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

Evolutionary studies in South American marsh rats (Rodentia: *Holochilus*)

Joyce Rodrigues do Prado

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

Piracicaba
2017

Joyce Rodrigues do Prado
Bachelor degree in Biological Sciences

Evolutionary studies in South American marsh rats (Rodentia: *Holochilus*)

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

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

Piracicaba
2017

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

Prado, Joyce Rodrigues do

Evolutionary studies in south american marsh rats (Rodentia: *Holochilus*)/ Joyce Rodrigues do Prado. - - Piracicaba, 2017.

218 p.

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

1. Oryzomyini 2. Clado D 3. Roedores 4. Sequenciamento de nova geração 5. Genômica 6. Morfometria 7. Biogeografia I. Título

To my grandmothers, Nadir and Geny,
for always seeing me better than I am.

ACKNOWLEDGEMENTS

The opportunity to thank those who were with me during this journey is the most rewarding part of this document. Fortunately, a lot of people helped me go through these 52 months of intense scientific work, and looked after me during the most difficult parts. Thanks to these people, I grew up as a researcher and as a human being.

I would like first thanks my beloved parents, Maria Lúcia and José Otávio, for supporting me in all the ways a person can be supported, for teaching me life values, for teaching me who I am. Mom and Dad, thank you for being the best parents in the world!

I am grateful to my advisor, Alexandre Reis Percequillo, for being by my side during my entire scientific life, always with patience, comprehension, and kindness. Thank you for all the teachings, advice, for always pushing me to seek new challenges, and for being an example of moral and principles. I will be eternally grateful to you for introducing me to the difficult, but very rewarding, profession of scientist.

I also want to make a special acknowledgment to Dr Lacey Knowles, for opennig the doors of her lab, in the University of Michigan, and receiving me with so much attention. Lacey, the time I spent in your lab changed the way in which I see science, and opened my eyes to a new and beautiful scientific world. I will always be grateful for this opportunity!!!

Among all the people that were part of this journey, I have special thanks to my friends from the Laboratório de Mamíferos (LaMa). It's a real pleasure work with people that are more than colleagues. Eli, Gaúcho, Pam, Saldanha, Aninha, Jeronymo, Roth, Edgar, Napister, Pipok, Inae, Leandro, Lidiani, Mariana, and Vanessa, thanks for all the discussions, advice, laugh, life claims, coffee time, fieldwork, and for supporting all the anguish of a thesis ending. Eli, thanks for the amazing trip through Europe visiting the collections. Jeronymo thanks for your tutorials, they are awesome, and of course thanks for reading the innumerable final versions of this document. I also want to thanks Gaúcho and Eli for reading and commenting on this text several times. Napister, thanks for being so present even so far, your help with my first chapter was indispensable. Aninha, thank you for help me with the biogeographic analyzes.

I am also grateful for the opportunity to work with so special people at the Knowles' Lab. Tristan, JP, Lucy, Anna, Libin, Carlos, Qixin, Rob, Mariah, Melissa, Renata, Dea, and Luciana, thank you guys for help me with all the molecular techniques and bioinformatics pipelines, and for teaching me so many things that I can not even list.

I have a special acknowledgment to the life for bringing me Renata, Dea, and Luciana. Girls, thank you for all the Friday's nights working at the lab, for all the help with analyzes, for the friendship, and discussions about science and life. You girls are for sure amongst the best things that the PhD brought me!!!

I would like to thank all the different people that helped me acquire samples, including researchers, museum curators and staffs: Mario de Vivo, Juliana Gualda, João Alves de Oliveira, Robert Voss, Chris Conroy, Jim Patton, Darrin Lunde, Kris Helgen, Bruce Patterson, Priscila Tucker, Cody Thompson, Nora Lange, Christiane Denys, Roberto Portela Miguez, Paula Jenkins, David L. Reed, Robert M. Timm, Frederick H. Sheldon, Joseph A. Cook, Santiago F. Burneo, David Flores, Sergio Lucero, Alexandre Christhoff, Flavia Lagiglia, Enrique Gonzalez, Anna Ludmilla Nascimento, Fernando Perini, Caroline Lang, Miguel Trefaut, Francois Catzeflis, Leonardo Machado, Valéria Fagundes, Victor Colombi, Yuri Leite, Leonora Costa, Maurício Graipel, Jorge Cherem, Ligia Tchaicka, Viviane Layne, Rogério Rossi, Pablo Gonçalves, Marco Aurélio Sábato, Bárbara Costa, Ana Paula Carmignotto.

