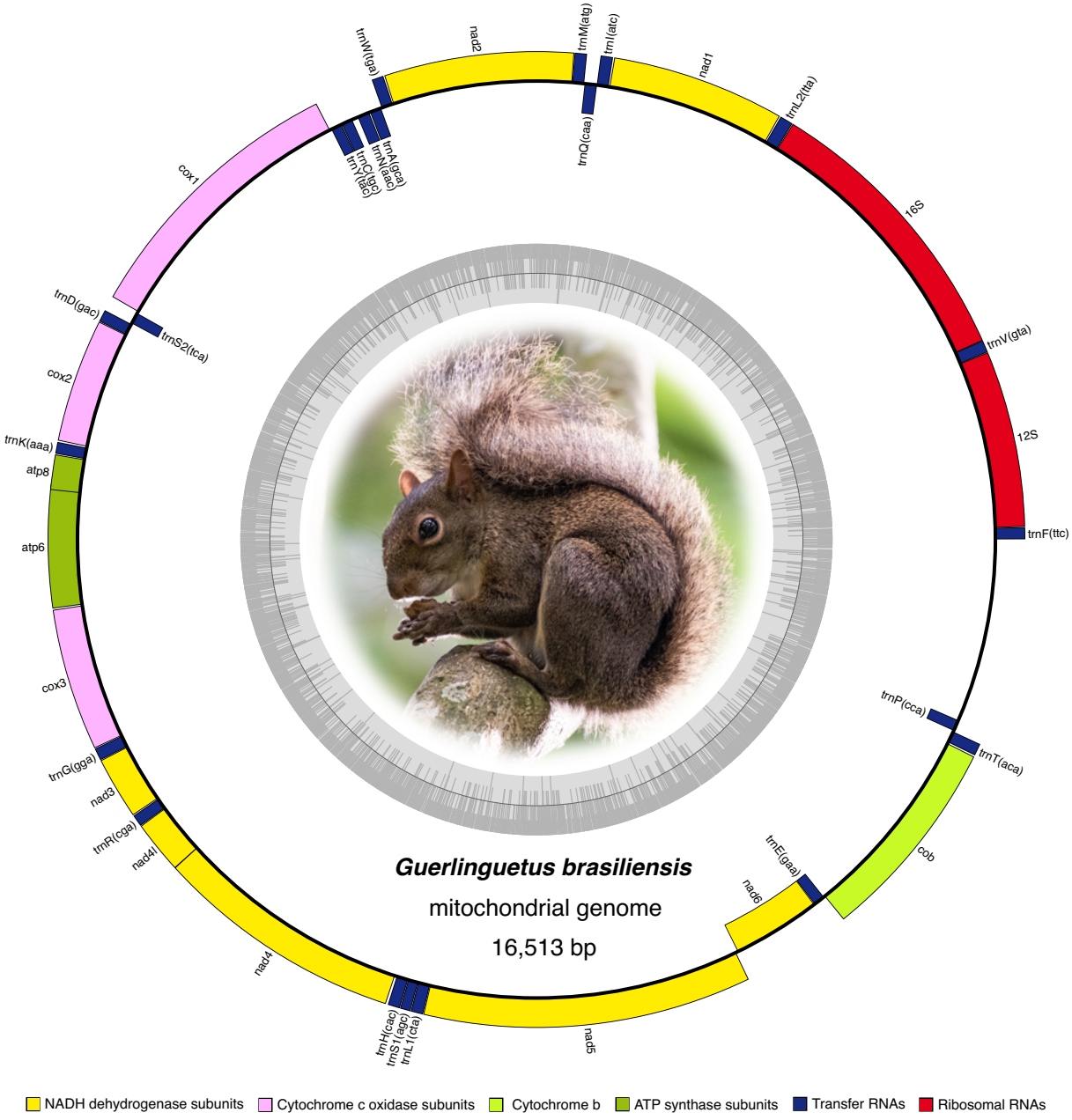


Figure 2. Circular mitochondrial genome map of *Guerlinguetus brasiliensis* depicting the gene organization in tree squirrels. The inner circle shows the GC content along the mitogenome. Photograph of *G. brasiliensis* by Pedro Peloso.

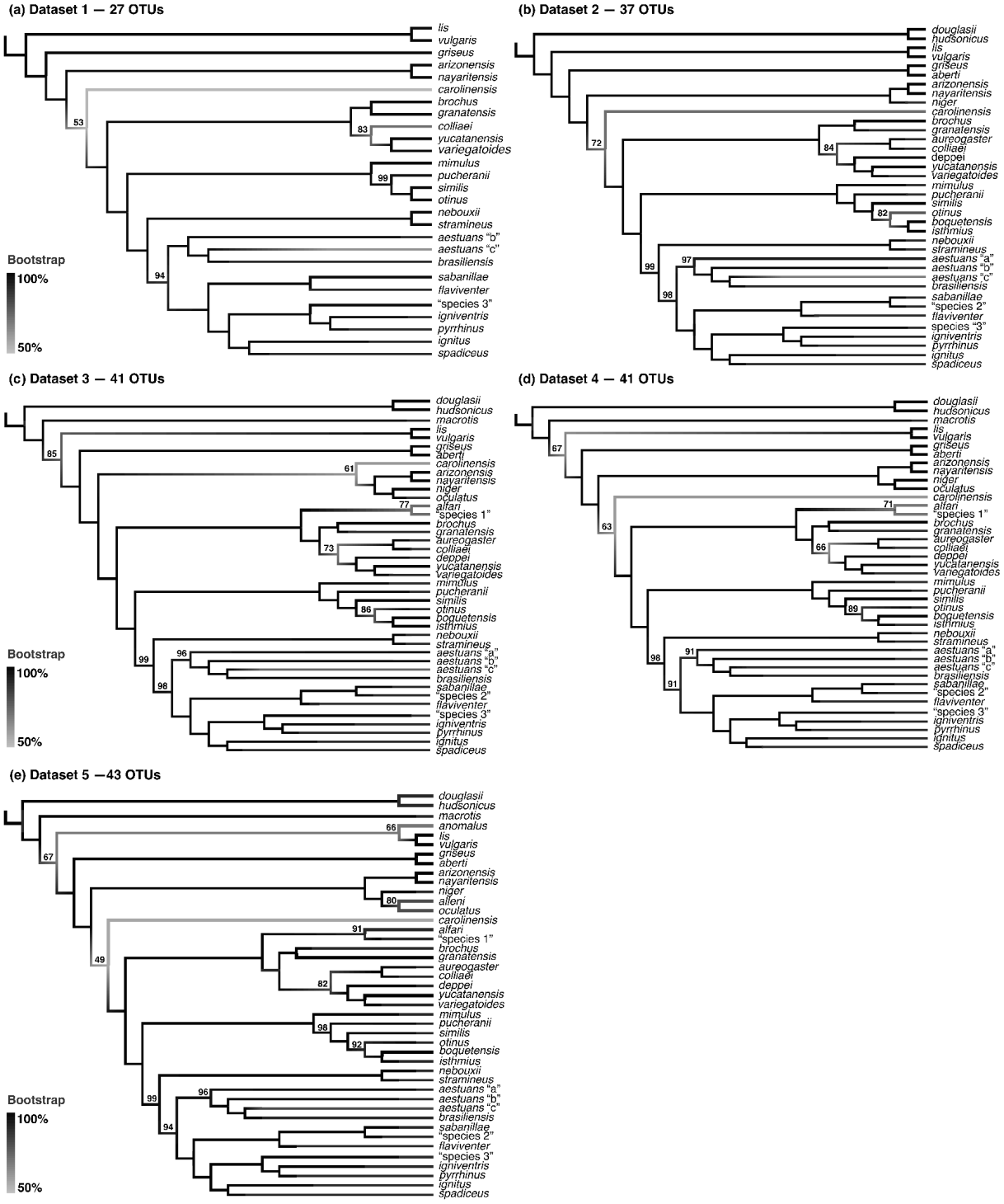


Figure 3. Simplified ML trees of Sciurini based on analyses of five distinct datasets. (a) Dataset 1—92 specimens with no missing data, (b) Dataset 2—162 specimens with <20% of missing data per sample, (c) Dataset 3—186 specimens with <40% of missing data per sample, (d) Dataset 4—210 specimens with <60% of missing data per sample, and (e) Dataset 5—232 specimens with <80% of missing data per sample. Additional details of each dataset are provided in Table 2. Numbers above branches indicate support values for all nodes that presented bootstrap frequencies below 100%.

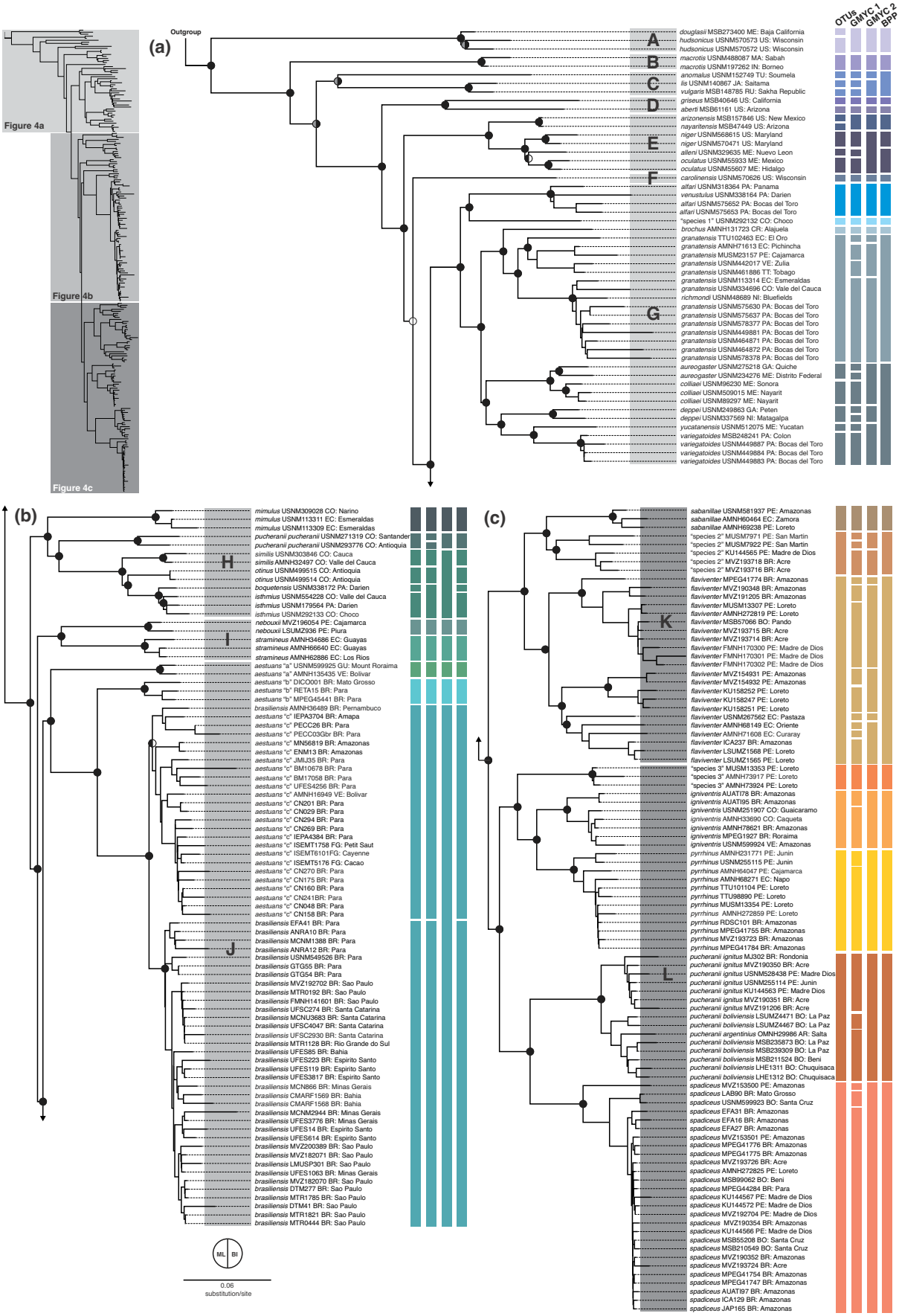


Figure 4 (parts a, b, and c). (See legend on next page)

Figure 4 (parts a, b, and c). Mitochondrial phylogenomic inference of Sciurini recovered by ML analysis of 232 specimens (Dataset 5). Nodal support from the ML bootstrap pseudo-replicates are indicated at each node along with Bayesian posterior probabilities (BI). White wedges indicate bootstrap values ≤ 50%, grey indicates bootstrap frequencies between 50% and 75%, and black indicates bootstrap frequencies ≥ 75%. For BI, white indicates PP < 0.95, whereas black indicates PP ≥ 0.95. Scale at the bottom represents substitutions per site. Letters A–L identify species groups discussed in the text. Except by putative unnamed species ("species 1–3"), terminals are named with specific epithets following [1], [2], and [47] (see methods for details), accompanied by museum voucher numbers and geographic information (country code and state/department). The first column represents Operacional Taxonomic Units (OTUs), recognized with base on original identifications and monophyly. The second to fourth columns correspond to the status of the delimitation of OTUs by different species delimitation methods: generalized mixed Yule-coalescent models using ultrametric trees generated with strict molecular clock (GMYC 1) and relaxed log-normal clock (GMYC 2); and Bayesian Phylogenetics and Phylogeography (BPP) analyses.

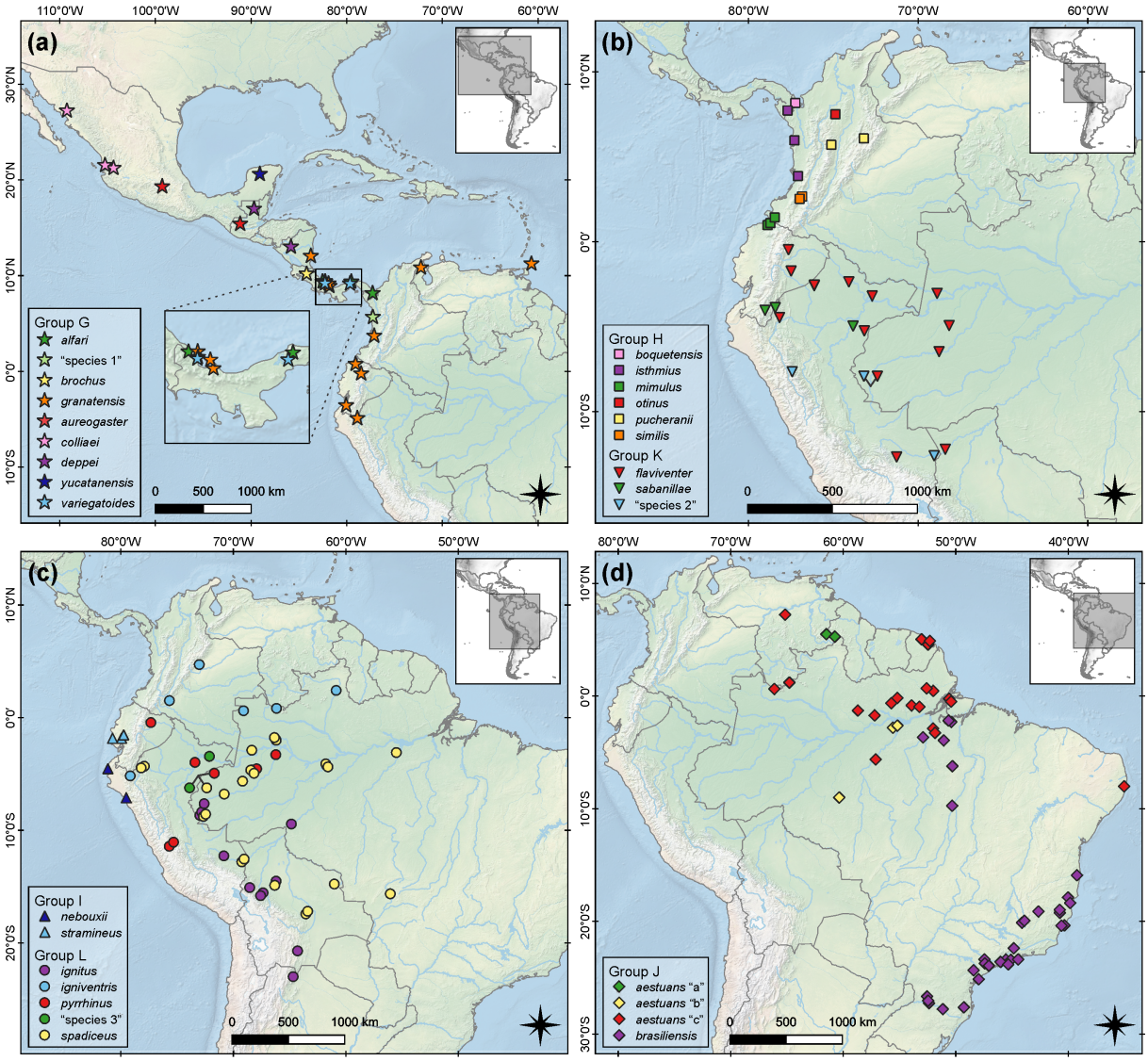


Figure 5. Collecting localities of samples composing Neotropical Sciurini lineages. Maps were generated using QGIS 3.4.4 (http://qgis.org).

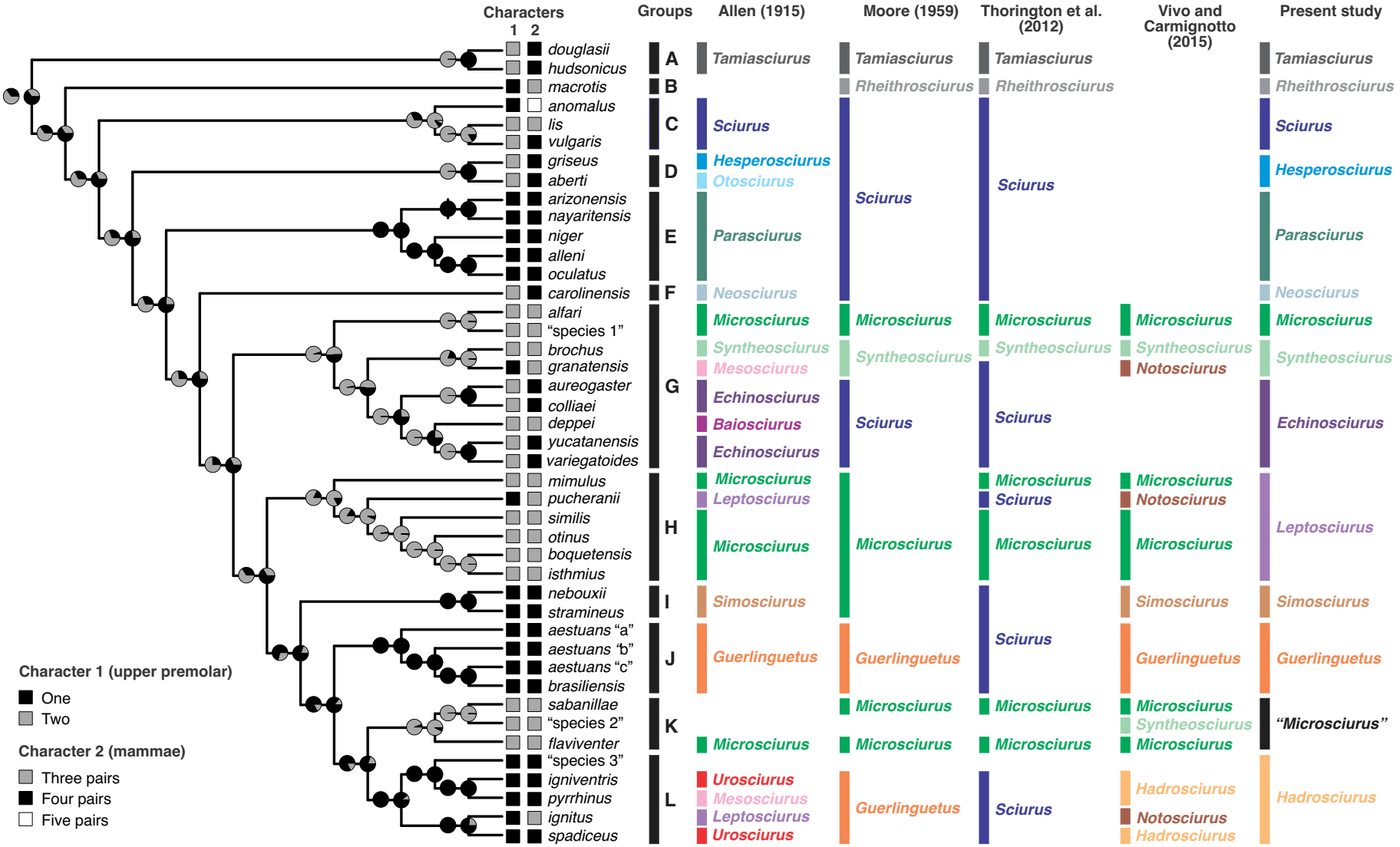
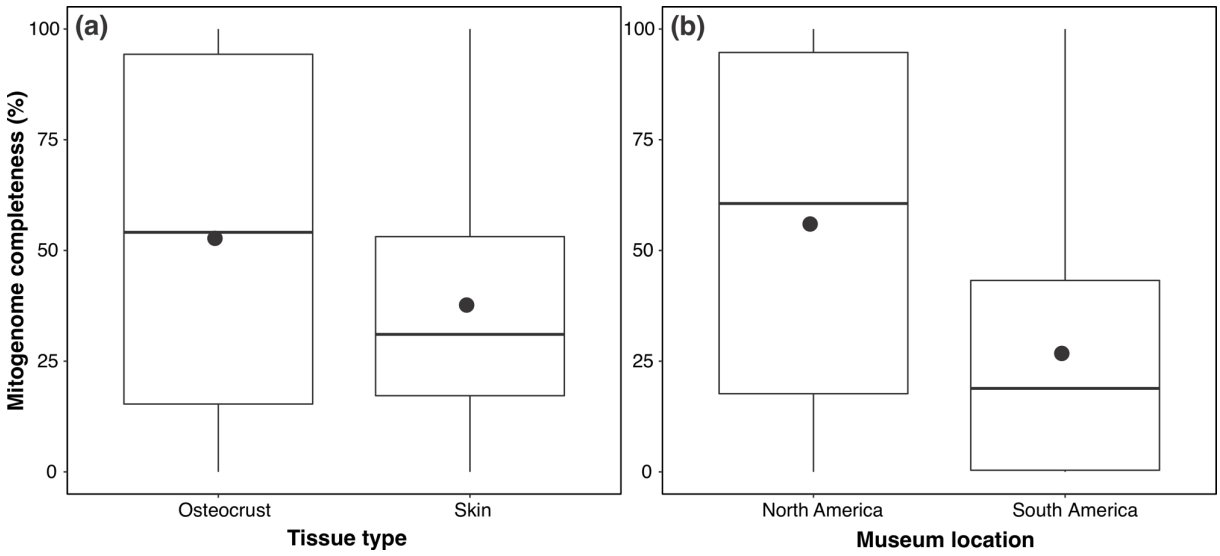


Figure 6. (See legend on next page)

Figure 6. Ancestral state reconstructions of morphological characters and comparison between phylogeny and classification schemes of Sciurini. The first column identifies species groups (A–L) discussed in the text, the second to fifth columns show previous classifications of Sciurini at genus level, and the last column depicts the tentative generic classification suggested by this study. Morphological characters are number of premolars (character 1) and number of pairs of mammae (character 2). See text for character definitions and scoring details. Pie diagrams at internal nodes represent estimated probabilities of alternative states.

Additional files

Additional file 1



Mitogenome recovery success (completeness) obtained from historical samples according to tissue type and museum location. (a) Tissue type: osteocrust samples (N = 70, X = 52.7%) and skin clips (N = 10, X = 37.6%). (b) Location of scientific collection: North America (N = 66, X = 55.9%) and South America (N = 14, X = 26.7%). Samples from both North America and South America collections are included on the tissue type comparison, while both osteocrusties and skin clips are included on the museum location comparison.

Additional file 2

Best-fitting models of sequence evolution used on BI analyses. Numbers between brackets are codon positions.

Subset	Best Model	Number of sites	Mitogenome Partitions
1	GTR+I+G	4578	CDS-ATP8 (2), tRNA-Glu, CDS-CYTB (1), CDS-NADH1 (1), rRNA-12S, tRNA-Ala, CDS-NADH3 (1), CDS-NADH5 (1), CDS-NADH6 (2), tRNA-Cys, tRNA-Asp, tRNA-Trp, CDS-NADH6 (1), tRNA-Ser-AGY, tRNA-Arg, tRNA-Leu-UUR, tRNA-Asn, tRNA-Pro, CDS-NADH4 (3), CDS-NADH2 (1), tRNA-Ser-UCN, tRNA-Phe, CDS-ATP8 (1), tRNA_Lys
2	SYM+I+G	1218	CDS-COII (1), CDS-COI (1), tRNA-Tyr, CDS-COIII (1), tRNA-Val, tRNA-Gly
3	GTR+I+G	1779	tRNA-lle, tRNA-Thr, rRNA-16S, tRNA-His
4	GTR+I+G	4095	CDS-COII (2), CDS-COI (2), CDS-NADH1 (2), CDS-NADH5 (2), CDS-NADH4 (1), CDS-ATP6 (3), CDS-NADH4L (1), CDS-ATP6 (1), CDS-CYTB (2), CDS-COIII (2), tRNA-Leu-CUN, CDS-NADH2 (2), tRNA-Gln, rep-origin, tRNA-Met, CDS-NADH4L (2), CDS-NADH3 (2)
5	GTR+I+G	3708	CDS-COIII (3), CDS-COI (3), CDS-NADH4 (2), CDS-CYTB (3), CDS-NADH6 (3), CDS-NADH2 (3), CDS-NADH4L (3), CDS-NADH3 (3), CDS-NADH5 (3), CDS-NADH1 (3), CDS-COII (3), CDS-ATP6 (2)
6	GTR+I+G	1133	D-loop, CDS-ATP8 (3)

Additional file 3

Summary of models tested to reconstruct the evolution of number of premolars, with respective AIC scores, delta values and AIC weights.

Model	AIC	Delta	AICw
Mk1-ER	42.872	0.000	0.686
Mk-ARD	44.435	1.563	0.314

Additional file 4

Summary of models tested to reconstruct the evolution of pairs of mammae, with respective AIC scores, delta values and AIC weights.

Model	AIC	Delta	AICw
Mk1-ER	65.321	3.599	0.126
Mk-SYM	61.722	0.000	0.763
Mk-ARD	65.581	3.859	0.111

Additional file 5

Catalog data of voucher material.

Voucher material for genetic data analyzed in this study (see Additional file 6) is preserved in the following scientific collections: American Museum of Natural History, USA (AMNH); Colección Boliviana de Fauna, Bolivia (CBF); Coleção de Mamíferos "Alexandre Rodrigues Ferreira", Universidade Estadual de Santa Cruz, Brazil (CMARF); Field Museum of Natural History, USA (FMNH); Instituto de Desenvolvimento Sustentável Mamirauá, Brazil (IDSM); Coleções Científicas Fauna do Amapá, Instituto de Pesquisas Científicas e Tecnológicas do Estado do Amapá, Brazil (IEPA); Institut des Sciences de l'Évolution Montpellier, France (ISEM); University of Kansas Natural History Museum, USA (KU); Coleção do Laboratório de Mamíferos, Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Brazil (LMUSP); Louisiana Museum of Natural History, USA (LSUMZ); Museu de Ciências Naturais, Fundação Zoobotânica do Rio Grande do Sul, Brazil (MCN-FZB); Coleção de Mastozoologia do Museu de Ciências Naturais, Pontifícia Universidade Católica de Minas Gerais, Brazil (MCN-M); Museu de Ciências Naturais, Universidade Luterana do Brasil, Brazil (MCNU); Museu Nacional da Universidade Federal do Rio de Janeiro, Brazil (MN); Museu Paraense Emilio Goeldi, Brazil (MPEG); Museum of Southwestern Biology, University of New Mexico, USA (MSB); Coleção de Tecidos de Vertebrados, Departamento de Zoologia, Instituto de Biociências, Universidade de São Paulo, Brazil (MTR); Museo de Historia Natural, Universidad Nacional Mayor de San Marcos, Peru (MUSM); Museum of Vertebrate Zoology, University of Califórnia, USA (MVZ); Museu de Zoologia da Universidade de São Paulo, Brazil (MZUSP); Sam Noble Oklahoma Museum of Natural History, USA (OMNH); Museum of Texas Tech University, USA (TTU); Coleção de Tecidos Animais, Universidade Federal do Espírito Santo, Brazil (UFES-CTA); Coleção Zoológica, Instituto de Biociências, Universidade Federal de Mato Grosso, Brazil (UFMT); Universidade Federal do Pará, Brazil (UFPA); Universidade Federal de Santa Catarina, Brazil (UFSC); Smithsonian National Museum of Natural History, USA (USNM).

Uncatalogued vouchers are identified (in Additional file 6) by the acronym of the museum where they are currently housed followed by field numbers (in parentheses) with the following prefixes: ANRA = A. Ravetta (MPEG); EFA = E. F. Abreu-Jr. (LMUSP); GTG = G. T. Garbino (MZUSP); and LHE = L. H. Emmons (CBF). The prefixes BM, DTM, ENM, ICA, JAP, and MJ correspond to field series at the LMUSP; CN, DICO, PECC, RDSC, and RETA correspond to field series at the MPEG; AUATI correspond to field series at the IDSM; LAB correspond to field series at the UFMT; and JMIJ correspond to field series at the UFPA.

Additional file 6

List of specimens successfully sequenced and analyzed in this study, with geographic information and GenBank accession numbers for complete mitochondrial genomes. Voucher numbers in bold refer to specimens from which we have used dried tissue (instead of ethanol-preserved tissue). Taxonomic identifications follow the new arrangement proposed here (see text for detailed explanation). The column "Group" refers to the major groups recognized within Sciurini, as recovered by our analyses (see Figures 3 and 4). See Catalog data of voucher material (Additional file 5) for explanations of voucher acronyms. (**Continued**)

Voucher	Species	Group	Country	State/ Department	Municipality/Locality	Latitude	Longitude	Accession mitogenome
KU 144565	*"Microsciurus"* "species 2"	K	Peru	Madre de Dios	Reserva Cuzco Amazonico, 14 km E of Puerto Maldonado	-12.600000	-69.054490	
MUSM 7922	*"Microsciurus"* "species 2"	K	Peru	San Martín	Pataz, Parque Nacional del Río Abiseo, Río El Susto ca. 30 km NE Pataz	-7.647143	-77.417009	
MUSM 7971	*"Microsciurus"* "species 2"	K	Peru	San Martín	Pataz, Parque Nacional del Río Abiseo, Vilcabamba del Pajate	-7.647143	-77.417009	
MVZ 193716	*"Microsciurus"* "species 2"	K	Brazil	Acre	Ocidente, right bank Rio Juruá	-8.566667	-72.800000	
MVZ 193718	*"Microsciurus"* "species 2"	K	Brazil	Acre	Ocidente, right bank Rio Juruá	-8.566667	-72.800000	
AMNH 272819	*"Microsciurus" flaviventer*	K	Peru	Loreto	Rio Galvez, Nuevo San Juan	-5.250000	-73.166670	MT259112
FMNH 170300	*"Microsciurus" flaviventer*	K	Peru	Madre de Dios	Quebrada Aguas Calientes, left bank, Rio Alto Madre de Dios, 2.75 km E Shintuya	-12.668330	-71.269000	
FMNH 170301	*"Microsciurus" flaviventer*	K	Peru	Madre de Dios	Quebrada Aguas Calientes, left bank, Rio Alto Madre de Dios, 2.75 km E Shintuya	-12.668330	-71.269000	
FMNH 170302	*"Microsciurus" flaviventer*	K	Peru	Madre de Dios	Quebrada Aguas Calientes, left bank, Rio Alto Madre de Dios, 2.75 km E Shintuya	-12.668330	-71.269000	
MPEG 41774	*"Microsciurus" flaviventer*	K	Brazil	Amazonas	Jutai, RDS Cujubim	-4.935490	-68.173360	MT259116
MSB 57066	*"Microsciurus" flaviventer*	K	Bolivia	Pando	Santa Rosa	-12.216670	-68.400000	
MUSM 13307	*"Microsciurus" flaviventer*	K	Peru	Loreto	Rio Galvez, Nuevo San Juan	-5.250000	-73.166670	MT259117
MVZ 190348	*"Microsciurus" flaviventer*	K	Brazil	Amazonas	Barro Vermelho, left bank Rio Juruá	-6.466667	-68.766667	
MVZ 191205	*"Microsciurus" flaviventer*	K	Brazil	Amazonas	Barro Vermelho, left bank Rio Juruá	-6.466667	-68.766667	MT259120

(**Continuation**)

MVZ 193714	*"Microsciurus" flaviventer*	K	Brazil	Acre	Flora [=Fazenda Santa Fé], left bank Rio Juruá	-8.600000	-72.850000	MT259121
MVZ 193715	*"Microsciurus" flaviventer*	K	Brazil	Acre	Ocidente, right bank Rio Juruá	-8.566667	-72.800000	MT259122
KU 158247	*"Microsciurus" flaviventer*	K	Peru	Loreto	Teniente Lopez	-2.583330	-76.116670	
KU 158251	*"Microsciurus" flaviventer*	K	Peru	Loreto	Teniente Lopez, 1.5 km N of	-2.557860	-76.116670	MT259114
KU 158252	*"Microsciurus" flaviventer*	K	Peru	Loreto	Teniente Lopez, 1.5 km N of	-2.557860	-76.116670	
MVZ 154931	*"Microsciurus" flaviventer*	K	Peru	Amazonas	Vicinity of Huampami (Aguaruna village), Rio Cenepa	-4.455630	-78.161230	MT259118
MVZ 154932	*"Microsciurus" flaviventer*	K	Peru	Amazonas	Vicinity of Huampami (Aguaruna village), Rio Cenepa	-4.455630	-78.161230	MT259119
AMNH 68149	*"Microsciurus" flaviventer*	K	Ecuador	Oriente	Rio Suno, abajo g marro	-0.474505	-77.640629	
AMNH 71608	*"Microsciurus" flaviventer*	K	Peru	Loreto	Rio Curaray	-2.366667	-74.083333	
LMUSP (ICA 237)	*"Microsciurus" flaviventer*	K	Brazil	Amazonas	Santo Antônio do Içá, margem direita do Rio Içá, oposto à Comunidade São Pedro	-3.038182	-68.879747	MT259113
LSUMZ-M 1565	*"Microsciurus" flaviventer*	K	Peru	Loreto	Quebrada Orán, ca. 5 km N Río Amazonas, 85 km NE Iquitos	-3.199000	-72.706000	
LSUMZ-M 1568	*"Microsciurus" flaviventer*	K	Peru	Loreto	Quebrada Orán, ca. 5 km N Río Amazonas, 85 km NE Iquitos	-3.199000	-72.706000	MT259115
USNM 267562	*"Microsciurus" flaviventer*	K	Ecuador	Pastaza	Sara-Yacu	-1.730000	-77.480000	
AMNH 60464[a]	*"Microsciurus" sabanillae*	K	Ecuador	Zamora-Chinchipe	Zamora, Sabanilla	-4.033333	-79.016667	
AMNH 69238	*"Microsciurus" sabanillae*	K	Peru	Loreto	Requena, Santa Rosa, upper Ucayali River	-4.966667	-73.833333	
USNM 581937	*"Microsciurus" sabanillae*	K	Peru	Amazonas	Cordillera del Condor, Valle Rio Comaina, camp at head of Valley below Table Mountain	-3.877000	-78.413000	MT259123
USNM 275218	*Echinosciurus aureogaster*	G	Guatemala	Quiche	Nebaj	15.406794	-91.147297	
USNM 509015	*Echinosciurus aureogaster*	G	Mexico	Nayarit	Estanzuela	21.235906	-104.383378	
USNM 89297	*Echinosciurus colliaei*	G	Mexico	Nayarit	San Blas	21.538690	-105.271038	
USNM 96230	*Echinosciurus colliaei*	G	Mexico	Sonora	Camoa, Rio Mayo	27.228668	-109.256309	MT240882
USNM 234276	*Echinosciurus colliaei*	G	Mexico	Distrito Federal	Desierto De Los Leones, 12 mi SE of City of Mexico	19.310727	-99.293345	
USNM 249863	*Echinosciurus deppei*	G	Guatemala	Peten	L. Del Sotz	16.982046	-89.692258	
USNM 337569	*Echinosciurus deppei*	G	Nicaragua	Matagalpa	Hla Tepeyac	13.017151	-85.834318	
MSB 248241	*Echinosciurus variegatoides*	G	Panama	Colón	Santa Rosa, Aguas Claras	9.166667	-79.666667	MT240884
USNM 449883	*Echinosciurus variegatoides*	G	Panama	Bocas Del Toro	Isla San Cristobal, Bocatorito	9.242064	-82.261056	

(**Continuation**)

USNM 449884	*Echinosciurus variegatoides*	G	Panama	Bocas Del Toro	Isla San Cristobal, Bocatorito	9.242064	-82.261056	MT264767
USNM 449887	*Echinosciurus variegatoides*	G	Panama	Bocas Del Toro	Tierra Oscura, 3.5 Km S. Tiger Key	9.196700	-82.275600	
USNM 512075	*Echinosciurus yucatanensis*	G	Mexico	Yucatan	Gramal	20.609951	-89.091315	MT240885
USNM 569025	*Glaucomys volans*	Outgroup	United States	Texas	Harrison County, Longhorn Army Ammunition Plant	32.665714	-94.143014	
USNM 569823	*Glaucomys volans*	Outgroup	United States	Pennsylvania	Berks County, Brecknoch Township, Gouglesville, Hirneisen Farm	40.250000	-76.020000	MT259089
AMNH 135435	*Guerlinguetus aestuans* "a"	J	Venezuela	Bolívar	Gran Sabana, Camarata Valley	5.500000	-61.500000	
USNM 599925	*Guerlinguetus aestuans* "a"	J	Guyana		Mount Roraima, N slope	5.283300	-60.750000	MT259077
MPEG 45441	*Guerlinguetus aestuans* "b"	J	Brazil	Pará	Santarém, Comunidade Alto Mentai	-2.800250	-55.583570	MT259078
MPEG (DICO 001)	*Guerlinguetus aestuans* "b"	J	Brazil	Mato Grosso	Reserva Extrativista Guariba-Roosevelt, margem direita Rio Roosevelt	-9.000556	-60.354167	
MPEG (RETA 15)	*Guerlinguetus aestuans* "b"	J	Brazil	Pará	Santarém, Comunidade de Capixauã	-2.612361	-55.192083	MT259079
AMNH 16949	*Guerlinguetus aestuans* "c"	J	Venezuela	Bolívar	Cedeno, Suapure	7.233333	-65.166667	
AMNH 36489	*Guerlinguetus aestuans* "c"	J	Brazil	Pernambuco	São Lourenço da Mata	-8.000000	-35.050000	MT259088
IEPA 3704	*Guerlinguetus aestuans* "c"	J	Brazil	Amapá	Porto Grande, margem direita do Rio Vila Nova, Floresta Estadual do Amapá	0.470000	-52.010000	MT259084
IEPA 4384	*Guerlinguetus aestuans* "c"	J	Brazil	Pará	Almeirim, margem direita do Rio Jari, Igarapé Pacanari	0.681910	-52.593170	
ISEM-T 1758	*Guerlinguetus aestuans* "c"	J	French Guiana	Petit Saut	Sinnamary, Petit Saut	5.050000	-53.050000	
ISEM-T 5176	*Guerlinguetus aestuans* "c"	J	French Guiana	Cacao	Roura, Cacao	4.576944	-52.468611	MT259083
ISEM-T 6101	*Guerlinguetus aestuans* "c"	J	French Guiana	Cayenne	Camp du Tigre, Cayenne	4.908333	-52.308333	
LMUSP (BM 10678)	*Guerlinguetus aestuans* "c"	J	Brazil	Pará	Vitória do Xingu, margem esquerda do Rio Xingu	-2.873760	-52.015519	
LMUSP (BM 17058)	*Guerlinguetus aestuans* "c"	J	Brazil	Pará	Vitória do Xingu, margem esquerda do Rio Xingu	-2.873760	-52.015519	
LMUSP (ENM 13)	*Guerlinguetus aestuans* "c"	J	Brazil	Amazonas	São Gabriel da Cachoeira, 5º PEF Maturaca	0.634605	-66.124988	MT259085
MN 56819	*Guerlinguetus aestuans* "c"	J	Brazil	Amazonas	Barcelos, Rio Katana-u	1.208611	-64.789167	MT259081
MPEG (CN 029)	*Guerlinguetus aestuans* "c"	J	Brazil	Pará	Faro, Flota de Faro, margem esquerda do Rio Nhamundá	-1.714011	-57.213300	
MPEG (CN 048)	*Guerlinguetus aestuans* "c"	J	Brazil	Pará	Faro, Flota de Faro, margem esquerda do Rio Nhamundá	-1.714011	-57.213300	
MPEG (CN 158)	*Guerlinguetus aestuans* "c"	J	Brazil	Pará	Alenquer, ESEC Grão-Pará, porção sul	-0.165489	-55.186400	
MPEG (CN 160)	*Guerlinguetus aestuans* "c"	J	Brazil	Pará	Alenquer, ESEC Grão-Pará, porção sul	-0.165489	-55.186400	MT259087

(**Continuation**)

MPEG (CN 175)	*Guerlinguetus aestuans* "c"	J	Brazil	Pará	Oriximiná, ESEC Grão-Pará, porção norte	-1.285419	-58.695900	
MPEG (CN 201)	*Guerlinguetus aestuans* "c"	J	Brazil	Pará	Oriximiná, ESEC Grão-Pará, porção norte	-1.285419	-58.695900	MT259086
MPEG (CN 241)	*Guerlinguetus aestuans* "c"	J	Brazil	Pará	Almerim, Reserva Biológica Maicuru	-0.828619	-53.931200	
MPEG (CN 269)	*Guerlinguetus aestuans* "c"	J	Brazil	Pará	Almerim, Flota Paru, margem direita do rio Paru de Leste	-0.943969	-53.236300	
MPEG (CN 270)	*Guerlinguetus aestuans* "c"	J	Brazil	Pará	Almerim, Flota Paru, margem direita do rio Paru de Leste	-0.943969	-53.236300	
MPEG (CN 294)	*Guerlinguetus aestuans* "c"	J	Brazil	Pará	Óbidos, ESEC Grão-Pará, porção central	-0.630281	-55.728500	
MPEG (PECC 03)	*Guerlinguetus aestuans* "c"	J	Brazil	Pará	Marajó-Afuá, Rio Cuieiras, PE Charapucú	-0.226960	-50.590000	
MPEG (PECC 26)	*Guerlinguetus aestuans* "c"	J	Brazil	Pará	Marajó-Afuá, Igarapé Torrão, rio Preto, PE Charapucú	-0.418840	-50.494876	MT259080
UFES-CTA 4256	*Guerlinguetus aestuans* "c"	J	Brazil	Pará	Vitória do Xingu, margem esquerda do Rio Xingu	-3.259346	-51.856208	
UFPA (JMIJ 35)	*Guerlinguetus aestuans* "c"	J	Brazil	Pará	Tapajós	-5.613410	-57.122430	MT259082
CMARF 1568	*Guerlinguetus brasiliensis*	J	Brazil	Bahia	Belmonte, Fazenda Ouro Verde	-15.895070	-39.238440	
CMARF 1569	*Guerlinguetus brasiliensis*	J	Brazil	Bahia	Belmonte, Fazenda Ouro Verde	-15.895070	-39.238440	
FMNH 141601	*Guerlinguetus brasiliensis*	J	Brazil	Sao Paulo	Ilha do Cardoso	-25.133333	-47.966667	MT174525
LMUSP 301	*Guerlinguetus brasiliensis*	J	Brazil	São Paulo	Sorocaba, APP Toyota	-23.374444	-47.470833	
LMUSP (DTM 277)	*Guerlinguetus brasiliensis*	J	Brazil	São Paulo	Caraguatatuba, Parque Estadual da Serra do Mar	-23.581412	-45.484681	MT174527
LMUSP (DTM 41)	*Guerlinguetus brasiliensis*	J	Brazil	São Paulo	Caraguatatuba, Parque Estadual da Serra do Mar	-23.581412	-45.484681	
LMUSP (EFA 41)	*Guerlinguetus brasiliensis*	J	Brazil	Pará	Pacajá, LT Xingu-Estreito	-3.929722	-51.069722	MT174526
MCN-M 866	*Guerlinguetus brasiliensis*	J	Brazil	Minas Gerais	Braúna, UHE Porto Estrela	-19.116389	-42.657778	
MCN-M 1388	*Guerlinguetus brasiliensis*	J	Brazil	Pará	Parauapebas, FLONA de Carajás	-6.220806	-50.298406	MT174522
MCN-M 2944	*Guerlinguetus brasiliensis*	J	Brazil	Minas Gerais	Brumadinho, Mina do Córrego do Feijão	-20.113931	-44.115419	
MCNU 3683	*Guerlinguetus brasiliensis*	J	Brazil	Santa Catarina	Anita Garibaldi, margem direita UHE Barra Grande	-27.788560	-51.154506	MT174521
MPEG (ANRA 10)	*Guerlinguetus brasiliensis*	J	Brazil	Pará	Portel, Igarapé Açaituba	-2.224211	-50.544875	MT174528
MPEG (ANRA 12)	*Guerlinguetus brasiliensis*	J	Brazil	Pará	Portel, Igarapé Quirino	-2.166682	-50.644508	
MTR-CIT 1128	*Guerlinguetus brasiliensis*	J	Brazil	Rio Grande do Sul	Itá	-27.256812	-52.392234	MT174518
MTR-CIT 1785	*Guerlinguetus brasiliensis*	J	Brazil	São Paulo	Biritiba Mirim	-23.596572	-46.026164	MT174517
MTR-CIT 1821	*Guerlinguetus brasiliensis*	J	Brazil	São Paulo	Juquitiba	-23.928496	-47.066410	MT174516