I always will be thankful to the Universidade de São Paulo (ESALQ/USP) for receiving me 13 years ago, as an undergrad student, and provide me the best environment to learn and grown as a professional and human being.

I am also indebted to Professor Maria Victoria Ballester, Giancarlo Conde Xavier Oliveira, Mara Casarin, and all the people from the PPGI-EA, thanks for all the administrative and bureaucratic support.

Professor Maria Victoria Ballester, thank you for accept to be the chair my evaluation committee.

I am grateful to Professor Alexandre Reis Percequillo, Jaime Bertoluci and the technician Gerson Romão for providing me a good infrastructure to work.

I would like to thanks my Thesis committee, Jaime Bertoluci, Giovanna Montingelli, Katia Ferraz, and Gabriel Marroig for all the support and help throughout the development of this work.

I especially thank the University of Michigan for the support, and also all the people there that for their generosity made my stay much more enjoyable: Robbin Murrell, Raquel Marchan, Jordan Bemmels, Pavel Klimov, Pamela Rojas, Susan Cheng, Anat Belasen and Jillian Myers.

A special thanks for the Brazilian community in Ann Arbor (MI), it was very important to feel Brazil close when I was abroad.

Thanks to the residents of the 820 Sylvan ave (Ann Arbor, MI), Anat, Erik, Melissa and Corey, for sharing a home with me and for being wonderful housemates.

I have special gratitude to Barbara Costa (Tafinha) and Gabriel Marroig for all the help analyzing my morphological and morphometric data. Tafinha, I would not have been able to finish chapter 1 without you!

I am extremely grateful to my family and friends, who share with me the best and the worst part of life. I will not dare list names here, but each of you know how luck I am for have you in my life. I would not have been able to get to this point without you! Thanks to all those who, even without understanding how academic life works, have always been by my side, supporting me and giving me

strength to move on. Also thanks to all those who know how academic life works and always encouraged me to continue and take one step further

Last, but not least, would like to thank FAPESP for the financial support (proc. FAPESP 2009/16009-1, 2012/24099-3 and 2014/22444-0).

"Go to nature.
Take the facts into your own hands.
Look, and see for yourself."

Louis Agassiz

SUMMARY

RESUMO

Estudos evolutivos dos ratos do brejo da América do Sul (Rodentia: *Holochilus*)

Uma abordagem interdisciplinar integrando micro e macroevolução, variação genômica, morfométrica e morfológica, sistemática, genética quantitativa e biogeografia foi empregada para investigar a história evolutiva do gênero *Holochilus* (Rodentia: Sigmodontinae). O gênero *Holochilus* apresenta espécies mal definidas, com problemas nomenclaturais e relações desconhecida. O número atual de espécies possivelmente não reflete a sua diversidade real e, até o momento, não foi realizado nenhum trabalho combinando evidências genéticas e morfométricas englobando toda a distribuição geográfica desse grupo. Este gênero pertence à tribo Oryzomyini, e juntamente com outros 14 gêneros (a diversidade genérica mais abrangente da tribo) formam o clado D. A relação filogenética interna dentro deste clado ainda é variável. Devido à sua ampla distribuição geográfica, *Holochilus* também representa uma peça chave no estudo da evolução dos oryzomíneos de formações abertas da América do Sul. Com base em uma amostragem abrangente, analisei padrões de variação morfométrica e genômica dentro de *Holochilus*, a fim de delimitar as espécies pertencentes a este gênero, bem como acessar a relação filogenética entre essas linhagens. Investiguei a variação sexual e ontogenética deste grupo, comparando populações naturais e de cativeiro, buscando entender o efeito das diferenças ambientais no padrão de variação e nas trajetórias ontogenéticas (Capítulo 1). Eu também avaliei e comparei a variação genômica entre três espécies de *Holochilus* a fim de verificar a influência dos biomas e das mudanças climáticas nas assinaturas genômicas das espécies (Capítulo 2). Em seguida eu apliquei uma abordagem baseada em modelos para delimitar as espécies (Capítulo 3). Finalmente, investigações adicionais foram realizadas para propor as relações filogenéticas entre os membros do clade D, fornecendo datas para os principais eventos de diversificação, e inferências sobre possíveis processos responsáveis pelo padrão biogeográfico atual, relacionado os mesmos com a ocupação florestal e áreas abertas (Capítulo 4). O dimorfismo sexual apresentou pequeno grau de variação entre as populações. A maior variação ontogenética é encontrada nas classes etárias mais jovens e mais velhas. Há também grandes diferenças nas trajetórias ontogenéticas entre as amostras, onde indivíduos da população cativeiro exibiram o menor grau de variação entre todas as classes etárias. A análise genética quantitativa mostrou que diferenças genômicas são observadas em todos os táxons e essa diferença está associada à geografia. Modelos de nichos ecológicos revelaram que os biomas com maiores áreas de estabilidade também apresentaram maior estruturação genômica, sugerindo que uma dimensão histórica impactou o isolamento/conectividade entre as populações. Os resultados também mostram que os biomas não só diferem geograficamente e ambientalmente (baseado em condições climáticas passadas), mas também mostram associação significativa entre o espaço ambiental e a variação genética que não está relacionada com a geografia. Adicionalmente, foi recuperado oito linhagens independentes dentro de *Holochilus*, e o arranjo filogenético parcialmente corrobora estudos anteriores. Finalmente, a filogenia proposta para o clado D apresentou algumas diferenças em comparação com outros estudos, e sugeriu que a maioria dos eventos cladogenéticos ocorreram durante o Pleistoceno, sendo a expansão dos ambientes abertos um importante motor de diversificação neste grupo.