(**Continuation**)

MTR-ITM 192	*Guerlinguetus brasiliensis*	J	Brazil	São Paulo	Piedade	-23.716279	-47.422659	MT174520
MTR-ITM 444	*Guerlinguetus brasiliensis*	J	Brazil	São Paulo	Juquitiba	-23.928496	-47.066410	MT174519
MVZ 182070	*Guerlinguetus brasiliensis*	J	Brazil	São Paulo	Ubatuba, Fazenda Capricórnio, 5 km N Ubatuba	-23.416667	-45.116667	MT174515
MVZ 182071	*Guerlinguetus brasiliensis*	J	Brazil	São Paulo	Ubatuba, Praia do Félix	-23.383333	-44.466667	
MVZ 192702	*Guerlinguetus brasiliensis*	J	Brazil	São Paulo	Capão Bonito, Base do Carmo, Fazenda Intervales	-24.333333	-48.416667	
MVZ 200389	*Guerlinguetus brasiliensis*	J	Brazil	São Paulo	São Sebastião, Fazenda da Toca, 2.4 km E, 0.8 km NE (by road) Ilhabela, Ilha de São Sebastião	-23.816667	-45.350000	
MZUSP (GTG 54)	*Guerlinguetus brasiliensis*	J	Brazil	Pará	Santana do Araguaia, Fazenda Fartura	-9.732500	-50.325278	MT174524
MZUSP (GTG 55)	*Guerlinguetus brasiliensis*	J	Brazil	Pará	Santana do Araguaia, Fazenda Fartura	-9.732500	-50.325278	MT174523
UFES-CTA 14	*Guerlinguetus brasiliensis*	J	Brazil	Espírito Santo	Vitória, Parque Estadual da Fonte Grande	-20.350000	-40.350000	
UFES-CTA 85	*Guerlinguetus brasiliensis*	J	Brazil	Bahia	Nova Viçosa, Fazenda Suécia	-17.878889	-40.026389	MT174514
UFES-CTA 119	*Guerlinguetus brasiliensis*	J	Brazil	Espírito Santo	Pancas, Córrego São Bento, Fazenda do Dr. Rolly Luís	-19.225833	-40.761667	MT174513
UFES-CTA 223	*Guerlinguetus brasiliensis*	J	Brazil	Espírito Santo	Águia Branca, Fazenda Pedra Redonda	-18.973333	-40.770556	
UFES-CTA 614	*Guerlinguetus brasiliensis*	J	Brazil	Espírito Santo	Viana, Pimenta	-20.379167	-40.468333	
UFES-CTA 1063	*Guerlinguetus brasiliensis*	J	Brazil	Minas Gerais	Itanhandu, Posses, 13 km SE Itanhandu	-22.383333	-44.850000	MT174512
UFES-CTA 3776	*Guerlinguetus brasiliensis*	J	Brazil	Minas Gerais	Belo Horizonte, Parque das Mangabeiras	-19.945590	-43.910533	MT174511
UFES-CTA 3817	*Guerlinguetus brasiliensis*	J	Brazil	Espírito Santo	Conceição da Barra, Floresta Nacional do Rio Preto	-18.355278	-39.844167	MT174510
UFSC 274	*Guerlinguetus brasiliensis*	J	Brazil	Santa Catarina	Leoberto Leal, A. Wagner	-27.621143	-49.292333	
UFSC 2930	*Guerlinguetus brasiliensis*	J	Brazil	Santa Catarina	Ipuaçu, AHE Quebra Queixo	-26.657551	-52.547408	
UFSC 4047	*Guerlinguetus brasiliensis*	J	Brazil	Santa Catarina	Xaxim, Arvoredo, PCH Arvoredo	-27.040977	-52.466292	MT174509
USNM 549526	*Guerlinguetus brasiliensis*	J	Brazil	Pará	Rio Xingu, east bank	-3.650000	-52.916667	
AMNH 73917	*Hadrosciurus* "species 3"	L	Peru	Loreto	Maynas, Orosa, Amazon River	-3.433333	-72.133333	
AMNH 73924	*Hadrosciurus* "species 3"	L	Peru	Loreto	Maynas, Orosa, Amazon River	-3.433333	-72.133333	
MUSM 13353	*Hadrosciurus* "species 3"	L	Peru	Loreto	Rio Galvez, Nuevo San Juan	-5.250000	-73.166670	
CBF (LHE 1311)	*Hadrosciurus ignitus*	L	Bolivia	Chuquisaca	El Limón, left bank of Rio Santa Marta	-20.700192	-64.300089	
CBF (LHE 1312)	*Hadrosciurus ignitus*	L	Bolivia	Chuquisaca	El Limón, left bank of Rio Santa Marta	-20.700192	-64.300089	
KU 144563	*Hadrosciurus ignitus*	L	Peru	Madre de Dios	Reserva Cuzco Amazonico, 14 km E of Puerto Maldonado	-12.600000	-69.054490	MT259124

(**Continuation**)

LMUSP (MJ 302)	*Hadrosciurus ignitus*	L	Brazil	Rondônia	Porto Velho, margem esquerda do Rio Madeira, Caiçara	-9.437526	-64.849635	
LSUMZ-M 4467	*Hadrosciurus ignitus*	L	Bolivia	La Paz	Prov. B. Saavedra, 83 km by road E Charazani, Cerro Asunta Pata	-15.083000	-68.550000	MT259125
LSUMZ-M 4471	*Hadrosciurus ignitus*	L	Bolivia	La Paz	Prov. B. Saavedra, 83 km by road E Charazani, Cerro Asunta Pata	-15.083000	-68.550000	
MSB 211524	*Hadrosciurus ignitus*	L	Bolivia	Beni	Totaisal, 1 km SW of Estacion Biologica Del Beni	-14.510000	-66.210000	MT259126
MSB 235873	*Hadrosciurus ignitus*	L	Bolivia	La Paz	Serronia Bella Vista	-15.683300	-67.500000	
MSB 239309	*Hadrosciurus ignitus*	L	Bolivia	La Paz	13.7 km by rd NE from La Reserva	-15.733330	-67.516670	MT259127
MVZ 190350	*Hadrosciurus ignitus*	L	Brazil	Acre	Igarapé Porongaba, right bank Rio Juruá	-8.666667	-72.783333	MT259128
MVZ 190351	*Hadrosciurus ignitus*	L	Brazil	Acre	Nova Vida, right bank Rio Juruá	-8.366667	-72.816667	
MVZ 191206	*Hadrosciurus ignitus*	L	Brazil	Acre	Cruzeiro do Sul, left bank Rio Juruá	-7.633333	-72.600000	
OMNH 29986/4031	*Hadrosciurus ignitus*	L	Argentina	Salta	Iruya, ca. 10 km de la intersección de la Ruta Prov. No. 18 el camino a Cortaderas	-22.990278	-64.690611	
USNM 255114	*Hadrosciurus ignitus*	L	Peru	Junín	La Merced	-11.050000	-75.316667	
USNM 528438	*Hadrosciurus ignitus*	L	Peru	Madre de Dios	Boca Rio Manu	-12.266667	-70.850000	
AMNH 33690	*Hadrosciurus igniventris*	L	Colombia	Caquetá	Morelia, La Murelia	1.516667	-75.683333	
AMNH 64047	*Hadrosciurus igniventris*	L	Peru	Cajamarca	Cutervo, Chaupe	-5.166667	-79.166667	
AMNH 78621	*Hadrosciurus igniventris*	L	Brazil	Amazonas	São Gabriel do Cachoeira, Tauá, Uaupés River	0.616667	-69.100000	
IDSM (AUATI 78)	*Hadrosciurus igniventris*	L	Brazil	Amazonas	RESEX Auatí-Paraná	-1.934167	-66.232500	
IDSM (AUATI 95)	*Hadrosciurus igniventris*	L	Brazil	Amazonas	RESEX Auatí-Paraná	-1.934167	-66.232500	MT259090
MPEG 1927	*Hadrosciurus igniventris*	L	Brazil	Roraima	Rio Mucajaí, sul de Boa Vista	2.420000	-60.870000	MT259091
USNM 251907	*Hadrosciurus igniventris*	L	Colombia		Guaicaramo	4.716667	-73.033333	
USNM 599924	*Hadrosciurus igniventris*	L	Venezuela	Amazonas	Neblina Base Camp, Rio Mawarinuma	0.830000	-66.170000	
AMNH 68271	*Hadrosciurus pyrrhinus*	L	Ecuador	Napo	Loreto, San José Nuevo	-0.433333	-77.333333	
AMNH 231771	*Hadrosciurus pyrrhinus*	L	Peru	Junín	Tarma, 2 mi NW San Ramon	-11.416667	-75.700000	
AMNH 272859	*Hadrosciurus pyrrhinus*	L	Peru	Loreto	Rio Galvez, Nuevo San Juan	-5.250000	-73.166670	
MPEG 41755	*Hadrosciurus pyrrhinus*	L	Brazil	Amazonas	Jutai, RDS Cujubim, P.1., transecto mata de terra firme	-4.654170	-68.323940	MT259092
MPEG 41784	*Hadrosciurus pyrrhinus*	L	Brazil	Amazonas	Jutai, RDS Cujubim, P.2., descendo margem esquerda rio Mutum	-4.935490	-68.173360	MT259093

(**Continuation**)

MPEG (RDSC 101)	*Hadrosciurus pyrrhinus*	L	Brazil	Amazonas	Jutai, RDS Cujubim, baixo rio Mutum, afluente dir. médio rio Jutai, afluente direito alto rio Solimoes, comunidade Pirarucu	-4.670524	-68.132023	MT259096
MUSM 13354	*Hadrosciurus pyrrhinus*	L	Peru	Loreto	Rio Galvez, Nuevo San Juan	-5.250000	-73.166670	MT259094
MVZ 193723	*Hadrosciurus pyrrhinus*	L	Brazil	Amazonas	Colocação Vira-Volta, left bank Rio Juruá on Igarapé Arabidi, affluent of Paranã Breu	-3.283333	-66.233333	MT259095
TTU 98890	*Hadrosciurus pyrrhinus*	L	Peru	Loreto	Maynas, Iquitos, 25 km S, Estacion Biologica Allpahuayo	-3.966670	-73.416670	
TTU 101104	*Hadrosciurus pyrrhinus*	L	Peru	Loreto	Maynas, Iquitos, 25 km S, Estacion Biologica Allpahuayo	-3.966670	-73.416670	
USNM 255115	*Hadrosciurus pyrrhinus*	L	Peru	Junín	La Merced	-11.050000	-75.316667	
AMNH 272825	*Hadrosciurus spadiceus*	L	Peru	Loreto	Rio Galvez, Nuevo San Juan	-5.250000	-73.166670	MT259097
IDSM (AUATI 97)	*Hadrosciurus spadiceus*	L	Brazil	Amazonas	RDS Mamirauá	-2.000278	-66.201111	
KU 144566	*Hadrosciurus spadiceus*	L	Peru	Madre de Dios	Reserva Cuzco Amazonico, 14 km E of Puerto Maldonado	-12.600000	-69.054490	MT259101
KU 144567	*Hadrosciurus spadiceus*	L	Peru	Madre de Dios	Reserva Cuzco Amazonico, 14 km E of Puerto Maldonado	-12.600000	-69.054490	
KU 144572	*Hadrosciurus spadiceus*	L	Peru	Madre de Dios	Reserva Cuzco Amazonico, 14 km E of Puerto Maldonado	-12.600000	-69.054490	
LMUSP (EFA 16)	*Hadrosciurus spadiceus*	L	Brazil	Amazonas	Anori, margem esquerda do Rio Purus, Comunidade do Caua-Cuiuanã, proximidades da Comunidade	-4.240372	-61.724023	
LMUSP (EFA 27)	*Hadrosciurus spadiceus*	L	Brazil	Amazonas	Anori, margem esquerda do Rio Purus, Comunidade do Caua-Cuiuanã, Igarapé do Cuiuanã, trilha do Acarizinho	-4.168714	-61.726627	MT259098
LMUSP (EFA 31)	*Hadrosciurus spadiceus*	L	Brazil	Amazonas	Anori, margem esquerda do rio Purus, Comunidade do Caua-Cuiuanã, mata atrás da Comunidade	-4.234045	-61.734642	
LMUSP (ICA 129)	*Hadrosciurus spadiceus*	L	Brazil	Amazonas	Santo Antônio do Içá, margem esquerda do Rio Içá, Comunidade Cuiauá ou Monte Tabor	-2.885465	-68.368681	MT259099
LMUSP (JAP 165)	*Hadrosciurus spadiceus*	L	Brazil	Amazonas	Japurá, margem direita do Rio Japurá, antiga Vila de Santa Fé, trilha da Canoa Virada	-1.764217	-66.357172	MT259100
MPEG 41747	*Hadrosciurus spadiceus*	L	Brazil	Amazonas	Jutai, RDS Cujubim, P.4, a 500m no transecto	-5.638010	-69.187770	MT259103
MPEG 41754	*Hadrosciurus spadiceus*	L	Brazil	Amazonas	Jutai, RDS Cujubim, P.1. Boca do Igarapé Sto. Antonio	-4.654170	-68.323940	MT259104
MPEG 41775	*Hadrosciurus spadiceus*	L	Brazil	Amazonas	Jutai, RDS Cujubim, P.2. Igapó, rio Mutum	-4.935490	-68.173360	MT259105
MPEG 41776	*Hadrosciurus spadiceus*	L	Brazil	Amazonas	Jutai, RDS Cujubim, P.2. Igapó, rio Mutum	-4.935490	-68.173360	MT259106
MPEG 44284	*Hadrosciurus spadiceus*	L	Brazil	Pará	Santarém, Comunidade de Boim	-3.093440	-55.522360	MT259107
MSB 55208	*Hadrosciurus spadiceus*	L	Bolivia	Santa Cruz	San Miguel Rincon	-17.416667	-63.566667	MT259108
MSB 99062	*Hadrosciurus spadiceus*	L	Bolivia	Beni	Totaisal	-14.881111	-66.328056	
MSB 210549	*Hadrosciurus spadiceus*	L	Bolivia	Santa Cruz	San Miguel Rincon	-17.383333	-63.533333	

(**Continuation**)

MVZ 153500	*Hadrosciurus spadiceus*	L	Peru	Amazonas	Chichijam Entsa [=Chichijam Creek], headwaters Rio Huampami, 3 hrs by trail N Huampami	-4.441670	-78.115020	
MVZ 153501	*Hadrosciurus spadiceus*	L	Peru	Amazonas	Kagka (Aguaruna village), Rio Kagka, tributary of Rio Comaina	-4.458980	-78.195050	
MVZ 190352	*Hadrosciurus spadiceus*	L	Brazil	Amazonas	Colocação Sabiá, left bank Rio Juruá	-6.783330	-70.816670	MT259109
MVZ 190354	*Hadrosciurus spadiceus*	L	Brazil	Amazonas	Colocação Sabiá, left bank Rio Juruá	-6.783330	-70.816670	
MVZ 192704	*Hadrosciurus spadiceus*	L	Peru	Madre de Dios	Albergue Cuzco Amazonico	-12.550000	-69.050000	MT259110
MVZ 193724	*Hadrosciurus spadiceus*	L	Brazil	Acre	Ocidente, right bank Rio Juruá	-8.566667	-72.800000	
MVZ 193726	*Hadrosciurus spadiceus*	L	Brazil	Acre	Ocidente, right bank Rio Juruá	-8.566667	-72.800000	MT259111
UFMT (LAB 90)	*Hadrosciurus spadiceus*	L	Brazil	Mato Grosso	Cuiabá, Bairro Coophema	-15.634458	-56.060244	MT259102
USNM 599923	*Hadrosciurus spadiceus*	L	Bolivia	Santa Cruz	El Refugio	-14.767222	-61.034722	
MSB 61161	*Hesperosciurus aberti*	D	United States	Arizona	Graham County, Grahm Mtns. Hospital Flats	32.665100	-109.911000	MT211954
MSB 40646	*Hesperosciurus griseus*	D	United States	California	Mariposa County, 8 mi N, 3 mi E Oakhurst	37.446031	-119.598744	MT211955
USNM 584420	*Hylopetes phayrei*	Outgroup	Myanmar	Mandalay	Pyin-Oo-Lwin (Maymyo), 6.8 mi. ENE, Yangon Monastary	22.067500	96.563500	
USNM 338172	*Leptosciurus boquetensis*	H	Panama	Darien	Cerro Mali	8.170000	-77.230000	
USNM 179564	*Leptosciurus isthmius*	H	Panama	Darien	Cana	7.730000	-77.680000	
USNM 292133	*Leptosciurus isthmius*	H	Colombia	Choco	Baudo Mountains, Rio Jurubida	5.970000	-77.280000	
USNM 554228	*Leptosciurus isthmius*	H	Colombia	Valle del Cauca	Beuhaventure, 6 Km N	3.870000	-77.080000	
USNM 113309	*Leptosciurus mimulus*	H	Ecuador	Esmeraldas	San Javier	1.070000	-78.780000	
USNM 113311	*Leptosciurus mimulus*	H	Ecuador	Esmeraldas	Carondelet	1.100000	-78.700000	MT259072
USNM 309028	*Leptosciurus mimulus*	H	Colombia	Narino	La Guayacana	1.430000	-78.450000	
USNM 499514	*Leptosciurus otinus*	H	Colombia	Antioquia	Zaragoza, 23 Km S, 22 Km W, at Providencia	7.500000	-74.866667	
USNM 499515	*Leptosciurus otinus*	H	Colombia	Antioquia	Zaragoza, 25 Km S, 22 Km W, at La Tirana	7.500000	-74.866667	MT259073
USNM 271319	*Leptosciurus pucheranii*	H	Colombia	Santander	Virolín, Santander Sur, 28 Km S of Charala, Diutama Road	6.083333	-73.200000	
USNM 293776	*Leptosciurus pucheranii*	H	Colombia	Antioquia	La Bodega, S Side Rio Negrito, Highway Sonson-Narino	5.700000	-75.116667	MT240886
AMNH 32497	*Leptosciurus similis*	H	Colombia	Cauca	El Tambo, La Gallera	2.583333	-76.916667	MT259074
USNM 303846	*Leptosciurus similis*	H	Colombia	Cauca	Cerro Munchique	2.530000	-76.980000	

(**Continuation**)

USNM 292132	*Microsciurus* "species 1"	G	Colombia	Choco	Rio Nuqui, Baudo Mountains, Base	5.670000	-77.270000	
USNM 318364	*Microsciurus alfari*	G	Panama	Panama	Candelaria Hydrographic Station, Rio Pequeni	9.370000	-79.530000	
USNM 338164	*Microsciurus alfari*	G	Panama	Darien	Cerro Tacarcuna	8.170000	-77.300000	
USNM 575652	*Microsciurus alfari*	G	Panama	Bocas Del Toro	Nuri	8.913025	-81.815461	
USNM 575653	*Microsciurus alfari*	G	Panama	Bocas Del Toro	Nuri	9.395119	-82.531526	
USNM 570626	*Neosciurus carolinensis*	F	United States	Wisconsin	Portage County, Stevens Point, 2633 Ellis Street, 54481	44.521300	-89.560900	MT240881
USNM 329635	*Parasciurus alleni*	E	Mexico	Nuevo Leon	El Potosi	24.874760	-100.237306	
MSB 157846	*Parasciurus arizonensis*	E	United States	New Mexico	Grant County, Gila River near Spar Canyon	33.023350	-108.535650	MT240880
MSB 47449	*Parasciurus nayaritensis*	E	United States	Arizona	Cochise County, Chiricahua Mts, 2 mi SW Paradise	31.914200	-109.242400	MT240883
USNM 568615	*Parasciurus niger*	E	United States	Maryland	Montgomery County, Potomac River, south side of Watkins Island	29.042200	-77.279200	
USNM 570471	*Parasciurus niger*	E	United States	Maryland	Dorchester County, Vienna, ca. 1.5 mi SSW on Elliots Island Road	38.469200	-75.845300	
USNM 55607	*Parasciurus oculatus*	E	Mexico	Hidalgo	Tulancingo	20.121236	-98.359043	
USNM 55933	*Parasciurus oculatus*	E	Mexico	Mexico	Volcano Toluca, N Slope	19.127644	-99.754869	
USNM 197262	*Rheithrosciurus macrotis*	B	Indonesia	Borneo	Kalimantan, Sungai Menganne	0.490000	117.570000	
USNM 488087	*Rheithrosciurus macrotis*	B	Malaysia	Sabah	Mount Kinabalu National Park, Ranau District, Poring	6.000000	116.680000	
USNM 152749	*Sciurus anomalus*	C	Turkey		Soumela	40.686465	39.655032	
USNM 140867	*Sciurus lis*	C	Japan	Saitama Prefecture	Musasi, Titibu	35.990278	139.076389	MT134013
MSB 148785	*Sciurus vulgaris*	C	Russia	Sakha Republic	Kenkeme River, 40 km W Yakutsk	62.070030	128.938310	MT211956
LSUMZ-M 936	*Simosciurus nebouxii*	I	Peru	Piura	Pariñas, 7 km N, 15 km E Talara	-4.533000	-81.150000	
MVZ 196054	*Simosciurus nebouxii*	I	Peru	Cajamarca	2.5 km N (by air) Monte Seco, Rio Zana	-7.110517	-79.516486	MT259075
AMNH 34686	*Simosciurus stramineus*	I	Ecuador	Guayas	Daule	-1.833333	-79.933333	
AMNH 62886	*Simosciurus stramineus*	I	Ecuador	Los Ríos	Vinces	-1.533333	-79.750000	MT259076
AMNH 66640	*Simosciurus stramineus*	I	Ecuador	Guayas	Santa Elena, Cerro Manglar Alto	-1.833333	-80.733333	
AMNH 131723[b]	*Syntheosciurus brochus*	G	Costa Rica	Alajuela	Poás, Volcan Poás	10.198092	-84.198861	MT240890
AMNH 71613	*Syntheosciurus granatensis*	G	Ecuador	Pichincha	Quito, La Carolina	-0.216667	-78.500000	MT240887

(**Continuation**)

MUSM 23157	*Syntheosciurus granatensis*	G	Peru	Cajamarca	1 km S Hito Jesus	-4.895000	-78.895278	
TTU 102463	*Syntheosciurus granatensis*	G	Ecuador	El Oro	Near to La Victoria, road from Arenillas to Puyango	-3.550000	-80.066670	
USNM 48689	*Syntheosciurus granatensis*	G	Nicaragua		Escondido River, 2.5 mi from Bluefields	12.051127	-83.760050	
USNM 113314	*Syntheosciurus granatensis*	G	Ecuador	Esmeraldas	Pambilar	0.766667	-79.083333	
USNM 334696	*Syntheosciurus granatensis*	G	Colombia	Valle del Cauca	Rio Raposo, Colombia Pacific Coast Virology Field Station	3.716667	-77.133333	
USNM 442017	*Syntheosciurus granatensis*	G	Venezuela	Zulia	Nr. Cerro Azul, 33 Km NW La Paz	10.850000	-72.250000	
USNM 461886	*Syntheosciurus granatensis*	G	Trinidad and Tobago	Tobago	Runnemede	11.250000	-60.700000	MT240888
USNM 449881	*Syntheosciurus granatensis*	G	Panama	Bocas Del Toro	Tierra Oscura, 3.5 Km S. Tiger Key	9.196700	-82.275600	
USNM 464871	*Syntheosciurus granatensis*	G	Panama	Bocas Del Toro	Isla Colón, La Gruta	9.400000	-82.266667	
USNM 464872	*Syntheosciurus granatensis*	G	Panama	Bocas Del Toro	Isla Colón, La Gruta	9.400000	-82.266667	
USNM 575630	*Syntheosciurus granatensis*	G	Panama	Bocas Del Toro	Nuri	8.913025	-81.815461	
USNM 575637	*Syntheosciurus granatensis*	G	Panama	Bocas Del Toro	Nuri	8.913025	-81.815461	MT240889
USNM 578377	*Syntheosciurus granatensis*	G	Panama	Bocas Del Toro	Peninsula Valiente, Punta Alegre	9.162550	-81.905133	
USNM 578378	*Syntheosciurus granatensis*	G	Panama	Bocas Del Toro	Peninsula Valiente, Punta Alegre	9.162550	-81.905133	
MSB 273400	*Tamiasciurus douglasii*	A	Mexico	Baja California	Sierra San Pedro Martir, 15 mi E Meling Ranch	30.905278	115.501111	
USNM 570572	*Tamiasciurus hudsonicus*	A	United States	Wisconsin	Portage County	44.333400	-89.381400	
USNM 570573	*Tamiasciurus hudsonicus*	A	United States	Wisconsin	Portage County	44.333400	-89.381400	

[a]Holotype of *Microsciurus sabanillae* Anthony, 1920.

[b]Holotype of *Syntheosciurus poasensis* (Goodwin, 1942), considered as junior-synonym of *S. brochus* by Thorington et al. (2012).

Additional file 7

State coding used on ancestral state reconstruction analyses. Number of upper premolars were coded as one (1) or two (2), and number of pairs of mammae were coded as three (3), four (4), or five (5). Taxonomic identifications follow the new arrangement proposed here (see text for detailed explanation). (**Continued**)

Species	Upper premolars	Pairs of mammae
"*Microsciurus*" "species 2"	2	3
"Microsciurus" flaviventer	2	3
"Microsciurus" sabanillae	2	3
Echinosciurus aureogaster	2	4
Echinosciurus colliaei	2	4
Echinosciurus deppei	2	3
Echinosciurus variegatoides	2	4
Echinosciurus yucatanensis	2	4
Guerlinguetus aestuans "a"	1	4
Guerlinguetus aestuans "b"	1	4
Guerlinguetus aestuans "c"	1	4
Guerlinguetus brasiliensis	1	4
Hadrosciurus "species 3"	1	4
Hadrosciurus ignitus	1	3
Hadrosciurus igniventris	1	4
Hadrosciurus pyrrhinus	1	4
Hadrosciurus spadiceus	1	4
Hesperosciurus aberti	2	4
Hesperosciurus griseus	2	4
Leptosciurus boquetensis	2	3
Leptosciurus isthmius	2	3
Leptosciurus mimulus	2	3
Leptosciurus otinus	2	3
Leptosciurus pucheranii	1	3
Leptosciurus similis	2	3
Microsciurus "species 1"	2	3
Microsciurus alfari	2	3
Neosciurus carolinensis	2	4
Parasciurus alleni	1	4
Parasciurus arizonensis	1	4
Parasciurus nayaritensis	1	4
Parasciurus niger	1	4

(**Continuation**)

Parasciurus oculatus	1	4
Rheithrosciurus macrotis	1	3
Sciurus anomalus	1	5
Sciurus lis	2	3
Sciurus vulgaris	2	4
Simosciurus nebouxii	1	4
Simosciurus stramineus	1	4
Syntheosciurus brochus	2	3
Syntheosciurus granatensis	1	3
Tamiasciurus douglasii	2	4
Tamiasciurus hudsonicus	2	4

3. CHAPTER 2: Spatiotemporal diversification of tree squirrels: is the South American invasion and speciation really that recent and fast?

Parasciurus niger (Credit: Pedro Peloso)

Published

Abreu-Jr, E.F., Pavan, S.E., Tsuchiya, M.T.N., Wilson, D., Percequillo, A.R., and Maldonado, J.E. 2020. Spatiotemporal Diversification of Tree Squirrels: Is the South American Invasion and Speciation Really That Recent and Fast? *Frontiers in Ecology and Evolution* 8(230): 1–14. doi: 10.3389/fevo.2020.00230

Spatiotemporal diversification of tree squirrels: is the South American invasion and speciation really that recent and fast?

Abstract

Tree squirrels (Sciurinae, Sciurini) represent a diverse radiation that successfully colonized Europe, Asia and the Americas during the Miocene-Pliocene, but information on their evolutionary history remains unclear. In the Neotropics, they have been shown to exhibit the highest rate of diversification amongst all arboreal squirrels, with strikingly high species accumulation rates in the past 3 My. In this study, we investigated the tempo and mode of diversification of tree squirrels using a mitogenome dataset that includes 43 Sciurini species. Our results corroborate the date of origin of the tribe Sciurini around 14 Mya (13.4-15.5) but suggest that their ancestral area was most likely in North America. This is in contrast to previous findings that suggested that the ancestors of this tribe occupied Eurasia. We estimated that cladogenetic events leading to the Eurasian lineages occurred twice at 10.5 and 9.7 Mya. Current North American genera originated in a temporal window from 6.2 to 2.3 Mya, and the origin of the Neotropical radiation was estimated to have occurred around 6 Mya in northwestern South America, in the Pacific dominion. Remarkably, our results indicate that tree squirrels entered South America at an earlier date than previously estimated. This could have happened either through a land corridor connecting the Caribbean islands or through the Panamanian land bridge. Most cladogenetic events in Eurasia and North America appear to have occurred either late in the Miocene or in the Pleistocene, while the majority of Neotropical cladogenetic events occurred along the Pliocene—right after the South American invasion. We found a fairly constant speciation rate for tree squirrels (averaging 0.29), which contrasts with the peak of lineage accumulation observed in the Pliocene. The absence of fluctuations in the diversification rate may be the result of several extinction events that were responsible for equalizing the number of lineages maintained over time. Finally, we conclude that the South American invasion was not as recent as previously inferred, but the diversification there was indeed very fast.

Keywords: Ancestral range, Founder-event speciation, Isthmus of Panama, Mitogenome, Neotropics, Sciurini.

3.1 Introduction

Tree squirrels (Sciurinae, Sciurini) represent a diverse radiation that successfully colonized forested landscapes of Europe, Asia and the Americas during the Miocene and Pliocene (Mercer and Roth, 2003; Pečnerová et al., 2015). The timing of the origin of the tribe is imprecise and has been estimated to have occurred from 19 to 13 million years ago (Mya; Mercer and Roth, 2003; Fabre et al., 2012; Pečnerová et al., 2015; Zelditch et al., 2015). In addition, the uncertainties regarding the location of their ancestral geographic range have also contributed to the blurred inferences of the evolutionary history of tree squirrels.

Pečnerová and Martínková (2012) suggested that Sciurini originated in the northern hemisphere, not distinguishing between Eurasia and North America. A similar result was found by Rocha et al. (2016), who proposed a Holarctic distribution for the ancestor of tree squirrels. Subsequently, Pečnerová et al. (2015) estimated a more restricted ancestral range of Sciurini, confined to the Palearctic region.

Discrepancies on inferences about the spatiotemporal diversification of Sciurini are likely a result of the limited datasets, in terms of geographic and taxonomic coverage, employed by previous analyses. None of the aforementioned studies included more than a third of the South American tree squirrel species and also missed some Central and North American taxa (sensu Thorington et al., 2012; Vivo and Carmignotto, 2015); therefore, the lack of so many taxa and presumed lineages might mislead phylogenetic and, consequently, biogeographic inferences. A recent comprehensive phylogenomic study conducted by Abreu-Jr et al. (2020), that included samples from 40 of the 43 recognized species of Sciurini (sensu Thorington et al., 2012; Vivo and Carmignotto, 2015; Hope et al., 2016), revealed an underestimated diversity on the genus and species group levels within the tribe and clarified several important points with regards to the phylogenetic relationships and systematics of tree squirrels. This study did not find support for the recognition of two species previously considered as valid—*Sciurus richmondi* sensu Thorington et al. (2012) and *Microsciurus venustus* sensu Vivo and Carmignotto (2015)— and found evidence for the recognition of six additional lineages that might represent species to be named or revalidated. This new taxonomic proposal suggested a diversity of 46 species within Sciurini which were tentatively organized in 14 genera coined by previous authors (see Abreu-Jr et al. 2020). The mitogenome dataset produced in that study is highly suitable to investigate the tempo and mode of evolution of tree squirrels, especially across the Neotropical region, where the taxonomic and geographic sampling were strengthened.

In the Neotropics, tree squirrels have been shown to exhibit the highest rate of diversification amongst all arboreal squirrels (a non-phylogenetic ecological category that includes members of four subfamilies; Roth and Mercer, 2008), with strikingly high species accumulation rates in the past three million years, presumably after the final closure of the Panamanian land bridge and the establishment of the Great American Biotic Interchange (GABI). The GABI is one of the most significant events of transcontinental fauna dispersal that occurred during the temporal window from the late Pliocene to the late Pleistocene,

wherein several mammal lineages migrated from North America to South America and the other way around (Simpson, 1983; Stehli and Webb, 1985; Woodburne, 2010). Among rodents, tree squirrels and sigmodontine rodents arguably shared a very successful history of diversification and colonization of South America accompanying the GABI (Patterson and Pascual, 1972; Simpson, 1983).

However, alternative scenarios proposed either a diversification at the generic level in tropical portions of North and Central America, followed by an invasion of South America, after the Panamanian land bridge arose; or an early entrance in South America, sometime in the Miocene, prior to the establishment of the Panamanian land bridge and the GABI (see detailed explanation in D'Elía, 2000). For sigmodontine rodents, current hypotheses suggest an early arrival in South America (Smith and Patton, 1999; Steppan et al., 2004a; Parada et al., 2013), in the Miocene, supporting the scenario proposed by Hershkovitz (1966, 1969) and Reig (1984). For squirrels, current hypotheses sustain the entrance of this group in South America during the GABI, after the Panamanian land bridge arose (Mercer and Roth, 2003), but results so far are still inconclusive, as most of the South American squirrel diversity remained unstudied.

These rodents are an ideal group of mammals that can help to address a wide range of biogeographic questions, such as the timing of the dispersal events across the Americas and colonization within South America, and the potential relevance of specific areas as sources of diversity within the Neotropics. The South American continent has a history of major geological (such as the Andes uplift and the Amazon basin formation) and climatic (such as the Pleistocene climate fluctuations) changes that occurred in the past 10 million years. These events gave rise to a great variety of environments and new habitats that leveraged the rapid diversification of several lineages of the tree of life (Hoorn et al., 2010).

In this study, we test current hypotheses on the tempo and mode of diversification of tree squirrels (Pečnerová and Martínková, 2012; Pečnerová et al., 2015) employing a comprehensive dataset (mitogenome data) and a diverse taxonomic coverage (including 43 Sciurini putative species; Abreu-Jr et al. 2020). We provided a time-scaled phylogeny to infer the timing of the origin and diversification of the main lineages, and performed biogeographic analyses to estimate ancestral ranges and evaluate the drivers of diversification at a global scale. In addition to the broad-scale analysis that includes all Sciurini species—and since the Neotropics is home of the most diverse assemblage of

species—we also performed a finer-scale analysis focused exclusively on the diversification processes of Neotropical taxa, to unravel their colonization history across Central and South Americas. Finally, we tested the hypothesis that tree squirrels represent a rapid-diversified radiation (Roth and Mercer, 2008) by estimating speciation rates and investigating diversification rate fluctuations through time.

3.2 Material and Methods

3.2.1 Dataset taxonomic composition

Our dataset is composed of 14 genera and 43 species of Sciurini delimited in a recent contribution (Abreu-Jr et al. 2020). Taxonomic identifications at the genus and species levels follow Abreu-Jr et al. (2020). At the species level, those authors kept the museum identifications for several specimens, especially those housed at the American Museum of Natural History (AMNH) and at the Smithsonian National Museum of Natural History (USNM), which had been made by some of the main authorities on tree squirrel taxonomy (e.g. R. W. Thorington and M. de Vivo). For material not previously identified, they performed identifications by examining the morphology of the vouchers, consulting original descriptions and other relevant literature. In cases of unviability to examine vouchers, original museum identifications were accepted if i) those identifications correspond to the known geographic distribution of the taxon in question, and ii) phylogenetic analyses of their DNA sequences were consistent with the museum identification. The currently valid species of tree squirrels not included in this study are: *Tamiasciurus fremonti* sensu Hope et al. (2016), *Microsciurus santanderensis* and *M. simonsi* sensu Vivo and Carmignotto (2015). We used as outgroup two representatives of the tribe Pteromyini (the sister-tribe of Sciurini): *Glaucomys volans* (USNM 569823) and *Hylopetes phayrei* (USNM 584420). A complete list of specimens analyzed in this study accompanied by geographic information and GenBank accession numbers for complete mitogenomes is provided as Supplementary Material (Table S1).

3.2.2 DNA matrix, alignment and saturation test

We selected the 13 mitochondrial protein-coding genes (CDS) from one sample of each of the 43 putative species of Sciurini to perform the analyses. The CDS matrix was aligned using MUSCLE (Edgar, 2004), with up to eight interactions. We investigated the presence of

saturation in our dataset by plotting transitions and transversions against nucleotide divergence in DAMBE7 (Xia, 2018).

3.2.3 Divergence times estimate

Divergence times were estimated using a Bayesian framework in BEAST 2.6.1 (Bouckaert et al., 2019). We applied a single nucleotide substitution model (GTR) with a gamma category of 4 to our dataset and implemented a lognormal relaxed clock with a Yule tree prior. The analysis was conducted by running 600 million generations of Markov chain Monte Carlo (MCMC), sampling every 60,000 generations. The results of MCMC runs were visualized in Tracer 1.7 (Rambaut et al., 2018) to confirm a minimum of 200 effective sample size for all parameters. A time-calibrated tree was generated with TreeAnnotator v2.6.0 (Bouckaert et al., 2019), considering 10% of burnin and selecting the maximum clade credibility as the target tree. We do not specify posterior limit for the nodes and we selected mean heights for the common ancestor heights. BEAST analysis was performed in the CIPRES Science Gateway (Miller et al., 2010).

We employed four calibrations points representing three distinct types of priors: (i) fossil record – the first known fossil record of *Tamiasciurus hudsonicus*, dated from the Irvingtonian (from 1.9 to 0.25 Mya; Steele, 1998), was implemented as uniform prior for the root of *Tamiasciurus*; (ii) fossil record – *Sciurus alleni* was first reported from the Late Pleistocene (from 0.126 to 0.0117 Mya; Jakway, 1958) and it was constrained as uniform prior for the split between *S. alleni* and *S. oculatus*; (iii) geological events plus fossil record – the crown of Nearctic *Sciurus* was constrained with uniform prior between 13.6 and 7.4 Mya, considering that *Sciurus* was present in North America since the late Miocene (Clarendonian; *Sciurus olsoni*; Emry et al., 2005) and this radiation supposedly arrived from Eurasia (Pečnerová et al., 2015), thus the colonization occurred before the initial opening of the Bering Strait about 7.4 Mya (Marincovich and Gladenkov, 1999); (iv) secondary estimation – we constrained the root of Sciurini with log-normal prior with offset of 13.4 Mya (M = 0.5 and S = 1.0), according to results of Upham et al. (2019), which is currently the most comprehensive phylogeny of mammals.

3.2.4 Ancestral range and diversification events estimate

We estimated the ancestral range of the phylogenetic nodes and investigated the historical events (e.g. vicariance, dispersal) that might be evoked to explain the diversification pattern of the tribe Sciurini. These analyses were performed employing a maximum likelihood framework using the R package BioGeoBEARS (Matzke, 2013). The time-scaled topology provided by BEAST with only one individual per species was used as the input tree. Six models of evolution—which distinctively incorporate vicariance, dispersal and extinction throughout the cladogenesis—were tested, as follows: (1) Dispersal-Extinction-Cladogenesis (DEC; Ree, 2005), (2) DEC + founder-event speciation ("jump"; DEC + J), (3) Dispersal-Vicariance Analyses (DIVALIKE; Ronquist, 1997), (4) DIVALIKE + J, (5) Bayesian inference of historical biogeography for discrete areas (BAYAREALIKE; Landis et al., 2013), and (6) BAYAREALIKE + J. The best-fitting model was selected based on the Akaike Information Criterion weights (AICw; Wagenmakers and Farrell, 2004).

These biogeographic analyses were performed in two geographic scales: global and Neotropical. The first included all Sciurini species and we investigated their colonization history across large continental areas; species were coded as occupying: (1) Eurasia, (2) Borneo, (3) North America, (4) Central America, and (5) South America. The second focused exclusively on the diversification of Neotropical taxa. For this finer-scale analysis, we followed the Neotropical biogeographical dominions proposed by Morrone (2014) to categorize the current geographical range of species, which were coded as occurring at: (1) the Mexican transition zone, (2) the Mesoamerican dominion, (3) the Pacific dominion, (4) the Boreal Brazilian dominion, (5) the South Brazilian dominion, (6) the South-eastern Amazonian dominion , and (7) the Parana dominion.

3.2.5 Lineage through time and speciation rates

We investigated the diversification pattern of Sciurini by constructing a lineage through time (LTT) plot using the R package Ape (Paradis and Schliep, 2019). Speciation rates and the presence and temporal location of putative diversification rate shifts were explored by employing the Bayesian Analysis of Macroevolutionary Mixtures (BAMM; Rabosky, 2014; Rabosky et al., 2017). We implemented a speciation-extinction model, run for 10 million generations and sampled every 2,000 generations in BAMM v2.5.0. We incorporated a sampling frequency of 0.93, considering that we have sampled about 93% of the extant

species of Sciurini (43 out of the 46 putative species; Abreu-Jr et al. 2020). Results of BAMM analysis were processed and visualized using the R package BAMMtools (Rabosky et al., 2014).

3.3 Results

3.3.1 Phylogeny of the tribe Sciurini

Our final mitogenome alignment was 11,278 bp long and included 45 species (43 ingroup and two outgroup). The mean sequence length per sample was 10,360 bp (ranging from 1,393 to 11,278 bp) and the overall matrix missing data (including gaps and undetermined characters) was 7.4%. We found no evidence for nucleotide substitution saturation in our sequences that could influence the phylogenetic signal (Supplementary Figure 1). The maximum clade credibility tree obtained from our BEAST analysis recovered similar relationships to those previously reported by Abreu-Jr et al. (2020), with all nodes strongly supported (PP ≥ 0.98), excepted by the node of the genus *Echinosciurus*, which did not receive a significant support (PP < 0.95).

3.3.2 Model selection in ancestral range estimates

The analyses performed with BioGeoBEARS at the global scale recovered DEC + J as the most suitable model to explain the diversification of Sciurini (LnL = -47.61, AICwt = 0.82; Table 1), considering distinct evolutionary events, such as dispersal, extinction and vicariance. The ancestral range estimation focused on the Neotropical taxa recovered that the model BAYAREALIKE+J better fit our data (LnL = -68.79, AICwt = 0.85; Table 2), favoring dispersal-extinction scenarios rather than vicariance events.

3.3.3 Spatiotemporal diversification of Sciurini at a global scale

The origin of the tribe Sciurini was estimated at 14.1 Mya (13.4-15.5; Figure 1; see also Supplementary Figure 2) and its ancestral geographic range was inferred with highest probability to North America (58.6%; Figure 1). An area including both North America and Eurasia was estimated with the second highest probability of occurrence for the ancestor of Sciurini (30.6%; Table S2). The first lineage to diverge within the tribe was the genus *Tamiasciurus*. All of the known species of this genus are from North America and its ancestor was also estimated to have occurred in North America. The subsequent diversification event,

leading to the Borneo endemic genus *Rheithrosciurus*, occurred at 10.5 Mya (9.5-11.7) and the ancestral range of this node was also in North America. Likewise, the ancestral ranges of the nodes leading to the Eurasian genus *Sciurus* (at 9.7 Mya), the North American genus *Hesperosciurus* (at 7.6 Mya), and the clade including both North American genera *Neosciurus* and *Parasciurus* (at 6.2 Mya), were most likely to be in North America. In summary, diversification events generating the major lineages from North America and Eurasia took place in the mid-late Miocene, whereas most speciation events involving extant species in those lineages occurred in the Pleistocene (Figure 1).

The large clade including exclusively Neotropical taxa (marked with a black star in Figure 1) originated with highest probability in South America at 5.9 Mya (5.2-6.5). Within this clade, the Central American radiation (including genera *Microsciurus*, *Syntheosciurus* and *Echinosciurus*) originated at 4.6 Mya (3.8-5.3) and the South American lineage (including genera *Leptosciurus*, *Simosciurus*, *Guerlinguetus*, "*Microsciurus*" and *Hadrosciurus*) at 5.0 Mya (4.3-5.6). Therefore, these Central and South American radiations, as well as most Neotropical genera, originated in the early-mid Pliocene and experienced a rapid diversification, with speciation events occurring mostly in the Pliocene and in the Plio-Pleistocene transition, and only a few during the Pleistocene (Figure 1).

The ancestral range estimation of the tribe Sciurini on a global scale (Figure 1) allows the inference of seven founder-event speciation and three anagenetic dispersals, while it does not indicate the occurrence of vicariance events during the evolutionary history of tree squirrels (see Supplementary Figures 3 and 4). The founder events (i.e. a jump of a lineage to a new area of occurrence outside of the ancestral range of the node; Matzke, 2013) were inferred for the ancestors of (1) the genus *Rheithrosciurus* from North America to Borneo, (2) the Eurasian genus *Sciurus* from North America to the Palearctic region, (3) the large Neotropical radiation from North America to South America, (4) the Central American clade from South America to Central America, (5) the clade *Parasciurus alleni* + *P. oculatus* from North America to Central America, (6) *Parasciurus nayaritensis* from North America to Central America and (7) *Microsciurus* "species 1" from Central America to South America. Anagenetic dispersals included (1) the colonization of Central America by *Hesperosciurus aberti*—although this species occurs in Central America mainly in the Sierra Madre Occidental and there is a discussion whether this region is part of the Nearctic region (Holt et al., 2013) or it belongs to the Neotropical region (Morrone, 2014)—, (2) colonization of

South America by *Syntheosciurus granatensis*, and (3) colonization of Central America by the ancestor of *Leptosciurus boquetensis* and *L. isthmius*.

3.3.4 Spatiotemporal diversification of Sciurini in the Neotropics

Our biogeographic analyses focused on Neotropical Sciurini suggested with highest probability (92.2%) that this large clade originated in the northwestern portion of South America in the Pacific dominion (Figure 2; Table S3). The Pacific dominion was also recovered with highest probability as the ancestral range of the Central American (93.6%) and South American (99.0%) major radiations. Within the Central American radiation, two genera originated in this same region: *Microsciurus* at 4.1 Mya (3.3-5) and *Syntheosciurus* at 3 Mya (2.1-3.7). The genus *Echinosciurus* originated at 3.7 Mya (3-4.4) in an area including the Mexican transition zone and the Mesoamerican dominion.

With regards to the South American radiation, the ancestors of the first two genera occupied the northwestern portion of the continent, west of the Andes, in the Pacific dominion: *Leptosciurus* that originated at 4.4 Mya (3.6-5.1) and *Simosciurus* that originated at 4 Mya (3.1-4.7). The following genera had their origins on the east side of the Andean cordillera, along the Amazon basin. The genus *Guerlinguetus* originated at 4.4 Mya (3.9-5.1) in the Boreal Brazilian dominion. The ancestor of the genera "*Microsciurus*" and *Hadrosciurus* occupied a large area including the Boreal Brazilian and the South Brazilian dominions. The origin of "*Microsciurus*" dated to 3.2 Mya (2.5-3.7) and the origin of *Hadrosciurus* was estimated to 3.9 Mya (3.6-4.9). In summary, the majority of the cladogenetic events observed for Neotropical tree squirrels took place in the northern portion of South America and southern Central America during the Pliocene. Two major dispersal events apparently led to the colonization of Mesoamerica by the genus *Echinoscirus* and the colonization of the east of the Andes by the most recent common ancestor of the genera *Guerlinguetus*, "*Microsciurus*" and *Hadrosciurus* (Figure 2).

The scenario recovered by this analysis allows the inference of four founder-event speciation and five anagenetic dispersal events during the diversification of the tribe Sciurini across the Neotropics, and, once again, no vicariance event was inferred (see Supplementary Figures 5 and 6). The founder-event speciation corresponds to the jump of the ancestors of (1) the genus *Echinosciurus* from the Pacific dominion to the Mexican transition zone and Mesoamerican dominion, (2) the genera *Guerlinguetus*, "*Microsciurus*" and *Hadrosciurus*

from the Pacific dominion to the Boreal Brazilian dominion, (3) (*Guerlinguetus aestuans* "b", (*G. aestuans* "c", *G. brasiliensis*) from the Boreal Brazilian dominion to the South Brazilian dominion, and (4) (*Guerlinguetus aestuans* "c", *G. brasiliensis*) from the South Brazilian dominion to Boreal Brazilian, South-eastern Amazon, and Parana dominions. Anagenetic dispersal occurred with the colonization of (1) the Pacific dominion by *Echinosciurus variegatoides*, (2) the South Brazilian dominion by the ancestor of "*Microsciurus*" and *Hadrosciurus*, (3) the Boreal Brazilian and Parana dominions by the ancestor of *Guerlinguetus aestuans* "c" and *G. brasiliensis*, (4) the Pacific dominion by "*Microsciurus*" *sabanillae*, and (5) the Pacific dominion by *Hadrosciurus igniventris*.

3.3.5 Diversification rates of the tribe Sciurini

The LTT plot indicated an initial low and constant net diversification rate from about 14 to 10.5 Mya, followed by a moderate increase from 10.5 to 5 Mya, and an acceleration on the diversification rate from 5 Mya to the present (Figure 3A). This period of rapid inflation on the number of lineages, along the Pliocene and Pleistocene, coincides with the South American invasion and colonization by tree squirrels (Figure 3A; red star). The Bayesian approach, however, showed a pattern of constant speciation rate through time, with no inflection point and with a slight decrease on the magnitude of speciation over time (Figure 3B). The BAMM analysis also recovered the null model "no shifts" on speciation rate as the most probable (0.66) to represent the diversification of Sciurini. The mean speciation rate of Sciurini was 0.29.

3.4 Discussion

The phylogenetic hypothesis presented here is concordant with the inferences depicted in Abreu-Jr et al. (2020), as we employed the same taxa and data source. We included in our analyses 43 of the 46 currently recognized species of tree squirrels (see Abreu-Jr et al. 2020). Previous biogeographic analyses have included up to 28 Sciurini species, and less than a third of South American tree squirrel species (Pečnerová and Martínková, 2012; Pečnerová et al., 2015). Therefore, this is the most comprehensive study investigating the biogeographic history of the tribe Sciurini, both in number of species and geographic coverage.

The models with best fitting scores on each biogeographic analysis are distinct in their criteria, as one detects vicariant events (DEC) and the other (BAYAREALIKE) does not

(Matzke, 2013). Both approaches employed to investigate the range of evolution of tree squirrels (analysis of the whole tribe and analysis focused on Neotropical taxa) did not recover vicariant events during the evolutionary history of Sciurini. These results suggest that vicariance was not a main driver of the diversification processes of tree squirrels at either global or local (Neotropical) scales (but extinctions might mask instances of vicariance; see discussion below). Dispersal events, on the other hand, were commonly recovered and founder-events were estimated to have happened more times than anagenetic dispersals in both approaches.

Our time estimation for the origin of Sciurini was similar to the result of Mercer and Roth (2003) and Upham et al. (2019), supporting a more recent date around 14 Mya compared to the estimation of Zelditch et al. (2015) and Pečnerová et al. (2015) around 18 and 19 Mya, respectively. This is also relatively more recent compared to the origin of the other tribes of the family Sciuridae that have been hypothesized to have happened before 18 Mya (Roth and Mercer, 2008). Our estimates suggested that the most recent common ancestor of Sciurini probably inhabited exclusively North America (near 60% of probability). This is in contrast to previous findings that suggested that the ancestors of this tribe originated in Eurasia or in a large area that included Eurasia and North America (Pečnerová and Martínková, 2012; Pečnerová et al., 2015; Rocha et al., 2016). Despite the higher probability for an exclusive North American origin, our results do not rule out the possibility of an origin that included both the Nearctic and Palearctic regions (supported as the ancestral range of Sciurini with about 30% of probability).

Pečnerová and Martínková (2012) suggested two alternative working hypotheses for the initial diversification and dispersal of tree squirrels. The first one postulates that the ancestor of the genus *Sciurus* dispersed from North America to Eurasia after diversification from *Tamiasciurus*. This ancestor diverged again in Eurasia and returned to North America via the Bering land bridge. The second hypothesis claims that ancestors of both *Tamiasciurus* and *Sciurus* originally occupied Eurasia and colonized the Americas in two distinct occasions, also crossing over Beringia. However, the results obtained by Pečnerová and Martínková (2012) do not allow them to choose conclusively between these two hypotheses and our results indicate a higher likelihood for a third distinct scenario. We found that the ancestors of *Tamiasciurus*, *Rheithrosciurus*, and the ancestor of the clade from which *Sciurus* diverged, occupied the Nearctic region with highest probability. Our results suggest that the ancestors

of *Rheithrosciurus* and *Sciurus* dispersed from North America to Eurasia in two independent events, through subsequent jump dispersal events some thousand years apart (Figure 4). The cladogenetic events leading to *Rheithrosciurus* dated at 10.5 Mya and to the Eurasian *Sciurus* at 9.7 Mya, which supports the colonization of the Palearctic region by these two lineages before the opening of the Bering Strait around 7.4 Mya (Marincovich and Gladenkov, 1999).

Rheithrosciurus macrotis, endemic to Borneo (Thorington et al., 2012), possibly colonized the island overland from southeast Asia and has been isolated there since its arrival. Borneo is connected to mainland areas through the Sunda Shelf, and the distribution of shallow marine carbonates and the depths of sea water on the Sunda Shelf indicate the presence of past dispersal routes from the continent (perhaps via Java) to Borneo (Hall, 2001). Unlike *Rheithrosciurus*, the genus *Sciurus* successfully spread throughout Asia and Europe, and its species are currently widespread in the Palearctic region (Thorington et al., 2012).

Besides *Tamiasciurus*, the remaining Nearctic genera, *Hesperosciurus*, *Parasciurus*, and *Neosciurus*, also had their ancestors occupying North America. The estimated origin of *Hesperosciurus* dated to the late Pliocene (around 4.8 Mya), *Parasciurus* to the late Pleistocene (around 2.3 Mya), and *Neosciurus* diverged in the late Miocene (around 6.2 Mya). Thus, despite exhibiting the same ancestral range, the North American groups originated in distinct periods, during a relatively large temporal window (about 3.9 Mya) compared to the South American genera (see below).

The ancestral range of the large Neotropical clade was in South America, which indicates that the colonization of Central America by Neotropical taxa was via movement north back from South America (Figure 1, Figure 4). Remarkably, our estimates indicate that this clade originated around 6 Mya, an earlier date than ever suggested. Previous studies had estimated that tree squirrels colonized South America around 3 Mya (Mercer and Roth, 2003; Roth and Mercer, 2008; Pečnerová et al., 2015) and they argued that the South American invasion was conditioned by the closure of the Isthmus of Panama.

The date of the Panamanian land bridge formation—allegedly the moment that triggered the GABI—has gained a lot of attention and discussion in the scientific literature in the past decade. Current biological data suggest an early emergence of the Isthmus, around 6-8 Mya, and indicate that the biotic turnover associated with GABI was much longer and

more complex than traditionally recognized, beginning perhaps as early as the Oligocene–Miocene transition (Bacon et al., 2015; Montes et al., 2015; Jaramillo et al., 2017; Molnar, 2017). However, comprehensive geological, paleontological, and also biological reviews pointed to scant evidence and analytical bias in some of the aforementioned studies, and sustained the conventional hypothesis of the formation of the Panamanian Isthmus *sensu stricto* around 3-2.8 Mya (Lessios, 2008, 2015; Leigh et al., 2014; Marko et al., 2015; O'Dea et al., 2016).

In addition to this ongoing discussion, Agnolin et al. (2019) proposed an alternative hypothesis to explain biotic interchanges between North and South America based on Pre-GABI land connections. The hypothesis from Agnolin et al. (2019) relies on paleontological and geological evidences for a large corridor of shallow waters and emergent spans of land connecting North and South America through the Caribbean islands during the Oligocene-Miocene, known as GAARlandia (Iturralde-Vinent and MacPhee, 1999). The presence of GAARlandia could help to explain the occurrence of early migrants (from the Miocene) in North and South America (see Pinto-Sánchez et al., 2012), considering the lack of strong geological evidence to sustain the early closure of the Panamanian Isthmus. Hershkovitz (1972), Savage (1974) and Reig (1984) also defended a much earlier entrance of land mammals in South America, in the early Miocene, prior to the GABI and the emergence of the land bridge, based on the patterns of diversity and endemism of sigmodontine rodents.

Our results for the first entrance of tree squirrels in South America could either fit this alternative route of colonization through GAARlandia or the traditional route recognized via the Panamanian Isthmus. Even if the complete establishment of the Isthmus occurred in the Pliocene, it might have been permeable to biota migration since the Oligocene (Molnar, 2008; Eizirik, 2012; Carrillo et al., 2014; Bloch et al., 2016). Marshall (1979) already defended that events of marine regression on Upper Miocene would have favored dispersion between South and Central/North America, via the Isthmus, suggesting its permeability. The timeframe of this invasion is also similar to what has been shown for sigmodontine rodents (Smith and Patton, 1999; Steppan et al., 2004a; Parada et al., 2013). To the best of our knowledge, no previous hypothesis ever suggested that these two distinct lineages of rodents may have shared a parallel and similar history of invasion and dispersal throughout South America and re-invasion of Central America. During a similar lapse of time, sigmodontine rodents appear to have been more successful in adapting to a greater variety

of ecosystems (including grasslands, shrublands, wetlands) and spreading as far as the temperate zone at the southern border of the continent, reaching a diversity of hundreds of species (Burgin et al., 2018). Squirrels, on the other hand, remained restricted to the mid-northern portion of the continent, inhabiting only dense and continuous forests (Vivo and Carmignotto, 2015), exhibiting much less species diversity.

The predominantly Central American radiation (including the genera *Microsciurus*, *Syntheosciurus* and *Echinosciurus*) is more likely to have originated in Central America around 4.6 Mya (probably southern Central America as suggested by the analyses focused on Neotropical taxa, see below), from a jump dispersal from South America. Also, most ancestors of the extant species within this radiation occurred in Central America. Two species, one from the genus *Microsciurus* (*M.* "species 1") and one from the genus *Syntheosciurus* (*S. granatensis*), colonized back to South America (Figure 1).

The major South American radiation is slightly older than the Central American one, with its date of origin estimated to around 5 Mya. Within this lineage all the ancestors occupied South America, except the most recent common ancestor of *Leptosciurus boquetensis* and *L. isthmius*, estimated to have inhabited Central America (Figure 1). Previous studies (Pečnerová and Martínková, 2012; Pečnerová et al., 2015) suggested a directional diversification process following the latitudinal gradient from North America to South America giving rise to all Neotropical taxa. Our results, supported by ancestral range estimation of Sciurini in both global and Neotropical scales, do not corroborate this hypothesis, and indicate a much more complex pattern, with distinct colonization events from South to Central America and vice-versa and perhaps through distinct dispersal routes.

Several Central American vertebrate lineages that originated along the Pliocene-Pleistocene boundary have been shown to exhibit their ancestral ranges in northern South America (Barrantes, 2009; Saldarriaga-Córdoba et al., 2017; Maestri et al., 2019). Savage (2002) proposed a two-step dispersal pulse hypothesis to explain amphibians and reptiles of South America inhabiting Middle America. The first colonization episode, according to him, occurred about 3.5 Mya when the sea level lowered and many taxa were able to invade Mesoamerica as far as Mexico. The second episode took place in the Pleistocene and the majority of taxa could not disperse further than southern Nicaragua or Costa Rica. The first episode described by Savage (2002) coincides with the diversification and dispersion of the genus *Echinosciurus* to the Mesoamerican dominion and Mexican transition zone (see Figure

2), and the second episode coincides with the dispersal into Central America by *Leptosciurus boquetensis* and *L. isthmius* (see Figure 1). Both episodes of Central American colonization can be plausibly explained via the traditional dispersal route across the Isthmus of Panama.

The biogeographic analysis focused on Neotropical taxa recovered the origin of this clade in the Pacific dominion, which ranges from southern Central America to northwestern South America. This was estimated to have occurred in the late Miocene (around 6 Mya), when the Panamanian land bridge was still emerging and during the final uplift stages of the Andean Cordillera (Hoorn et al., 2010; Stange et al., 2018). This region was also the ancestral range of four genera: *Microsciurus* and *Syntheosciurus* (composed of Central and South American species), *Leptosciurus* (composed of species occupying mainly the northwestern portion of South America), and *Simosciurus* (composed of two species occurring in the coast of Ecuador and Peru). Within the mainly Central American radiation, a founder-event speciation is inferred, leading to the colonization of the Mexican transition zone and Mesoamerican dominion by the ancestor of the genus *Echinosciurus* around 3.7 Mya. Moreover, among these species, *E. variegatoides* colonized back to the Pacific dominion through anagenetic dispersal. Some groups of sigmodontine rodents also had their origin estimated to be in the late Miocene or Pliocene in northern South America, suggesting that this region has played an important role on the origin and diversification of forest rodents (Maestri et al., 2019). In the mid-Pliocene, when several of the cladogenetic events of the Central American radiation have been estimated, important geological events such as the development of the Sierra Madre oriental and occidental and the volcanism of the Trans-Mexican Volcanic Belt were taking place (Sedlock et al., 1993). This intense geologic activity culminating with the formation and expansion of highland forests might have played a crucial role on the diversification of several biotic components (see Morrone, 2010), perhaps including the tree squirrels.

The ancestors of the remaining South American genera (*Guerlinguetus*, *"Microsciurus"*, *Hadrosciurus*) occupied areas east of the Andes with highest probability. This was the point (approximately 4.8 Mya) in the evolutionary history of Sciurini that tree squirrels crossed over the Andes and through a founder-event colonized large forested areas in the eastern portion of South America. *Guerlinguetus* showed the highest probability of ancestral occupancy in the Boreal Brazilian dominion, which includes the northern bank of the Amazon basin and the Guiana Shield. The ancestors of the genera "*Microsciurus*" and

Hadrosciurus occupied an area including the Boreal Brazilian and South Brazilian dominions. Within *Guerlinguetus*, the ancestor of three species (*G. aestunas* "b", *G. aestunas* "c" and *G. brasiliensis*) colonized the South-eastern Amazonian dominion through a founder-event speciation. Through another founder event speciation, the ancestor of *G. aestuans* "c" + *G. brasiliensis* colonized the South-eastern Amazonian dominion. Subsequently, it expanded its range, reaching the Boreal Brazilian dominion and, for the first and only time in the tree squirrel diversification, the Atlantic Forest on the coast of Brazil (see Supplementary Figure 6 for details). South-eastern Amazon and Atlantic Forest were the last two areas occupied by tree squirrels in South America.

The timing of Atlantic Forest colonization by tree squirrels (around 1 Mya) is comparable to the timing of the entrance of some species of sigmodontine rodents (such as species in the genera *Euryoryzomys*, *Hylaeamys* and *Oligoryzomys*; Steppan and Schenk, 2017; Maestri et al., 2019). At least two main distinct spatiotemporal routes have been proposed between the Amazon and the Atlantic Forest (Costa, 2003; Batalha-Filho et al., 2013; Ledo and Colli, 2017): older (middle to late Miocene) connections between southeastern Atlantic Forest and western Amazonia (the SE-NW bridge), and younger (Pliocene and Pleistocene) connections across northeastern Brazil (the NE bridge). As our results indicate a young invasion, squirrels might have used NE routes, possibly favored by interglacial periods of the Pleistocene, as the NE bridge hypothesis assumes.

Within the genera "*Microsciurus*" and *Hadrosciurus*, most species remained restricted to the Amazon basin and adjacent areas (along Boreal Brazilian and South Brazilian dominions). Only one species in each genus ("*Microsciurus*" *sabanillae* and *Hadrosciurus igniventris*) dispersed to the eastern slopes of the Andean cordillera and invaded the border of the Pacific dominion (Figure 2).

All South American genera originated from 4.4 to 3.2 Mya. This very short temporal window (about 1.2 My) overlaps with the South American Land Mammal Ages (SALMAs) of the Montehermosan (from about 6.8 to 4 Mya) and the Chapadmalalan (from about 4 to 3.4 Mya) (Flynn and Swisher III, 1995). At this time, the Andean uplift was complete and the Amazon basin was in its final stages of formation with a large supply of Andean-derived sediments (Hoorn et al., 2010). These massive geological changes produced a great variety of new ecosystems in South America that favored the diversification of many groups of organisms.

South American genera usually do not overlap their distribution, except in a small area on the northern bank of the Amazon river, where one species of *Guerlinguetus* (*G. aestuans* "c") is sympatric with two species of *Hadrosciurus* (*H. igniventris* and *H. spadiceus*), and in the western Amazon basin, where species from "*Microsciurus*" and *Hadrosciurus* are mostly sympatric. It is important to mention that these last two genera are the most morphologically dissimilar among all South American forms. Species of "*Microsciurus*" are small squirrels with head and body length ranging from 125 to 180 mm, and species from *Hadrosciurus* are large, with head and body length ranging from 220 to 325 mm (Vivo and Carmignotto, 2015). It is also noteworthy that along lowland areas in the western Amazon basin, species of "*Microsciurus*" are more commonly found in non-flooded upland forests, while species of *Hadrosciurus* inhabit mainly seasonally inundated floodplain forests (Emmons and Feer, 1997). Besides tree squirrels, South America is the home of another remarkable radiation of Sciuridae, the Neotropical pygmy squirrel, *Sciurillus pusillus*. This singular radiation diverged early in the history of the family (about 35 Mya) and is included in its own subfamily, Sciurillinae (Mercer and Roth, 2003; Steppan et al., 2004b; Roth and Mercer, 2008). The diversification and colonization of South America by the Neotropical pygmy squirrels are enigmatic pieces on the evolutionary history of this group.

Regarding speciation events generating the 43 extant species of Sciurini analyzed here, three species originated in the Miocene, 17 in the Pliocene, and 23 in the Pleistocene. The events in the Miocene were: the diversification of the Borneo endemic species *Rheithrosciurus macrotis*, the Eurasian *Sciurus anomalus*, and the North American *Neosciurus carolinensis*. During the Pliocene, the great majority of speciation events took place in South America, where ten species (from all genera) originated; in Central America five species originated (also from all genera) and in North America only one speciation event occurred (within the genus *Hesperosciurus*). In the Pleistocene, eight speciation events occurred in South America (within the genera *Leptosciurus*, *Guerlinguetus*, "*Microsciurus*", and *Hadrosciurus*), six in North America (within the genera *Tamiasciurus* and *Parasciurus*), and six in Central America (within the genera *Echinosciurus* and *Lepstosciurus*). In summary, most speciation events in North America occurred late in the Pleistocene, and in South America the majority occurred during the Pliocene or in the Pliocene-Pleistocene transition. Moreover, most cladogenetic events in South and Central America took place in the

Pliocene. In North America and Eurasia, cladogenetic events occurred mainly early in the Miocene or late in the Pleistocene (Figure 1).

Pleistocene climatic fluctuations were probably a determining factor in the diversification process within North American lineages, for instance across the genera *Tamiasciurus* (see Arbogast et al., 2001; Chavez et al., 2011) and *Parasciurus*. In the Neotropics, Pleistocene fluctuations apparently played a secondary role, as dozens of speciation events occurred before the intensification of climatic cycles. It seems that the Central and South American lineages rapidly adapted to the new (and perhaps vacant) environments and experienced explosive diversification. This is quite different from the history of other rodent radiations in South America, rats and spiny rats, where most of the speciation events took place in the Pleistocene (Fabre et al., 2012; Steppan and Schenk, 2017).

Our analyses revealed a mostly constant (slightly decreasing) speciation rate through time for the tribe Sciurini. However, we observed a peak of lineage accumulation since the beginning of the Pliocene, which coincides with the Neotropical invasion, followed by a rapid diversification in this area. The constancy of the speciation rate (with no shifts during the Pliocene) might indicate that several extinctions may have occurred during the evolutionary history of tree squirrels. In this way, the number of species that went extinct possibly equalized the evident increase in species accumulation since the Pliocene and, therefore, the number of species maintained over time is similar to the number of extinct species. This could also strengthen the prevalence of founder-event speciation and overlook the presence of vicariant events in the ancestral range estimates of Sciurini, as extinctions might be expected to eliminate instances of vicariance and inflate instances of “jump” dispersal. Moreover, our analysis yielded a mean diversification rate around 0.29 for tree squirrels, which is slightly higher than the estimate of Zelditch et al. (2015) and similar to the diversification rate of muroid rodents in the same time frame (Steppan and Schenk, 2017). Finally, we conclude that the South American invasion was not as recent as previously inferred, but the speciation was indeed very fast, mostly restricted to a short period of time in the Pliocene, much earlier than for other groups of rodents.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

EFA and SEP conceived the project; EFA analyzed the data; EFA, SEP, ARP, and JEM interpreted and discussed the results; EFA wrote the paper with contributions from SEP, MTNT, DEW, ARP, JEM. SEP, ARP, DEW, JEM secured funding for this project.

Funding

This work was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico through a doctoral fellowships to EFA (147145/2016-3, 203692/2017-9, and 165553/2017-0) and through a productivity scholarship to ARP (304156/2019-1); by the American Society of Mammalogists through the Latin American Student Field Research Award to EFA, providing support for a field expedition to the Rio Purus, BR; by the American Museum of Natural History through a Collection Study Grant, allowing EFA to collect morphological data from tree squirrels housed at the AMNH; by the Smithsonian Institution through postdoctoral fellowships to SEP and MTNT (Smithsonian Women's Committee), and through research funds provided to SEP, MTNT, DEW and JEM, covering costs associated to laboratory procedures and DNA sequencing; by the National Geographic Society through an Explorers Grant to SEP (NGS-381R-18), providing support to visit and collect samples housed at the MUSM and for a field expedition to Parque Nacional del Rio Abiseo, PE, and through Support for Women + Dependent Care to SEP (NGS 3758), allowing her to visit and collect additional data in North American museums; and by the Fundação de Amparo à Pesquisa do Estado de São Paulo to ARP (09/16009-1), providing support for several field expeditions to the Amazon, BR.

Acknowledgments

We are thankful to the following curators and collection support staff, who loaned us preserved or destructive historical tissue samples of valuable specimens included in this study: R.S. Voss, E. Hoeger, and N. Duncan (AMNH); B.D. Patterson and A.W. Ferguson (FMNH); R.M. Timm and M. Eifler (KU); J.A. Oliveira and M. Weksler (MN); J. Silva Junior

(MPEG); M. Campbell, J. Cook and J. Dunnum (MSB); V. Pacheco (MUSM); C. Conroy, J.L. Patton, and F.J.M. Pascal (MVZ); D.P. Lunde, and N.R. Edmison (USNM). We also thank L. Parker, M. Venkatraman, S. Castañeda and N. McInerney, who provided crucial help during lab work.

Contribution to the Field Statement

Tree squirrels in the tribe Sciurini show the highest diversification rates amongst all arboreal squirrels. This is the first mitogenomic study to include a dense taxonomic sampling (93% of valid species) in order to unveil the tempo and mode of this dramatic diversification. Our biogeographic analyses point to North America as the ancestral range of the tribe Sciurini, and not Eurasia as previously suggested. The origin of the Neotropical radiation was estimated to have occurred in South America around 6 Mya. Remarkably, our results indicate that tree squirrels entered South America earlier than previously estimated. Our results also do not support the findings from previous studies with sparser taxonomic sampling, which proposed that a directional diversification process gave rise to all Neotropical taxa following a latitudinal gradient from North to South America. In contrast, they reveal a more complex pattern, with distinct colonization events from South to Central America and vice-versa. Our estimates indicate a fairly constant speciation rate for Sciurini, which contrasts with the peak of lineage accumulation observed in the Pliocene. Finally, we conclude that the South American invasion was not as recent as previously inferred, but the diversification there was indeed very fast.

Data Availability Statement

The mitogenomic dataset analyzed in this study is available in the Figshare Repository (https://doi.org/10.6084/m9.figshare.12288185.v1). GenBank accession numbers for all complete mitogenomes are provided in the Supplementary Table S1.

References

Abreu-Jr, E. F., Pavan, S. E., Tsuchiya, M. T. N., Wilson, D. E., Percequillo, A. R., and Maldonado, J. E. (2020). Museomics of tree squirrels: a dense taxon sampling of mitogenomes reveals hidden diversity, phenotypic convergence, and the need of a taxonomic overhaul. *BMC Evol. Biol.* 20, 1–25. doi:10.1186/s12862-020-01639-y.

Agnolin, F. L., Chimento, N. R., and Lucero, S. O. (2019). Pre-gabi biotic connections between the Americas: an alternative model to explain the "less-splendid isolation" of South America. *Rev. Geológica América Cent.*, 91–106. doi:10.15517/RGAC.V61I0.40089.

Arbogast, B. S., Browne, R. A., and Weigl, P. D. (2001). Evolutionary Genetics and Pleistocene Biogeography of North American Tree Squirrels (*Tamiasciurus*). *J. Mammal.* 82, 302–319.

Bacon, C. D., Silvestro, D., Jaramillo, C., Smith, B. T., Chakrabarty, P., and Antonelli, A. (2015). Biological evidence supports an early and complex emergence of the Isthmus of Panama. *Proc. Natl. Acad. Sci. U. S. A.* 112, 6110–6115. doi:10.1073/pnas.1423853112.

Barrantes, G. (2009). The role of historical and local factors in determining species composition of the highland avifauna of Costa Rica and Western Panamá. *Rev. Biol. Trop.* 57, 333–349. doi:10.15517/rbt.v57i0.21360.

Batalha-Filho, H., Fjeldså, J., Fabre, P-H., and Miyaki, C. Y. (2013). Connections between the Atlantic and the Amazonian forest avifaunas represent distinct historical events. *J. Ornithol.* 154, 41–50. doi:10.1007/s10336-012-0866-7.

Bloch, J. I., Woodruff, E. D., Wood, A. R., Rincon, A. F., Harrington, A. R., Morgan, G. S., et al. (2016). First North American fossil monkey and early Miocene tropical biotic interchange. *Nature* 533, 243–246. doi:10.1038/nature17415.

Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., et al. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLOS Comput. Biol.* 15, e1006650. doi:10.1371/journal.pcbi.1006650.

Burgin, C. J., Colella, J. P., Kahn, P. L., and Upham, N. S. (2018). How many species of mammals are there? *J. Mammal.* 99, 1–14. doi:10.1093/jmammal/gyx147.

Carrillo, J. D., Forasiepi, A., Jaramillo, C., and Sánchez-Villagra, M. R. (2014). Neotropical mammal diversity and the great American biotic interchange: Spatial and temporal variation in South America's fossil record. *Front. Genet.* 5, 1–11. doi:10.3389/fgene.2014.00451.

Chavez, A. S., Saltzberg, C. J., and Kenagy, G. J. (2011). Genetic and phenotypic variation across a hybrid zone between ecologically divergent tree squirrels (*Tamiasciurus*). *Mol. Ecol.* 20, 3350–3366. doi:10.1111/j.1365-294X.2011.05184.x.

Costa, L. P. (2003). The historical bridge between the Amazon and the forest of brazil a study of molecular phylogeography with small mammals. *J. Biogeogr.* 30, 71–86. doi:10.1046/j.1365-2699.2003.00792.x.

D'Elía, G. (2000). Comments on recent advances in understanding sigmodontine phylogeny and evolution. *Mastozool. Neotrop.* 7, 47–54.

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi:10.1093/nar/gkh340.

Eizirik, E. (2012). "A molecular view on the evolutionary history and biogeography of Neotropical carnivores (Mammalia, Carnivora)," in *Bones, Clones, and Biomes: The History and Geography of Recent Neotropical Mammals*, eds. B. D. Patterson, and L. P. Costa (University of Chicago Press), 123–142.

Emmons, L. H., and Feer, F. (1997). *Neotropical Rainforest Mammals: a field guide*. Second. Chicago and London: The University of Chicago Press.

Emry, R. J., Korth, W. W., and Bell, M. A. (2005). A Tree Squirrel (Rodentia, Sciuridae, Sciurini) from the Late Miocene (Clarendonian) of Nevada. *J. Vertebr. Paleontol.* 25, 228–235.

Fabre, P.-H., Hautier, L., Dimitrov, D., and Douzery, E. J. (2012). A glimpse on the pattern of rodent diversication: a phylogenetic approach. *BMC Evol. Biol.* 12, 88. doi:10.1186/1471-2148-12-88.

Flynn, J. J., and Swisher III, C. C. (1995). "Cenozoic South American land mammal ages: correlation to global geochronologies," in *Geochronology, time scales and global stratigraphic correlation*, eds. W. Berggren, D. V. Kent, M.-P. Aubry, and J. Hardenbol (Society of Economic Paleontologists and Mineralogists), 317–334. doi:10.2110/pec.95.04.0317.

Hall, R. (2001). "Cenozoic reconstructions of SE Asia and the SW Pacific: Changing patterns of land and sea," in *Faunal and Floral Migrations and Evolution in SE Asia-Australasia*, eds. I. Metcalfe, J. M. B. Smith, M. Morwood, and I. D. Davidson (Lisse: A.A. Balkema (Swets & Zeitlinger Publishers)), 35–56.

Hershkovitz, P. (1966). "Mice, land bridges and Latin American faunal interchange," in *Ectoparasites of Panama*, eds. R. L. Wenzel, and V. J Tipton (Field Museum of Natural History, Chicago), 725–751.

Hershkovitz, P. (1969). The Recent Mammals of the Neotropical Region: A Zoogeographic and Ecological Review. *Quaterly Rev. Biol.* 44, 1–70.

Hershkovitz, P. (1972). "The recent mammals of the Neotropical region: a zoogeographic and ecological review," in *Evolution, mammals and Southern continents*, eds. A. Keast, F. C. Erk, and B. P. Glass (State University of New York Press), 311–431.

Holt, B. G., Lessard, J.-P., Borregaard, M. K., Fritz, S. A., Araújo, M. B., Dimitrov, D., et al. (2013). An Update of Wallace's Zoogeographic Regions of the World. *Science.* 339, 74–78. doi:10.1126/science.1228282.

Hoorn, C., Wesselingh, F. P., ter Steege, H., Bermudez, M. A., Mora, A., Sevink, J., et al. (2010). Amazonia Through Time: Andean Uplift, Climate Change, Landscape Evolution, and Biodiversity. *Science* 330, 927–931. doi:10.1126/science.1194585.

Hope, A. G., Malaney, J. L., Bell, K. C., Salazar-Miralles, F., Chavez, A. S., Barber, B. R., et al. (2016). Revision of widespread red squirrels (genus: *Tamiasciurus*) highlights the complexity of speciation within North American forests. *Mol. Phylogenet. Evol.* 100, 170–182. doi:10.1016/j.ympev.2016.04.014.

Iturralde-Vinent, M. A., and MacPhee, R. D. E. (1999). Paleogeography of the Caribbean region: Implications for cenozoic biogeography. *Bull. Am. Museum Nat. Hist.* 238, 1–95.

Jakway, G. E. (1958). Pleistocene Lagomorpha and Rodentia from the San Josecito Cave, Nuevo Leon, Mexico. *Trans. Kansas Acad. Sci.* 61, 313–327.

Jaramillo, C., Montes, C., Cardona, A., Silvestro, D., Antonelli, A., and Bacon, C. D. (2017). Comment (1) on "Formation of the Isthmus of Panama" by O'Dea et al . *Sci. Adv.* 3, e1602321. doi:10.1126/sciadv.1602321.

Landis, M. J., Matzke, N. J., Moore, B. R., and Huelsenbeck, J. P. (2013). Bayesian analysis of biogeography when the number of areas is large. *Syst. Biol.* 62, 789–804. doi:10.1093/sysbio/syt040.

Ledo, R. M. D., and Colli, G. R. (2017). The historical connections between the Amazon and the Atlantic Forest revisited. *J. Biogeogr.* 44, 2551–2563. doi:10.1111/jbi.13049.

Leigh, E. G., O'Dea, A., and Vermeij, G. J. (2014). Historical biogeography of the isthmus of Panama. *Biol. Rev.* 89, 148–172. doi:10.1111/brv.12048.

Lessios, H. A. (2008). The Great American Schism: Divergence of Marine Organisms After the Rise of the Central American Isthmus. *Annu. Rev. Ecol. Evol. Syst.* 39, 63–91. doi:10.1146/annurev.ecolsys.38.091206.095815.

Lessios, H. A. (2015). Appearance of an early closure of the Isthmus of Panama is the product of biased inclusion of data in the metaanalysis. *Proc. Natl. Acad. Sci. U. S. A.* 112, E5765. doi:10.1073/pnas.1514719112.

Maestri, R., Upham, N. S., and Patterson, B. D. (2019). Tracing the diversification history of a Neogene rodent invasion into South America. *Ecography* 42, 683–695. doi:10.1111/ecog.04102.

Marshall, L. G. (1979). A model for paleobiogeography of South American Cricetine rodents. *Paleobiology* 5, 126–132.

Marincovich, L., and Gladenkov, A. Y. (1999). Evidence for an early opening of the Bering Strait. *Nature* 397, 149–151. doi:10.1038/16446.

Marko, P. B., Eytan, R. I., and Knowlton, N. (2015). Do large molecular sequence divergences imply an early closure of the Isthmus of Panama? *Proc. Natl. Acad. Sci. U. S. A.* 112, E5766. doi:10.1073/pnas.1515048112.

Matzke, N. J. (2013). Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Front. Biogeogr.* 5, 242–248. doi:10.21425/f5fbg19694.

Mercer, J. M., and Roth, V. L. (2003). The Effects of Cenozoic Global Change on Squirrel Phylogeny. *Science* 299, 1568–1572. doi:10.1126/science.1079705.

Miller, M. A., Pfeiffer, W., and Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *Proceedings of the Gateway Computing Environments Workshop*, 14 Nov. 2010, 1–8.

Molnar, P. (2008). Closing of the Central American Seaway and the ice age: A critical review. *Paleoceanography* 23, 1–15. doi:10.1029/2007PA001574.

Molnar, P. (2017). Comment (2) on "Formation of the Isthmus of Panama" by O'Dea et al . *Sci. Adv.* 3, e1602320. doi:10.1126/sciadv.1602320.

Montes, C., Cardona, A., Jaramillo, C., Pardo, A., Silva, J. C., Valencia, V., et al. (2015). Middle Miocene closure of the Central American Seaway. *Science* 348, 226–229. doi:10.1126/science.aaa2815.

Morrone, J. J. (2010). Fundamental biogeographic patterns across the Mexican transition zone: An evolutionary approach. *Ecography (Cop.).* 33, 355–361. doi:10.1111/j.1600-0587.2010.06266.x.

Morrone, J. J. (2014). *Biogeographical regionalisation of the neotropical region*. *Zootaxa* 3782, 1–110. doi:10.11646/zootaxa.3782.1.1.

O'Dea, A., Lessios, H. A., Coates, A. G., Eytan, R. I., Restrepo-Moreno, S. A., Cione, A. L., et al. (2016). Formation of the Isthmus of Panama. *Sci. Adv.* 2, e1600883. doi:10.1126/sciadv.1600883.

Parada, A., Pardiñas, U. F. J., Salazar-Bravo, J., D'Elía, G., and Palma, R. E. (2013). Dating an impressive Neotropical radiation: Molecular time estimates for the Sigmodontinae (Rodentia) provide insights into its historical biogeography. *Mol. Phylogenet. Evol.* 66, 960–968. doi:10.1016/j.ympev.2012.12.001.

Paradis, E., and Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35, 526–528. doi:https://doi.org/10.1093/bioinformatics/bty633.

Patterson, B., and Pascual R. (1972). "The fossil mammal fauna of South America" in *Evolution, mammals and southern continents*, eds. A. Keast, F. C. Erk, and B. P. Glass (State University of New York Press), 247–309.

Pečnerová, P., and Martínková, N. (2012). Evolutionary history of tree squirrels (Rodentia, Sciurini) based on multilocus phylogeny reconstruction. *Zool. Scr.* 41, 211–219. doi:10.1111/j.1463-6409.2011.00528.x.

Pečnerová, P., Moravec, J. C., and Martínková, N. (2015). A skull might lie: Modeling ancestral ranges and diet from genes and shape of tree squirrels. *Syst. Biol.* 64, 1074–1088. doi:10.1093/sysbio/syv054.

Pinto-Sánchez, N. R., Ibáñez, R., Madriñán, S., Sanjur, O. I., Bermingham, E., and Crawford, A. J. (2012). The Great American Biotic Interchange in frogs: Multiple and early colonization of Central America by the South American genus Pristimantis (Anura: Craugastoridae). *Mol. Phylogenet. Evol*. 62, 954–972. doi:10.1016/j.ympev.2011.11.022.

Rabosky, D. L. (2014). Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS One* 9. doi:10.1371/journal.pone.0089543.

Rabosky, D. L., Grundler, M., Anderson, C., Title, P., Shi, J. J., Brown, J. W., et al. (2014). BAMMtools: An R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods Ecol. Evol.* 5, 701–707. doi:10.1111/2041-210X.12199.

Rabosky, D. L., Mitchell, J. S., and Chang, J. (2017). Is BAMM Flawed? Theoretical and Practical Concerns in the Analysis of Multi-Rate Diversification Models. *Syst. Biol.* 66, 477–498. doi:10.1093/sysbio/syx037.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., and Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* 67, 901–904. doi:10.1093/sysbio/syy032.

Ree, R. H. (2005). Detecting the historical signature of key innovations using stochastic models of character evolution and cladogenesis. *Evolution (N. Y).* 59, 257–265. doi:10.1111/j.0014-3820.2005.tb00986.x.

Reig, O. A. (1984). Distribuição geográfica e história evolutiva dos roedores muroideos sulamericanos (Cricetidae: Sigmodontinae). *Rev. Bras. Genética* II, 336–335.

Rocha, R. G., Leite, Y. L. R., Costa, L. P., and Rojas, D. (2016). Independent reversals to terrestriality in squirrels (Rodentia: Sciuridae) support ecologically mediated modes of adaptation. *J. Evol. Biol.* 29, 2471–2479. doi:10.1111/jeb.12975.

Ronquist, F. (1997). Dispersal-vicariance analysis: A new approach to the quantification of historical biogeography. *Syst. Biol.* 46, 195–203. doi:10.1093/sysbio/46.1.195.

Roth, V. L., and Mercer, J. M. (2008). Differing rates of macroevolutionary diversification in arboreal squirrels. *Curr. Sci.* 95, 857–861.

Saldarriaga-Córdoba, M., Parkinson, C. L., Daza, J. M., Wüster, W., and Sasa, M. (2017). Phylogeography of the Central American lancehead *Bothrops asper* (Serpentes: Viperidae). *PLoS One* 12, 1–20. doi:10.1371/journal.pone.0187969.

Savage, J. M. (1974). The Isthmian Link and the evolution of neotropical mammals. *Los Angeles County Mus. Contr. Sci.* 260, 1–51.

Savage, J. M. (2002). *The Amphibians and Reptiles of Costa Rica: A Herpetofauna between Two Continents, between Two Seas*. Chicago and London: University of Chicago Press.

Sedlock, R. L., Ortega-Gutierrez, F., Speed, R. C., and Ortega-gutierrez, F. (1993). *Tectonostratigraphic terranes and tectonic evolution of Mexico*. Boulder: The Geological Society of America doi:10.1130/SPE278-p1.

Simpson, G. G. (1983). *Splendid Isolation: The Curious History of South American Mammals*. Yale University Press.

Smith, M. F., and Patton, J. L. (1999). Phylogenetic Relationships and the Radiation of Sigmodontine Rodents in South America: Evidence from Cytochrome b. 6, 89–128.

Stange, M., Sánchez-Villagra, M. R., Salzburger, W., and Matschiner, M. (2018). Bayesian divergence-time estimation with genome-wide single-nucleotide polymorphism data of sea catfishes (Ariidae) supports miocene closure of the Panamanian Isthmus. *Syst. Biol.* 67, 681–699. doi:10.1093/sysbio/syy006.

Steele, M. A. (1998). Tamiasciurus hudsonicus. *Mamm. Species* 586, 1–9. doi:10.1007/978-1-4020-6754-9_9813.

Stehli, F. G., and Webb, S. D. eds. (1985). *The Great American Biotic Interchange*. Boston, MA: Springer US doi:10.1007/978-1-4684-9181-4.

Steppan, S. J., Adkins, R. M., and Anderson, J. (2004). Phylogeny and divergence-date estimates of rapid radiations in muroid rodents based on multiple nuclear genes. *Syst. Biol.* 53, 533–553. doi:10.1080/10635150490468701.

Steppan, S. J., and Schenk, J. J. (2017). *Muroid rodent phylogenetics: 900-Species tree reveals increasing diversification rates*. doi:10.1371/journal.pone.0183070.

Thorington, R. W., Koprowski, J. L., Steele, M. A., and Whatton, J. (2012). *Squirrels of the World*. Baltimore: The Johns Hopkins University Press.

Upham, N. S., Esselstyn, J. A., and Jetz, W. (2019). Inferring the mammal tree: Species-level sets of phylogenies for questions in ecology, evolution, and conservation. *PLOS Biol.* 17, e3000494. doi:10.1371/journal.pbio.3000494.

Vivo, M., and Carmignotto, A. P. (2015). “Family Sciuridae G. Fischer, 1817,” in *Mammals of South America, Volume 2, Rodents*, eds. J. L. Patton, U. F. J. Pardiñas, and G. D’Elía (Chicago and London: The University of Chicago Press), 1–48.

Wagenmakers, E. J., and Farrell, S. (2004). AIC model selection using Akaike weights. *Psychon. Bull. Rev.* 11, 192–196. doi:10.3758/BF03206482.

Woodburne, M. O. (2010). The Great American Biotic Interchange: Dispersals, Tectonics, Climate, Sea Level and Holding Pens. *J. Mamm. Evol.* 17, 245–264. doi:10.1007/s10914-010-9144-8.

Xia, X. (2018). DAMBE7: New and improved tools for data analysis in molecular biology and evolution. *Mol. Biol. Evol.* 35, 1550–1552. doi:10.1093/molbev/msy073.

Zelditch, M. L., Li, J., Tran, L. A. P., and Swiderski, D. L. (2015). Relationships of diversity, disparity, and their evolutionary rates in squirrels (Sciuridae). *Evolution (N. Y).* 69, 1284–1300. doi:10.1111/evo.12642.

Tables

Table 1. Results of the six biogeographic models implemented by BioGeoBEARS on the ancestral range estimation of Sciurini in a global scale. Fits of alternative models were compared using AIC values and AIC weights (AICwt). The model that provided the best fit to our data is in bold.

Model	LnL	Parameters	d	e	j	AIC	AICwt
DEC	-52.54	2	0.018	0.007	0	109.1	0.016
DEC+J	**-47.61**	**3**	**0.007**	**1.00E-12**	**0.023**	**101.2**	**0.820**
DIVALIKE	-54.74	2	0.022	1.00E-12	0	113.5	0.002
DIVALIKE+J	-49.39	3	0.009	1.00E-12	0.025	104.8	0.140
BAYAREALIKE	-73.16	2	0.020	0.078	0	150.3	1.80E-11
BAYAREALIKE+J	-51.33	3	0.006	1.00E-07	0.031	108.7	0.020

Table 2. Results of the six biogeographic models implemented by BioGeoBEARS on the ancestral range estimation of Sciurini in a Neotropical scale. Fits of alternative models were compared using AIC values and AIC weights (AICwt). The model that provided the best fit to our data is in bold.