Palavras-chave: Oryzomyini; Clado D; Roedores; Sequenciamento de nova geração; Genômica; Morfometria; Biogeografia

ABSTRACT

Evolutionary studies in the South America marsh rats (Rodentia:*Holochilus*)

An interdisciplinary approach integrating micro and macroevolution, genomic, morphometric and morphological variation, systematics, quantitative genetics, and biogeography was employed to investigate the evolutionary history of the genus *Holochilus* (Rodentia: Sigmodontinae). *Holochilus* presents poorly defined species, with nomenclatural problems and phylogenetic relationships on species level unknown. The current species number possibly does not reflect its real diversity, and no work combining genetic and morphometric evidences from all its geographic range was performed. This genus belongs to the tribe Oryzomyini, and along with other 14 genera constitute the Oryzomyini clade D, the most comprehensive generic diversity of the tribe, occupying distinct environments. The internal phylogenetic relationship within this clade is still unclear and variable. Due to its broad geographic distribution, *Holochilus* also represents a key piece on the study of the evolution of oryzomines of open formations of South America. Based on a comprehensive sampling, I analyzed patterns of morphometric and genomic variation within *Holochilus*, in order to delimit the species belonging to this genus, as well as access the phylogenetic relationship between these lineages. I investigated the sexual and ontogenetic variation in this group, comparing natural and captive populations, seeking for understand the effect of the environmental differences in the pattern of variation and ontogenetic trajectories (Chapter 1). I also evaluated and compared the genomic variation among three species of *Holochilus* to verify the influence of the biomes and the climatic changes in the genomic signatures (Chapter 2). I applied a model-based approach to delimit species (Chapter 3). And finally, additional investigations were made to propose the phylogenetic relationship between members of clade D, and provide date intervals for the main diversifications events, as well as the possible process responsible for the biogeographic pattern current observed related with the forest and open areas occupation (Chapter 4). Sexual dimorphism exhibited small degree of variation among populations. The greater ontogenetic variation is found in the younger age classes, but oldest individuals also show larger degree of differentiation. There are also great differences in the ontogenetic trajectories among samples, where individuals from the captive population exhibited the lower degree of variation between all age classes. The quantitative genetic analysis showed that genomic differences are observed across the taxa, and it was associated with geography. Ecological niche models revealed that biomes with larger areas of stability also presented more genomic structure, suggesting that historical dimension impacted population isolation/connectivity. Results also shows that biomes not only differ geographically and environmentally (based on past climatic conditions), but also show significant association between the environmental space and the genetic variation that is not related with geography. Eight independent lineages within *Holochilus* were recovered, and the phylogenetic arrangement partially corroborates previous studies. Finally, the phylogeny proposed for the clade D presented some differences in comparisons with other previously reported, and suggest that most of the cladogenetic events happened during the Pleistocene, being the expansion of open environments an important driver of diversification in this group.

Keywords: Oryzomyini; Clade D; Rodents; Next generation sequencing; Genomics; Morphometrics; Biogeography

1. INTRODUCTION

1.1. Systematic Studies and Evolution

"The world is so full of a number of things – that it can be extremely, even hopelessly, confusing". Simpson made this statement in his book about animal taxonomy over fifty years ago (1962:6), and even nowadays – after great improvements of methodologies to access, identify, describe and delimit biodiversity – this statement remains very actual.