Model	LnL	Parameters	d	e	j	AIC	AICwt
DEC	-73.30	2	0.025	2.00E-08	0	150.6	0.025
DEC+J	-72.82	3	0.023	1.00E-12	0.008	151.6	0.015
DIVALIKE	-79.16	2	0.031	4.10E-09	0	162.3	7.20E-05
DIVALIKE+J	-77.19	3	0.026	1.00E-12	0.012	160.4	0.0002
BAYAREALIKE	-71.81	2	0.013	0.160	0	147.6	0.110
BAYAREALIKE+J	**-68.79**	**3**	**0.012**	**0.038**	**0.014**	**143.6**	**0.850**

Figures

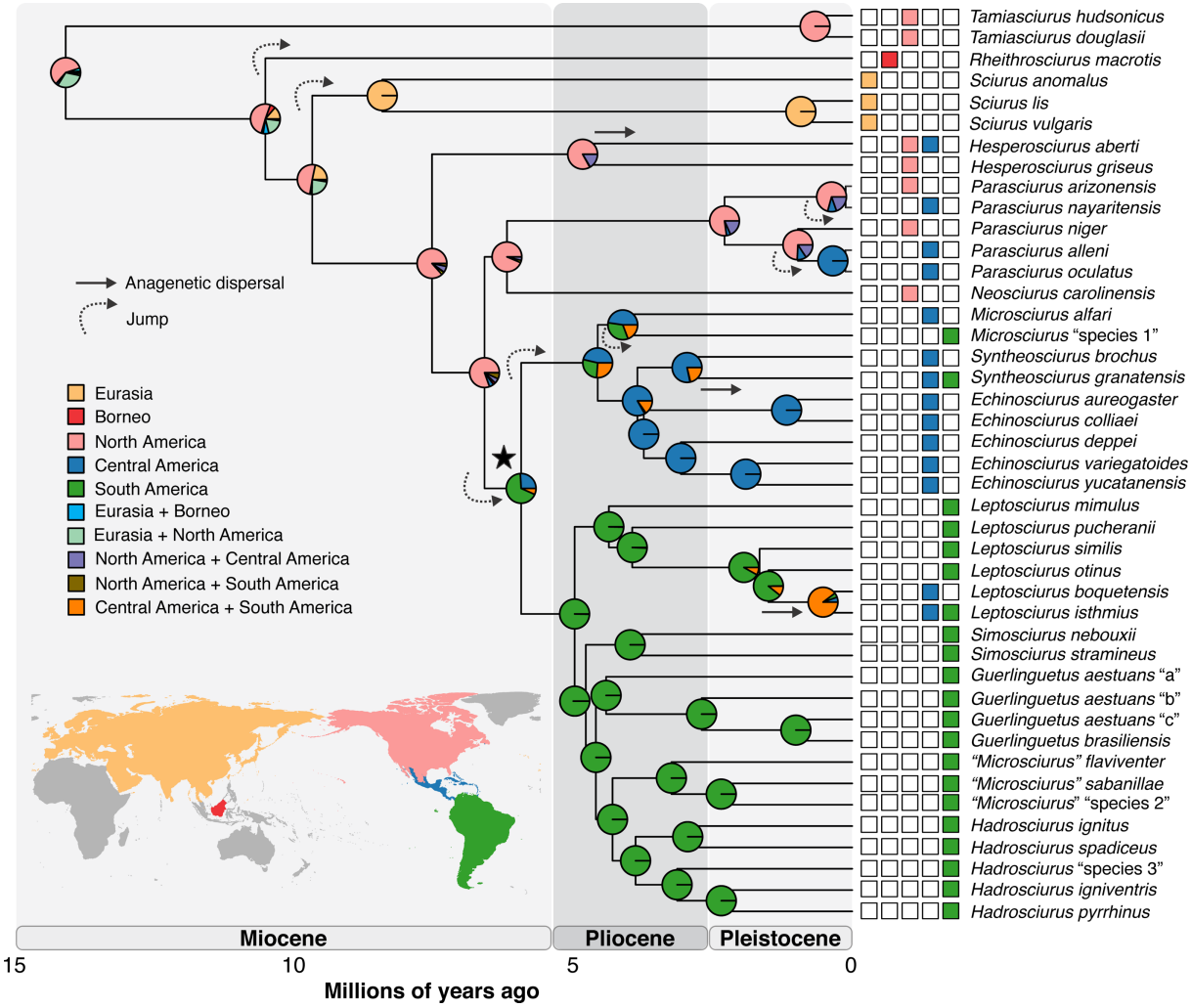


Figure 1. Bayesian maximum clade credibility chronogram based on 13 mitochondrial protein-coding genes (11,278 bp) of 43 species of Sciurini and two species of Pteromyini used as outgroup (not shown). Ancestral ranges for Sciurini estimated by DEC + J model in the global scale analysis are illustrated. Colors in key boxes and map correspond to coloring of internal nodes (estimated probabilities of ancestral ranges) and terminal boxes (current distribution of species). Strait arrows denote anagenetic dispersal events and curved arrows designate founder-event speciation (jump). Black star identifies the node of Neotropical clade.

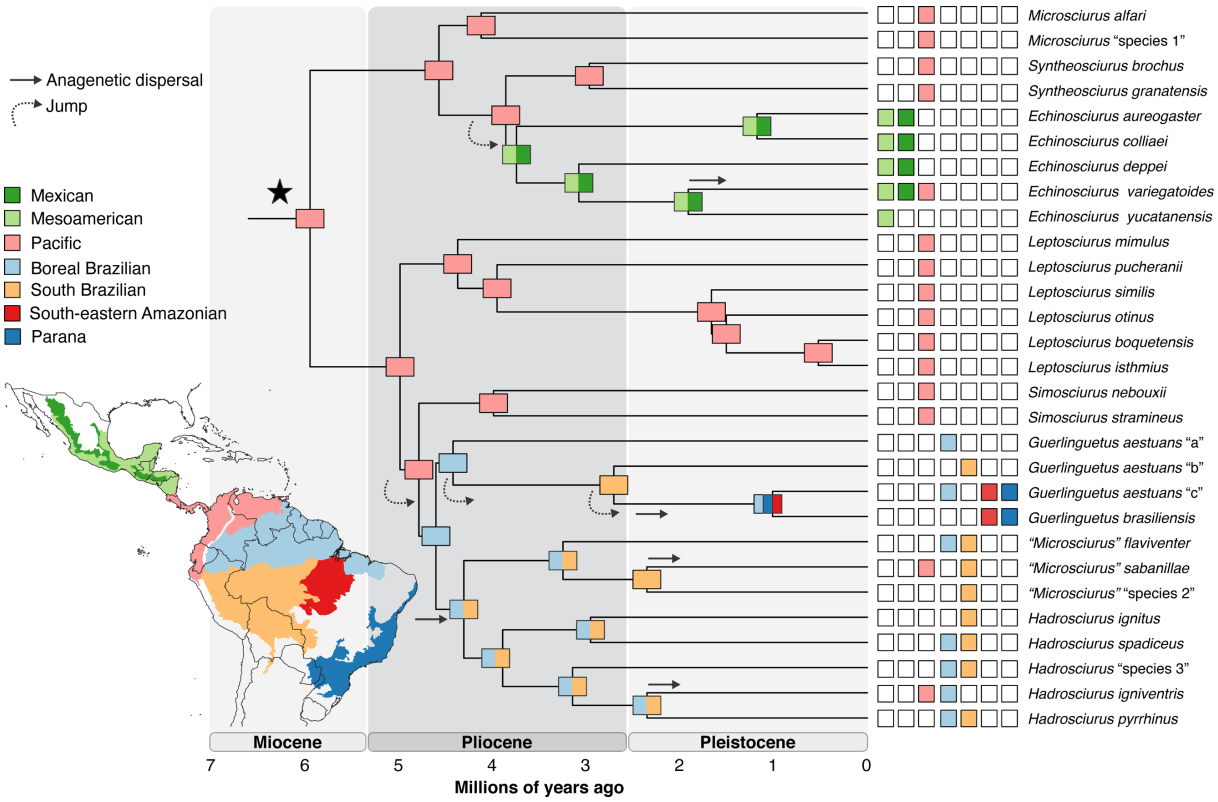


Figure 2. The most probable biogeographical scenario for Neotropical Sciurini based on the set of ancestral ranges that received the highest likelihood in the Neotropical scale analysis, as inferred by BAYAREALIKE + J model. Colors in key boxes and map correspond to coloring of internal tree nodes (most probable ancestral ranges) and terminal taxa (current distribution of species). Strait arrows denote anagenetic dispersal events and curved arrows designate founder-event speciation (jump).

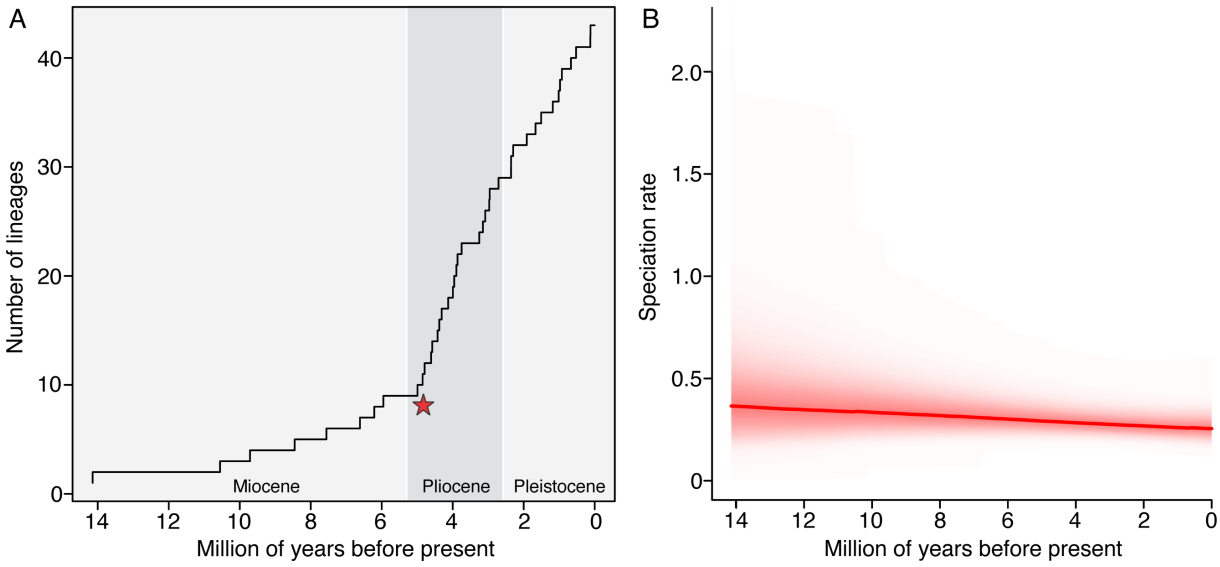


Figure 3. Speciation plots of Sciurini. (a) Lineage-through-time (LTT) plot showing a rapid increase in lineage accumulation after the South American invasion (red star). (b) BAMM analysis demonstrating the mostly constant speciation rate of Sciurini through time; solid red line shows median estimate and thin red lines (cloud) indicate probability distribution.

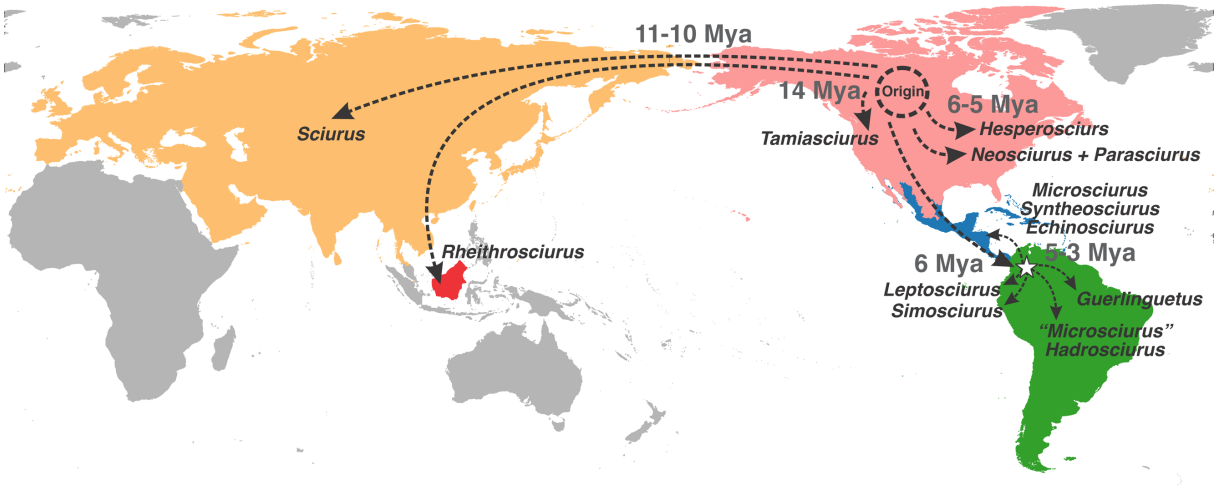
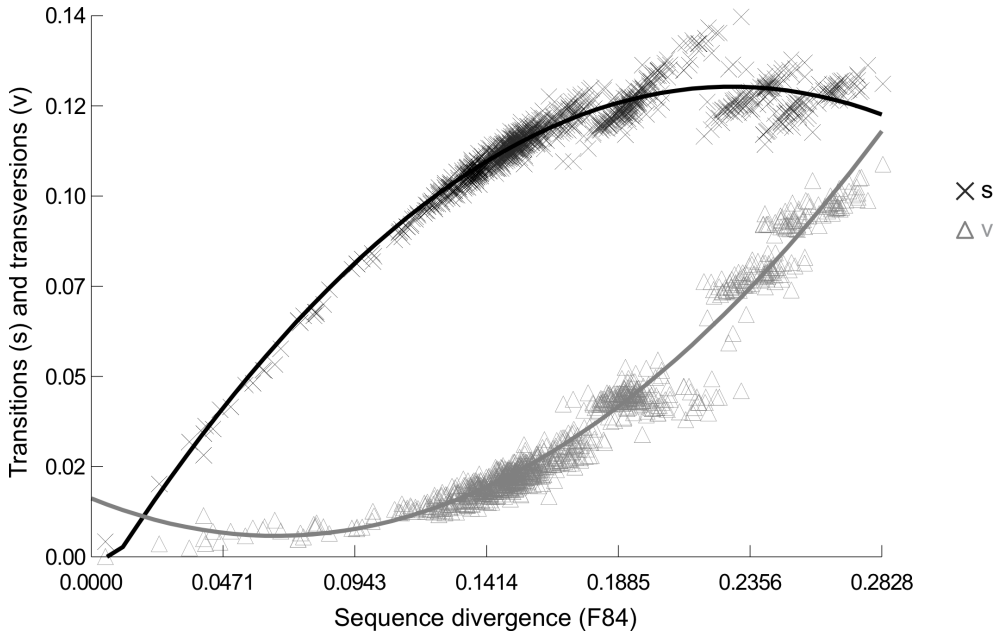


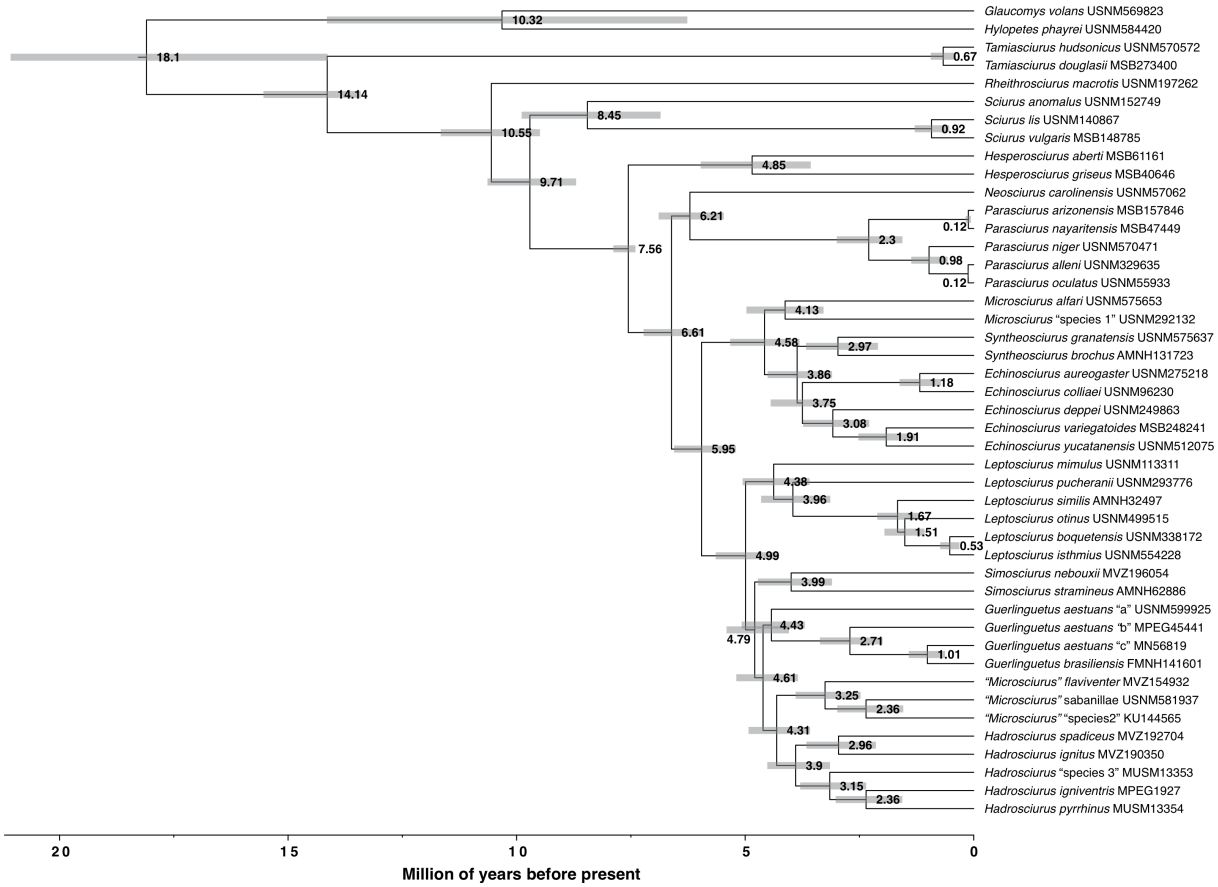
Figure 4. Map summarizing the inferred colonization scenario for the tribe Sciurini showing the timing and the place of its origin and diversification in North America, the two independent dispersions into Eurasia, the invasion of South America, the re-invasion of Central America, and the diversification in South America.

Supplementary Material

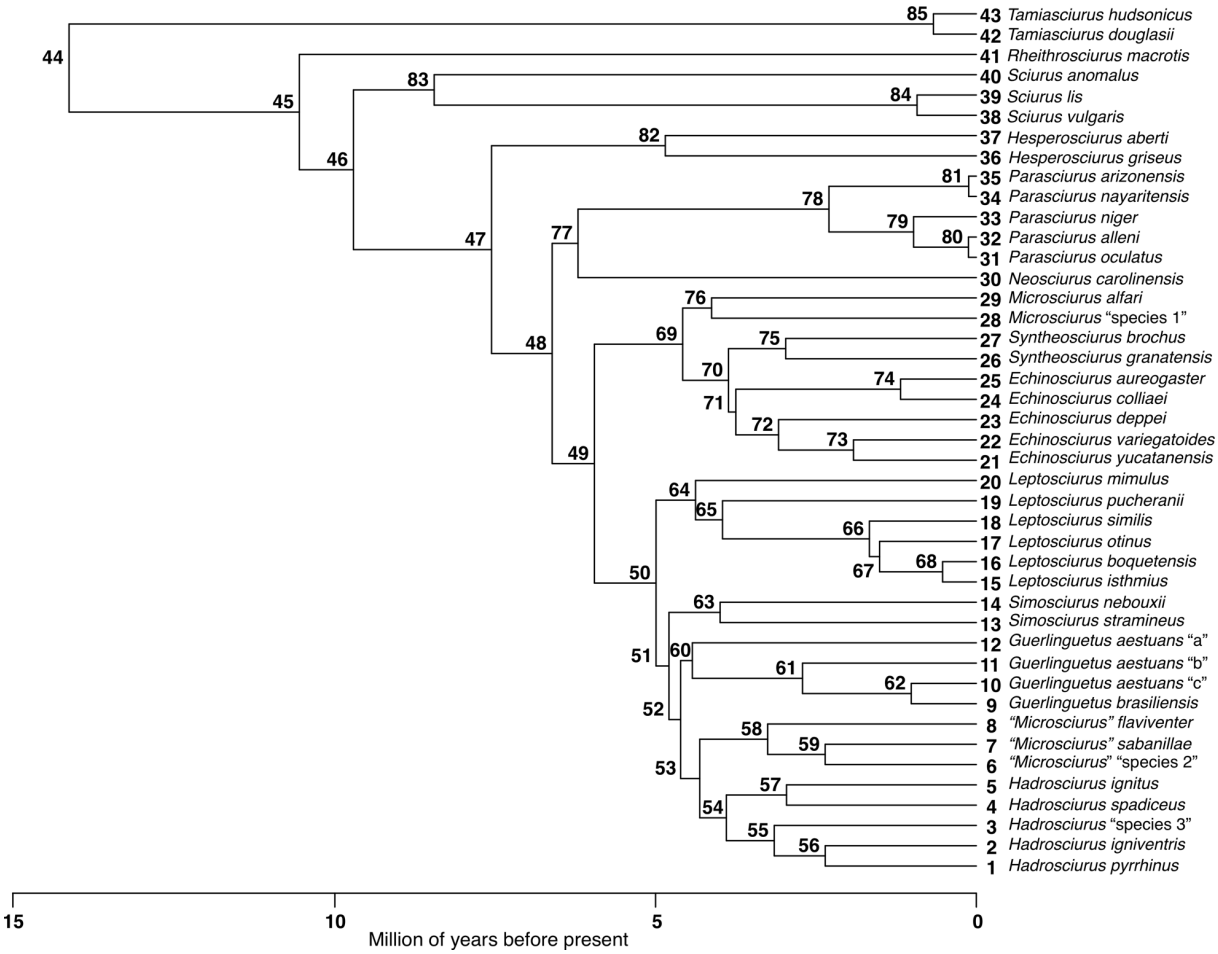
Supplementary Figures



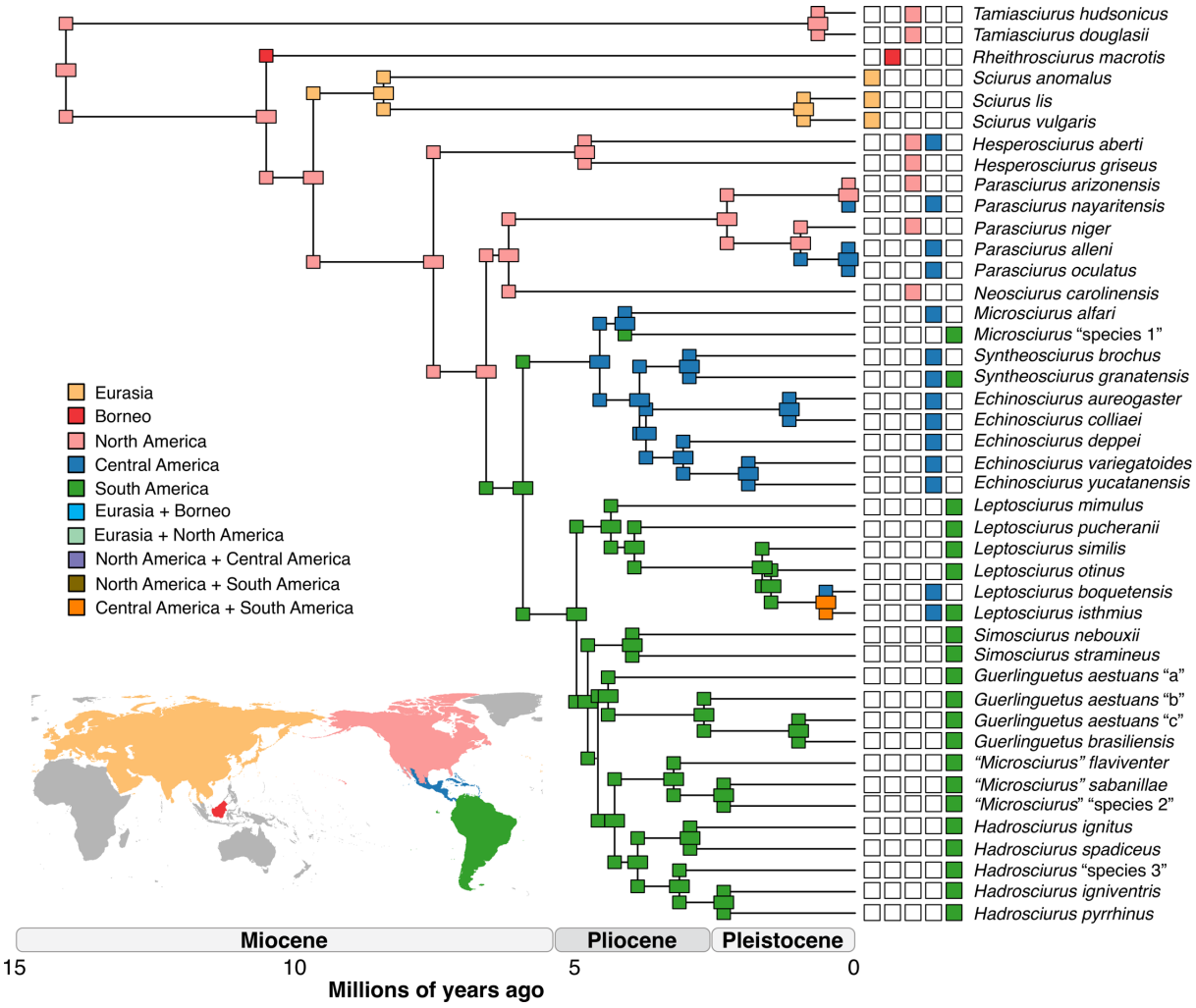
Supplementary Figure 1. Nucleotide substitution saturation plot. The number of transitions (s) and transversions (v) in the mitochondrial genome dataset is plotted against the F84 distances.



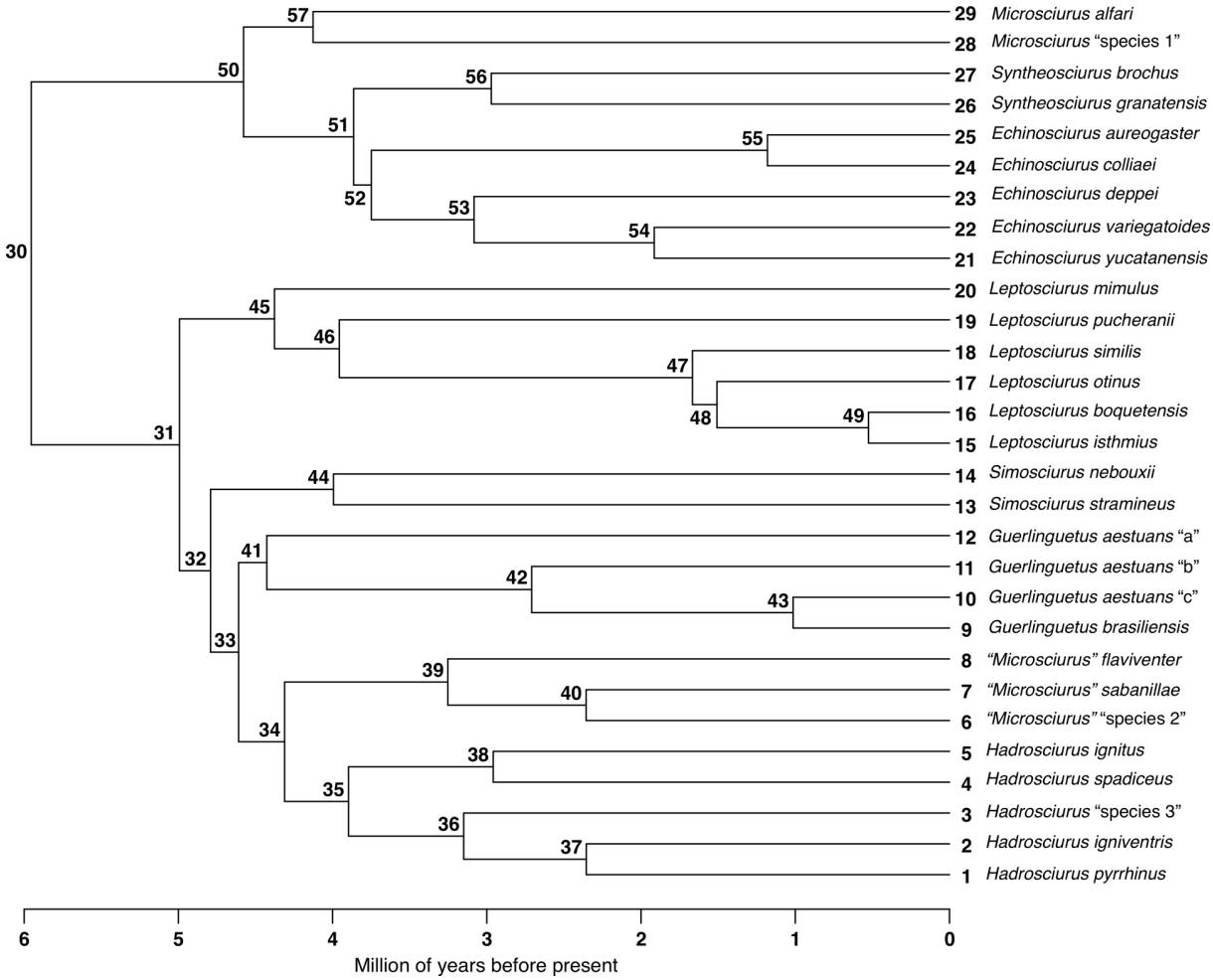
Supplementary Figure 2. Bayesian maximum clade credibility chronogram based on 13 mitochondrial protein-coding genes (11,278 bp) of 43 species of Sciurini and two species of Pteromyini used as outgroup. The mean heights along with the 95% highest posterior density intervals of dates are shown at each node.



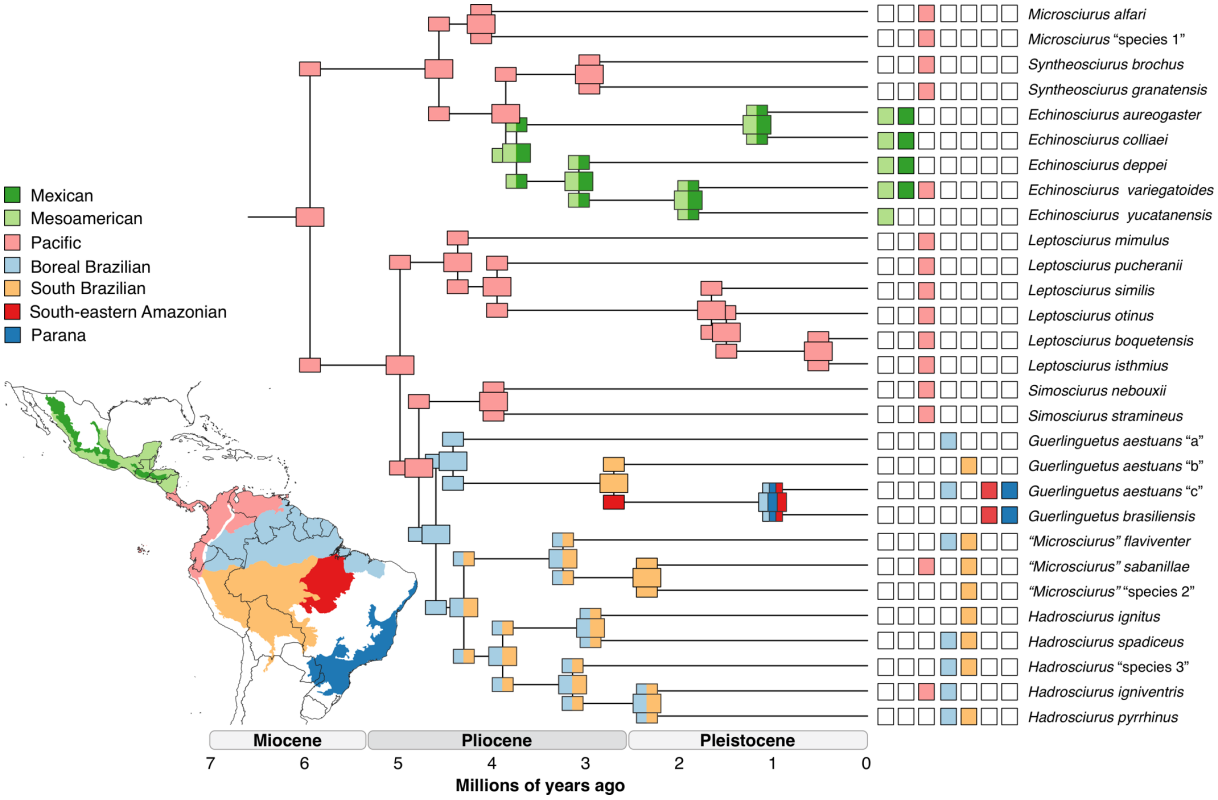
Supplementary Figure 3. Chronogram employed on the range evolution estimation of the tribe Sciurini on a global scale, showing the numbers of the nodes.



Supplementary Figure 4. The most probable biogeographical scenario for Sciurini based on the set of ancestral ranges that received the highest likelihood in the global scale analysis, as inferred by DEC + J model. Colors in key boxes and map correspond to coloring of boxes within the tree (most probable ancestral ranges) and at terminals (current distribution of species).



Supplementary Figure 5. Chronogram employed on the range evolution estimation of Sciurini on a Neotropical scale, showing the numbers of the nodes.



Supplementary Figure 6. The most probable biogeographical scenario for Sciurini based on the set of ancestral ranges that received the highest likelihood in the Neotropical scale analysis, as inferred by BAYAREALIKE + J model. Colors in key boxes and map correspond to coloring of boxes within the tree (most probable ancestral ranges) and at terminals (current distribution of species).

Supplementary Tables

Supplementary Table 1. List of specimens analyzed in this study, with geographic information and GenBank accession numbers for complete mitochondrial genomes. Taxonomic identifications follow Abreu-Jr et al. (2020). Voucher acronyms refer to the following scientific collections: American Museum of Natural History, USA (AMNH); Field Museum of Natural History, USA (FMNH); University of Kansas Natural History Museum, USA (KU); Museu Nacional da Universidade Federal do Rio de Janeiro, Brazil (MN); Museu Paraense Emilio Goeldi, Brazil (MPEG); Museum of Southwestern Biology, University of New Mexico, USA (MSB); Museo de Historia Natural, Universidad Nacional Mayor de San Marcos, Peru (MUSM); Museum of Vertebrate Zoology, University of California, USA (MVZ); Smithsonian National Museum of Natural History, USA (USNM). (**Continued**)

Voucher	Species	Locality	GenBank
MSB 273400	*Tamiasciurus douglasii*	Mexico: Baja California	
USNM 570572	*Tamiasciurus hudsonicus*	United States: Wisconsin	
USNM 197262	*Rheithrosciurus macrotis*	Indonesia: Borneo	
USNM 152749	*Sciurus anomalus*	Turkey: Soumela	
USNM 140867	*Sciurus lis*	Japan: Saitama Prefecture	MT134013
MSB 148785	*Sciurus vulgaris*	Russia: Sakha Republic	MT211956
MSB 61161	*Hesperosciurus aberti*	United States: Arizona	MT211954
MSB 40646	*Hesperosciurus griseus*	United States: California	MT211955
USNM 329635	*Parasciurus alleni*	Mexico: Nuevo Leon	
MSB 157846	*Parasciurus arizonensis*	United States: New Mexico	MT240880
MSB 47449	*Parasciurus nayaritensis*	United States: Arizona	MT240883
USNM 570471	*Parasciurus niger*	United States: Maryland	
USNM 55933	*Parasciurus oculatus*	Mexico: Mexico	
USNM 570626	*Neosciurus carolinensis*	United States: Wisconsin	MT240881
USNM 275218	*Echinosciurus aureogaster*	Guatemala: Quiche	
USNM 96230	*Echinosciurus colliaei*	Mexico: Sonora	MT240882
USNM 249863	*Echinosciurus deppei*	Guatemala: Peten	
MSB 248241	*Echinosciurus variegatoides*	Panama: Colón	MT240884
USNM 512075	*Echinosciurus yucatanensis*	Mexico: Yucatan	MT240885
USNM 292132	*Microsciurus* “species 1”	Colombia: Choco	
USNM 575653	*Microsciurus alfari*	Panama: Bocas Del Toro	
AMNH 131723	*Syntheosciurus brochus*	Costa Rica: Alajuela	MT240890
USNM 575637	*Syntheosciurus granatensis*	Panama: Bocas Del Toro	MT240889
USNM 338172	*Leptosciurus boquetensis*	Panama: Darien	
USNM 554228	*Leptosciurus isthmius*	Colombia: Valle del Cauca	
USNM 113311	*Leptosciurus mimulus*	Ecuador: Esmeraldas	MT259072

Supplementary Table 1. (**Continuation**)

USNM 499515	*Leptosciurus otinus*	Colombia: Antioquia	MT259073
USNM 293776	*Leptosciurus pucheranii*	Colombia: Antioquia	MT240886
AMNH 32497	*Leptosciurus similis*	Colombia: Cauca	MT259074
MVZ 196054	*Simosciurus nebouxii*	Peru: Cajamarca	MT259075
AMNH 62886	*Simosciurus stramineus*	Ecuador: Los Ríos	MT259076
USNM 599925	*Guerlinguetus aestuans* "a"	Guyana: Mount Roraima	MT259077
MPEG 45441	*Guerlinguetus aestuans* "b"	Brazil: Pará	MT259078
MN 56819	*Guerlinguetus aestuans* "c"	Brazil: Amazonas	MT259081
FMNH 141601	*Guerlinguetus brasiliensis*	Brazil: Sao Paulo	MT174525
KU 144565	*"Microsciurus"* "species 2"	Peru: Madre de Dios	
MVZ 154932	*"Microsciurus" flaviventer*	Peru: Amazonas	MT259119
USNM 581937	*"Microsciurus" sabanillae*	Peru: Amazonas	MT259123
MUSM 13353	*Hadrosciurus* "species 3"	Peru: Loreto	
MVZ 190350	*Hadrosciurus ignitus*	Brazil: Acre	MT259128
MPEG 1927	*Hadrosciurus igniventris*	Brazil: Roraima	MT259091
MUSM 13354	*Hadrosciurus pyrrhinus*	Peru: Loreto	MT259094
MVZ 192704	*Hadrosciurus spadiceus*	Peru: Madre de Dios	MT259110
USNM 569823	*Glaucomys volans*	United States: Pennsylvania	MT259089
USNM 584420	*Hylopetes phayrei*	Myanmar: Mandalay	

Supplementary Table 2. Alternative states of ancestral range estimated by DEC + J model in the global scale analyses. Node numbers are the same depicted in the phylogeny of Supplementary Figure 3. (**Continued**)

Nodes	Highest probability		Second highest probability	
	Area	Probability	Area	Probability
44	North America	58.6	Eurasia + North America	30.6
45	North America	51.1	Eurasia + North America	19.2
46	North America	50.7	Eurasia + North America	22.8
47	North America	86.6	North America + Central America	6.6
48	North America	80.4	North America + Central America	6.5
49	South America	67.5	Central America	25.9
50	South America	99.6	Central America + South America	0.4
51	South America	100.0		
52	South America	100.0		
53	South America	100.0		
54	South America	100.0		
55	South America	100.0		
56	South America	100.0		
57	South America	100.0		
58	South America	100.0		
59	South America	100.0		
60	South America	100.0		
61	South America	100.0		
62	South America	100.0		
63	South America	100.0		
64	South America	99.5	Central America + South America	0.5
65	South America	99.3	Central America + South America	0.7
66	South America	91.8	Central America + South America	8.2
67	South America	89.1	Central America + South America	10.7
68	Central America + South America	89.7	Central America	5.4
69	Central America	46.4	South America	27.5
70	Central America	83.2	Central America + South America	15.3
71	Central America	100.0		
72	Central America	100.0		
73	Central America	100.0		
74	Central America	100.0		
75	Central America	78.3	Central America + South America	21.2
76	Central America	47.7	South America	33.1
77	North America	93.6	North America + Central America	3.8
78	North America	77.2	North America + Central America	18.4

Supplementary Table 2. (**Continuation**)

79	North America	74.2	North America + Central America	15.7
80	Central America	100.0		
81	North America	70.8	North America + Central America	19.6
82	North America	83.0	North America + Central America	16.5
83	Eurasia	100.0		
84	Eurasia	100.0		
85	North America	100.0		

Supplementary Table 3. Alternative states of ancestral range estimated by BAYAREALIKE + J model in the Neotropical scale analyses. Node numbers are the same depicted in the phylogeny of Supplementary Figure 5. Areas acronyms are: Mexican transition zone (MEX), Mesoamerican dominion (MES), Pacific dominion (PAC), Boreal Brazilian dominion (BBR), South Brazilian dominion (SBR), South-eastern Amazonian dominion (SEA), and Parana dominion (PAR).

Nodes	Highest probability		Second highest probability	
	Area	Probability	Area	Probability
30	PAC	92.2	MES + MEX	3.9
31	PAC	99.0	BBR + SBR	0.7
32	PAC	97.4	BBR	1.1
33	BBR	61.3	SBR	25.4
34	BBR + SBR	74.1	SBR	22.0
35	BBR + SBR	89.0	SBR	10.5
36	BBR + SBR	98.0	SBR	0.9
37	BBR + SBR	93.6	BBR + PAC + SBR	4.7
38	BBR + SBR	88.1	SBR	11.6
39	BBR + SBR	71.8	SBR	25.8
40	SBR	77.4	BBR + SBR	16.9
41	BBR	63.8	SBR	21.8
42	SBR	29.9	BBR	13.3
43	BBR + PAR + SEA	52.3	PAR + SEA	46.6
44	PAC	99.8	BBR + PAC	0.03
45	PAC	100.0		
46	PAC	100.0		
47	PAC	100.0		
48	PAC	100.0		
49	PAC	100.0		
50	PAC	93.6	MES + MEX	5.2
51	PAC	86.7	MES + MEX	11.3
52	MES + MEX	71.8	MES	19.6
53	MES + MEX	84.2	MES	13.5
54	MES + MEX	79.6	MES	14.0
55	MES + MEX	99.7	MES	0.2
56	PAC	99.8	MES + PAC	0.1
57	PAC	99.8	MES + PAC	0.04

4. CHAPTER 3: Ultraconserved Elements resolve the squirrel tree at deep time scales but deliver lack of phylogenomic consistency in a rapid Neotropical radiation

Echinosciurus aureogaster (Credit: Pedro Peloso)

Ultraconserved Elements resolve the squirrel tree at deep time scales but deliver lack of phylogenomic consistency in a rapid Neotropical radiation

Abstract

Ultraconserved Elements have been proved useful to resolve challenging phylogenies, helping to unpuzzle conflicting relationships and to unravel the evolutionary history of several branches of the Tree of Live. In the Neotropical region, two distinct radiation of squirrels are found: the Neotropical pigmy squirrel (only representative of the subfamilie Sciurillinae) and the tree squirrels (members of the tribe Sciurini, subfamilie Sciurinae). The relationship of the pigmy squirrel with other sciurids remains unclear. For tree squirrels, a recent phylogenetic hypothesis was published and brought some resolution on the genus and species ranks, but it was based exclusively on mitogenome data, which not always agrees with inferences employing nuclear markers. In the present study, we sequenced thousands of UCE loci from 184 modern and historical samples, including representatives of all Sciuridae subfamilies, with especial emphasis to the Neotropical radiations. The recovery of UCE loci and the mean length per UCE locus were significantly impacted by the sample age. None of our 18 phylogenetic inferences, performed with a concatenated-based method (RAxML) and two coalescent-based methods (ASTRAL-III and SVDquartets), resulted in identical topologies. In general, RAxML outperforms the coalescent methods, providing the topologies less dissimilar among each other and with the highest medians of nodal support. We observed that matrices including larger numbers of loci provided more similar topologies and higher values of nodal support, and we also detected that filtering for information content did not increased topological consistency nor improved the overall nodal support for any inference method. Our analyzes estimated with strong support the relationship among the five subfamilies of Sciuridae, and also provided consistent and well-supported results for the relationships among the deepest branches of Sciurini. For the Neotropical radiation, which experienced a rapid diversification, conflicting relationships at both genus- and species-level were estimated upon inference method. Some inconsistences were also recovered with regards to the mitogenomic hypothesis previously published. Even with the tremendous amounts of genome-scale molecular data available to estimate phylogenies, for some recalcitrant clades —which seems to be the case of the Neotropical tree squirrels—we might never reach a phylogenetic consensus.

Keywords: Sciurillinae, Sciurini, Historical samples, Phylogeny, Filtering approaches, Mito-nuclear discordance.

4.1 Introduction

Targeted capture of Ultraconserved Elements (UCEs) has been shown a powerful technique to sample and sequence thousands of loci across the nuclear genome (Faircloth et al. 2012). It has been useful to resolve challenging phylogenies, unpuzzling conflicting relationships and unraveling the evolutionary history of several branches of the Tree of Live (Crawford et al. 2012; Faircloth et al. 2013; Hawkins et al. 2016; Baca et al. 2017; Streicher and Wiens

2017; Oliveros et al. 2019). UCEs are regions of the genome largely conserved among divergent taxa (Faircloth et al. 2012), and their great advantage on phylogenetic studies is that they contain —besides the highly conserved core regions— flanking regions that are sufficiently variable among closer taxa (Faircloth et al. 2012; McLean et al. 2019). Therefore, UCEs can be targeted at more inclusive taxonomic categories and be informative at multiple evolutionary timescales (McLean et al. 2019).

However, there is an ongoing debate on how to curate and properly analyze UCE data, as recent phylogenetic studies have shown the sensitivity of UCE datasets in providing distinct topologies upon taxa sampling strategy, data filtering options, and inference methods (Hosner et al. 2016; Streicher et al. 2016; Platt et al. 2018). Regarding data filtering, some approaches are available (see Streicher et al. 2016; Platt et al. 2018; McLean et al. 2019) and two of the most common are: (1) the matrix completeness, which determines the percentage of taxa that has to be present in a given locus to be included in the final matrix; and (2) the information content, which can be expressed by the percentage of informative sites per UCE locus. Both approaches ultimately define the number of loci to be analyzed. It has been demonstrated that increasing the overall amount of information (e.g. number of loci in the datasets) would maximize the robustness of the phylogenetic inferences, in terms of both statistical support and topological consistency, especially when using concatenated-based methods of inference (Platt et al. 2018; Dornburg et al. 2019; McLean et al. 2019). On the other hand, the use of a filtering strategy based on the information content is strongly recommended, especially for datasets to be analyzed with coalescent methods. These methods are sensitive to imprecise gene trees provided by loci with low information content, as those loci can support alternative gene topologies with equal probabilities of topologies derived from loci with greater information and higher phylogenetic signal (Hahn and Nakhleh 2016; Mirarab et al. 2016; Blom et al. 2017). Filtering UCE loci based on information content can also benefit inferences of rapid diversified radiations, which —due to the presence of large zones of low phylogenetic signal— represent some of the most difficult phylogenetic problems to resolve (Smith et al. 2014; McLean et al. 2019).

Two distinct radiations of squirrels (family Sciuridae) are found in the Neotropics. One of those is represented by the monotypic Neotropical pygmy squirrel, *Sciurillus pusillus*, which is included in its own subfamily, Sciurillinae (Thorington et al. 2012). This enigmatic squirrel diverged early in the history of the family (about 35 Million years ago [Mya]; Mercer

and Roth 2003; Upham et al. 2019), and its phylogenetic affinities remain unclear, as controversial relationships have been suggested (Mercer and Roth 2003; Steppan et al. 2004; Fabre et al. 2012; Zelditch et al. 2015; Upham et al. 2019). It is noteworthy that no phylogenetic study published so far included more than two samples of *S. pusillus*, and so the genetic diversity within this Amazon widespread taxon (Vivo and Carmignotto 2015) is also unknown.

The other radiation, the tribe Sciurini (subfamily Sciurinae), is a speciose group of Holarctic and Neotropical forest dwelling squirrels that originated around 14 Mya, most likely in North America (Mercer and Roth 2003; Thorington et al. 2012; Abreu-Jr et al. 2020a). Tree squirrels have been shown to exhibit strikingly high rates of diversification (Roth and Mercer 2008), especially in the Neotropics, where they experienced an explosive diversification with dramatic species accumulation after the South America invasion around 6 Mya (Abreu-Jr et al. 2020a). Until recently, the genetic diversity and the molecular systematics of tree squirrel were poorly known, as the few published studies were very limited in terms of taxonomic coverage, particularly regarding to Neotropical taxa, and also relayed on quite small and heterogeneous datasets, composed of few mitochondrial and nuclear genes (Oshida and Masuda 2000; Villalobos and Cervantes-Reza 2007; Oshida et al. 2009; Pečnerová and Martínková 2012; Villalobos and Gutierrez-Espeleta 2014; Pečnerová et al. 2015; Aghbolaghi et al. 2019, 2020).

Abreu-Jr et al. (2020b) published heretofore the most comprehensive phylogenetic study for tree squirrels, including data from 232 historical and modern specimens, representing 43 of the 46 putative species of the tribe Sciurini (sensu Thorington et al. 2012; Vivo and Carmignotto 2015; Hope et al. 2016). In this contribution, they clarified several important points on the phylogenetic relationships and systematics of tree squirrels, and proposed a new tentative supra-specific classification, employing 14 generic names used by previous authors (e.g. Allen 1915; Moore 1959; Vivo and Carmignotto 2015). The phylogenetic hypothesis presented by Abreu-Jr et al. (2020b) includes a dense taxonomic and geographic sampling and was based exclusively on mitochondrial genome data. Despite the consistent hypothesis established by these authors, the evidence for pervasive natural selection, uniparental inheritance and the lack of recombination on the mitochondrial genome make it susceptible to evolutionary processes distinct to the nuclear genome (Edwards et al. 2005; Edwards and Bensch 2009), and thus the history derived from

mitogenomic data not always agrees with the history provided by a nuclear dataset. Even the ability of mitogenomes to provide strong-supported topologies do not guarantee that the mitochondrial tree represents the taxa tree (Platt et al. 2018).

In this study, we provide a nuclear genome-wide perspective of the Neotropical squirrels (Sciurilinnae and Sciurinae: Sciurini) phylogeny, employing over 3,700 UCE loci sequenced from 184 historical and modern museum specimens. Our taxa sampling includes, besides the Neotropical forms, representatives of all Sciuridae subfamilies. For historical samples we investigated if the sample age (i.e. the year in which the specimen was collected) influenced the number of UCE loci recovered and also the average length of those loci. We examined the impact of two dataset filtering approaches (taxa representativeness per UCE loci and proportion of variable sites per UCE loci) in order to verify how matrix completeness and informative content could affect the phylogenetic consistency. We also assessed the sensitivity of three conceptually distinct optimality criteria (a concatenated method in RAxML, a coalescent gene-tree based method in ASTRAL-III, and a coalescent site-based method in SVDquartets) to estimate the phylogeny of squirrels.

4.2 Material and methods

4.2.1 Sampling

We obtained frozen or ethanol-preserved tissues (hereafter modern samples) from 188 specimens and we sampled remains of muscular tissue adherent to skulls or skin clips (hereafter historical samples) from 94 dry museum specimens. Historical samples were gathered in order to complement missing species or relevant geographic variants, with special effort on Neotropical taxa. Sampling from dry museum specimens followed rigorous procedures to avoid cross-contamination, as described in McDonough et al. (2018) and Abreu-Jr et al. (2020b). Voucher material from all samples are housed in 27 scientific collections from South America, North America, and Europe (see Supplementary Catalog data of voucher material). Our sampled material includes the only representative of Sciurillinae (*Sciurillus pusillus*), one species of Ratufinae (*Ratufa* sp.), two samples of Xerinae (*Marmota monax*), and 45 species of Sciurinae, being two from the tribe Pteromyini (*Glaucomys volans* and *Hylopetes phayrei*) and 43 from the tribe Sciurini. Additionally, in order to include representatives of all Sciuridae subfamilies, we obtained UCE data available at GenBank for three species of Callosciurinae (*Callosciurus adamsi*, *Exilisciurus exilis*, and

Lariscus insignis; Hawkins et al. 2016). We also sampled two individuals of *Aplodontia rufa* (family Aplodontiidae, sister to Sciuridae; Upham et al. 2019) to be used as outgroup. A complete list of specimens included in this study accompanied by catalogue and geographic data and other relevant information is provided as Supplementary Table S1. Taxonomic identifications at the genus and species levels for the tribe Sciurini follow Abreu-Jr et al. (2020b) and for the remaining taxa we follow Thorington et al. (2012).

4.2.2 DNA extraction and library preparation

DNA extractions of modern samples were performed using the DNeasy® Blood & Tissue kit, following manufacture's protocol (Qiagen Inc.). DNA of historical samples was extracted using a standard phenol-chloroform protocol (see detailed protocol in McDonough et al. 2018 and also additional information in Abreu-Jr et al. 2020b) in an isolated ancient DNA facility at the Smithsonian's Center for Conservation Genomics (CCG). Total DNA concentrations were measured using a Qubit 2.0 fluorometer (Thermo Fisher Scientific). DNA extracted from ethanol-preserved tissues was sheared with QSonica Q800R using 25 % of amplitude and 5 min of on/off pulse. Sheared DNA was visualized on agarose gel to confirm the resulting fragment size around 300 bp. Approximately 500 ng of sheared modern DNA and about 35 µl of historical DNA (regardless of concentration) were purified using 5x SPRI magnetic beads (Rohland and Reich 2012). Library preparations were performed using the KAPA LTP Library Preparation Kit (Roche Sequencing) following the manufacturer's protocol. Nextera-style indices and KAPA HiFi Hotstart ReadyMix (Roche Sequencing) were used for indexing PCRs (iPCR). The iPCR profile followed Abreu-Jr et al. (2020b). Amplified samples were purified using 1.8x SPRI magnetic beads, quantified with a Qubit 2.0 fluorometer (Thermo Fisher Scientific) and visualized on a 1.5% agarose gel.

4.2.3 UCE capture and sequencing

Libraries were multiplexed in equimolar ratios for UCE enrichment. For historical samples we pooled up to four libraries and for modern samples up to eight libraries. No historical samples were pooled with modern samples to avoid biased enrichment. Multiplexed libraries were then dried out and re-diluted in 7 µl of nuclease-free water. UCEs enrichments were performed using myBaits UCE Tetrapods 5Kv1 kit (Arbor Biosciences), containing a probe set of around 5,000 UCE loci. Post-capture amplifications were performed using KAPA

HiFi Hotstart ReadyMix (Roche Sequencing), with the following profile: initial denaturation at 98 °C for 2 min, a final extension at 72 °C for 7 min, and 15 (for modern samples) or 16 (for historical samples) cycles of amplification, with denaturation at 98 °C for 20 sec, annealing at 60 °C for 30 sec, and extension at 72 °C for 30 sec. A 1.8x SPRI magnetic beads cleanup was performed subsequently.

Cleaned amplification products were quantified using a Qubit 2.0 fluorometer (Thermo Fisher Scientific) and visualized on a Bioanalyzer (Agilent) with high sensitivity kits. Equimolar pooling of samples was performed based on the concentration (ng/µl) and on the average size (bp) of amplified fragments. Due to the high concentration of dimmers obtained, especially for historical samples, we size-selected the fragments of DNA to be sequenced in between 200 and 550 bp using a Pippin Prep (Sage Science). The total of 199 samples were split in three distinct Illumina runs: one on a Hi-Seq 2500 125 PE at the DNA Sequencing Center at the Brigham Young University, Utah; and two on a Hi-Seq 4000 150 PE at the Vincent J. Coates Genomics Sequencing Laboratory at the University of California, Berkeley. Sequencing reactions included Illumina Free Adapter Blocking Reagent to prevent index hopping.

4.2.4 Data processing

We followed the PHYLUCE 1.6 pipeline (Faircloth et al. 2012; Faircloth 2016) adapted to the Smithsonian Institution High Performance Cluster (SI/HPC) to process target-enriched UCE data (see https://github.com/SmithsonianWorkshops/Targeted_Enrichment/). We cleaned adapter contamination and low-quality bases using Illumiprocessor 2.0 (Faircloth 2013; Bolger et al. 2014). Reads were assembled into contigs with Trinity (Grabherr et al. 2011) and contigs were matched to the uce-5k-probe-set (Faircloth et al. 2012). We selected only samples for which at least 1,000 UCE loci were enriched to be included in the subsequent alignments. UCE loci were aligned using MAFFT 7 (Katoh and Standley 2013; Nakamura et al. 2018) considering edge-trimming only. We did not use the default internal trimming of PHYLUCE, Gblocks (Castresana 2000; Talavera and Castresana 2007), because this approach has been shown to produce less accurate results in the context of phylogenetic reconstructions (Tan et al. 2015). In the resulting alignments, we finally excluded samples with over 70% of missing data, as in preliminary analyses (results not shown) those samples

appeared with extreme long branches in ML topologies and also their phylogenetic affinities were inconsistently recovered by coalescent methods.

4.2.5 UCE sequencing success for historical museum samples

We investigated if the sample age influenced our success in obtaining UCE data for historical museum specimens, by exploring the relationship between sample age and two metrics of sequencing success: the number of UCE loci recovered per sample and the mean length of UCE locus per sample. We performed linear regression models using the function "lm" from the package stats v3.6.2 in RStudio 1.1.463 (RStudio, Inc.).

4.2.6 UCE data filtering approaches

We applied two locus-based strategies of data filtering. Firstly, we generated data matrices with distinct percentages of taxa (=samples) representativeness per UCE loci (also referred as matrix completeness). This option controls the percentage of missing taxa allowed per locus to be included in the data matrices. Allowing lower values of taxa representatives per locus (i.e. higher values of missing taxa per locus), the number of UCEs presented in the final data matrices increases substantially, but it also leverages an increase on the matrices overall missing data. We generated three distinct datasets, each one including UCEs that were enriched in at least 70% (1,013 UCEs), 60% (3,276 UCEs), and 50% (3,713 UCEs) of samples, using the "phyluce_align_get_only_loci_with_min_taxa" function (Faircloth 2016). Our main goal with this strategy was to evaluate if more strict datasets (with fewer loci, but with lower rates of overall missing data) would perform differently from more flexible datasets (with a larger number of loci, but also with higher rates of overall missing data).

These datasets were the basis for the subsequent filtering strategy focusing on the information content. We use the function "phyluce_align_get_informative_sites" to create three new datasets including only UCEs with 5% or more of variable sites. This reduced the number of UCE loci in each matrix to: 484 loci in the 70% taxa representativeness matrix, 1,630 loci in the 60% matrix, and 1,838 loci in the 50% matrix. The percentage of variable sites was quantified relatively to the locus length. According to McLean et al. (2019) using the proportion of variable sites as a metric to quantify information content is very desired and advantageous because it captures intrinsic properties of loci (e.g. substitution rate and numbers of parsimony informative sites) and do not require previous inferences of genes or

species trees that could introduce additional errors. The step-by-step workflow of our data filtering approach, along with our taxon sampling and alignment strategies and inference methods used, is presented as Supplementary Figure S1.

4.2.7 Phylogenetic inferences

We performed phylogenetic inferences through a concatenated approach (Maximum Likelihood [ML] method) and also using two distinct coalescent approaches, a two-step summary or locus-based coalescent method (species tree inferred from individual gene trees obtained from ML analyses) and a site-based method. For the ML analyses concatenated matrices were generated using the function “phyluce_align_format_nexus_files_for_raxml” (Faircloth 2016). ML analyses were performed in RAxML 8.2.7 (Stamatakis 2014) by running ten independent searches under the GTR nucleotide substitution model with gamma-distributed rate heterogeneity (GTRGAMMA). The best-scoring ML trees were selected to draw the bootstrap support values obtained from 100 replicates using the “thorough standard bootstrap” optimization option. All RAxML analyses were conducted on the Smithsonian Institution High Performance Cluster (https://doi.org/10.25572/SIHPC).

To perform locus-based coalescent analyses we employed ASTRAL-III (Zhang et al. 2018). Individual gene trees were inferred by running ML analyses in RAxML 8.2.7 (Stamatakis 2014) as described above, but with 1,000 bootstrap replicates. The gene trees with the highest likelihood scores were used as input trees in ASTRAL-III. Inferences in ASTRAL-III were performed with default parameters. Site-based coalescent analyses were performed using SVDquartets (Chifman and Kubatko 2014) implemented in PAUP* 4.0a166 (Swofford 2003). We used as input source concatenated matrices partitioned by locus. We run exhaustive quartet sampling amylases —which evaluates all possible quartets— to search for the best three for each dataset, and then we run a new set of analyses assessing 1,000,000 randomly sampled quartets and performing 100 standard bootstrap replicates. Nodal support was calculated from the bootstrap replicates and added to the exhaustive quartet sampling trees using the function “addConfidences” from package phangorn (Schliep 2011).

4.2.8 Topological comparisons

We performed within-method comparisons of the overall nodal support obtained for the six distinct datasets (matrices with 50, 60, and 70% of completeness, and with or without filter for informative sites), using boxplots created with the function "geom_boxplot" from packed ggplot2 (Wickham 2016). We examined and quantified discordances among the inferred phylogenetic trees —within and among methods— based on the Robinson-Foulds (RF) distance metric, calculated using the function "RF.dist" from package phangorn (Schliep 2011). Distances were normalized to enable comparisons among methods. To visualize those differences, we firstly performed pairwise comparisons of trees, creating heat maps with the function "levelplot" from lattice (Sarkar 2008). Subsequently, we used the Metric Multidimensional Scaling (MDS, aka Principal Coordinates Analysis, PCoA) to summarize tree distances ("dudi.pco" function in adegenet; Jombart 2008) and visualize them in a two-dimensional space using ggplot2 (Wickham 2016).

4.3 Results

4.3.1 Sequencing summary, UCE assembly and success for historical samples

We sequenced 282 samples, 94 historical museum specimens and 188 modern specimens with frozen or ethanol-preserved tissues. The mean paired-end reads recovered for historical samples was 2,584,400 (ranging from 32,012 to 15,776,174) and for modern samples was 3,344,948 (ranging from 4,077 to 12,726,177). Trinity assembled an average of 15,501 contigs per historical sample (ranging from 8 to 226,333) and 41,704 contigs per modern sample (ranging from 101 to 317,895). Unique contigs matched an average of 720 UCE loci per historical sample (ranging from 0 to 2,419) with mean length of 274 bp, and an average of 2,459 UCE loci per modern sample (ranging from 18 to 3,385) with mean length of 454 bp (see more detailed information on Supplementary Table S2). Modern samples performed significantly better than historical samples both in number of UCE loci enriched per sample and also on the UCE's mean length ($W = 995$, $p < 0.001$, and $W = 867$, $p < 0.001$, respectively). We observed that specimen age was a significant predictor of the number of UCE loci recovered for historical samples ($R^2 = 0.1072$, $p < 0.01$; Figure 1A) and that the average UCE locus length was also correlated with specimen age ($R^2 = 0.0544$, $p < 0.05$; Figure 1B).

We adopted a threshold of a minimum of 1,000 UCE loci sequenced and a maximum of 70% of overall missing data per sample to be included in the phylogenetic analyzes. Our goal with this was to ensure robustness and reliability to our inferences, dropping potentially problematic samples with unsatisfactory amount of data. Therefore, from the initial 282 samples sequenced by us, our final taxon set included 181 samples: 12 historical (representing 12.8% of the initial set) and 169 modern samples (representing 89.9% of the original sample set). The remaining 101 samples (82 historical and 19 modern) were not included in the downstream analyses. Most of these samples represented additional specimens from taxa included in the phylogenetic analyses, but some of those were unique representatives of one monotypic genus (*Rheithrosciurus macrotis*) and of ten other species, two from the genus *Sciurus* (*S. anomalus* and *S. lis*), two from *Parasciurus* (*P alleni* and *P. oculatus*), four from *Echinosciurus* (*E. aureogaster*, *E. colliaei*, *E. deppei* and *E. yucatanensis*), one from *Syntheosciurus* (*S. brochus*), and one from *Simosciurus* (*S. stramineus*).

4.3.2 Topological incongruences and the influence of data filtering strategies

We generated six distinct datasets that were analyzed with three methods of inference each, making a total of 18 phylogenetic analyses performed in this study (see Supplementary Figure S1). None of these analyzes resulted in identical topologies. RAxML provided the less divergent topologies based on the mean RF distance (0.24 ± 0.09), while SVDquartets recovered the most dissimilar trees (0.34 ± 0.13); the mean RF distance among ASTRAL-III topologies was intermediate (0.31 ± 0.11). Differences among topologies inferred by each method can be visualized on the scatterplots of the first two principal coordinates from PCoAs and on heat maps, both based on a pairwise RF matrix (Figure 2).

We also performed a PCoA including all the 18 trees. In the scatterplot depicted on Figure 3, RAxML and SVDquartets phylogenetic reconstructions are clustered close to each other —excepted from the inferences based on the 70% matrices from both methods (see details below)—, regardless the discrepancies on their conceptual frameworks. ASTRAL-III, which used RAxML trees built for each UCE to perform a locus-based coalescent analysis, provided the most distinctive topologies, not overlapping with any other tree, neither from SVDquartets or RAxML.

Across all methods of inference, matrices with 50 and 60% of completeness recovered more similar trees among each other, whereas 70% matrices provided the most

unique topologies (Figures 2 and 3). This includes trees from both unfiltered datasets and datasets containing only UCE loci with at least 5% of informative sites. Furthermore, using a filter for information content did not help to increase topological cohesion, as the two-dimensional space occupied by topologies from datasets filtered by information content is even bigger than the space occupied by topologies resulting from unfiltered datasets (see Figures 2A, 2C, and 2E). Same pattern can be observed in the heat maps, where the dissimilarities among filtered datasets are higher than the values among filtered ones (see Figures 2B, 2D, and 2F).

Regarding the overall nodal support, among methods comparison show that the median bootstrap values obtained from RAxML analyses were higher than the nodal support obtained for the coalescent methods (Figure 4). Within method comparisons demonstrate that 50 and 60% matrices recovered higher medians and shorter interquartile ranges than 70% matrices (Figure 4). Moreover, in all methods of inference datasets with no filter for information content performed slightly better than datasets with ≥ 5% of informative sites. In summary, we observed that datasets including more UCE loci (50 and 60% matrices with no filter for informative sites) provided the most similar topologies among each other and also the highest values of nodal support.

4.3.3 Phylogeny of Sciuridae: relationships among subfamilies and the genetic diversity of Sciurillinae

We recovered the family Sciuridae as a monophyletic group with strong support (bootstrap > 95% or local posterior probability > 0.95) in all phylogenetic reconstructions, regardless the dataset or inference method. From the total of 18 phylogenetic analyses, 11 recovered identical relationships for the five currently recognized subfamilies of Sciuridae. All concatenated matrices, excepted by one, analyzed with RAxML recovered Ratufinae as the most external lineage within Sciuridae, followed by Sciurillinae, and a clade composed of Sciurinae sister to Xerinae + Callosciurinae (Figure 5). These relationships received full support (bootstrap = 100%) in those analyses. Only in the RAxML analysis performed with the most restricted dataset (70% matrix with filter for informative sites; including 484 UCE loci), Ratufinae appeared as sister-group of Sciurillinae, but with no significant support.

SVDquartets analyses with 50% and 60% data matrices (regardless the percentage of informative sites) recovered identical inter-subfamilies relationships as the most frequently

observed topology from RAxML; however, in SVDquartets reconstructions the position of Sciurillinae was significantly supported only in the analyses performed with the 60% data matrix, without filter for information content. Both 70% matrices recovered an alternative scenario, with Sciurillinae closer to the root of Sciuridae and Ratufine diverging afterward, but, likewise in RAxML, this alternative hypothesis did not receive statistical support.

ASTRAL-III suggested the most conflicting evolutionary hypotheses for the initial diversification of Sciuridae. Analyzes performed with unfiltered datasets regarding the information content, recovered Ratufinae and Sciurillinae as sister-groups. In two analyses with datasets including only loci with at least 5% of informative sites, Ratufinae was recovered as the sister clade to all other subfamilies within Sciuridae; and in the last inference (also filtered by information content) Sciurillinae appeared as sister to all other lineages. However, none ASTRAL-III analyses recovered significant support (i.e. local posterior probability > 0.95) for these relationships. Regarding to the remaining subfamilies, ASTRAL-III recovered identical relationships as RAxML and SVDquartets, and also with full support (local posterior probability = 1).

The subfamily Sciurillinae currently includes exclusively the monotypic *Sciurillus pusillus*. Our analyzes of UCE data based on both concatenated and coalescent approaches, consistently recovered two fully supported clades within Sciurillinae (see Figure 5). The first clade is composed of three samples from eastern Amazon, being two from Pará (localities [1] and [3] in Figure 6), Brazil, and one from French Guiana; the second clade, includes four samples from two localities in Peru, western Amazon (localities [4] and [5] in Figure 6).

4.3.4 Phylogeny of Sciurini: generic relationships, species monophyly and affinities

Our phylogenetic inferences included 163 specimens of tree squirrels, representing 13 genera and 32 species (sensu Abreu-Jr et al. 2020b). As showed above discrepant topologies were recovered upon dataset and inference method. To simplify the presentation and discussion of the UCE phylogenomic hypotheses for Sciurini, the results presented below (and further discussed) correspond to analyses performed with the data matrices considering 50% of taxa representativeness per locus and with no filter for informative sites —which corresponds to the most inclusive dataset with 3,713 UCE loci—, as these matrix performed better or equally better to the others, in term of both phylogenetic consistency and nodal support (see the above section “Topological incongruences and the influence of

data filtering strategies" for details). In cases that analyses performed with a different dataset from the abovementioned provided distinct topologies and more reliable resolution (i.e. higher nodal support) for a recalcitrant lineage, appropriated observation was made.

The tribe Sciurini was recovered as a monophyletic group, sister to the tribe Pteromyini, in all phylogenetic reconstructions with full support (Figure 7). Concatenated and coalescent-based approaches were coincident in estimating the initial diversification of Sciurini (Figure 7), showing the North American genus *Tamiasciurus* (represented by *T. douglasii* and *T. hudsonicus*) as the sister lineage to a clade containing all other genera, followed by the Eurasian genus *Sciurus* (represented by *S. vulgaris*). The subsequent cladogenetic events led to the North American genera *Hesperosciurus* (composed of *H. aberti* and *H. griseus*) and *Neosciurus* (represented by *N. carolinensis*) + *Parasciurus* (composed of *P. niger* sister to *P. nayaritensis* + *P. arizonensis*).

Regarding Neotropical taxa, all methods suggested similar relationships for the Central American genera *Echinosciurus* (represented by *E. variegatoides*) and *Syntheosciurus* (represented by *S. granatensis*) and for *Microsciurus alfari*, but different affinities for *Microsciurus* "species 1"; consequently, the genus *Microsciurus* was also not recovered as monophyletic in our inferences of UCE data (see Figure 7). Only RAxML recovered with significant support the position of both species of *Microsciurus*. In these analyses, *Microsciurus alfari* appeared as sister to all South American taxa, including *Microsciurus* "species 1" (from Colombia) —that was suggested as the first South American lineage to diverge (Figure 7A).

Among the South American genera, the placement and monophyly of *Leptosciurus* also varied upon inference method. RAxML and SVDquartets suggested this genus diverging after *Microsciurus* "species 1" and sister to all other South American genera; in the analyses with the 50% matrix week support values were obtained for the position of *Leptosciurus* (Figures 7A and 7C), but moderate values were recovered based on 60% matrices. In ASTRAL-III analyses *Leptosciurus* was estimated as sister to a clade formed by *Microsciurus* "species 1" plus all South American genera, however, the position of *Microsciurus* "species 1", as already mentioned, did not received significant support in ASTRAL-III analyses (Figure 7B). The relationships among *Leptosciurus* species were similarly suggested by RAxML and SVDquartets, although, the placement of *Leptosciurus pucheranii* as sister to *L. mimulus* was

recovered by SVDquartets only with the 70% matrix. Most alternative interspecific relationships recovered by ASTRAL-III were not statistically supported (Figure 7).

The genus *Simosciurus* (represented by *S. nebouxii*) was recovered as sister to a clade composed of *Guerlinguetus*, *Hadrosciurus*, and "*Microsciurus*" in all analyses (Figure 7). The relationships among these three genera, although, were not identical across all methods. In the concatenated approach "*Microsciurus*" appeared as sister to a clade composed of *Hadrosciurus* and *Guerlinguetus*, while in both coalescent analyses the position of "*Microsciurus*" switch with *Guerlinguetus* (Figure 7). Interspecific relationships within "*Microsciurus*" were identical in all methods. In the genus *Hadrosciurus*, only *H. spadiceus* changed its position upon inference method: RAxML recovered this species as sister to a clade composed of *Hadrosciurus* "species 3", *H. igniventris*, and *H. pyrrhinus*, whereas both coalescent methods suggested *H. spadiceus* as sister to *H. ignitus*, but with no significant support (Figure 7). Within *Guerlinguetus*, RAxML and SVDquartets recovered identical interspecific relationships and with strong support values (Figures 7A and 7C), the alternative hypothesis provided by ASTRAL-III received weaker support (Figures 7B).

Most current recognized species of tree squirrels —for which we included multiple samples—, were recovered as monophyletic entities in our phylogenetic inferences based on UCE data , with only two exceptions: (1) *Leptosciurus isthmius*, in which one individual (USNM554228 from Colombia) appeared as sister to (*L. similis* (*L. otinus* (*L. isthmus*, *L boquetensis*))) and another (USNM292133 from Panama) as sister to *L. boquetensis* in RAxML and SVDquartets analyses or as sister to *L. similis* in ASTRAL-III analyses (Figure 7; see also Figure 8 for RAxML topology, and Supplementary Figures S2 and S3 for ASTRAL-III and SVDquartets trees, respectively); and (2) *Guerlinguetus aestuans* "c" —a putative unnamed species recognized in Abreu-Jr et al. (2020b)— represented here by 25 specimens that composed two distinct and well supported clades in RAxML (Figure 8) and SVDquartets (Supplementary Figure S3) analyses, or comprised three clades (one of those not statically supported) in ASTRAL-III analyses (Supplementary Figure S2).

Some species of Neotropical tree squirrels are composed of divergent lineages that consistently appeared geographic structured in our inferences. For example, *Syntheosciurus granatensis* includes one clade composed of specimens from Ecuador and another clustering specimens from Panama (Figure 8; see also Supplementary Figures S2 and S3). "*Microsciurus*" *flaviventer* is strongly supported in all inferences as including four lineages;

the branch lengths leading to these lineages is similar or even longer, for instance, to the branch that defined "*Microsciurus*" *sabanillae* in RAxML analyses (Figures 8). *Hadrosciurus ignitus* in RAxML (Figure 8) and SVDquartets (Supplementary Figure S3) analyses exhibit two well-supported lineages: one including four specimens from Brazil and Peru; and the second, composed of seven specimens from Bolivia.

4.4 Discussion

4.4.1 The role of historical samples in the Neotropical squirrels UCE phylogeny

In a previous contribution, Abreu-Jr et al. (2020b) showed the pivotal role of historical samples in the phylogenetic inferences of the tribe Sciurini, based on mitochondrial genome data. About a third of the samples used in the analyses of Abreu-Jr et al. (2020b) were historical, and those samples allowed them to clarify several long-standing taxonomic issues of the tribe. Here, we also attempt to sequence several historical samples, but we were able to include in the phylogenetic inferences about only 12% of the dry museum specimens sampled, which represent less than 7% of the total samples in our analyses. Even so, the sampling of historical specimens allowed us to include in the UCE phylogenetic inferences 10 nominal taxa of tree squirrels (about a third of Sciurini species analyzed in this study), with no modern samples available. Additionally, we were able to include in our phylogenetic analyses an important historical specimen of *Sciurillus pusillus* from French Guiana, which represents our closest sample to the type locality of this species.

Sample age did not affect the recovery of mitochondrial genome data for historical samples of tree squirrels in the study of Abreu-Jr et al. (2020b). Unlikely, our enrichment of UCE data for those same samples was negatively affected by the year in which the specimen was collected (Figure 1). This result corroborates previous studies showing that, as opposite of mitogenomes, the number of UCE loci recovered and also the average length of UCE loci decreases with specimen age (Heintzman et al. 2014; McCormack et al. 2016; McDonough et al. 2018). The inferior success in obtaining nuclear versus mitochondrial data from historical samples might be expected considering the lower number of copies of the nuclear genome in a regular mammal cell compared to the number of mitochondrial genome copies. Also, the decrease in the mean length of UCE loci is likely a result of extensive DNA fragmentation, which is accelerated with age if the DNA is not properly stored (McDonough et al. 2018). McCormack et al. (2016) suggested that extra sequencing could help to improve the number

of UCE loci recovered for historical samples, but not average locus length, which is a critical issue particularly for inferences performed with coalescent methods (see more details in the next section).

A suggestion that natural history museums should cut off a fragment of unique specimens —representing taxa with no fresh tissue available— for cryogenic preservation has been raised in the recent literature (see McCormack et al. 2016). Despite the potential damage to a voucher material, this procedure could be helpful to deaccelerate the DNA degradation, particularly for specimens collected in the near past and that were not submitted to severe chemical treatments (e.g. with arsenic). As soon as this action is taken, higher would be the likelihood of obtaining good-quality nuclear DNA for irreplaceable historical samples.

4.4.2 Impact of UCE data filtering and discordance among inference methods

Filtering approaches applied to genome-wide datasets is a trending topic in current phylogenetic studies, as the recovery of conflicting evolutionary histories —upon taxa and/or data sampling, and inference method— has become very common, especially for recent radiations subjected to incomplete lineage sorting, early lineage isolation or introgressive hybridization (Hosner et al. 2016; Blom et al. 2017; Platt et al. 2018). There is no consensus in the recent literature on what filtering strategy could favor a particular method of inference, as the biological proprieties of the loci being analyzed play a determinant role. Therefore, the best practice to estimate the phylogeny of recalcitrant lineages should include examining the influence of a variety of filtering parameters and to use conceptually distinct inference methods (Platt et al. 2018; McLean et al. 2019).

We applied two locus-based approaches for data filtering, i.e. matrix completeness and percentage of informative sites, and both strategies directly had an impact on the number of UCE loci in our datasets. We created matrices varying the percentages of sample representativeness per loci (e.g. 70, 60 and 50%), expecting that the inclusion of a greater number of loci, even with the consequent reduction on the number of samples per locus, would positively affect the robustness of our inferences, based on previous studies (e.g. Hosner et al. 2016; Streicher et al. 2016; Platt et al. 2018). Our results confirmed that datasets containing more loci (60% and 50% matrices), even with more missing data as well, performed better than the stricter dataset (70% matrix), showing higher medians and lower

variation on the nodal support values (Figure 4). Increasing the number of loci in UCE datasets has been shown to be effective to maximize the statistical support and also to improve topological consistency (Platt et al. 2018; McLean et al. 2019). Our results on smaller RF distances among topologies obtained from more inclusive datasets also corroborate this last assumption (see Figure 2).

Regarding the second filtering strategy, the initial 50, 60, and 70% matrices were curated to include only UCE loci with a minimum of 5% of informative sites. The 50% matrix from 3,713 UCE loci ended up including 1,838 loci, the 60% matrix reduced its contents from 3,276 loci to 1,630 loci, and the 70% matrix remained with only 484 loci. Comparisons of unfiltered with filtered datasets based on information content revealed that this approach did not effectively affected the largest datasets (50 and 60% matrices), but greater impacted the 70% matrix, as for the latter even more dissimilar topologies were recovered upon this filter (Figures 2 and 3). Therefore, filtering the datasets by information content did not increased the topological cohesion in our study. This is in contrast to other studies in which filtering loci by information content has helped to improved topological concordance among inference methods (Hosner et al. 2016; Manthey et al. 2016; McLean et al. 2019). Moreover, the overall nodal support of the inferences also did not vary conspicuously regarding the information content (Figure 4). This result was consistent among all inference methods. Since coalescent methods are sensitive to imprecise gene trees —for example, resulted from very conserved locus with low rates of informative sites (Hahn and Nakhleh 2016; Mirarab et al. 2016; Blom et al. 2017)— we expected that ASTRAL-III and SVDquartets analyses would benefit from datasets with higher rates of informative sites. The disadvantage of dropping so many loci (with less than 5% of informative sites) was perhaps greater than the advantage of summarizing fewer but supposedly more precise gene trees.

RAxML and SVDquartets analyses inferred more similar topologies, while ASTRAL-III suggested the most unique trees (Figure 3). ASTRAL-III, as a two-step coalescent method that uses a collection of gene trees to summarize a species tree (Zhang et al. 2018), is likely disfavored by the limited length of some UCE loci recovered. According to Chou et al. (2015), gene trees estimation on short alignments can lead to high estimation error. This could be more pronounced considering that the core regions of the UCEs are frequently more successfully sequenced and they are much less variable than the flanking regions (Faircloth et al. 2012); therefore, longer sequences including longer portions of the flanking regions

generally produce more resolved gene trees (Platt et al. 2018). The mean length of our UCE loci was 462 bp. Previous studies, analyzing UCEs with similar mean length or even loci with longer average length, also recovered inconsistent and heterogeneous results from ASTRAL inferences (e.g. Bryson et al. 2016; Baca et al. 2017; Streicher and Wiens 2017), corroborating that UCE loci are relatively short markers —in general with small percentages of informative characters— making it difficult to produce well-resolved gene trees (Platt et al. 2018).

The other coalescent-based method, SVDquartets, differs from ASTRAL-III on its conception, as is takes multi-locus unlinked single-site data to infer quartet trees for subsets of four terminals, and then combines the quartet trees onto a species tree (Chifman and Kubatko 2014). Since this site-based summary method does not require to represent each gene by a single tree, it is less sensitive to individual gene tree estimation error than locus-based coalescent methods, as ASTRAL-III (Chou et al. 2015). This may help to explain why SVDquartets resulted in topologies largely diverging from ASTRAL-III. Disagreement between coalescent methods is frequently reported and for inferences of recalcitrant clades it can be persistent (Chou et al. 2015; Linkem et al. 2016; Streicher et al. 2016; McLean et al. 2019).

The last inference method, RAxML, uses a concatenated alignment of all loci to estimate a species tree based on the ML algorithm (Stamatakis 2014). This method outperforms the coalescent analyzes in our study, based on the higher medians of nodal support for all datasets and on the lower RF distance among the inferred trees. The advantages of RAxML are also related to the possibly limited number of informative characters per locus and consequently reduced phylogenetic signal for gene tree estimation (Platt et al. 2018). Concatenation methods recurrently provide more accurate and cohesive topologies over coalescent methods when analyzing large data matrices that prioritize overall character sampling over maximizing taxon-sampling per locus (Streicher et al. 2016; Platt et al. 2018).

4.4.3 UCE phylogenomic hypothesis for the relationships among subfamilies of Sciuridae

This is the first phylogenomic study investigating the relationships among the five currently recognized subfamilies of Sciuridae (sensu Thorington et al. 2012). Previous studies, based on a few mitochondrial and nuclear genes (e.g. Mercer and Roth 2003; Steppan et al. 2004) or on supermatrices with up to 31 genes (e.g. Fabre et al. 2012; Zelditch et al. 2015; Upham

et al. 2019), have estimated conflicting relationships for those subfamilies, especially regarding the position of the basal lineages represented by the Neotropical pigmy squirrel (Sciurillinae) and the Asian giant tree squirrels (Ratufinae). Mercer and Roth (2003) recovered Sciurillinae as the first lineage to diverge within Sciuridae; in Fabre et al. (2012) Sciurillinae diverged after Ratufinae; and in some other inferences Sciurillinae and Ratufinae appear as sister subfamilies (Steppan et al. 2004; Zelditch et al. 2015; Upham et al. 2019).

Our analyses based on thousands of UCE loci confidentially inferred a similar hypothesis as the one previously suggested by Fabre et al. (2012) for the initial diversification of Sciuridae. All of our inferences with significant nodal support from both concatenation (RAxML) and coalescence (SVDquartets) estimated the Asian giant squirrels as the sister group of all extant squirrels (see Figure 5). Alternative relationships were suggested mostly by ASTRAL-III, however, those inferences were not statistically supported. The relationships among the remaining subfamilies were identically inferred and with strong nodal support in the 18 phylogenetic analyses performed in this study. UCE datasets suggest Sciurinae (including Sciurini and Pteromyini) as the sister group of a clade composed of Callosciurini (southern Asian tree squirrels) + Xerini (African and Holarctic ground squirrels and African tree squirrels) (Figure 5). These phylogenetic affinities are the most recurrently inferred for these radiations (Steppan et al. 2004; Fabre et al. 2012; Upham et al. 2019).

4.4.4 UCE phylogenomic hypotheses for Sciurini and the mito-nuclear discordance

Molecular systematics of the tribe Sciurini have been subject of many phylogenetic studies, but most of those with emphasis on geographic-restricted taxa, such as Eurasian (Oshida and Masuda 2000; Oshida et al. 2009; Aghbolaghi et al. 2019, 2020), North American (Hope et al. 2016), or Mesoamerican (Villalobos and Cervantes-Reza 2007; Villalobos and Gutierrez-Espeleta 2014) species. Some other studies included a broader geographic coverage of tree squirrels (Fabre et al. 2012; Pečnerová and Martínková 2012; Pečnerová et al. 2015; Zelditch et al. 2015), but only in Abreu-Jr et al. (2020b) the South American forms were widely represented. Phylogenetic inferences of Abreu-Jr et al. (2020b) consistently recovered Sciurini organized in 12 well-supported major clades. These author tentatively proposed the use of 14 generic names, employed by previous authors (e.g. Allen 1915; Moore 1959; Vivo and Carmignotto 2015), to represent the genus-level diversity of the group. The species-level diversity was estimated to 46 in the analyses of Abreu-Jr et al. (2020b), an increase of three

species from the last taxonomic arrangements (Thorington et al. 2012; Vivo and Carmignotto 2015).

In our analyzes of UCE data we were not able to include all nominal taxa sampled in the mitogenome analyses of Abreu-Jr et al. (2020b). Our taxon sampling comprised 32 putative species, representing 13 of the 14 currently valid genera of Sciurini. Still, a broader taxonomic representativeness compared to any other previous study including nuclear markers (e.g. Fabre et al. 2012; Pečnerová et al. 2015; Zelditch et al. 2015). We observed two distinct patterns on the UCE analyses performed with three conceptual distinct inference methods. For the older and less diverse (monotypic or with few species) lineages —including North American (*Tamiasciurus*, *Hesperosciurus*, *Neosciurus*, and *Parasciurus*) and Eurasian (*Sciurus*) genera— that have been estimated to diversify between 14 and 7 Mya (Abreu-Jr et al. 2020a), and also for two Central American genera (*Echinosciurus* and *Syntheosciurus*), RAxML, ASTRAL-III and SVDquartets, recovered identical relationships with no ambiguities and full support (Figure 7). Furthermore, this portion of the tree squirrel phylogeny is not only consistent among UCE inferences, but it also largely corroborates the mitogenomic hypothesis. Except for *Echinosciurus* and *Syntheosciurus*, that in the analyses with mitogenomic data appeared as sister-groups (see more details below), the relationships among the remaining above-mentioned genera are identical in the inferences based on mitochondrial and nuclear data (see Abreu-Jr et al. 2020b).

Regarding the remaining Central and South American taxa —which experienced a more recent and an extremely rapid diversification after the South American invasion around 6 Mya (Abreu-Jr et al. 2020a)— numerous conflicting relationships were detected at both genus- and species-level. *Microsciurus* (which in the concept of Abreu-Jr et al. (2020b) includes one Central American [*M. alfari*] and one South American [*Microsciurus* "species 1"] species) was the only genus of Sciurini consistently not recovered as monophyletic in our analyzes of UCE data. Moreover, these two species were also recovered with distinct phylogenetic affinities upon inference method (see Figure 7). In the mitogenome analyses, *Microsciurus* was found as sister to *Echinosciurus* + *Syntheosciurus*, with the three genera composing a large predominantly Central American clade (Abreu-Jr et al. 2020b). This is perhaps the most disputing point between the UCE and the mitogenome phylogenies, and also the most poorly represented in term of taxonomic coverage in the UCE analyses, as one species of *Syntheosciurus* and four species of *Echinosciurus* were missing in the last.

Leptosciurus (at least in RAxML and SVDquartets [excepted for the position of *L. pucheranii*] analyses) and *Simosciurus* (in all inferences) were recovered with identical relationships as in the mitogenome inferences (Abreu-Jr et al. 2020b). Within *Leptosciurus* the interspecific relationships suggested by the highest supported hypothesis (from RAxML, which are similar to SVDquartets) differ from the mitogenomic hypothesis (Abreu-Jr et al. 2020b) by not recovering *L. pucheranii* and *L. mimulus* as sister taxa (only ASTRAL-III did; Figure 7B), and suggesting *L. isthmius* as non-monophyletic. As a matter of fact, the two specimens of *L. isthmius* represented in our UCE datasets are the most divergent in the mitochondrial dataset, which indicates that further analyses (e.g. with phenotypic data) are needed to evaluate the species limits within this taxon.

The relationships among the last South American genera, *Guerlinguetus*, *Hadrosciurus*, and "*Microsciurus*", were identically inferred by the coalescent methods (see Figure 7B and 7C), which are in full agreement with the mitogenomic hypothesis (Abreu-Jr et al. 2020b). Within *Guerlinguetus*, RAxML and SVDquartets corroborate the interspecific relationships previously inferred with mitochondrial data, and for *Hadrosciurus,* species affinities suggested by both coalescent methods are identical to the mitogenome inferences (Abreu-Jr et al. 2020b). Within "*Microscirus*" no disagreement was found among UCE inferences and neither between nuclear and mitochondrial data.

Most species-level lineages recognized by the mitogenome data were also confirmed as so in our analyses with UCE data (Figure 8; Supplementary Figures S2 and S3). Besides *Leptosciurus isthmius*, the only other case of non-monophyly among all species of tree squirrels included in this study is *Guerlinguetus aestuans* "c". This taxa, according to Abreu-Jr et al. (2020b), represents an unrecognized species composed of specimens from the Atlantic Forest, in northeastern Brazil, and from a broad area in the Amazon basin, including mostly the north bank of the Amazon river in Brazil (Pará and Amazonas states), a small area in the southern bank of the Amazon river (in Pará), besides the French Guiana and Venezuela (Figure 9A). The very same specimens were recovered as two well-supported and not sister clades in our analyses of UCE data (Figures 8 and 9B). Discrepant species limits for those individuals in this region might attest for a putative shortfall on the mitogenome dataset. Although conflicting species recognition could simply reflect error in phylogenetic estimation (Huang et al. 2010), it also may be driven by genuine discordance between nuclear and mitochondrial marker types that, by extension, could be associated with several biological

factors (Platt et al. 2018). In the case of *Guerlinguetus aestuans* "c", introgressive hybridization in the mitochondrial DNA is a candidate explanation for the recognition of a single clade including specimens from distinct nuclear lineages.

A recent study, using distinct sources of genomic data (e.g. mtDNA, exons, introns, and UCEs) to evaluate species limits on a complex of cryptic entities, empirically demonstrated that gene flow and introgression may produce phylogenetic structures with high levels of divergence among mitochondrial clades that reassemble different species (Chan et al. 2020). Those authors also raised the question whether high levels of cryptic diversity —which is frequently observed in many Neotropical rodent genera, e.g. *Akodon* (Gonçalves et al. 2005), *Hylaeamys* (Brennand et al. 2013), *Proechimys* (Da Costa et al. 2016)— are actually cryptic species or highly admixed and structured lineages. D'Elía et al. (2019) provided a review on how hybridization and introgression have been addressed on rodent systematics; clearly, there is a poorly understanding on the presence and magnitude of these phenomena in the group, especially regarding Neotropical taxa. A very concerning implication, highlighted by D'Elía et al. (2019), is that the disregard of introgression in taxonomic studies may be biasing species delimitation due to the widespread use of mtDNA data and the more restricted use of nuclear markers. High levels of introgression have been found in some lineages of mammals from South America, for example in felids (Trigo et al. 2013) and canids (Tchaicka et al. 2016). Further investigation is needed to examine the presence of introgression and to test alternative hypotheses on the causes of the discrepant results found between genetic markers in the genus *Guerlinguetus*.