The discipline concerned in studying the biodiversity is referred to as systematics (MAYR; ASHLOCK, 1991). The term *systematics* comes from the Greek word *systema* and it is defined as the scientific study of organisms, its diversity and the relationship between them (SIMPSON, 1962). Modern systematics is closely related with evolution, which relies on phylogenies. Evolution can be studied from two different perspectives, involving different time scales: the microevolution and the macroevolution. The basic tasks of systematic studies are inserted in these two approaches, sometimes delimiting species (microevolution), and sometimes describing patterns and process of speciation in different time scale (macroevolution; WILEY; LIEBERMAN, 2011).

The accuracy of the diversity estimation depends on different factors, among them: (1) the species concept applied; (2) the biology of the organisms; (3) the quantity and quality of data available; and (4) the analytical techniques employed (WILEY; LIEBERMAN, 2011). Probably one of the main controversies among who works with systematic is their different view on the nature of species (WILEY; MAYDEN, 2000). Most species concepts agree that species are lineages (MAYDEN, 1997; de QUEIROZ, 2005a, 2005b), but the discordance begins with the criteria that delimit those lineages. For example, some authors advocate that species must represent reciprocal monophyletic lineages (ROSEN, 1979; DONOGHUE, 1985), while other studies suggest that independent lineages can be delimited long before reciprocal monophyly (KNOWLES; CARSTENS, 2007).

Moreover, for the most species concept (e.g., the evolutionary species concept) there are still problems about how reasonably apply them, and what is needed to delimit different species confidently. This difficulty is, in many aspects, exemplified by the designation of subspecies. Subspecies is defined as "an aggregate of phenotypically similar populations of a species inhabiting a geographic subdivision of the range of that species and differing taxonomically from other populations of that species" (MAYR; ASHLOCK, 1991; pp 43). Many authors have used the term subspecies arbitrarily, for distinguishable entities that they considered less distinct than species. For example, Hershkovitz (1955) arranged all described *taxa* of the sigmodontine rodent genus *Holochilus* in a single species with 10 subspecies, justified by small morphological variation between them. According to Frost et al., (1992) if subspecies are discoverable (i.e., temporarily isolated sublineages), then they clearly are elements of evolutionary biology; if they are invented (i.e., artifacts of diagnoses

to fit our notions of process), they are an impediment to the formulation of historical hypotheses. Fortunately, nowadays it is possible to evaluate the statistical support for taxonomic differences between species and subspecies, throughout model-based statistical approaches. This can help to understand if these differences are justified biologically (HUANG; KNOWLES, 2016).

Consistent systematic research should employ several distinct datasets to recognize and define species or higher hierarchical groups – a procedure that has being employed for some mammalogists since the 1980's (CHIQUITO; D'ELÍA; PERCEQUILLO, 2014). Nowadays, it is a consensus that the best approach to access biodiversity is to assemble as much evidence as possible, in an integrative approach (EDWARDS; KNOWLES, 2014; CARSTENS et al., 2013). Although molecular data is being the main source of information for delimit species (both by the increase of the availability of DNA sequences and by the recent analytical developments; OLAVE; SOLÁ; KNOWLES, 2014), there is evidence that genetic methods alone can describe diversity inadequately (HARRINGTON; NEAR, 2012; CARSTENS et al., 2013). New methods, such as the multi-locus coalescent species delimitation, the multivariate and clustering techniques (that uses genetic, morphological and ecological data), the ecological niche modeling (ENM), or even the multivariate tolerance regions to test for discontinuities or gaps in morphology, have been recently used in an integrative way in new taxonomic studies (EDWARDS; KNOWLES, 2014). As preconized by Mayr, Simpson and other luminaries of systematics, as much information about the biology of the target group (such as life history, geographical distribution, morphology and behavior) is available the more supported is the species delimitation.

Besides the need of integration between different data sources, we also have to cope with limitations on sample size, type of markers and the methods to integrate them to establish hypothesis. Sampling is a critical issue in evolutionary studies, and can lead to ambiguous, inconclusive, or negative results when it is not appropriated (NAZARENO et al., 2017). Species delimitation based solely on morphology has rarely been used after the advances of the molecular biology. In some morphological studies, species are recognized by discontinuous and non-overlapping patterns of variation in characters presented in individuals from geographically distinctly populations (MAYR; ASHLOCK, 1991). To follow this workflow a good coverage on the presumed organism distribution, as well as enough sample size per operational unit is crucial for a high statistical power. However, Simpson et al. (2003) stated that the ideal sample need to be adequate, homogeneous, and unbiased, they belive that in zoology the sample size is usually fixed in practice by what we can obtain. Regarding to molecular approaches, we are also limited by sample availability (most important historical series available at the most representative museums throughout the world do not have tissue samples, including the type series). Moreover, Olave et al. (2014) shows that purely genetic methods require significant amounts of data (well beyond that which has typically been used) to delimit rapid radiations diverging along ecological and morphological aspects. Also, it is important to find an

optimal range of sampling loci effort, individuals, and sequence lengths for a given speciation history (CAMARGO et al., 2012).

These points highlight the limitation of traditional genetic markers (e.g., small fragments of mtDNA or a few nuclear loci) in the resolution of phylogenetic relationships; this limitation may also obstruct our skills to distinguish between alternative evolutionary scenarios. However, the vast amounts of genome wide polymorphism data that now can be generated can provide the requisite resolution for addressing fine-scale phylogeny structure (WAGNER et al. 2013). Genomic data together with the recent development of quantitative methods can bring objectivity and reproducibility into species delimitation (CAMARGO; SITES 2013; OLAVE et al., 2017).

Among several methodologies to access genomic data, there is the double digest restriction associated DNA sequencing (ddRADseq; PETERSON et al., 2012). This technique produces abundant and anonymous data from the whole genome, highly useful for phylogenetic inference (EATON; REE, 2013; HIPP et al., 2014). Among the many advantages of the ddRADseq are the relatively low costs, minimal required starting material, and independence of a reference genome (SHAFER et al., 2016). The ddRADseq has being broadly used to standard population genomic approaches, and demographic and phylogenetic analyses (e.g. TAKAHASHI; NAGATA; SOTA, 2014; BEMMELS et al., 2016; MASSATTI; KNOWLES, 2016; SHAFER et al., 2016).

Collecting, generating and analyzing the necessary amount of genetic data (not mention others types of data as ecological, karyological, etc.) are financial and time consuming (even with the decrease of these resources in the Next generation Sequencing, NGS). Additionally, the process of analyzing huge amount of information is still a challenging, and regarding to the genetic data, most of them are computationally demanding and some methods cannot be applied to phylogenomic datasets yet (DOMINGOS et al., 2017). However, the development of methodologies to analyze this kind of data is increasing fast and several programs, specifically designed for analyzing NGS data, are being produced.

Finally, it is important to highlight that the importance of systematics studies extends beyond immediate systematics goals, as species delimitation and descriptions of patterns and process of speciation. Systematic and taxonomic treatments have profound implications for other fields, ranging from studies in evolution and ecology to biogeography and conservation biology, which makes the systematics an interdisciplinary field of research. The new evolutionary approaches are currently implemented in various disciplines such as the developmental and psychological sciences, the sociocultural and linguistic sciences, physics or medicine, and in turn necessitate scholars to cross-disciplinary boundaries. Moreover, evolutionary studies integrated with species distribution models, demographic and ancestral areas analyzes, can provide a firmer foundation for management and conservations priorities.

1.2. Model Organism

1.2.1. Subfamily Sigmodontinae

Assembling 84 genera and 390 species (D'ELIA; PARDIÑAS, 2015), the subfamily Sigmodontinae is part of complex and diverse groups of mammals that correspond to about 10% of the mammal species in the world and about 20% of the rodents (WILSON; REEDER, 2005). The sigmodontine rodents are distributed predominantly in South America, but also occur in Central and North America (MUSSER; CARLETON, 2005).

The current taxonomic composition of this subfamily is the result of intense reformulation in recent decades due to a large volume of published data (e.g., PATTON; HAFNER, 1983; CARLETON; MUSSER, 1989; VOSS, 1991, 1993; VOSS; CARLETON, 1993; SMITH; PATTON, 1993, 1999; MUSSER et al., 1998; CARLETON; OLSON, 1999; PATTON; SILVA; MALCOLM, 2000; LANGGUTH; BONVICINO, 2002; D'ELÍA, 2003; D'ELÍA; GONZALEZ; PARDIÑAS, 2003; WEKSLER, 2003, 2006; GÓMEZ-LAVERDE; ANDERSON; GARCÍA, 2004; MUSSER; CARLETON, 2005; WEKSLER; PERCEQUILLO; VOSS, 2006; PERCEQUILLO; HINGST-ZAHER; BONVICINO, 2008; PERCEQUILLO; WEKSLER; COSTA, 2011), but although this composition is relatively stable, the phylogenetic relationships within the subfamily remains uncertain (D'ELÍA; PARDIÑAS, 2015).