The remaining mito-nuclear discordances on the Sciurini phylogeny may attest for additional phenomena acting on the tree squirrel evolution. For example, causes of conflicting species trees include gene duplication, horizontal transfer, and incomplete lineage sorting. With regards to the latter, it has been shown to be a pivotal phenomenon affecting the phylogenetic estimation of some branches of the mammal Tree of Live, as the rapid radiations (Platt et al. 2018; McLean et al. 2019). Incomplete lineage sorting is influenced by the speciation and fixation rates, i.e. fast speciation rates and/or relatively long time to fixation will increase the probability of occurrence of incomplete lineage sorting (Platt et al. 2018). The Neotropical radiation of tree squirrels, for which conflicting evolutionary histories were inferred by UCE and mitogenome datasets, underwent an explosive diversification with remarkable high species accumulation curves in the past 6 Mya

(Abreu-Jr et al. 2020a). With that in mind, it is very likely that at least the Neotropical radiation was subjected to intense incomplete lineage sorting. This would not just help to explain the mito-nuclear discordances, but also the extremely variable UCE phylogenetic results upon dataset filtering, as the phylogenetic signal in several loci could be low due to the lack of synapomorphies caused by incomplete lineage sorting (see McLean et al. 2019). Lastly, the lower speciation rates along the initial diversification of Sciurini (Abreu-Jr et al. 2020a), would make it less susceptible to incomplete lineage sorting, and it might reflects the robustness of our inferences for the deeper branches of the tree squirrels tree.

4.5 Conclusion

The recovery of UCE loci and the mean length per UCE locus were negatively impacted by the sample age in our study. A regular mammal cell contains from a hundred to a thousand more copies of the mitochondrial DNA than to the nuclear DNA. This makes the sequencing of nuclear markers much more sensitive and prone to be affected by the sample quality. Several historical specimens sampled in the present study were dropped from the phylogenetic inferences due to insufficient amount of UCE data sequenced. Yet, our analyzes were performed with a robust taxonomic and geographic coverage. We observed that data matrices including larger numbers of loci (matrices with 50 and 60% of taxa representativeness per loci) provided more similar topologies and higher values of nodal support. We also detected that filtering for information content (including only loci with a minimum of 5% of informative sites in the datasets) did not increased the topological consistency nor improved the nodal support for any inference method. In general, RAxML outperforms the coalescent methods, providing the topologies less dissimilar among each other and with the highest medians of nodal support. Our analyzes estimated with strong support the relationship among the five currently recognized subfamilies of Sciuridae, and also provided consistent and well-supported results for the relationships among the deepest branches of Sciurini. For the Neotropical radiation, conflicting relationships at both genu- and species-level were estimated. Some inconsistences were also recovered with regards to the mitogenomic hypothesis previously published (Abreu-Jr et al. 2020b). Introgressive hybridization and incomplete lineage sorting are amongst the most plausible phenomena acting on the tree squirrels evolution and which could help to explain the mito-nuclear discordances observed for the rapid Neotropical radiation of tree squirrels. Finally, even with

the tremendous amounts of genome-scale molecular data available to estimate phylogenies, for some recalcitrant clades —which seems to be the case of the Neotropical tree squirrels— we might never reach a phylogenetic consensus.

Acknowledgments

We would like to express our gratitude to the following curators and collection support staff, who loaned us preserved tissue samples, allowed us to analyze and obtain destructive samples from valuable historical specimens under their care, or provided essential information regarding vouchers: R.S. Voss, E. Hoeger, and N. Duncan (AMNH); I. Moya (CBF); M.R. Alvarez (CMARF); B.D. Patterson and A.W. Ferguson (FMNH); J. Valsecchi and I. Junqueira (IDSM); C.R. Silva and A.F. Sobrinho (IEPA); R.M. Timm and M. Eifler (KU); D.L. Dittmann and R.T. Brumfield (LSU-MZM); E. Pasa (MCN-FZB); C.G. Costa (MCN-M); A.U. Christoff, F.B. Peters and A.M. Gandini (MCNU); J.A. Oliveira and M. Weksler (MN); J. Silva Junior (MPEG); M. Campbell, J. Cook and J. Dunnum (MSB); M.T. Rodrigues and B.M.S. Batista (MTR); V. Pacheco (MUSM); C. Conroy, J.L. Patton, and F.J.M. Pascal (MVZ); M. Vivo, L.F. Silveira and J.G. Barros (MZUSP); J.K. Braun and B.S. Coyner (OMNH); H. Garner, R.D. Bradley, and C. Phillips (TTU); L.P. Costa, Y.L.R. Leite and M.P. Nascimento (UFES-CTA); R. Rossi and T. Semedo (UFMT); A.C. Mendes Oliveira (UFPA); J. Cherem, M. Graipel and E.C. Grisard (UFSC); D.P. Lunde, S.C. Peurach, N.R. Edmison, I. Rochon, M. Krol, J. Jacobs, C. Ludwig, C. Huddleston, D. DiMichele, M. Braun, L.H. Emmons, and A.L. Gardner (USNM); F. Catzeflis (ISEM); A. Ravetta, G.T. Garbino and L.P. Godoy provided tissue samples or allowed us to sample recently collected or uncatalogued specimens. We thank L. Parker, M. Venkatraman, S. Castañeda and N. McInerney, who provided crucial help during lab work. We are also thankful to J.R. Prado and J. Dallapicola, who have provided analytical support and have shared scripts and insights on computational procedures. We are grateful to R. Dikow, who have always been supportive with any matter regarding the Smithsonian HPC. We are deeply indebted to G. Libardi, who generously made the squirrels drawings depicted on Figure 5.

This work had financial support from the Conselho Nacional de Desenvolvimento Científico e Tecnológico through doctoral fellowships to EFA (147145/2016-3, 203692/2017-9, and 165553/2017-0) and through a productivity scholarship to ARP (304156/2019-1); from the American Society of Mammalogists through the Latin American Student Field Research

Award to EFA, providing support for a field expedition to the Rio Purus, BR; from the American Museum of Natural History through a Collection Study Grant, allowing EFA to collect morphological data from tree squirrels housed at the AMNH; from the Smithsonian Institution through postdoctoral fellowships to SEP and MTNT (Smithsonian Women's Committee), and through research funds provided to SEP, MTNT, DEW and JEM, covering costs associated to laboratory procedures and DNA sequencing; from the National Geographic Society through an Explorers Grant to SEP (NGS-381R-18), providing support to visit and collect samples housed at the MUSM and for a field expedition to Parque Nacional del Rio Abiseo, PE, and through Support for Women + Dependent Care to SEP (NGS 3758), allowing her to visit and collect additional data in North American museums; and from the Fundação de Amparo à Pesquisa do Estado de São Paulo to ARP (09/16009-1), providing support for several filed expeditions to the Amazon, BR.

References

Abreu-Jr E.F., Pavan S.E., Tsuchiya M.T.N., Wilson D.E., Percequillo A.R., Maldonado J.E. 2020a. Spatiotemporal Diversification of Tree Squirrels: Is the South American Invasion and Speciation Really That Recent and Fast? Front. Ecol. Evol. 8:1–14.

Abreu-Jr E.F., Pavan S.E., Tsuchiya M.T.N., Wilson D.E., Percequillo A.R., Maldonado J.E. 2020b. Museomics of tree squirrels: a dense taxon sampling of mitogenomes reveals hidden diversity, phenotypic convergence, and the need of a taxonomic overhaul. BMC Evol. Biol. 20:1–25.

Aghbolaghi M.A., Ahmadzadeh F., H. Kiabi B., Keyghobadi N. 2020. Evolutionary history of the Persian squirrel (*Sciurus anomalus*): It emerged on the Eurasian continent in the Miocene. Zool. Anz. 287:17–24.

Aghbolaghi M.A., Ahmadzadeh F., Kiabi B., Keyghobadi N. 2019. The permanent inhabitant of the oak trees: Phylogeography and genetic structure of the Persian squirrel (*Sciurus anomalus*). Biol. J. Linn. Soc. 127:197–212.

Allen J.A. 1915. Review of the south american Sciuridae. Bull. Am. Museum Nat. Hist. 34:147–309.

Baca S.M., Alexander A., Gustafson G.T., Short A.E.Z. 2017. Ultraconserved elements show utility in phylogenetic inference of Adephaga (Coleoptera) and suggest paraphyly of 'Hydradephaga.' Syst. Entomol. 42:786–795.

Blom M.P.K., Bragg J.G., Potter S., Moritz C. 2017. Accounting for uncertainty in gene tree estimation: Summary-coalescent species tree inference in a challenging radiation of Australian lizards. Syst. Biol. 66:352–366.

Bolger A.M., Lohse M., Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. 30:2114–2120.

Brennand P.G.G., Langguth A., Percequillo A.R. 2013. The genus *Hylaeamys* Weksler, Percequillo, and Voss 2006 (Rodentia: Cricetidae: Sigmodontinae) in the Brazilian Atlantic Forest: geographic variation and species definition. J. Mammal. 94:1346–1363.

Bryson R.W., Faircloth B.C., Tsai W.L.E., McCormack J.E., Klicka J. 2016. Target enrichment of thousands of ultraconserved elements sheds new light on early relationships within New World sparrows (Aves: Passerellidae). Auk. 133:451–458.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17:540–552.

Chan K.O., Hutter C.R., Wood P.L., Grismer L.L., Das I., Brown R.M. 2020. Gene flow creates a mirage of cryptic species in a Southeast Asian spotted stream frog complex. Mol. Ecol.:0–1.

Chifman J., Kubatko L. 2014. Quartet inference from SNP data under the coalescent model. Bioinformatics. 30:3317–3324.

Chou J., Gupta A., Yaduvanshi S., Davidson R., Nute M., Mirarab S., Warnow T. 2015. A comparative study of SVDquartets and other coalescent-based species tree estimation methods. BMC Genomics. 16:1–11.

Da Costa M.J.R., Do Amaral P.J.S., Pieczarka J.C., Sampaio M.I., Rossi R. V., Mendes-Oliveira A.C., Noronha R.C.R., Nagamachi C.Y. 2016. Cryptic Species in *Proechimys goeldii* (Rodentia, Echimyidae)? A Case of Molecular and Chromosomal Differentiation in Allopatric Populations. Cytogenet. Genome Res. 148:199–210.

Crawford N.G., Faircloth B.C., McCormack J.E., Brumfield R.T., Winker K., Glenn T.C. 2012. More than 1000 ultraconserved elements provide evidence that turtles are the sister group of archosaurs. Biol. Lett. 8:783–786.

D'Elía G., Fabre P.H., Lessa E.P. 2019. Rodent systematics in an age of discovery: Recent advances and prospects. J. Mammal. 100:852–871.

Dornburg A., Su Z., Townsend J.P. 2019. Optimal Rates for Phylogenetic Inference and Experimental Design in the Era of Genome-Scale Data Sets. Syst. Biol. 68:145–156.

Edwards S., Bensch S. 2009. Looking forwards or looking backwards in avian phylogeography? A comment on Zink and Barrowclough 2008. Mol. Ecol. 18:2930–2933.

Edwards S. V., Kingan S.B., Calkins J.D., Balakrishnan C.N., Bryan Jennings W., Swanson W.J., Sorenson M.D. 2005. Speciation in birds: Genes, geography, and sexual selection. Proc. Natl. Acad. Sci. 102:6550–6557.

Fabre P.-H., Hautier L., Dimitrov D., Douzery E.J. 2012. A glimpse on the pattern of rodent diversication: a phylogenetic approach. BMC Evol. Biol. 12:88.

Faircloth B.C. 2013. Illumiprocessor: a trimmomatic wrapper for parallel adapter and quality trimming.

Faircloth B.C. 2016. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics. 32:786–788.

Faircloth B.C., McCormack J.E., Crawford N.G., Harvey M.G., Brumfield R.T., Glenn T.C. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Syst. Biol. 61:717–726.

Faircloth B.C., Sorenson L., Santini F., Alfaro M.E. 2013. A Phylogenomic Perspective on the Radiation of Ray-Finned Fishes Based upon Targeted Sequencing of Ultraconserved Elements (UCEs). PLoS One. 8:1–7.

Gonçalves P.R., Myers P., Vilela J.F., Oliveira J.A. 2005. Systematics of species of the genus *Akodon* (Rodentia: Sigmodontinae) in southeastern Brazil and implications for the biogeography of the Campos de Altitude. Misc. Publ. Museum Zool. Univ. Michigan. 197:1–24.

Grabherr M.G., Haas B.J., Yassour M., Levin J.Z., Thompson D.A., Amit I., Adiconis X., Fan L., Raychowdhury R., Zeng Q., Chen Z., Mauceli E., Hacohen N., Gnirke A., Rhind N., di Palma F., Birren B.W., Nusbaum C., Lindblad-Toh K., Friedman N., Regev A. 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat. Biotechnol. 29:644–652.

Hahn M.W., Nakhleh L. 2016. Irrational exuberance for resolved species trees. Evolution (N. Y). 70:7–17.

Hawkins M.T.R., Leonard J.A., Helgen K.M., McDonough M.M., Rockwood L.L., Maldonado J.E. 2016. Evolutionary history of endemic Sulawesi squirrels constructed from UCEs and mitogenomes sequenced from museum specimens. BMC Evol. Biol. 16:1–16.

Heintzman P.D., Elias S.A., Moore K., Paszkiewicz K., Barnes I. 2014. Characterizing DNA preservation in degraded specimens of *Amara alpina* (Carabidae: Coleoptera). Mol. Ecol. Resour. 14:606–615.

Hope A.G., Malaney J.L., Bell K.C., Salazar-Miralles F., Chavez A.S., Barber B.R., Cook J.A. 2016. Revision of widespread red squirrels (genus: *Tamiasciurus*) highlights the complexity of speciation within North American forests. Mol. Phylogenet. Evol. 100:170–182.

Hosner P.A., Faircloth B.C., Glenn T.C., Braun E.L., Kimball R.T. 2016. Avoiding missing data biases in phylogenomic inference: An empirical study in the landfowl (Aves: Galliformes). Mol. Biol. Evol. 33:1110–1125.

Huang H., He Q., Kubatko L.S., Knowles L.L. 2010. Sources of error inherent in species-tree estimation: Impact of mutational and coalescent effects on accuracy and implications for choosing among different methods. Syst. Biol. 59:573–583.

Jombart T. 2008. Adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics. 24:1403–1405.

Katoh K., Standley D.M. 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. 30:772–780.

Linkem C.W., Minin V.N., Leaché A.D. 2016. Detecting the anomaly zone in species trees and evidence for a misleading signal in higher-level skink phylogeny (Squamata: Scincidae). Syst. Biol. 65:465–477.

Manthey J.D., Campillo L.C., Burns K.J., Moyle R.G. 2016. Comparison of target-capture and restriction-site associated DNA sequencing for phylogenomics: A test in cardinalid tanagers (Aves, Genus: Piranga). Syst. Biol. 65:640–650.

McCormack J.E., Tsai W.L.E., Faircloth B.C. 2016. Sequence capture of ultraconserved elements from bird museum specimens. Mol. Ecol. Resour. 16:1189–1203.

McDonough M.M., Parker L.D., Rotzel McInerney N., Campana M.G., Maldonado J.E. 2018. Performance of commonly requested destructive museum samples for mammalian genomic studies. J. Mammal. 99:789–802.

McLean B.S., Bell K.C., Allen J.M., Helgen K.M., Cook J.A. 2019. Impacts of Inference Method and Data set Filtering on Phylogenomic Resolution in a Rapid Radiation of Ground Squirrels (Xerinae: Marmotini). Syst. Biol. 68:298–316.

Mercer J.M., Roth V.L. 2003. The Effects of Cenozoic Global Change on Squirrel Phylogeny. Science. 299:1568–1572.

Mirarab S., Bayzid M.S., Warnow T. 2016. Evaluating summarymethods for multilocus species tree estimation in the presence of incomplete lineage sorting. Syst. Biol. 65:366–380.

Moore J.C. 1959. Relationships among the living squirrels of the Sciurinae. Bull. Am. Museum Nat. Hist. 118:157–206.

Nakamura T., Yamada K.D., Tomii K., Katoh K. 2018. Parallelization of MAFFT for large-scale multiple sequence alignments. Bioinformatics. 34:2490–2492.

Oliveros C.H., Field D.J., Ksepka D.T., Barker F.K., Aleixo A., Andersen M.J., Alström P., Benz B.W., Braun E.L., Braun M.J., Bravo G.A., Brumfield R.T., Chesser R.T., Claramunt S., Cracraft J., Cuervo A.M., Derryberry E.P., Glenn T.C., Harvey M.G., Hosner P.A., Joseph L., Kimball R.T., Mack A.L., Miskelly C.M., Peterson A.T., Robbins M.B., Sheldon F.H., Silveira L.F., Smith B.T., White N.D., Moyle R.G., Faircloth B.C. 2019. Earth history and the passerine superradiation. Proc. Natl. Acad. Sci. 116:7916–7925.

Oshida T., Arslan A., Noda M. 2009. Phylogenetic relationships among the old world *Sciurus* squirrels. Folia Zool. 58:14–25.

Oshida T., Masuda R. 2000. Phylogeny and Zoogeography of Six Squirrel Species of the Genus *Sciurus* (Mammalia, Rodentia), Inferred from Cytochrome b Gene Sequences. Zoolog. Sci. 17:405–409.

Pečnerová P., Martínková N. 2012. Evolutionary history of tree squirrels (Rodentia, Sciurini) based on multilocus phylogeny reconstruction. Zool. Scr. 41:211–219.

Pečnerová P., Moravec J.C., Martínková N. 2015. A skull might lie: Modeling ancestral ranges and diet from genes and shape of tree squirrels. Syst. Biol. 64:1074–1088.

Platt R.N., Faircloth B.C., Sullivan K.A.M., Kieran T.J., Glenn T.C., Vandewege M.W., Lee T.E., Baker R.J., Stevens R.D., Ray D.A. 2018. Conflicting Evolutionary Histories of the Mitochondrial and Nuclear Genomes in New World Myotis Bats. Syst. Biol. 67:236–249.

Rohland N., Reich D. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Res. 22:939–946.

Roth V.L., Mercer J.M. 2008. Differing rates of macroevolutionary diversification in arboreal squirrels. Curr. Sci. 95:857–861.

Sarkar D. 2008. Lattice. New York, NY: Springer New York.

Schliep K.P. 2011. phangorn: Phylogenetic analysis in R. Bioinformatics. 27:592–593.

Smith B.T., Harvey M.G., Faircloth B.C., Glenn T.C., Brumfield R.T. 2014. Target capture and massively parallel sequencing of ultraconserved elements for comparative studies at shallow evolutionary time scales. Syst. Biol. 63:83–95.

Stamatakis A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30:1312–1313.

Steppan S.J., Storz B.L., Hoffmann R.S. 2004. Nuclear DNA phylogeny of the squirrels (Mammalia: Rodentia) and the evolution of arboreality from c-myc and RAG1. Mol. Phylogenet. Evol. 30:703–719.

Streicher J.W., Schulte J.A., Wiens J.J. 2016. How Should Genes and Taxa be Sampled for Phylogenomic Analyses with Missing Data? An Empirical Study in Iguanian Lizards. Syst. Biol. 65:128–145.

Streicher J.W., Wiens J.J. 2017. Phylogenomic analyses of more than 4000 nuclear loci resolve the origin of snakes among lizard families. Biol. Lett. 13.

Swofford D. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sunderland: Sinauer Associates.

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

Tan G., Muffato M., Ledergerber C., Herrero J., Goldman N., Gil M., Dessimoz C. 2015. Current methods for automated filtering of multiple sequence alignments frequently worsen single-gene phylogenetic inference. Syst. Biol. 64:778–791.

Tchaicka L., de Freitas T.R.O., Bager A., Vidal S.L., Lucherini M., Iriarte A., Novaro A., Geffen E., Garcez F.S., Johnson W.E., Wayne R.K., Eizirik E. 2016. Molecular assessment of the phylogeny and biogeography of a recently diversified endemic group of South American canids (Mammalia: Carnivora: Canidae). Genet. Mol. Biol. 39:442–451.

Thorington R.W., Koprowski J.L., Steele M.A., Whatton J. 2012. Squirrels of the World. Baltimore: The Johns Hopkins University Press.

Trigo T.C., Schneider A., De Oliveira T.G., Lehugeur L.M., Silveira L., Freitas T.R.O., Eizirik E. 2013. Molecular data reveal complex hybridization and a cryptic species of Neotropical wild cat. Curr. Biol. 23:2528–2533.

Upham N.S., Esselstyn J.A., Jetz W. 2019. Inferring the mammal tree: Species-level sets of phylogenies for questions in ecology, evolution, and conservation. PLOS Biol. 17:e3000494.

Villalobos F., Cervantes-Reza F. 2007. Phylogenetic relationships of Mesoamerican species of the genus *Sciurus* (Rodentia: Sciuridae). Zootaxa. 40:31–40.

Villalobos F., Gutierrez-Espeleta G. 2014. Mesoamerican tree squirrels evolution (Rodentia: Sciuridae): a molecular phylogenetic analysis. Rev. Biol. Trop. 62:649–657.

Vivo M., Carmignotto A.P. 2015. Family Sciuridae G. Fischer, 1817. In: Patton J.L., Pardiñas U.F.J., D'Elía G., editors. Mammals of South America, Volume 2, Rodents. Chicago and London: The University of Chicago Press. p. 1–48.

Wickham H. 2016. ggplot2. Cham: Springer International Publishing.

Zelditch M.L., Li J., Tran L.A.P., Swiderski D.L. 2015. Relationships of diversity, disparity, and their evolutionary rates in squirrels (Sciuridae). Evolution (N. Y). 69:1284–1300.

Zhang C., Rabiee M., Sayyari E., Mirarab S. 2018. ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics. 19:15–30.

Tables

Table 1. Summary characteristics of the UCE datasets analyzed in this study.

	Matrix completeness	Filter for information content	Number of UCE loci included	Total base-pairs	Overall missing data
Dataset 1	50%	No	3,713	1,676,395	45.61%
Dataset 2	60%	No	3,276	1,491,314	44.45%
Dataset 3	70%	No	1,013	465,529	40.23%
Dataset 4	50%	≥ 5% of IS	1,838	817,097	45.05%
Dataset 5	60%	≥ 5% of IS	1,630	731,957	43.99%
Dataset 6	70%	≥ 5% of IS	484	220,730	40.05%

Figures

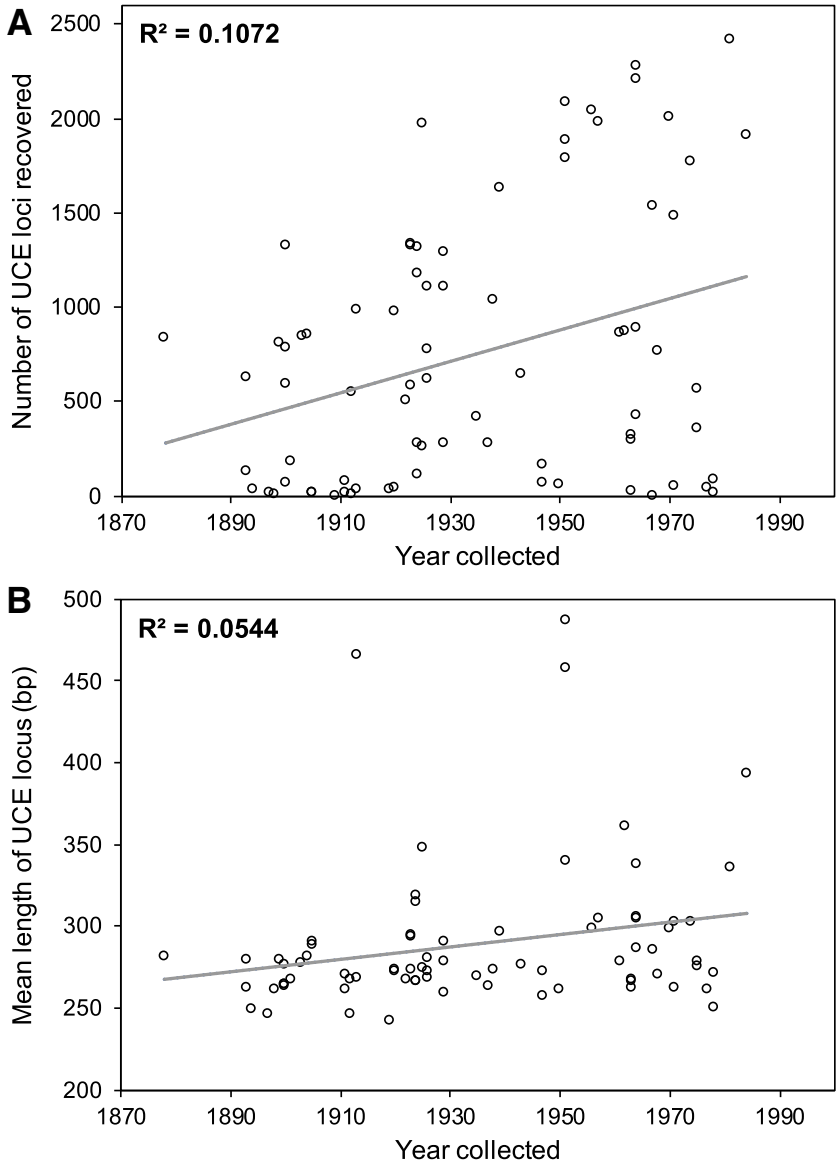


Figure 1. Relationship between historical sample age and (**A**) number of UCE loci recovered per sample, and (**B**) mean length of UCE locus per sample. Gray lines represent linear regressions based on 95 historical samples. Both analyses were statistically significant ($p < 0.05$).

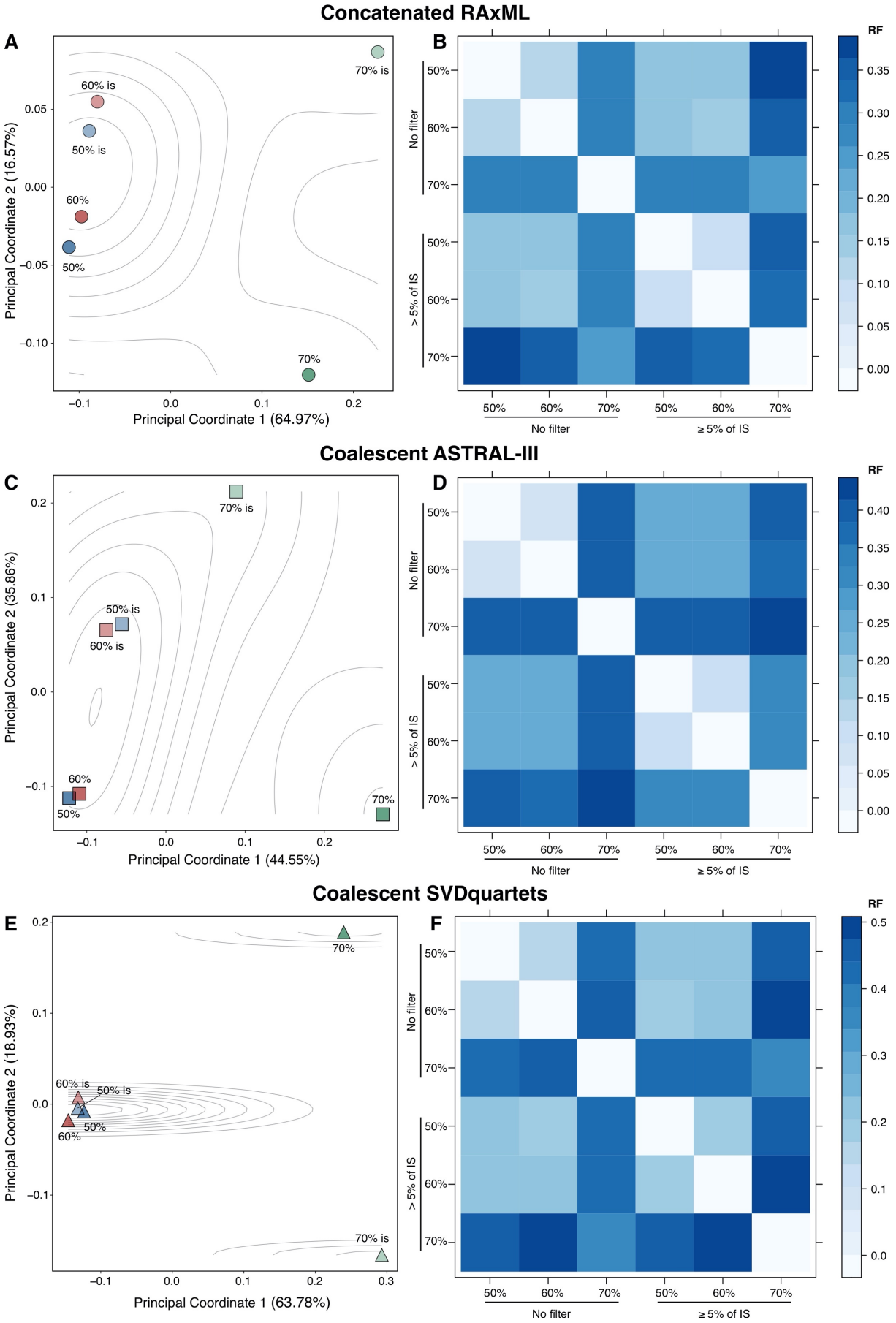


Figure 2. (See legend on next page)

Figure 2. Differences among topologies inferred for the six datasets using a concatenated method (RAxML [**A** and **B**]) and two coalescent approaches (ASTRAL-III [**C** and **D**] and SVDquartets [**E** and **F**]). (**A**, **C**, **E**) Results of Principal Coordinate Analyses (PCoA) based on a matrix of Robinson–Foulds distances among topologies. Blue markers correspond to topologies resulting from 50% data matrices, red markers to 60% data matrices, and green markers to 70% data matrices. Dark shades represent datasets not filtered for information content and light shades represent dataset including only loci with ≥ 5% informative sites. Isoclines highlight the density of trees. (**B**, **D**, **F**) Heat maps of normalized pairwise Robinson–Foulds distances among datasets. Darker shads correspond higher dissimilarity between pairs of trees. “IS” means “informative sites”.

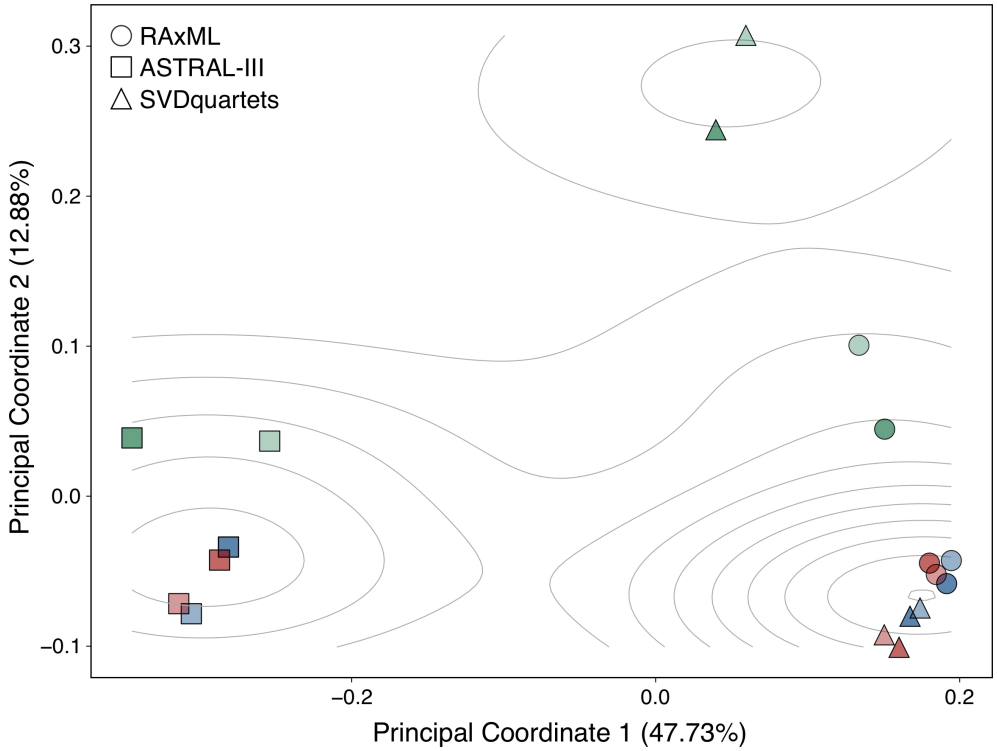


Figure 3. Scatterplot depicting results of Principal Coordinate Analyses (PCoA) based on a matrix of Robinson–Foulds distances among all trees inferred in this study. Blue markers correspond to topologies resulting from 50% data matrices, red markers to 60% data matrices, and green markers to 70% data matrices. Dark shades represent datasets not filtered for information content and light shades represent dataset including only loci with ≥ 5% informative sites. Isoclines highlight the density of trees.

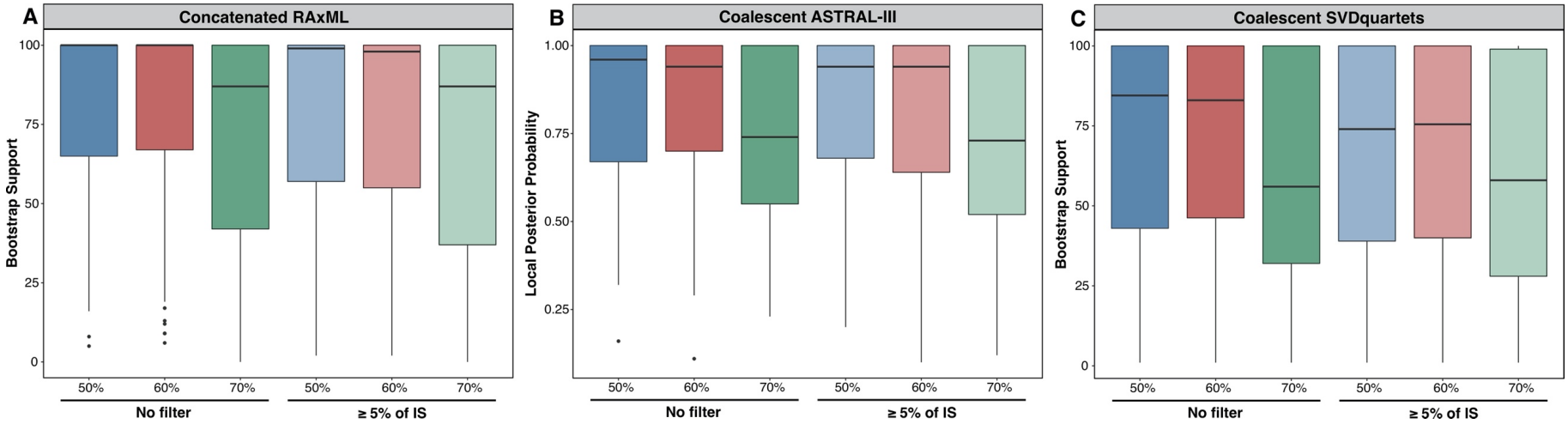


Figure 4. Boxplots showing changes in overall support values of trees inferred for each dataset using three distinct methods: RAxML (**A**), ASTRAL-III (**B**), and SVDquartets (**C**).

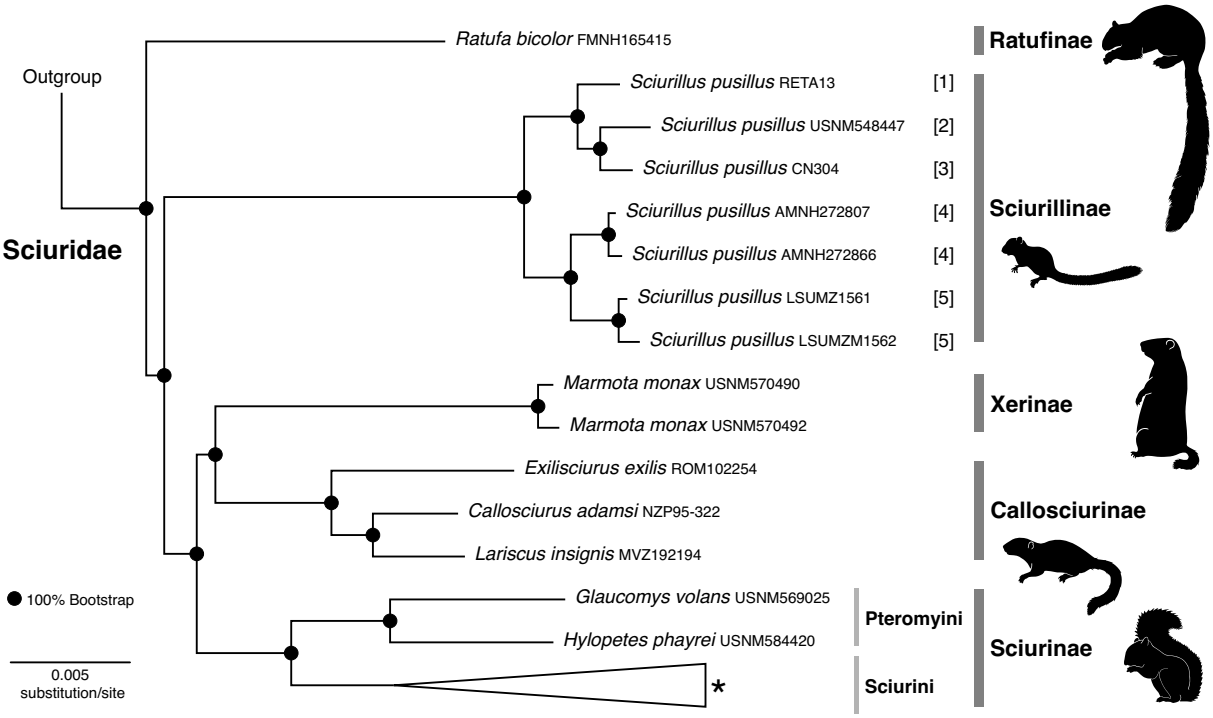


Figure 5. Phylogenetic hypothesis for the inter-subfamilies relationships of Sciuridae based on a Maximum Likelihood analysis (RAxML) performed with a concatenated matrix of 3,713 UCE loci (50% taxa representativeness per locus and no filter for information content). Samples are identified by the binomial, followed by the voucher catalog number. Museum acronyms are described in the Supplementary Catalog Data of Voucher material. Numbers after catalog numbers of *Sciurillus pusillus* specimens represent localities depicted on the map of Figure 6. *Terminals of the tribe Sciurini were collapsed. The detailed phylogenetic hypothesis for Sciurini is presented in the Figure 8.

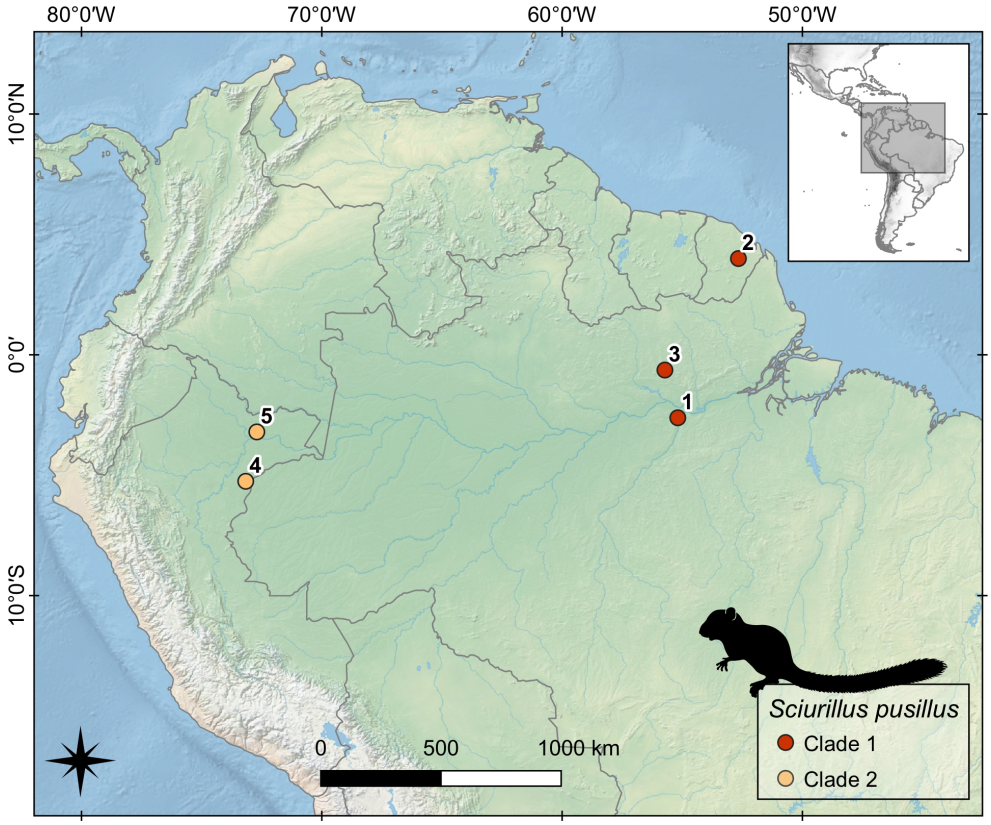


Figure 6. Map of collecting localities of samples of *Sciurillus pusillus* in the Amazon basin. Localities numbers correspond to specimens in the topology depicted on Figure 5.

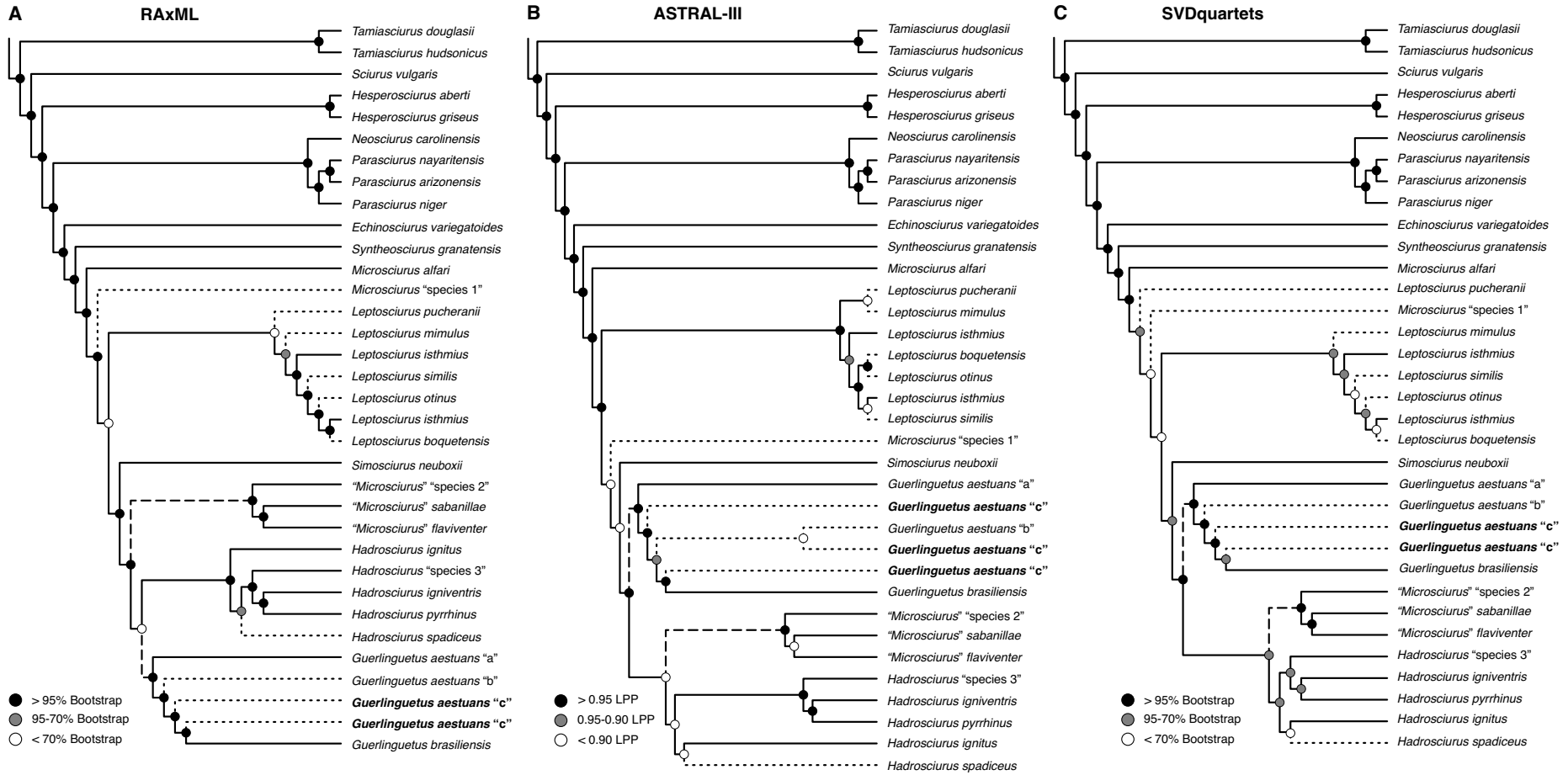


Figure 7. Phylogenetic hypotheses for the species-level relationships of the tribe Sciurini recovered by RAxML (**A**), ASTRAL-III (**B**), and SVDquartets (**C**). Analyses were performed with matrices of 3,713 UCE loci (50% taxa representativeness per locus and no filter for information content). Taxonomic identifications at both genus- and species-level follow Abreu-Jr et al. (2020b). Dotted branches correspond to conflicting species relationships and dashed branches represent distinct generic affinities recovered upon inference method.

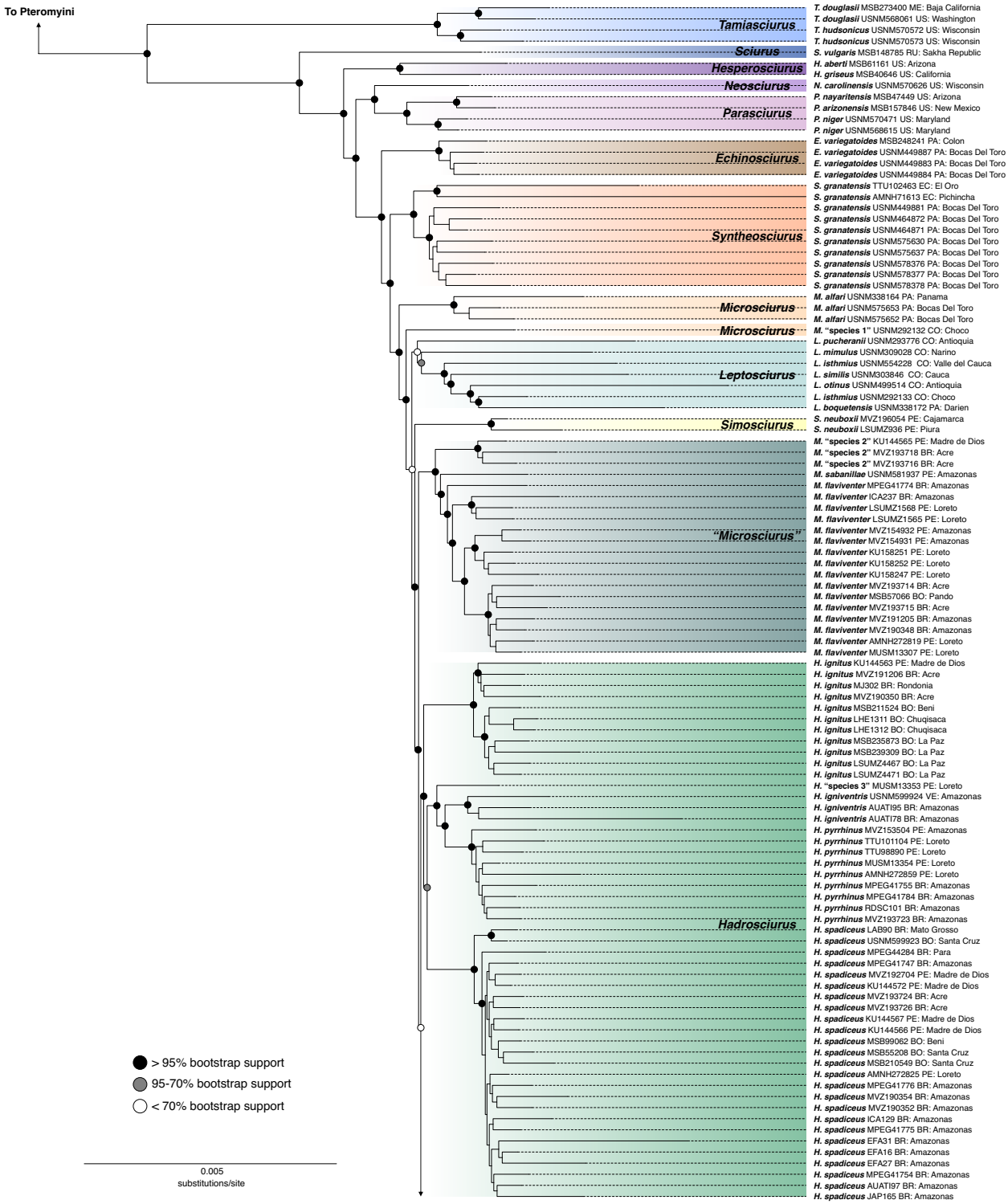


Figure 8. (**continued**) Detailed phylogenetic hypothesis of the tribe Sciurini based on a Maximum Likelihood analysis (RAxML) performed with a concatenated matrix of 3,713 UCE loci (50% taxa representativeness per locus and no filter for information content). Terminals are named with binomials (following Abreu-Jr et al. 2020b), accompanied by museum voucher numbers and geographic information (country code and state/department). Scale at the bottom represents substitutions per site.

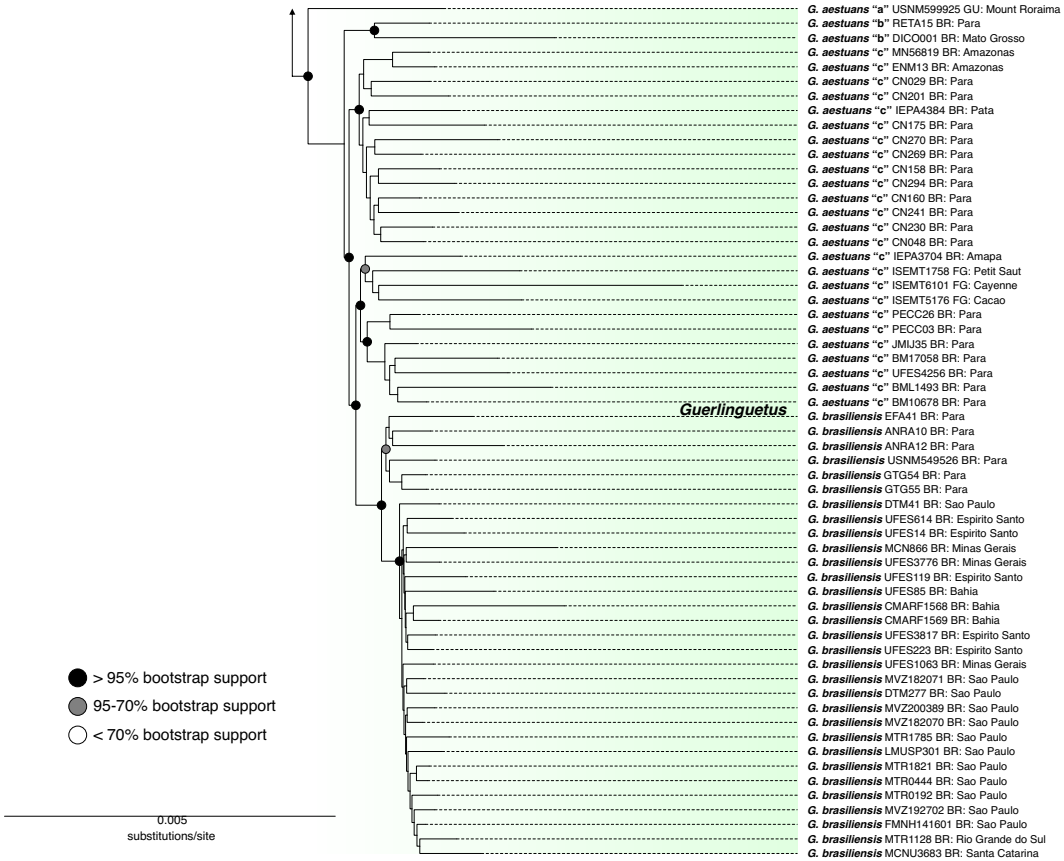


Figure 8. (**continuation**)

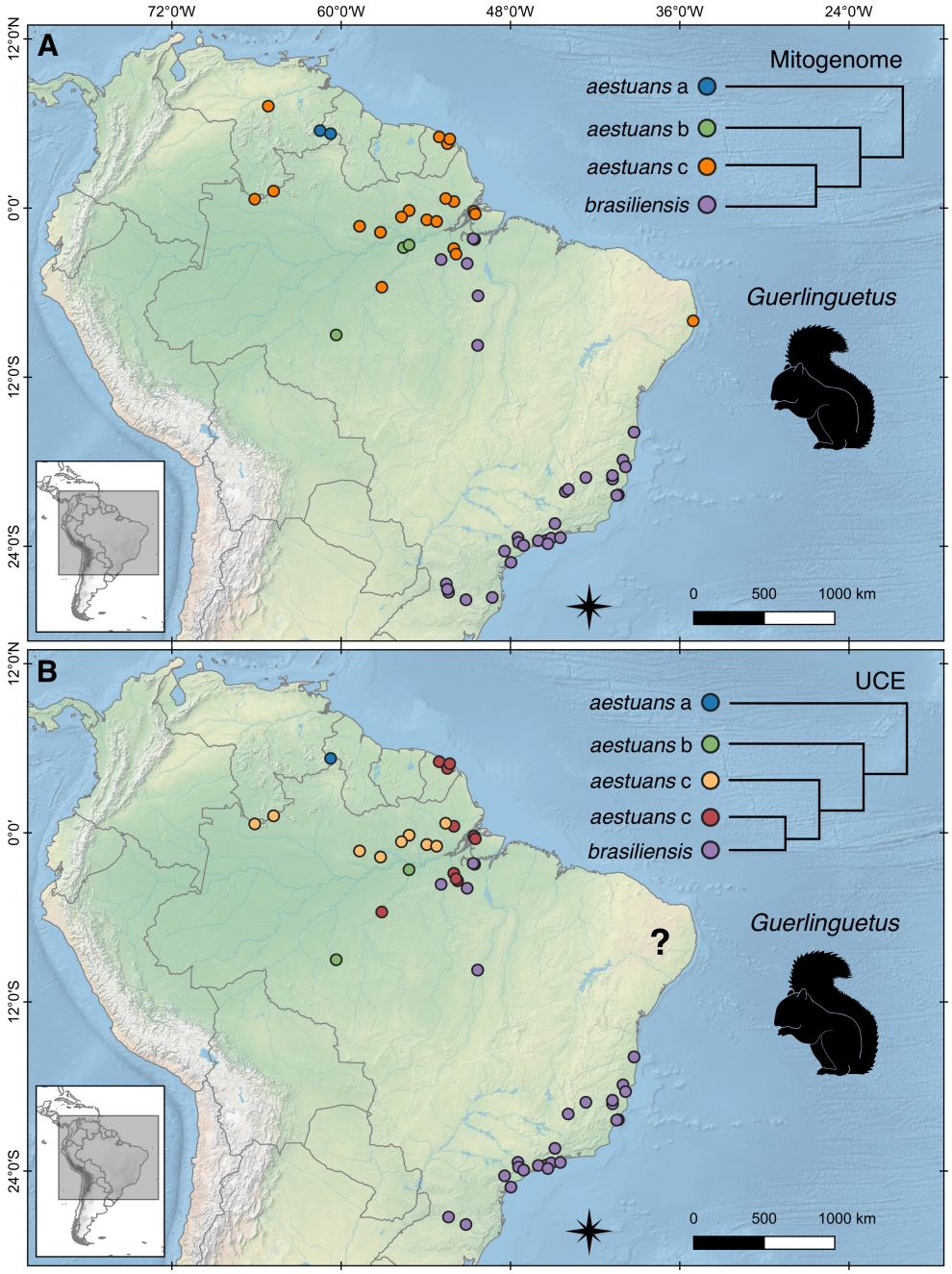


Figure 9. Species limits suggested by the mitogenome data (from Abreu-Jr et al. 2020b) and UCE data (this study) for samples assigned to the genus *Guerlinguetus*. (**A**) Map of collecting localities of *Guerlinguetus* samples included in the analyses of Abreu-Jr et al. (2020b); samples are identified according to the species limits suggested by the mitogenomic hypothesis, which is simplified at the top-right corner of map A. (**B**) Map of collecting localities of *Guerlinguetus* samples included in the present analyses with UCE data; samples are identified according to the species limits suggested by the highest supported UCE inferences (RAxML and SVDquartets), the UCE hypothesis is simplified at the top-right corner of map B.

Supplementary Material

Supplementary Catalog Data of Voucher Material.

Voucher material for modern and historical samples analyzed in this study (listed in Supplementary Table S1) is preserved in the following scientific collections: American Museum of Natural History, USA (AMNH); Colección Boliviana de Fauna, Bolivia (CBF); Coleção de Mamíferos “Alexandre Rodrigues Ferreira”, Universidade Estadual de Santa Cruz, Brazil (CMARF); Field Museum of Natural History, USA (FMNH); Instituto de Desenvolvimento Sustentável Mamirauá, Brazil (IDSM); Coleções Científicas Fauna do Amapá, Instituto de Pesquisas Científicas e Tecnológicas do Estado do Amapá, Brazil (IEPA); Institut des Sciences de l’Évolution Montpellier, France (ISEM); University of Kansas Natural History Museum, USA (KU); Coleção do Laboratório de Mamíferos, Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Brazil (LMUSP); Louisiana Museum of Natural History, USA (LSUMZ); Museu de Ciências Naturais, Fundação Zoobotânica do Rio Grande do Sul, Brazil (MCN-FZB); Coleção de Mastozoologia do Museu de Ciências Naturais, Pontifícia Universidade Católica de Minas Gerais, Brazil (MCN-M); Museu de Ciências Naturais, Universidade Luterana do Brasil, Brazil (MCNU); Museu Nacional da Universidade Federal do Rio de Janeiro, Brazil (MN); Museu Paraense Emilio Goeldi, Brazil (MPEG); Museum of Southwestern Biology, University of New Mexico, USA (MSB); Coleção de Tecidos de Vertebrados, Departamento de Zoologia, Instituto de Biociências, Universidade de São Paulo, Brazil (MTR); Museo de Historia Natural, Universidad Nacional Mayor de San Marcos, Peru (MUSM); Museum of Vertebrate Zoology, University of Califórnia, USA (MVZ); Museu de Zoologia da Universidade de São Paulo, Brazil (MZUSP); Sam Noble Oklahoma Museum of Natural History, USA (OMNH); Museum of Texas Tech University, USA (TTU); Coleção de Tecidos Animais, Universidade Federal do Espírito Santo, Brazil (UFES-CTA); Coleção Zoológica, Instituto de Biociências, Universidade Federal de Mato Grosso, Brazil (UFMT); Universidade Federal do Pará, Brazil (UFPA); Universidade Federal de Santa Catarina, Brazil (UFSC); Smithsonian National Museum of Natural History, USA (USNM). Uncatalogued vouchers are identified (in Supplementary Table S1) by the acronym of the museum where they are currently housed followed by field numbers (in parentheses) with the following prefixes: ANRA = A. Ravetta (MPEG); EFA = E. F. Abreu-Jr. (LMUSP); GTG = G. T. Garbino (MZUSP); and LHE = L. H. Emmons (CBF). The prefixes BM, DTM, ENM, ICA, JAP, and MJ correspond to field series at the LMUSP; CN, DICO, PECC, RDSC, and RETA correspond to field series at the MPEG; AUATI correspond to field series at the IDSM; LAB correspond to field series at the UFMT; and JMIJ correspond to field series at the UFPA.

Supplementary Table S1. List of specimens included in the UCE phylogenetic analyzes with geographic information. Taxonomic identifications follow Abreu-Jr et al. (2020b) for representatives of the tribe Sciurini (subfamily Sciurinae) and Thorington et al. (2012) for the remaining taxa. See Supplementary Catalog Data of Voucher Material for explanations of voucher acronyms. (**continued**)

Voucher	Species	Country	State/ Department	Municipality/Locality	Latitude	Longitude
MVZ 193716	"*Microsciurus*" "species 2"	Brazil	Acre	Ocidente, right bank Rio Juruá	-8.566667	-72.800000
MVZ 193718	"*Microsciurus*" "species 2"	Brazil	Acre	Ocidente, right bank Rio Juruá	-8.566667	-72.800000
KU 144565	"*Microsciurus*" "species 2"	Peru	Madre de Dios	Reserva Cuzco Amazonico, 14 km E of Puerto Maldonado	-12.600000	-69.054490
MSB 57066	*"Microsciurus" flaviventer*	Bolivia	Pando	Santa Rosa	-12.216670	-68.400000
MVZ 193714	*"Microsciurus" flaviventer*	Brazil	Acre	Flora [=Fazenda Santa Fé], left bank Rio Juruá	-8.600000	-72.850000
MVZ 193715	*"Microsciurus" flaviventer*	Brazil	Acre	Ocidente, right bank Rio Juruá	-8.566667	-72.800000
MVZ 190348	*"Microsciurus" flaviventer*	Brazil	Amazonas	Barro Vermelho, left bank Rio Juruá	-6.466667	-68.766667
MVZ 191205	*"Microsciurus" flaviventer*	Brazil	Amazonas	Barro Vermelho, left bank Rio Juruá	-6.466667	-68.766667
MPEG 41774	*"Microsciurus" flaviventer*	Brazil	Amazonas	Jutai, RDS Cujubim	-4.935490	-68.173360
LMUSP (ICA 237)	*"Microsciurus" flaviventer*	Brazil	Amazonas	Santo Antônio do Içá, margem direita do Rio Içá, oposto à Comunidade São Pedro	-3.038182	-68.879747
MVZ 154931	*"Microsciurus" flaviventer*	Peru	Amazonas	Vicinity of Huampami (Aguaruna village), Rio Cenepa	-4.455630	-78.161230
MVZ 154932	*"Microsciurus" flaviventer*	Peru	Amazonas	Vicinity of Huampami (Aguaruna village), Rio Cenepa	-4.455630	-78.161230
LSUMZ-M 1565	*"Microsciurus" flaviventer*	Peru	Loreto	Quebrada Orán, ca. 5 km N Río Amazonas, 85 km NE Iquitos	-3.199000	-72.706000
LSUMZ-M 1568	*"Microsciurus" flaviventer*	Peru	Loreto	Quebrada Orán, ca. 5 km N Río Amazonas, 85 km NE Iquitos	-3.199000	-72.706000
KU 158247	*"Microsciurus" flaviventer*	Peru	Loreto	Teniente Lopez	-2.583330	-76.116670
KU 158251	*"Microsciurus" flaviventer*	Peru	Loreto	Teniente Lopez, 1.5 km N of	-2.557860	-76.116670
KU 158252	*"Microsciurus" flaviventer*	Peru	Loreto	Teniente Lopez, 1.5 km N of	-2.557860	-76.116670
AMNH 272819	*"Microsciurus" flaviventer*	Peru	Loreto	Rio Galvez, Nuevo San Juan	-5.250000	-73.166670
MUSM 13307	*"Microsciurus" flaviventer*	Peru	Loreto	Rio Galvez, Nuevo San Juan	-5.250000	-73.166670
USNM 581937	*"Microsciurus" sabanillae*	Peru	Amazonas	Cordillera del Condor, Valle Rio Comaina, camp at head of Valley below Table Mountain	-3.877000	-78.413000

Supplementary Table S1. (**continuation**)

MVZ 201331	*Aplodontia rufa*	United States	California	Mariposa County, Meadow 600 m NW Monroe Meadows, Yosemite National Park	37.665960	-119.667980
MVZ 231446	*Aplodontia rufa*	United States	California	Mono County, Valentine Reserve, Mammoth Lakes	37.629667	-119.005333
USNM 449883	*Echinosciurus variegatoides*	Panama	Bocas Del Toro	Isla San Cristobal, Bocatorito	9.242064	-82.261056
USNM 449884	*Echinosciurus variegatoides*	Panama	Bocas Del Toro	Isla San Cristobal, Bocatorito	9.242064	-82.261056
USNM 449887	*Echinosciurus variegatoides*	Panama	Bocas Del Toro	Tierra Oscura, 3.5 Km S. Tiger Key	9.196700	-82.275600
MSB 248241	*Echinosciurus variegatoides*	Panama	Colón	Santa Rosa, Aguas Claras	9.166667	-79.666667
USNM 569025	*Glaucomys volans*	United States	Texas	Harrison County, Longhorn Army Ammunition Plant	32.665714	-94.143014
USNM 599925	*Guerlinguetus aestuans* "a"	Guyana		Mount Roraima, N slope	5.283300	-60.750000
MPEG (DICO 001)	*Guerlinguetus aestuans* "b"	Brazil	Mato Grosso	Reserva Extrativista Guariba-Roosevelt, margem direita Rio Roosevelt	-9.000556	-60.354167
MPEG (RETA 15)	*Guerlinguetus aestuans* "b"	Brazil	Pará	Santarém, Comunidade de Capixauã	-2.612361	-55.192083
IEPA 3704	*Guerlinguetus aestuans* "c"	Brazil	Amapá	Porto Grande, margem direita do Rio Vila Nova, Floresta Estadual do Amapá	0.470000	-52.010000
MN 56819	*Guerlinguetus aestuans* "c"	Brazil	Amazonas	Barcelos, Rio Katana-u	1.208611	-64.789167
LMUSP (ENM 13)	*Guerlinguetus aestuans* "c"	Brazil	Amazonas	São Gabriel da Cachoeira, 5º PEF Maturaca	0.634605	-66.124988
MPEG (CN 158)	*Guerlinguetus aestuans* "c"	Brazil	Pará	Alenquer, ESEC Grão-Pará, porção sul	-0.165489	-55.186400
MPEG (CN 160)	*Guerlinguetus aestuans* "c"	Brazil	Pará	Alenquer, ESEC Grão-Pará, porção sul	-0.165489	-55.186400
IEPA 4384	*Guerlinguetus aestuans* "c"	Brazil	Pará	Almeirim, margem direita do Rio Jari, Igarapé Pacanari	0.681910	-52.593170
MPEG (CN 269)	*Guerlinguetus aestuans* "c"	Brazil	Pará	Almerim, Flota Paru, margem direita do rio Paru de Leste	-0.943969	-53.236300
MPEG (CN 270)	*Guerlinguetus aestuans* "c"	Brazil	Pará	Almerim, Flota Paru, margem direita do rio Paru de Leste	-0.943969	-53.236300
MPEG (CN 230)	*Guerlinguetus aestuans* "c"	Brazil	Pará	Almerim, Reserva Biológica Maicuru	-0.828619	-53.931200
MPEG (CN 241)	*Guerlinguetus aestuans* "c"	Brazil	Pará	Almerim, Reserva Biológica Maicuru	-0.828619	-53.931200
UFMT (BML 1493)	*Guerlinguetus aestuans* "c"	Brazil	Pará	Belo Monte	-3.400000	-51.750000
MPEG (CN 029)	*Guerlinguetus aestuans* "c"	Brazil	Pará	Faro, Flota de Faro, margem esquerda do Rio Nhamundá	-1.714011	-57.213300
MPEG (CN 048)	*Guerlinguetus aestuans* "c"	Brazil	Pará	Faro, Flota de Faro, margem esquerda do Rio Nhamundá	-1.714011	-57.213300
MPEG (PECC 26)	*Guerlinguetus aestuans* "c"	Brazil	Pará	Marajó-Afuá, Igarapé Torrão, rio Preto, PE Charapucú	-0.418840	-50.494876

Supplementary Table S1. (**continuation**)

MPEG (PECC 03)	*Guerlinguetus aestuans* "c"	Brazil	Pará	Marajó-Afuá, Rio Cuieiras, PE Charapucú	-0.226960	-50.590000
MPEG (CN 294)	*Guerlinguetus aestuans* "c"	Brazil	Pará	Óbidos, ESEC Grão-Pará, porção central	-0.630281	-55.728500
MPEG (CN 175)	*Guerlinguetus aestuans* "c"	Brazil	Pará	Oriximiná, ESEC Grão-Pará, porção norte	-1.285419	-58.695900
MPEG (CN 201)	*Guerlinguetus aestuans* "c"	Brazil	Pará	Oriximiná, ESEC Grão-Pará, porção norte	-1.285419	-58.695900
UFPA (JMIJ 35)	*Guerlinguetus aestuans* "c"	Brazil	Pará	Tapajós	-5.613410	-57.122430
LMUSP (BM 10678)	*Guerlinguetus aestuans* "c"	Brazil	Pará	Vitória do Xingu, margem esquerda do Rio Xingu	-2.873760	-52.015519
LMUSP (BM 17058)	*Guerlinguetus aestuans* "c"	Brazil	Pará	Vitória do Xingu, margem esquerda do Rio Xingu	-2.873760	-52.015519
UFES-CTA 4256	*Guerlinguetus aestuans* "c"	Brazil	Pará	Vitória do Xingu, margem esquerda do Rio Xingu	-3.259346	-51.856208
ISEM-T 5176	*Guerlinguetus aestuans* "c"	French Guiana	Cacao	Roura, Cacao	4.576944	-52.468611
ISEM-T 6101	*Guerlinguetus aestuans* "c"	French Guiana	Cayenne	Camp du Tigre, Cayenne	4.908333	-52.308333
ISEM-T 1758	*Guerlinguetus aestuans* "c"	French Guiana	Petit Saut	Sinnamary, Petit Saut	5.050000	-53.050000
CMARF 1568	*Guerlinguetus brasiliensis*	Brazil	Bahia	Belmonte, Fazenda Ouro Verde	-15.895070	-39.238440
CMARF 1569	*Guerlinguetus brasiliensis*	Brazil	Bahia	Belmonte, Fazenda Ouro Verde	-15.895070	-39.238440
UFES-CTA 85	*Guerlinguetus brasiliensis*	Brazil	Bahia	Nova Viçosa, Fazenda Suécia	-17.878889	-40.026389
UFES-CTA 223	*Guerlinguetus brasiliensis*	Brazil	Espírito Santo	Águia Branca, Fazenda Pedra Redonda	-18.973333	-40.770556
UFES-CTA 3817	*Guerlinguetus brasiliensis*	Brazil	Espírito Santo	Conceição da Barra, Floresta Nacional do Rio Preto	-18.355278	-39.844167
UFES-CTA 119	*Guerlinguetus brasiliensis*	Brazil	Espírito Santo	Pancas, Córrego São Bento, Fazenda do Dr. Rolly Luís	-19.225833	-40.761667
UFES-CTA 614	*Guerlinguetus brasiliensis*	Brazil	Espírito Santo	Viana, Pimenta	-20.379167	-40.468333
UFES-CTA 14	*Guerlinguetus brasiliensis*	Brazil	Espírito Santo	Vitória, Parque Estadual da Fonte Grande	-20.350000	-40.350000
UFES-CTA 3776	*Guerlinguetus brasiliensis*	Brazil	Minas Gerais	Belo Horizonte, Parque das Mangabeiras	-19.945590	-43.910533
MCN-M 866	*Guerlinguetus brasiliensis*	Brazil	Minas Gerais	Braúna, UHE Porto Estrela	-19.116389	-42.657778
UFES-CTA 1063	*Guerlinguetus brasiliensis*	Brazil	Minas Gerais	Itanhandu, Posses, 13 km SE Itanhandu	-22.383333	-44.850000
LMUSP (EFA 41)	*Guerlinguetus brasiliensis*	Brazil	Pará	Pacajá, LT Xingu-Estreito	-3.929722	-51.069722
MPEG (ANRA 10)	*Guerlinguetus brasiliensis*	Brazil	Pará	Portel, Igarapé Açaituba	-2.224211	-50.544875
MPEG (ANRA 12)	*Guerlinguetus brasiliensis*	Brazil	Pará	Portel, Igarapé Quirino	-2.166682	-50.644508

Supplementary Table S1. (**continuation**)

USNM 549526	*Guerlinguetus brasiliensis*	Brazil	Pará	Rio Xingu, east bank	-3.650000	-52.916667
MZUSP (GTG 54)	*Guerlinguetus brasiliensis*	Brazil	Pará	Santana do Araguaia, Fazenda Fartura	-9.732500	-50.325278
MZUSP (GTG 55)	*Guerlinguetus brasiliensis*	Brazil	Pará	Santana do Araguaia, Fazenda Fartura	-9.732500	-50.325278
MTR-CIT 1128	*Guerlinguetus brasiliensis*	Brazil	Rio Grande do Sul	Itá	-27.256812	-52.392234
MCNU 3683	*Guerlinguetus brasiliensis*	Brazil	Santa Catarina	Anita Garibaldi, margem direita UHE Barra Grande	-27.788560	-51.154506
FMNH 141601	*Guerlinguetus brasiliensis*	Brazil	Sao Paulo	Ilha do Cardoso	-25.133333	-47.966667
MTR-CIT 1785	*Guerlinguetus brasiliensis*	Brazil	São Paulo	Biritiba Mirim	-23.596572	-46.026164
MVZ 192702	*Guerlinguetus brasiliensis*	Brazil	São Paulo	Capão Bonito, Base do Carmo, Fazenda Intervales	-24.333333	-48.416667
LMUSP (DTM 277)	*Guerlinguetus brasiliensis*	Brazil	São Paulo	Caraguatatuba, Parque Estadual da Serra do Mar	-23.581412	-45.484681
LMUSP (DTM 41)	*Guerlinguetus brasiliensis*	Brazil	São Paulo	Caraguatatuba, Parque Estadual da Serra do Mar	-23.581412	-45.484681
MTR-CIT 1821	*Guerlinguetus brasiliensis*	Brazil	São Paulo	Juquitiba	-23.928496	-47.066410
MTR-ITM 444	*Guerlinguetus brasiliensis*	Brazil	São Paulo	Juquitiba	-23.928496	-47.066410
MTR-ITM 192	*Guerlinguetus brasiliensis*	Brazil	São Paulo	Piedade	-23.716279	-47.422659
MVZ 200389	*Guerlinguetus brasiliensis*	Brazil	São Paulo	São Sebastião, Fazenda da Toca, 2.4 km E, 0.8 km NE (by road) Ilhabela, Ilha de São Sebastião	-23.816667	-45.350000
LMUSP 301	*Guerlinguetus brasiliensis*	Brazil	São Paulo	Sorocaba, APP Toyota	-23.374444	-47.470833
MVZ 182070	*Guerlinguetus brasiliensis*	Brazil	São Paulo	Ubatuba, Fazenda Capricórnio, 5 km N Ubatuba	-23.416667	-45.116667
MVZ 182071	*Guerlinguetus brasiliensis*	Brazil	São Paulo	Ubatuba, Praia do Félix	-23.383333	-44.466667
MUSM 13353	*Hadrosciurus* "species 3"	Peru	Loreto	Rio Galvez, Nuevo San Juan	-5.250000	-73.166670
MSB 211524	*Hadrosciurus ignitus*	Bolivia	Beni	Totaisal, 1 km SW of Estacion Biologica Del Beni	-14.510000	-66.210000
CBF (LHE 1311)	*Hadrosciurus ignitus*	Bolivia	Chuquisaca	El Limón, left bank of Rio Santa Marta	-20.700192	-64.300089
CBF (LHE 1312)	*Hadrosciurus ignitus*	Bolivia	Chuquisaca	El Limón, left bank of Rio Santa Marta	-20.700192	-64.300089
MSB 239309	*Hadrosciurus ignitus*	Bolivia	La Paz	13.7 km by rd NE from La Reserva	-15.733330	-67.516670
LSUMZ-M 4467	*Hadrosciurus ignitus*	Bolivia	La Paz	Prov. B. Saavedra, 83 km by road E Charazani, Cerro Asunta Pata	-15.083000	-68.550000
LSUMZ-M 4471	*Hadrosciurus ignitus*	Bolivia	La Paz	Prov. B. Saavedra, 83 km by road E Charazani, Cerro Asunta Pata	-15.083000	-68.550000

Supplementary Table S1. (**continuation**)

MSB 235873	*Hadrosciurus ignitus*	Bolivia	La Paz	Serronia Bella Vista	-15.683300	-67.500000
MVZ 191206	*Hadrosciurus ignitus*	Brazil	Acre	Cruzeiro do Sul, left bank Rio Juruá	-7.633333	-72.600000
MVZ 190350	*Hadrosciurus ignitus*	Brazil	Acre	Igarapé Porongaba, right bank Rio Juruá	-8.666667	-72.783333
LMUSP (MJ 302)	*Hadrosciurus ignitus*	Brazil	Rondônia	Porto Velho, margem esquerda do Rio Madeira, Caiçara	-9.437526	-64.849635
KU 144563	*Hadrosciurus ignitus*	Peru	Madre de Dios	Reserva Cuzco Amazonico, 14 km E of Puerto Maldonado	-12.600000	-69.054490
IDSM (AUATI 78)	*Hadrosciurus igniventris*	Brazil	Amazonas	RESEX Auatí-Paraná	-1.934167	-66.232500
IDSM (AUATI 95)	*Hadrosciurus igniventris*	Brazil	Amazonas	RESEX Auatí-Paraná	-1.934167	-66.232500
USNM 599924	*Hadrosciurus igniventris*	Venezuela	Amazonas	Neblina Base Camp, Rio Mawarinuma	0.830000	-66.170000
MVZ 193723	*Hadrosciurus pyrrhinus*	Brazil	Amazonas	Colocação Vira-Volta, left bank Rio Juruá on Igarapé Arabidi, affluent of Paranã Breu	-3.283333	-66.233333
MPEG (RDSC 101)	*Hadrosciurus pyrrhinus*	Brazil	Amazonas	Jutai, RDS Cujubim, baixo rio Mutum, afluente dir. médio rio Jutai, afluente direito alto rio Solimoes, comunidade Pirarucu	-4.670524	-68.132023
MPEG 41755	*Hadrosciurus pyrrhinus*	Brazil	Amazonas	Jutai, RDS Cujubim, P.1., transecto mata de terra firme	-4.654170	-68.323940
MPEG 41784	*Hadrosciurus pyrrhinus*	Brazil	Amazonas	Jutai, RDS Cujubim, P.2., descendo margem esquerda rio Mutum	-4.935490	-68.173360
MVZ 153504	*Hadrosciurus pyrrhinus*	Peru	Amazonas	Headwaters of Rio Huampami, ca. 3 hrs by canoe N Huampami (Aguaruna village), Rio Cenepa	-4.312950	-78.128000
TTU 101104	*Hadrosciurus pyrrhinus*	Peru	Loreto	Maynas, Iquitos, 25 km S, Estacion Biologica Allpahuayo	-3.966670	-73.416670
TTU 98890	*Hadrosciurus pyrrhinus*	Peru	Loreto	Maynas, Iquitos, 25 km S, Estacion Biologica Allpahuayo	-3.966670	-73.416670
AMNH 272859	*Hadrosciurus pyrrhinus*	Peru	Loreto	Rio Galvez, Nuevo San Juan	-5.250000	-73.166670
MUSM 13354	*Hadrosciurus pyrrhinus*	Peru	Loreto	Rio Galvez, Nuevo San Juan	-5.250000	-73.166670
MSB 99062	*Hadrosciurus spadiceus*	Bolivia	Beni	Totaisal	-14.881111	-66.328056
USNM 599923	*Hadrosciurus spadiceus*	Bolivia	Santa Cruz	El Refugio	-14.767222	-61.034722
MSB 210549	*Hadrosciurus spadiceus*	Bolivia	Santa Cruz	San Miguel Rincon	-17.383333	-63.533333
MSB 55208	*Hadrosciurus spadiceus*	Bolivia	Santa Cruz	San Miguel Rincon	-17.416667	-63.566667
MVZ 193724	*Hadrosciurus spadiceus*	Brazil	Acre	Ocidente, right bank Rio Juruá	-8.566667	-72.800000
MVZ 193726	*Hadrosciurus spadiceus*	Brazil	Acre	Ocidente, right bank Rio Juruá	-8.566667	-72.800000

Supplementary Table S1. (**continuation**)

LMUSP (EFA 27)	*Hadrosciurus spadiceus*	Brazil	Amazonas	Anori, margem esquerda do Rio Purus, Comunidade do Caua-Cuiuanã, Igarapé do Cuiuanã, trilha do Acarizinho	-4.168714	-61.726627
LMUSP (EFA 31)	*Hadrosciurus spadiceus*	Brazil	Amazonas	Anori, margem esquerda do rio Purus, Comunidade do Caua-Cuiuanã, mata atrás da Comunidade	-4.234045	-61.734642
LMUSP (EFA 16)	*Hadrosciurus spadiceus*	Brazil	Amazonas	Anori, margem esquerda do Rio Purus, Comunidade do Caua-Cuiuanã, proximidades da Comunidade	-4.240372	-61.724023
MVZ 190352	*Hadrosciurus spadiceus*	Brazil	Amazonas	Colocação Sabiá, left bank Rio Juruá	-6.783330	-70.816670
MVZ 190354	*Hadrosciurus spadiceus*	Brazil	Amazonas	Colocação Sabiá, left bank Rio Juruá	-6.783330	-70.816670
LMUSP (JAP 165)	*Hadrosciurus spadiceus*	Brazil	Amazonas	Japurá, margem direita do Rio Japurá, antiga Vila de Santa Fé, trilha da Canoa Virada	-1.764217	-66.357172
MPEG 41754	*Hadrosciurus spadiceus*	Brazil	Amazonas	Jutai, RDS Cujubim, P.1. Boca do Igarapé Sto. Antonio	-4.654170	-68.323940
MPEG 41775	*Hadrosciurus spadiceus*	Brazil	Amazonas	Jutai, RDS Cujubim, P.2. Igapó, rio Mutum	-4.935490	-68.173360
MPEG 41776	*Hadrosciurus spadiceus*	Brazil	Amazonas	Jutai, RDS Cujubim, P.2. Igapó, rio Mutum	-4.935490	-68.173360
MPEG 41747	*Hadrosciurus spadiceus*	Brazil	Amazonas	Jutai, RDS Cujubim, P.4, a 500m no transecto	-5.638010	-69.187770
IDSM (AUATI 97)	*Hadrosciurus spadiceus*	Brazil	Amazonas	RDS Mamirauá	-2.000278	-66.201111
LMUSP (ICA 129)	*Hadrosciurus spadiceus*	Brazil	Amazonas	Santo Antônio do Içá, margem esquerda do Rio Içá, Comunidade Cuiauá ou Monte Tabor	-2.885465	-68.368681
UFMT (LAB 90)	*Hadrosciurus spadiceus*	Brazil	Mato Grosso	Cuiabá, Bairro Coophema	-15.634458	-56.060244
MPEG 44284	*Hadrosciurus spadiceus*	Brazil	Pará	Santarém, Comunidade de Boim	-3.093440	-55.522360
MVZ 192704	Hadrosciurus spadiceus	Peru	Madre de Dios	Albergue Cuzco Amazonico	-12.550000	-69.050000
KU 144566	*Hadrosciurus spadiceus*	Peru	Madre de Dios	Reserva Cuzco Amazonico, 14 km E of Puerto Maldonado	-12.600000	-69.054490
KU 144567	*Hadrosciurus spadiceus*	Peru	Madre de Dios	Reserva Cuzco Amazonico, 14 km E of Puerto Maldonado	-12.600000	-69.054490
KU 144572	*Hadrosciurus spadiceus*	Peru	Madre de Dios	Reserva Cuzco Amazonico, 14 km E of Puerto Maldonado	-12.600000	-69.054490
AMNH 272825	*Hadrosciurus spadiceus*	Peru	Loreto	Rio Galvez, Nuevo San Juan	-5.250000	-73.166670
MSB 61161	*Hesperosciurus aberti*	United States	Arizona	Graham County, Grahm Mtns. Hospital Flats	32.665100	-109.911000
MSB 40646	*Hesperosciurus griseus*	United States	California	Mariposa County, 8 mi N, 3 mi E Oakhurst	37.446031	-119.598744
USNM 584420	*Hylopetes phayrei*	Myanmar	Mandalay	Pyin-Oo-Lwin (Maymyo), 6.8 mi. ENE, Yangon Monastary	22.067500	96.563500
USNM 338172	*Leptosciurus boquetensis*	Panama	Darien	Cerro Mali	8.170000	-77.230000

Supplementary Table S1. (**continuation**)

USNM 292133	*Leptosciurus isthmius*	Colombia	Choco	Baudo Mountains, Rio Jurubida	5.970000	-77.280000
USNM 554228	*Leptosciurus isthmius*	Colombia	Valle del Cauca	Beuhaventure, 6 Km N	3.870000	-77.080000
USNM 309028	*Leptosciurus mimulus*	Colombia	Narino	La Guayacana	1.430000	-78.450000
USNM 499514	*Leptosciurus otinus*	Colombia	Antioquia	Zaragoza, 23 Km S, 22 Km W, at Providencia	7.500000	-74.866667
USNM 293776	*Leptosciurus pucheranii*	Colombia	Antioquia	La Bodega, S Side Rio Negrito, Highway Sonson-Narino	5.700000	-75.116667
USNM 303846	*Leptosciurus similis*	Colombia	Cauca	Cerro Munchique	2.530000	-76.980000
USNM 570490	*Marmota monax*	United States	Virginia	Warren County, Front Royal, National Zoological Park, Conservation Research Center, Hayfields	38.886376	-78.167332
USNM 570492	*Marmota monax*	United States	Virginia	Warren County, Front Royal, National Zoological Park, Conservation Research Center, Hayfields	38.886376	-78.167332
USNM 292132	*Microsciurus* "species 1"	Colombia	Choco	Rio Nuqui, Baudo Mountains, Base	5.670000	-77.270000
USNM 575652	*Microsciurus alfari*	Panama	Bocas Del Toro	Nuri	8.913025	-81.815461
USNM 575653	*Microsciurus alfari*	Panama	Bocas Del Toro	Nuri	9.395119	-82.531526
USNM 318364	*Microsciurus alfari*	Panama	Panama	Candelaria Hydrographic Station, Rio Pequeni	9.370000	-79.530000
USNM 570626	*Neosciurus carolinensis*	United States	Wisconsin	Portage County, Stevens Point, 2633 Ellis Street, 54481	44.521300	-89.560900
MSB 157846	*Parasciurus arizonensis*	United States	New Mexico	Grant County, Gila River near Spar Canyon	33.023350	-108.535650
MSB 47449	*Parasciurus nayaritensis*	United States	Arizona	Cochise County, Chiricahua Mts, 2 mi SW Paradise	31.914200	-109.242400
USNM 570471	*Parasciurus niger*	United States	Maryland	Dorchester County, Vienna, ca. 1.5 mi SSW on Elliots Island Road	38.469200	-75.845300
USNM 568615	*Parasciurus niger*	United States	Maryland	Montgomery County, Potomac River, south side of Watkins Island	29.042200	-77.279200
FMNH 165415	*Ratufa bicolor*	NA	NA	NA	NA	NA
MPEG (CN 304)	*Sciurillus pusillus*	Brazil	Pará	Óbidos, ESEC Grão-Pará, porção central (locality [3], in Figure 6)	-0.630281	-55.728500
MPEG (RETA 13)	*Sciurillus pusillus*	Brazil	Pará	Santarém, Comunidade de Capixauã (locality [1], in Figure 6)	-2.612361	-55.192083
USNM 548447	*Sciurillus pusillus*	French Guiana		River Arataye, Inini (locality [2], in Figure 6)	4.000000	-52.670000
LSUMZ-M 1561	*Sciurillus pusillus*	Peru	Loreto	Quebrada Orán, ca. 5 km N Río Amazonas, 85 km NE Iquitos (locality [5], in figure 6)	-3.199000	-72.706000

Supplementary Table S1. (**continuation**)

LSUMZ-M 1562	*Sciurillus pusillus*	Peru	Loreto	Quebrada Orán, ca. 5 km N Río Amazonas, 85 km NE Iquitos (locality [5], in figure 6)	-3.199000	-72.706000
AMNH 272807	*Sciurillus pusillus*	Peru	Loreto	Rio Galvez, Nuevo San Juan (locality [4], in Figure 6)	-5.250000	-73.166670
AMNH 272866	*Sciurillus pusillus*	Peru	Loreto	Rio Galvez, Nuevo San Juan (locality [4], in Figure 6)	-5.250000	-73.166670
MSB 148785	*Sciurus vulgaris*	Russia	Sakha Republic	Kenkeme River, 40 km W Yakutsk	62.070030	128.938310
MVZ 196054	*Simosciurus nebouxii*	Peru	Cajamarca	2.5 km N (by air) Monte Seco, Rio Zana	-7.110517	-79.516486
LSUMZ-M 936	*Simosciurus nebouxii*	Peru	Piura	Pariñas, 7 km N, 15 km E Talara	-4.533000	-81.150000
TTU 102463	*Syntheosciurus granatensis*	Ecuador	El Oro	Near to La Victoria, road from Arenillas to Puyango	-3.550000	-80.066670
AMNH 71613	*Syntheosciurus granatensis*	Ecuador	Pichincha	Quito, La Carolina	-0.216667	-78.500000
USNM 464871	*Syntheosciurus granatensis*	Panama	Bocas Del Toro	Isla Colón, La Gruta	9.400000	-82.266667
USNM 464872	*Syntheosciurus granatensis*	Panama	Bocas Del Toro	Isla Colón, La Gruta	9.400000	-82.266667
USNM 575630	*Syntheosciurus granatensis*	Panama	Bocas Del Toro	Nuri	8.913025	-81.815461
USNM 575637	*Syntheosciurus granatensis*	Panama	Bocas Del Toro	Nuri	8.913025	-81.815461
USNM 578376	*Syntheosciurus granatensis*	Panama	Bocas Del Toro	Peninsula Valiente, Punta Alegre	9.162550	-81.905133
USNM 578377	*Syntheosciurus granatensis*	Panama	Bocas Del Toro	Peninsula Valiente, Punta Alegre	9.162550	-81.905133
USNM 578378	*Syntheosciurus granatensis*	Panama	Bocas Del Toro	Peninsula Valiente, Punta Alegre	9.162550	-81.905133
USNM 449881	*Syntheosciurus granatensis*	Panama	Bocas Del Toro	Tierra Oscura, 3.5 Km S. Tiger Key	9.196700	-82.275600
MSB 273400	*Tamiasciurus douglasii*	Mexico	Baja California	Sierra San Pedro Martir, 15 mi E Meling Ranch	30.905278	115.501111
USNM 568061	*Tamiasciurus douglasii*	United States	Washington	Island County, Whidbey Island, Oak Harbor, Nas	48.293019	-122.643203
USNM 570572	*Tamiasciurus hudsonicus*	United States	Wisconsin	Portage County	44.333400	-89.381400
USNM 570573	*Tamiasciurus hudsonicus*	United States	Wisconsin	Portage County	44.333400	-89.381400

Supplementary Table S2. Summary of sequencing success, assembly statistics, and resulting UCE data for samples analyzed in this study. Samples in bold were included in the phylogenetic analyses. (**continued**)