Over the past 15 years, phylogenetic analyzes employing both morphological and molecular characters have grouped sigmodontines in different tribes (e.g., BRAUN, 1993; VOSS; CARLETON, 1993; SMITH; PATTON 1993, 1999; STEPPAN, 1995; D'ELÍA; GONZALEZ; PARDIÑAS, 2003; D'ELÍA; PARDIÑAS; MYERS; 2005; D'ELÍA et al., 2006a, 2006b; PACHECO, 2003; WEKSLER, 2003, 2006; PERCEQUILLO; WEKSLER; COSTA, 2011; PARDIÑAS; TETA; SALAZAR-BRAVO, 2015; SALAZAR-BRAVO et al., 2016), causing many reconsiderations on their boundaries and contents (see D'ELÍA; PARDIÑAS, 2015 for a review). Currently there are eleven recognized tribes, namely, Abrotrichini, Akodontini, Andinomyini, Euneomyini, Ichthyomyini, Oryzomyini, Phyllotini, Reithrodontini, Sigmodontini, Thomasomyini, and Wiedomyini (sensu SALAZAR-BRAVO et al., 2016).

The tribe Oryzomyini is the most diverse within the radiation of Sigmodontinae subfamily, with 34 genera and 130 species, and also the most widely distributed (WEKSLER, 2015). In this tribe, most genera can be classified into three general patterns of distribution – Trans-Andean, Andean and Cis-Andean – and through these areas they inhabit forests, savannas, swamps, grasslands and semi-arid environments (MUSSER; CARLETON, 2005; WEKSLER; PERCEQUILLO; VOSS, 2006; PRADO; PERCEQUILLO, 2013). Accordingly to Hershkovitz (1962, 1966, 1972) the tribe had only representatives of forest environments with pentalophodont molars (with five ridges or lophos), but

subsequently Voss and Carleton (1993) included in the tribe open dweller genera with molars with four ridges or lophs (tetralophodont).

Phylogenetic studies on the tribe Oryzomyini expanded considerably over recent years (WEKSLER, 2003, 2006; TURVEY et al., 2010; PERCEQUILLO; WEKSLER; COSTA, 2011; PINE; TIMM; WEKSLER, 2012). Although its composition, which was originally defined and diagnosed by Voss and Carleton (1993), is often corroborated by phylogenetic studies (STEPPAN, 1995; WEKSLER, 2003, 2006; D'ELÍA et al., 2007; PERCEQUILLO; WEKSLER; COSTA, 2011; PINE; TIMM; WEKSLER, 2012), the relationships among the four major Oryzomyini clades, A, B, C and D (see WEKSLER, 2006) and among the genera within these clades are not clear. Additionally, there is still some variation in the number of species in each genus (WEKSLER, 2006; PERCEQUILLO; WEKSLER; COSTA, 2011; PINE; TIMM; WEKSLER, 2012), and there is also some doubts about the existence and position of *Microakodontomys* and the alfaroi-chapmani-melanotis group that may be erected as a new genus (WEKSLER, 2015). Regarding the species group exists a large volume of data being described and revaluated (MUSSER et al., 1998; LANGGUTH; BONVICINO, 2002; BONVICINO, 2003; EMMONS; PATTON, 2005; PERCEQUILLO; HINGST-ZAHER; BONVICINO, 2008; TAVARES; PESSÔA; GONÇALVES, 2011; PARDIÑAS et al., 2013; SILVA et al., 2015; UTURUNCU; PACHECO, 2016; PRADO; PERCEQUILLO, 2017), and the relationships among taxa of species group are uncertain and not adequately addressed, as well as the time and diversification mechanisms are inconclusive for most genera and species (STEPPAN, 1995; WEKSLER, 2006; PARADA; D'ELÍA; PALMA, 2015).

The clade D of the tribe Oryzomyini is a typical example of the issues presented previously. This clade has the most comprehensive generic diversity of the entire tribe, comprising 15 genera and 41 species. Their representatives are distributed throughout South America, inhabiting both forest environments and open areas (PRADO; PERCEQUILLO, 2013). The clade D is also the clade with largest number of monotypic genera in the tribe; there are seven genera with only one species (WEKSLER, 2015). Is interesting to note that most of these monotypic genera are forest habitants, except by *Lundomys* and *Pseudoryzomys*. Within the more diversified taxa of clade D, the most genera are open areas dwellers except by *Melanomys* and *Nectomys* (PRADO; PERCEQUILLO, 2013). In the most recent phylogeny for clade D (MACHADO et al., 2013) the tetralophodont genera of clade D appears as paraphyletic, but these oryzomyines have usually been recovered as a monophyletic group in previous studies (CARLETON; OLSON, 1999; WEKSLER, 2003, 2006; PERCEQUILLO; WEKSLER; COSTA, 2011; PINE; TIMM; WEKSLER, 2012). Actually, the sister group relationship of the genus *Holochilus* varies depending on the gene and the method used in the phylogenetic reconstruction (MACHADO et al., 2013).

Holochilus is one of the tetralophodont genera that belong to clade D. This genus has a confusing taxonomy and not established phylogenetic history. Although currently allocated unambiguously in the tribe Oryzomyini (WEKSLER, 2003, 2006; TURVEY et al., 2010;

PERCEQUILLO; WEKSLER; COSTA, 2011; PINE; TIMM; WEKSLER, 2012) in the recent past (REIG, 1986) this genus was allocated in tribe Sigmodontini along with *Sigmodon*, a pastoral genera that shares with *Holochilus* simplified tetralophodont molars. Moreover, its phylogenetic position regarding other genera of the tribe Oryzomyini is still controversial. For example, in WEKSLER (2003, 2006) phylogenetic analyzes, *Holochilus* is represented by *H. brasiliensis* and *H. chacarius* and was defined as a member of clade D, appearing as sister group of *Lundomys* and *Pseudoryzomys*. Another previously constructed phylogenetic hypothesis for the Oryzomyini tribe (CARLETON; OLSON, 1999) shows the genus *Noronhomys* (now extinct) as sister group of *Holochilus,* and the genus *Lundomys* more associated with them than with *Pseudoryzomys*. In the phylogenetic analysis conducted by Voss and Carleton (1993) based on morphological data, two equally parsimonious trees were obtained, with *Pseudoryzomys* or *Lundomys* also appearing as sister group of *Holochilus*. The phylogenetic tree recovered by Steppan (1995) also supports the close relationship between *Holochilus* and *Pseudoryzomys*, but in this analysis *Holochilus* is sister group of *Zygodontomys* and this clade is sister group of *Pseudoryzomys,* and *Lundomys* not entered in the analysis. The relationship among these three genera is also variable in more recent studies (WEKSLER, 2003, 2006; TURVEY et al., 2010; PERCEQUILLO; WEKSLER; COSTA, 2011; PINE; TIMM; WEKSLER, 2012; MACHADO et al., 2013).

1.2.2. The genus *Holochilus*

The rodents of genus *Holochilus* (Fig. 1) are oryzomines of large body size, semi-aquatic, and mainly inhabiting marshes, swamps, grasslands, wetlands and other open areas of South America, with a predominantly herbivorous diet (MASSOIA 1971, 1976; FORMOSO; SAUTHIER; PARDIÑAS, 2010).

The weight of adult specimens can range from 160-455g, the head and body size from 130-220mm, and the tail length from 115-183mm. The dorsal coloration is tawny to brownish and the ventral pelage varies from ochraceous to white. They present interdigital pads and natatory fringes. The skull is heavily built, with a short and broad rostrum and squared braincase. The mandible is robust and deep. Upper incisors are opisthodont, and the molars are flat crowned and moderately hypsodont (GONÇALVES; TETA; BONVICINO, 2015). *Holochilus*, also has a robust single pair of preputial glands, with baculum slightly shorter than glands (HOOPER; MUSSER, 1964). The stomach is hemiglandular-unilocular (VOSS; CARLETON, 1993), the cecum is large (GONÇALVES; TETA; BONVICINO, 2015) and a gall bladder is absent (DOMÍNGUEZ-BELLO; ROBINSON, 1991).

Based on morphological (MASSOIA, 1981; VOSS; CARLETON, 1993) and molecular features (D'ELÍA et al., 2015) *Holochilus* is divided in two main distinct groups: the *brasiliensis* group and the *sciureus* group.

The *brasiliensis* group comprises large animals with dense and soft fur, tail as long as head-and-body, hindfoot with no hypothenar pad on the plantar epithelium, eight pairs of mammae (one inguinal, one abdominal, one postaxial, and one pectoral), distinct postorbital ridges, a long palate (sense HERSHKOVITZ, 1962), subsquamosal fenestra always distinct and patent, masseteric crests forming a long single crest that extends beyond the anterior root of M1 and margin the mental foramen, molars with the main cusps arranged in an intermediate pattern between an opposite and alternated configuration, the hypoflexus of M1 and M2 extend deeper beyond the midline, and the presence of vestigial mesoloph/id structures in M1 and M2.