Sample ID	Sample type	Total reads	Contigs	UCEs	UCEs mean length	Total bp
MVZ 182071	Modern	3,128,739	23,367	3,385	547	1,850,845
MTR 0444	Modern	4,606,692	65,399	3,339	547	1,825,701
MSB 40646	Modern	3,973,147	59,775	3,262	458	1,494,350
MTR 0192	Modern	2,976,892	50,863	3,253	509	1,654,232
MVZ 192704	Modern	4,917,538	85,728	3,253	461	1,498,726
MVZ 182070	Modern	5,470,087	55,462	3,239	572	1,852,127
MVZ 191205	Modern	4,905,780	66,363	3,189	444	1,417,123
MVZ 190348	Modern	1,193,220	12,872	3,185	471	1,498,580
GTG 54	Modern	7,247,417	102,927	3,183	464	1,475,847
MTR 1821	Modern	9,973,602	263,300	3,177	529	1,679,266
MSB 47449	Modern	3,007,275	51,620	3,168	410	1,298,239
MTR 1128	Modern	3,957,715	60,339	3,152	434	1,367,327
MVZ 192702	Modern	2,912,144	24,372	3,142	521	1,637,233
MSB 148785	Modern	3,340,511	34,642	3,137	500	1,567,954
USNM 599923	Modern	3,525,628	24,621	3,136	624	1,956,359
MVZ 193715	Modern	5,458,719	77,746	3,135	396	1,243,009
MSB 157846	Modern	4,205,295	40,409	3,129	396	1,238,960
MVZ 193726	Modern	2,572,335	25,714	3,129	510	1,595,166
KU 144572	Modern	2,715,423	23,185	3,124	637	1,990,613
TTU 98890	Modern	2,503,794	19,341	3,113	428	1,331,575
UFES 3776	Modern	3,037,960	37,389	3,092	426	1,318,099
CN 160	Modern	2,603,351	42,448	3,079	500	1,540,551
UFES 3817	Modern	2,931,172	25,566	3,076	425	1,308,246
GTG 55	Modern	7,146,873	116,881	3,072	491	1,509,638
UFES 1063	Modern	3,845,796	64,868	3,070	461	1,414,686
UFES 223	Modern	1,712,279	15,431	3,061	468	1,432,467
AMNH 272825	Modern	4,489,324	72,042	3,059	640	1,959,231
MVZ 193724	Modern	1,193,046	13,171	3,057	534	1,632,933
MVZ 193714	Modern	6,225,285	60,793	3,042	450	1,367,419
MN 56819	Modern	3,713,372	35,118	3,039	395	1,199,892
MSB 248241	Modern	4,471,567	29,821	3,028	393	1,189,360
USNM 575637	Modern	3,669,033	116,413	3,026	396	1,198,025
LSUMZ 1565	Modern	5,401,990	29,084	3,017	594	1,791,194
CN 304	Modern	4,307,356	39,539	3,008	469	1,409,650
AMNH 272807	Modern	3,534,263	52,228	2,996	601	1,799,521

Supplementary Table S2. (**continuation**)

MVZ 190350	Modern	4,840,887	31,640	2,990	385	1,150,408
MVZ 193723	Modern	6,218,943	41,664	2,987	420	1,253,338
USNM 449883	Modern	4,586,727	58,277	2,969	493	1,462,368
MTR 1785	Modern	6,523,770	75,084	2,965	485	1,439,464
MVZ 191206	Modern	1,401,360	12,396	2,943	392	1,154,546
USNM 575652	Modern	1,550,929	17,506	2,930	415	1,214,761
TTU 101104	Modern	2,460,062	30,491	2,926	385	1,126,868
AUATI 97	Modern	1,778,962	18,022	2,922	539	1,576,066
CMARF 1569	Modern	1,070,447	11,742	2,918	460	1,342,669
USNM 549526	Modern	1,754,647	14,288	2,916	531	1,547,774
USNM 578376	Modern	1,547,336	17,413	2,908	404	1,175,120
USNM 464871	Modern	2,764,660	23,142	2,900	490	1,422,061
CN 048	Modern	3,713,017	38,172	2,876	490	1,408,663
AMNH 272866	Modern	3,251,521	19,514	2,873	537	1,542,717
MVZ 193716	Modern	2,318,421	17,625	2,871	451	1,294,481
FMNH 141601	Modern	10,002,974	249,826	2,870	603	1,731,908
JMIJ 35	Modern	5,998,352	32,536	2,870	533	1,529,772
LAB 90	Modern	3,687,748	58,371	2,869	472	1,353,895
USNM 575630	Modern	2,085,391	17,368	2,868	386	1,105,796
MSB 273400	Modern	1,500,869	14,951	2,850	385	1,097,543
UFES 119	Modern	3,610,396	41,107	2,848	432	1,229,599
AUATI 95	Modern	6,212,490	153,023	2,844	464	1,320,929
RETA 15	Modern	4,948,019	54,457	2,842	477	1,354,954
IEPA 3704	Modern	2,443,068	45,422	2,838	389	1,103,658
MVZ1 54932	Modern	10,964,880	265,917	2,828	594	1,679,663
MVZ1 96054	Modern	3,810,236	28,855	2,816	553	1,556,779
USNM 568615	Modern	4,593,084	29,158	2,801	583	1,633,957
PECC 26	Modern	1,256,514	25,055	2,798	504	1,410,106
MVZ 200389	Modern	2,915,868	23,950	2,796	473	1,322,176
CN 029	Modern	3,721,076	29,101	2,784	500	1,390,820
MCNU 3683	Modern	3,194,106	46,996	2,777	414	1,148,426
USNM 575653	Modern	6,317,687	31,145	2,774	496	1,376,057
USNM 464872	Modern	3,203,761	25,426	2,769	473	1,308,830
LHE 1312	Modern	3,218,926	25,494	2,764	474	1,309,968
USNM 570471	Modern	1,860,400	19,612	2,761	531	1,465,682
KU 144565	Modern	2,044,600	21,327	2,758	585	1,612,450
USNM 581937	Modern	2,210,572	29,332	2,750	494	1,357,942
MVZ 154931	Modern	3,495,753	24,152	2,729	468	1,277,037
KU 158247	Modern	2,601,550	21,521	2,727	442	1,206,378
RETA 13	Modern	2,663,045	20,453	2,720	584	1,587,723

Supplementary Table S2. (**continuation**)

MPEG 41776	Modern	1,530,837	20,220	2,713	442	1,199,420
KU 158252	Modern	2,454,398	22,946	2,712	592	1,606,585
MPEG 44284	Modern	2,461,478	19,573	2,707	451	1,220,324
LHE 1311	Modern	1,588,340	13,741	2,706	477	1,290,634
LSUMZ 4471	Modern	4,732,635	34,335	2,706	587	1,589,352
UFES 85	Modern	2,683,160	33,324	2,701	395	1,067,887
MPEG 41775	Modern	2,911,961	23,542	2,697	493	1,330,594
EFA 16	Modern	3,443,641	29,222	2,692	562	1,512,040
CN 270	Modern	809,705	9,584	2,689	378	1,017,759
IEPA 4384	Modern	922,634	14,000	2,687	394	1,058,791
MPEG 41784	Modern	3,442,151	55,846	2,678	476	1,275,409
MSB 211524	Modern	11,268,206	260,010	2,677	615	1,646,868
EFA 41	Modern	3,350,171	45,164	2,675	435	1,164,793
CN 269	Modern	3,716,970	24,039	2,674	518	1,384,265
MPEG 41774	Modern	2,551,946	22,740	2,674	502	1,342,179
CN 175	Modern	981,148	10,071	2,672	372	993,747
MPEG 41747	Modern	2,091,943	30,262	2,668	512	1,365,514
MPEG 41755	Modern	4,182,898	47,644	2,667	431	1,148,791
MUSM 13354	Modern	5,688,506	103,696	2,667	579	1,543,178
CN 241	Modern	1,703,969	14,731	2,665	393	1,048,104
MSB 99062	Modern	4,261,029	25,942	2,663	595	1,585,092
MVZ 231446	Modern	497,608	7,012	2,650	366	970,598
USNM 569025	Modern	5,499,439	29,159	2,642	561	1,482,405
ISEMT 5176	Modern	2,798,491	76,052	2,637	359	946,087
LSUMZ 1568	Modern	8,440,496	212,163	2,636	567	1,495,647
MUSM 13307	Modern	6,944,452	104,813	2,607	604	1,573,841
MVZ 190352	Modern	4,291,497	65,163	2,598	452	1,175,222
MSB 239309	Modern	4,006,028	52,946	2,590	410	1,063,186
MSB 57066	Modern	1,023,431	11,770	2,587	433	1,119,174
MSB 210549	Modern	2,699,729	19,113	2,586	437	1,130,409
UFES 14	Modern	2,164,847	17,593	2,584	411	1,063,035
ICA 129	Modern	3,700,263	36,376	2,559	420	1,073,583
ANRA 12	Modern	1,729,236	14,817	2,557	379	968,461
ANRA 10	Modern	4,466,787	65,060	2,548	534	1,360,122
MUSM 13353	Modern	3,320,337	65,643	2,531	554	1,402,992
LSUMZ 1562	Modern	6,885,973	156,777	2,526	482	1,216,952
LSUMZ 4467	Modern	3,650,151	26,277	2,520	552	1,389,808
USNM 578377	Modern	4,232,670	25,110	2,516	456	1,147,192
UFES 614	Modern	1,990,289	17,889	2,514	445	1,118,722
MJ 302	Modern	4,819,423	29,749	2,513	551	1,385,042

Supplementary Table S2. (**continuation**)

KU 158251	Modern	3,122,581	26,090	2,511	580	1,456,921
AMNH 272859	Modern	1,043,961	12,223	2,505	385	965,383
LSUMZ 936	Modern	1,331,304	14,788	2,501	419	1,048,344
USNM 584420	Modern	4,338,913	30,181	2,493	530	1,321,936
DTM 41	Modern	5,230,294	32,613	2,490	518	1,290,550
BM 10678	Modern	5,689,463	35,309	2,488	565	1,406,653
AMNH 272819	Modern	3,467,831	39,979	2,476	504	1,248,813
LMUSP 301	Modern	3,694,192	24,495	2,459	516	1,269,477
MSB 61161	Modern	426,121	13,144	2,457	359	883,061
LSUMZ 1561	Modern	7,505,020	40,625	2,456	561	1,378,186
CN 158	Modern	2,798,011	24,363	2,455	455	1,117,888
MPEG 41754	Modern	12,726,177	317,895	2,442	531	1,296,211
ISEMT 1758	Modern	1,789,870	16,060	2,425	373	903,776
USNM 554228	Historical	5,022,145	15,656	2,419	335	811,385
CN 230	Modern	2,805,022	19,429	2,416	461	1,113,322
MSB 235873	Modern	3,143,939	28,684	2,408	497	1,197,945
MVZ 201331	Modern	3,978,361	50,609	2,404	374	899,403
USNM 570573	Modern	2,807,250	24,797	2,400	439	1,054,797
USNM 449887	Modern	4,101,196	25,228	2,368	461	1,091,651
MCN 866	Historical	2,784,582	5,839	2,355	321	754,786
JAP 165	Modern	3,020,575	51,342	2,343	406	951,280
KU 144566	Modern	2,187,959	22,071	2,341	553	1,293,705
MVZ 190354	Modern	3,614,061	34,655	2,328	418	973,169
ISEMT 6101	Modern	772,346	9,300	2,302	331	762,791
MSB 55208	Modern	7,341,709	44,917	2,296	497	1,141,342
CN 294	Modern	1,939,813	43,393	2,288	432	988,938
USNM 338172	Historical	3,415,027	9,806	2,276	304	692,622
TTU 102463	Modern	868,017	25,241	2,266	345	782,151
USNM 570492	Modern	5,578,858	65,120	2,259	478	1,079,871
CN 201	Modern	2,440,526	50,198	2,235	460	1,028,724
ICA 237	Modern	2,404,937	45,166	2,232	397	885,644
DTM 277	Modern	2,907,195	23,059	2,228	446	994,748
USNM 338164	Historical	1,972,371	10,597	2,205	337	743,305
BM 17058	Modern	2,259,744	20,603	2,180	445	970,527
USNM 570572	Modern	3,895,133	30,202	2,178	492	1,072,646
CMARF 1568	Modern	1,135,577	12,823	2,153	354	762,230
AUATI 78	Modern	1,195,468	15,158	2,136	329	702,411
KU 144567	Modern	5,297,374	36,158	2,131	528	1,124,114
USNM 599925	Modern	2,514,965	22,048	2,131	475	1,011,478
MVZ 193718	Modern	3,237,530	26,993	2,105	408	857,958

Supplementary Table S2. (**continuation**)

USNM 599924	Modern	2,730,738	20,184	2,085	412	859,980
USNM 293776	Historical	10,868,228	90,520	2,084	340	708,112
BML 1493	Modern	776,562	6,832	2,082	319	664,124
MVZ 153504	Modern	2,697,264	26,324	2,052	425	872,153
USNM 303846	Historical	2,323,413	6,607	2,036	299	607,799
RDSC 101	Modern	4,580,317	37,241	2,020	500	1,010,541
USNM 499514	Historical	2,125,589	6,602	2,004	298	597,751
USNM 309028	Historical	2,281,971	8,186	1,975	305	602,002
AMNH 71613	Historical	15,776,174	226,333	1,969	348	684,804
USNM 449884	Modern	5,541,516	67,263	1,957	467	914,655
PECC 03	Modern	81,152	3,803	1,939	330	640,826
USNM 568061	Modern	1,607,530	20,713	1,931	414	798,828
DICO 001	Modern	1,201,020	18,848	1,922	373	716,341
USNM 548447	Historical	4,058,078	10,734	1,909	393	750,924
ENM 13	Modern	8,799,346	34,701	1,893	476	901,106
USNM 292132	Historical	7,346,701	37,059	1,884	487	916,592
USNM 570490	Modern	2,897,616	21,931	1,869	492	920,088
EFA 31	Modern	1,592,241	18,662	1,847	374	690,995
USNM 449881	Modern	3,112,614	31,870	1,840	418	769,759
USNM 570626	Modern	2,761,920	26,409	1,818	466	847,716
FMNH 165415	Modern	2,377,290	23,108	1,796	436	782,874
USNM 292133	Historical	7,067,957	22,596	1,790	458	819,501
USNM 461886	Historical	7,639,404	94,524	1,768	302	534,244
UFES 4256	Modern	1,939,063	26,875	1,752	403	706,200
MVZ 153501	Modern	530,795	4,426	1,659	301	499,508
KU 144563	Modern	8,063,468	212,716	1,656	459	760,907
OMNH 29986	Modern	1,075,087	13,712	1,636	323	528,814
AMNH 135435	Historical	1,409,182	8,583	1,630	296	482,500
UFSC 2930	Historical	2,251,202	6,049	1,612	299	481,822
EFA 27	Modern	2,656,893	27,783	1,557	403	627,842
USNM 409798	Historical	2,357,580	7,413	1,535	286	438,450
MA 130	Modern	959,452	10,150	1,512	299	451,811
UFSC 4047	Historical	2,948,418	6,544	1,498	341	510,127
USNM 560645	Modern	2,750,453	26,612	1,493	347	518,219
USNM 488087	Historical	5,169,169	20,370	1,484	302	448,050
USNM 578378	Modern	1,578,516	20,327	1,471	406	597,411
MPEG 45441	Modern	2,459,938	31,981	1,389	451	626,101
AMNH 66640	Historical	1,542,884	9,700	1,335	293	390,941
MA 172	Modern	490,493	7,940	1,334	282	375,662
AMNH 64047	Historical	2,298,942	4,327	1,326	294	389,715

Supplementary Table S2. (**continuation**)

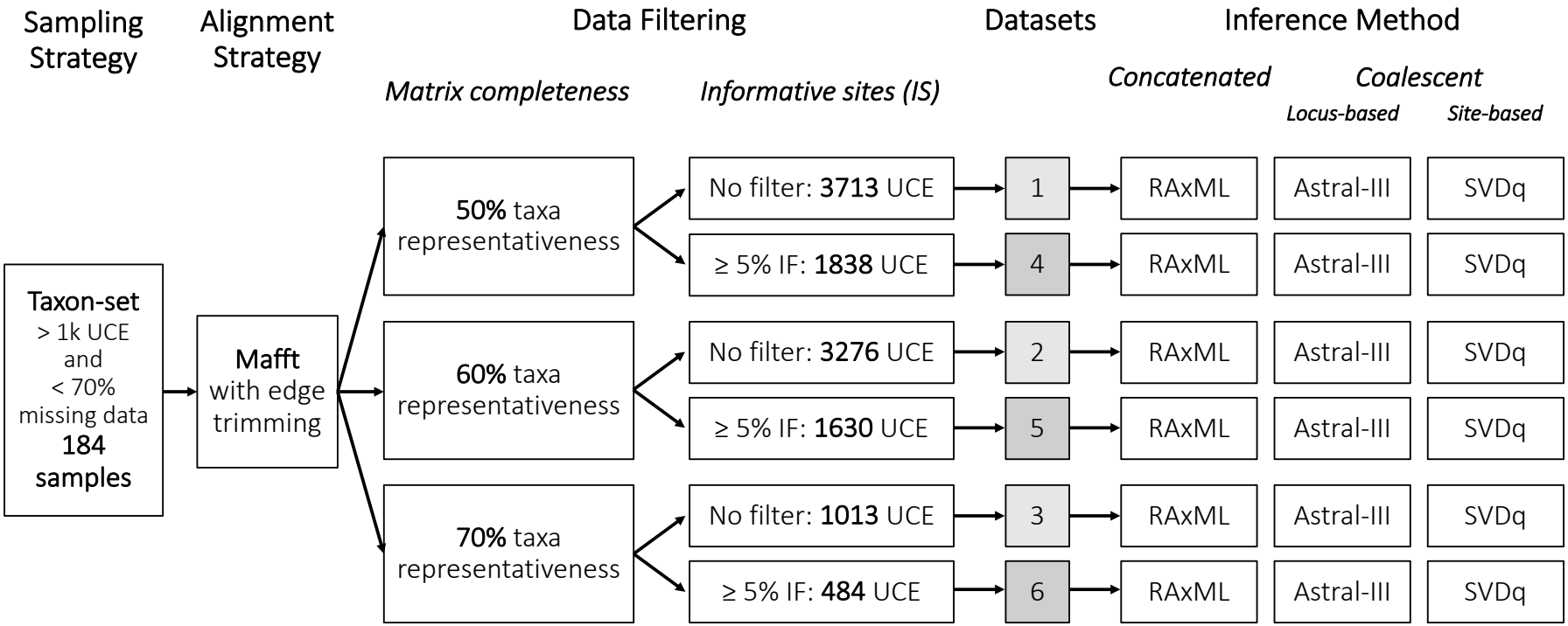
USNM 113314	Historical	5,486,985	29,678	1,324	276	365,810
AMNH 68271	Historical	1,255,190	10,109	1,314	319	419,002
AMNH 78621	Historical	3,116,695	15,370	1,287	291	373,884
AMNH 69238	Historical	2,297,514	16,505	1,171	315	368,628
USNM 255114	Historical	4,177,904	12,093	1,104	279	307,493
AMNH 73917	Historical	3,356,978	7,866	1,102	280	308,954
UFES 907	Modern	1,025,050	13,988	1,079	286	308,726
MUSM 23157	Historical	1,222,629	6,454	1,066	273	290,692
AMNH 131723	Historical	6,761,483	31,393	1,037	273	283,454
AMNH 34686	Historical	1,304,505	7,328	985	268	264,186
AMNH 60464	Historical	4,034,527	6,643	977	272	265,642
MVZ 153500	Modern	340,204	6,368	960	306	293,666
USNM 569823	Modern	7,383,765	45,402	913	505	461,235
MVZ 190351	Modern	372,312	3,027	893	289	258,461
AMNH 231771	Historical	2,733,829	18,895	890	287	255,122
USNM 334696	Historical	6,426,351	27,224	872	361	315,005
USNM 318364	Historical	2,308,616	11,099	861	278	239,575
USNM 140867	Historical	3,602,970	9,661	853	281	239,568
AMNH 36489	Historical	7,420,042	51,318	843	277	233,891
USNM 13006	Historical	354,870	5,762	831	281	233,477
MCNM 1388	Modern	3,493,628	32,967	814	393	319,499
USNM 96230	Historical	3,073,118	30,815	808	279	225,387
USNM 113311	Historical	14,630,632	37,073	784	264	207,267
MUSM 7920	Historical	689,181	4,070	780	270	210,301
AMNH 73924	Historical	1,303,553	10,590	769	272	209,390
FZB 3514	Modern	212,344	2,496	764	276	210,932
USNM 442017	Historical	681,039	1,802	762	270	205,702
MCNM 2944	Modern	4,246,201	33,359	685	355	242,979
USNM 271319	Historical	2,465,474	2,977	643	276	177,381
USNM 55933	Historical	2,357,656	30,932	626	279	174,636
USNM 251907	Historical	2,439,939	5,221	618	268	165,704
AMNH 16949	Historical	3,709,551	19,211	591	264	156,315
USNM 244931	Historical	763,479	6,193	578	273	157,914
USNM 509034	Historical	416,829	3,996	567	275	155,997
AMNH 33690	Historical	2,220,433	4,946	549	267	146,847
AMNH 62886	Historical	4,887,808	10,578	499	267	133,218
IEPA 3276	Modern	95,989	863	446	269	119,774
USNM 337569	Historical	5,796,968	61,507	426	306	130,186
AMNH 103733	Historical	419,662	1,378	412	269	110,945
USNM 512075	Historical	3,804,910	19,747	353	262	92,604

Supplementary Table S2. (**continuation**)

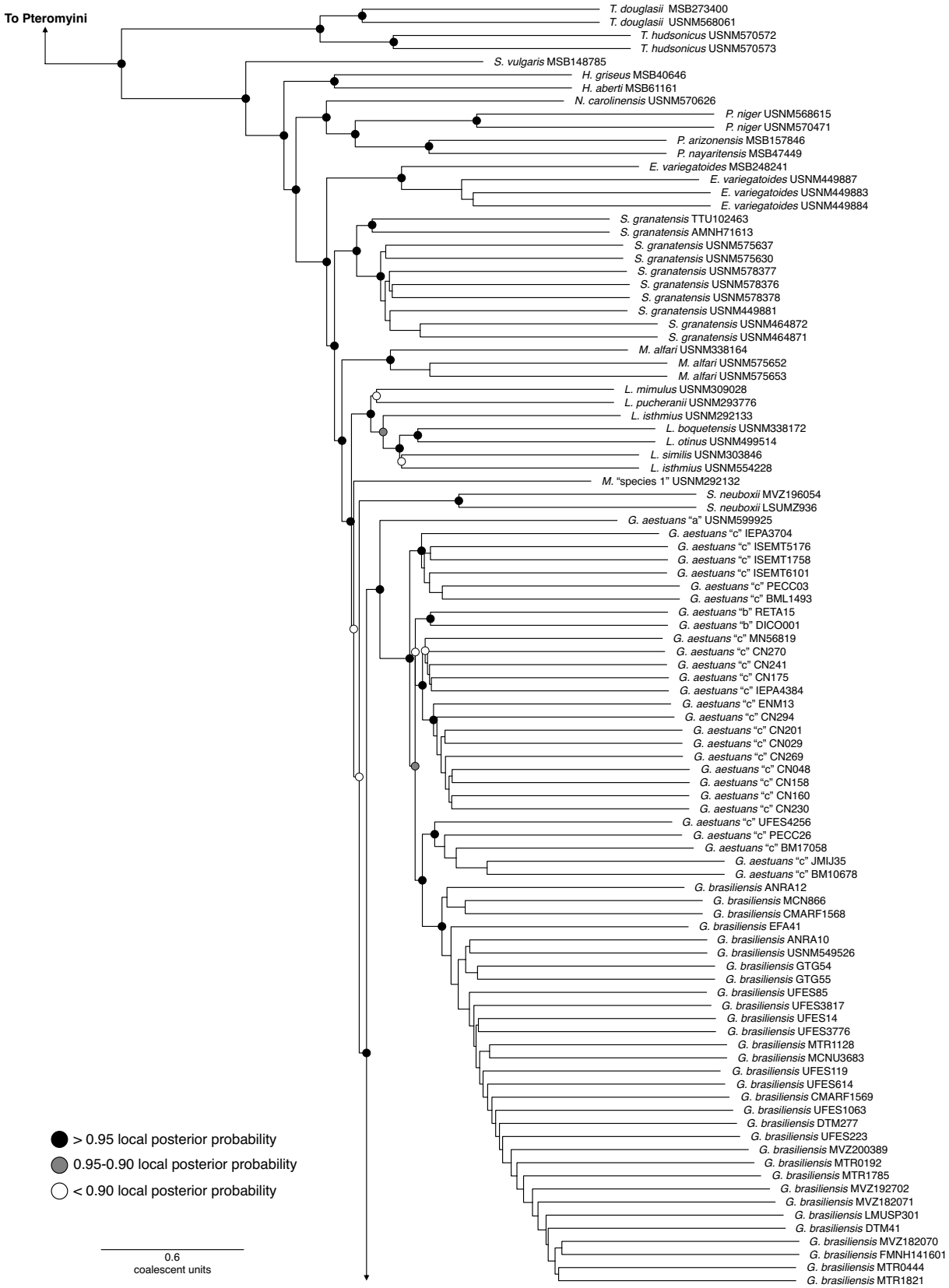
USNM 509015	Historical	1,677,587	5,352	351	278	97,560
USNM 329635	Historical	230,850	2,239	318	266	84,526
MPEG 1927	Historical	12,686,402	172,779	303	267	80,948
USNM 337731	Historical	740,001	1,664	289	267	77,036
USNM 255115	Historical	1,562,709	5,798	278	259	72,132
USNM 267562	Historical	2,432,528	6,143	274	263	71,964
AMNH 68149	Historical	459,788	2,162	273	266	72,729
AMNH 71608	Historical	2,512,080	11,850	255	274	69,821
USNM 48689	Historical	604,420	7,477	251	271	68,105
MUSM 7922	Historical	2,487,018	8,016	221	269	59,463
MUSM 7971	Historical	904,723	4,205	209	271	56,736
USNM 107930	Historical	119,842	1,616	179	267	47,798
USNM 275218	Historical	1,308,825	3,713	162	273	44,148
USNM 55607	Historical	150,977	1,766	124	262	32,545
AMNH 69237	Historical	171,132	583	106	266	28,154
USNM 528438	Historical	1,100,740	7,659	84	271	22,733
FMNH 170300	Modern	146,723	5,549	80	276	22,113
AMNH 32497	Historical	1,762,390	3,143	75	261	19,556
USNM 282755	Historical	172,401	307	67	257	17,201
USNM 113309	Historical	2,319,084	2,337	65	263	17,076
USNM 296782	Historical	230,975	2,115	60	261	15,632
USNM 499515	Historical	897,372	2,312	44	262	11,541
USNM 234276	Historical	302,962	2,118	37	273	10,108
USNM 545120	Historical	177,952	614	37	261	9,675
AMNH 11290	Historical	971,683	1,193	33	249	8,226
FMNH 170302	Modern	51,122	1,536	32	265	8,495
MUSM 23833	Modern	421,138	23,751	31	327	10,148
UFSC 5725	Modern	4,077	101	30	252	7,545
USNM 197262	Historical	4,872,424	13,059	30	465	13,957
USNM 249863	Historical	452,387	2,971	30	242	7,249
MCN 766	Historical	51,070	77	21	260	5,469
USNM 339054	Historical	1,769,210	35,385	21	262	5,503
FMNH 170301	Modern	28,563	1,145	18	269	4,834
USNM 152749	Historical	313,605	2,041	17	288	4,893
USNM 89297	Historical	44,489	441	14	246	3,439
USNM 152748	Historical	44,231	471	12	290	3,482
USNM 528437	Historical	149,148	234	12	250	2,994
USNM 177565	Historical	45,649	666	11	270	2,972
AMNH 14210	Historical	680,188	145	4	261	1,044
USNM 179564	Historical	193,812	320	2	246	491

Supplementary Table S2. (**continuation**)

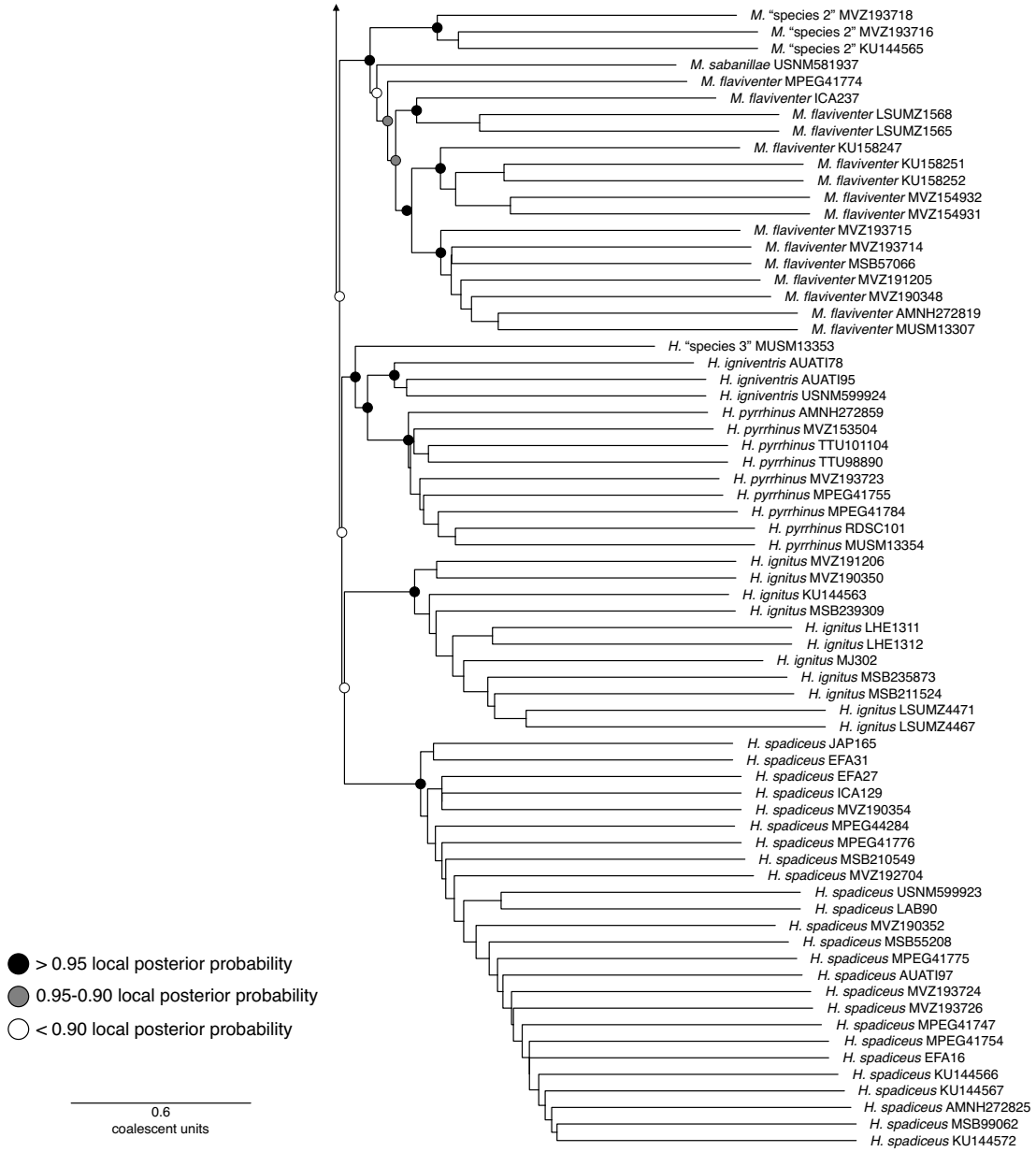
MUSM 7921	Historical	32,012	140	1	367	367
UFSC 274	Historical	52,497	33	1	320	320
MCN 1682	Historical	941,464	11,252	0	0	0
MCN 3340	Historical	738,453	5,597	0	0	0
MPEG 21808	Historical	210,188	63	0	0	0
USNM 168217	Historical	112,608	511	0	0	0
USNM 392859	Historical	65,371	8	0	0	0



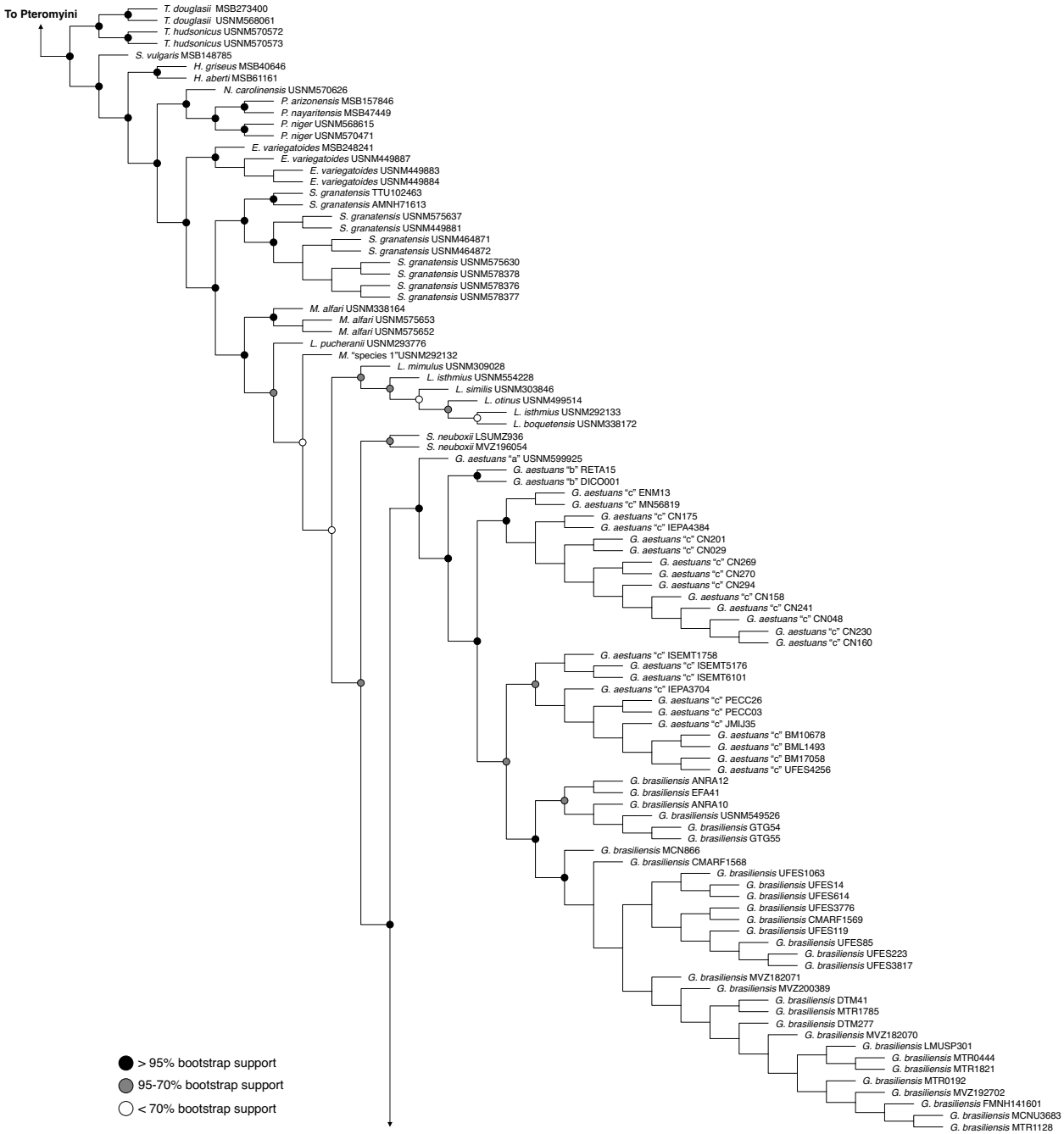
Supplementary Figure S1. Step-by-step workflow of our taxa sampling and alignment strategies, UCE data filtering approach, and inference methods used in this study.



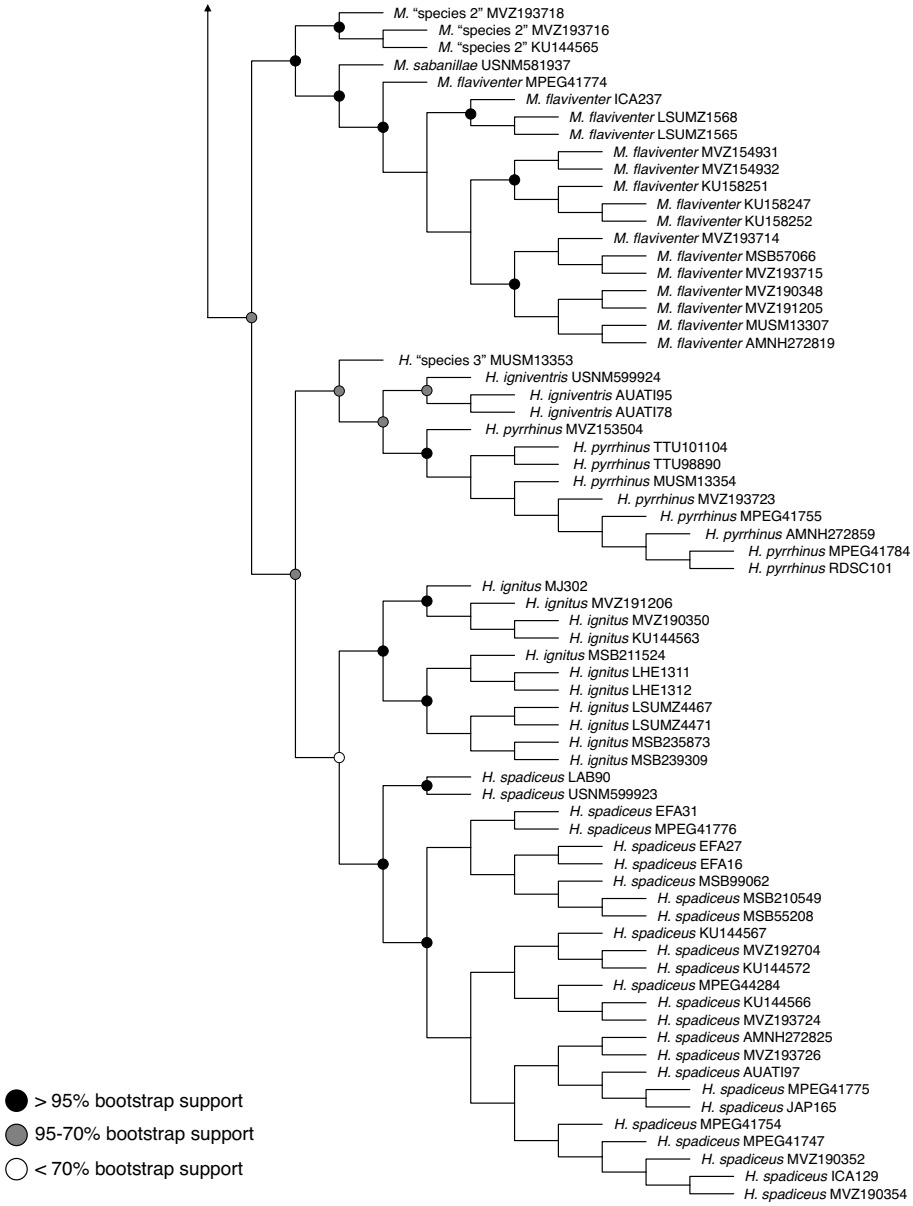
Supplementary Figure S2. (continued) (See legend on next page)



Supplementary Figure S2. Detailed phylogenetic hypothesis of the tribe Sciurini based on a coalescent analysis (ASTRAL-III) performed with a matrix of 3,713 UCE loci (50% taxa representativeness per locus and no filter for information content). Terminals are named with binomials (following Abreu-Jr et al. 2020b), accompanied by museum voucher numbers. Scale at the bottom represents coalescent unities.



Supplementary Figure S3. (continued) Detailed phylogenetic hypothesis of the tribe Sciurini based on a coalescent analysis (SVDquartets) performed with a matrix of 3,713 UCE loci (50% taxa representativeness per locus and no filter for information content). Terminals are named with binomials (following Abreu-Jr et al. 2020b), accompanied by museum voucher numbers.



Supplementary Figure S3. (continuation)

5. CONCLUDING REMARKS

The scarcity of preserved tissue samples in scientific collections, along with difficulties in assessing phylogenetically informative morphological traits, resulted in a long-lasting poor knowledge on the phylogeny of Neotropical squirrels. This scenario has not undergone conspicuous improvements even on the last 20 years —a period that testified huge changes on the systematics, evolution and biogeography of mammals due to the advance and popularization of molecular methods. As a consequence, highly discrepant taxonomic arrangements were proposed for the group over time. In the present thesis, the inclusion of historical samples was crucial to provide a comprehensive dataset, at both number of species and geographic coverage, that allowed the establishment of well resolved and supported phylogenetic hypotheses, which helped to elucidate several long-standing taxonomic, evolutionary, and biogeographic issues in the Neotropical squirrel radiations.

The recovery of mitochondrial genome data was not influenced by the sample age, and complete mitogenomes were obtained for samples as old as 120 years. The sampling of historical specimens enabled the addition of 18 nominal taxa —for which ethanol-preserved tissues were not available— in the mitogenomic analyses. On the other hand, the recovery of UCE loci and the mean length of the UCE loci were negatively impacted by the sample age. A regular mammal cell contains from a hundred to a thousand more copies of the mitochondrial DNA than of the nuclear DNA, which makes the sequencing of nuclear markers much more sensitive and prone to be affected by the sample quality. The difficult in obtaining nuclear data for historical specimens led to a reduction of the taxonomic sampling, but even so, the UCE phylogenetic inferences included 10 nominal taxa represented exclusively by historical samples. These results highlighted the potential of old museum specimens—including holotypes and other taxonomically important material— for phylogenetic and evolutionary inferences of elusive lineages of the Tree of Life.

The mitogenome phylogeny —which included 232 samples, representing almost all current valid species of tree squirrels— allowed the recognition of 12 species-groups well-supported within Sciurini. The comparison of the phylogenetic results with proposed classification schemes for the tribe illustrated that none of the taxonomic arrangements ever proposed fully corresponded to the phylogenetic structure reported, with only a few of the currently recognized genera recovered as monophyletic. Therefore, based on this

taxonomically inclusive mitophylogenomic hypothesis a preliminary and tentative nomenclatural designation for the taxa at the genus-group level was advanced, employing 13 names used in previous taxonomic classifications. By investigating the evolution of two morphological traits widely employed in the taxonomy of the group —number of upper premolars and number of pairs of mammae— it was clear that alternative conditions of both characters must have evolved multiple times throughout the evolutionary history of tree squirrels. This result helped to explain the incongruence between phylogenetic results and classificatory schemes presented through time.

UCE phylogenetic inferences largely supported the recognition of the genera proposed based on the mitogenomic data, excepted by *Microsciurus* (composed of *M. alfari* and *Microsciurus* "species 1") that was not recovered as monophyletic in the UCE analyses. However, relationships among some Sciurini genera, particularly those from South America, were not identically recovered by the mitogenomic hypothesis and the UCE inferences. As a matter of fact, conflicting relationships at both genus- and species-levels were estimated in the UCE inferences upon optimality criteria. Introgressive hybridization and incomplete lineage sorting are amongst the most plausible phenomena acting on the tree squirrel evolution and which could help to explain the mito-nuclear discordances observed for the rapid-diversified South American radiation.

Regarding the species recognition of tree squirrels, mitogenome- and UCE-based phylogenetic reconstructions revealed that most currently valid species are highly supported as monophyletic groups and only a few taxa appeared nested within others. Nevertheless, we found evidence that the diversity of Neotropical tree squirrels is currently underestimated. Species delimitation analyses performed with the mitogenomic dataset revealed at least six lineages that might represent taxa to be named or revalidated. Additionally, UCE inferences suggested two other lineages that merit taxonomic reassessment. In summary, this study hypothesize that the tribe Sciurini comprises 14 genera and at least 46 species —of which 43 species were sampled here and three were not included in the present study, but are provisionally treated as valid—, a more diverse estimate than recent catalogues. *Sciurus*, formerly the most diverse genus in the tribe, harbors only three species, while the genera *Leptosciurus* (with six species), *Hadrosciurus*, *Parasciurus* and *Echinosciurus* (all with five species each), are the most diverse within this radiation; the only monotypic genus is *Rheithrosciurus*. The Neotropical region harbors eight

genera and 29 species. However, it is important to mention that a detailed taxonomic investigation is necessary to carefully evaluate the applicability of the genus-level names and to provide diagnoses and or descriptions to them, as well as to evaluate the species-level taxonomy.

Analyses of UCE data estimated with strong support the relationship among the five currently recognized subfamilies of Sciuridae and provided consistent and well-supported results for the position of Sciurillinae. UCE inferences also showed two geographically structured clades within *Sciurillus pusillus*, suggesting that further taxonomic investigation is needed for this taxon. UCE datasets, as mentioned above, provided conflicting results for more recent and rapid branches of the squirrel tree, as the Neotropical radiation of Sciurini, whereas for deeper branches they proved to be effective and reliable to estimate relationships, regardless of filtering strategy and/or inference method. In general, UCE matrices including larger numbers of loci (matrices with 50 and 60% of taxa representativeness per loci) provided more similar topologies and higher values of nodal support. Filtering for information content (including only loci with a minimum of 5% of informative sites in the UCE datasets) did not increased the topological consistency nor improved the nodal support for any inference method. RAxML outperforms the coalescent methods, providing the topologies less dissimilar among each other and with the highest medians of nodal support.

It seems that, even employing a tremendous amount of genome-scale data to estimate phylogenies, for some recalcitrant clades we might never reach a phylogenetic consensus. In the case of the Neotropical tree squirrels, this may reflect the dramatic diversification of the group, that was rapid and much more complex than previously hypothesized. Biogeographic analyses, performed with the taxonomically inclusive mitogenomic dataset, estimated the origin of the Neotropical radiation to have occurred around 6 Mya in northwestern South America, and that the majority of Neotropical cladogenetic and speciation events occurred along the Pliocene, right after the South American invasion. Surprisingly, the peak of lineage accumulation detected in the Pliocene did not produce fluctuations in the diversification rate observed for tree squirrels, which may attest for the presence of several extinction events helping to equalize the number of lineages maintained over time.

This contribution employed modern techniques of genomic data generation and state-of-the-art conceptual and methodological approaches of data analyses to bring light to a charismatic but obscure group of rodents. There were consistent advances on the recognition of the genetic diversity and on the knowledge of the phylogenetic relationships among Sciuridae subfamilies and within the Neotropical radiations, Sciurillinae and Sciurini. Also, there was important progress on the taxonomy at both genus and species levels, on the optimization of some morphological traits, and on the biogeographic history of the tribe Sciurini. I am confident that these advances will represent a solid foundation for future research on this group, from population genetics to macroevolution, and for more clear assessments on the conservation of these unique sciurid radiations.