The *sciureus* group includes medium-sized animals with short and close fur, tail shorter than head-and-body, hindfoot with a very small hypothenar pad sometimes present, either eight or ten pairs of mammae (one inguinal, one abdominal, one postaxial, one pectoral, and sometimes one thoracic), subsquamosal fenestra often obstructed by an expanded hamular process or by an internal crest or septum of the periotic, a mesopterygoid fossa often extending to or slightly between the third molars, less pronounced postorbital ridges, masseteric crests coalescing in a single crest in some populations or coalescing at the level of the anterior root of M1 in others, molar with the principal cusps arranged in alternating pattern, hypoflexus of M1 and M2 extend beyond the midline, but no as deeper than in the brasiliensis group, mesoloph/id typically absent in all molars, and the presence of a expanded paracone.

Holochilus presents morphological specializations to the semi-aquatic life, as strong bones, broader muscular ligaments, short femur and long posterior foot, fusiform body and webbed feet. Strong bones benefit semiaquatic animals as they allow supporting body size and muscle stress caused by the difficulty of moving in water (SAMUELS; VAN VALKENBURGH, 2008; GANDINI, 2014).

Regarding to reproduction, *Holochilus* starts mating with three months of life, the fertile period of females lasts about four days and the gestation lasts on average 21 days (LIRA et al., 2016) with approximately 2 to 10 per litter (TWIGG, 1962; LIRA et al., 2016). In addition, results found for breeding of *Holochilus* in captivity, seem to indicate that in a favorable environment, with sufficient space for survival and abundant food, there is no time frame for its reproduction (KAWAZOE; PINTO, 1983). These rodents often make nests 20 to 50 cm above the water surface, with a helicoidal form, and the individuals gnaw material, using big leaves as an external cover and stem fiber and some grasses in the inside to build the nests. The nests have a mean size of 22.5 cm length, 13 cm width and 18.7 cm height (SAUTHIER; ABBA; SAUTHIER, 2010).

Holochilus has a diet consisting basically of leaves and other items of plant origin (HERSHKOVITZ, 1955), and, frequently attack sugar cane, corn and rice fields, causing great damages to these crops (TWIGG, 1962; MASSOIA, 1974; MARTINO; AGUILERA, 1989). The herbivorous diet is confirmed by the gut morphology, that according to Domínguez-Bello and Robinson (1991) suggest great specialization for an herbivorous diet. *Holochilus* lost the gall bladder, which can be attributed to the low fat content of plant diets and presents a high metabolic rate,

surpassing data for other placental mammals (DOMÍNGUEZ-BELLO; ROBINSON, 1991). The mesoloph/id (crest connected to the median mure localized between the paracone and metacone in the upper molars and between the metaconid and entoconid in the lower molars) is reduced or lost, and this trait together with the lamination of molar cusps and a reduction in the number of molar folds (condition designate as tetralophodont) reinforce the herbivorous specialization of the genus (HERSHKOVITZ, 1962).

Figure 1. External morphology of genus *Holochilus*. Note the long and dense pelage, the venter clearly distinct from dorsal and lateral pelage, the very large hind foot, the long tail, the short and densely haired pinnae Pictures took by P.R.O. Roth in Rio Grande do Sul state (Brazil).

Currently this genus includes six valid species (sensu GONÇALVES; TETA; BONVICINO, 2015), namely *H. brasiliensis*, *H. sciureus, H. vulpinus, H. venezuelae, H. chacarius* and *H. vulpinus. H. brasiliensis* occurs in southeastern and southern Brazil, from Espírito Santo state to coastal Rio Grande do Sul state, including the eastern border of Minas Gerais, and west into the Humid Chaco Paraguayan; *H. chacarius* is distributed through the lowlands of the Paraguayan and northern Argentinean Chaco until the Brazilian boarder; *H. lagigliai* is known only from late Holocene records in the Argentinean Mendoza Province (FERNÁNDEZ et al., 2016); *H. sciureus* occupies the lowlands of the northeastern, central and northern part of Brazil, Guyanas, Bolivia and Peru; *H. venezuelae* occurs from the lower Orinoco basin to Lago Maracaibo; *H. vulpinus* is distributed through Paraguay and east-central part of Argentina throughout Uruguay and Rio Grande do Sul state in Brazil (GONÇALVES; TETA; BONVICINO, 2015; Fig. 2).

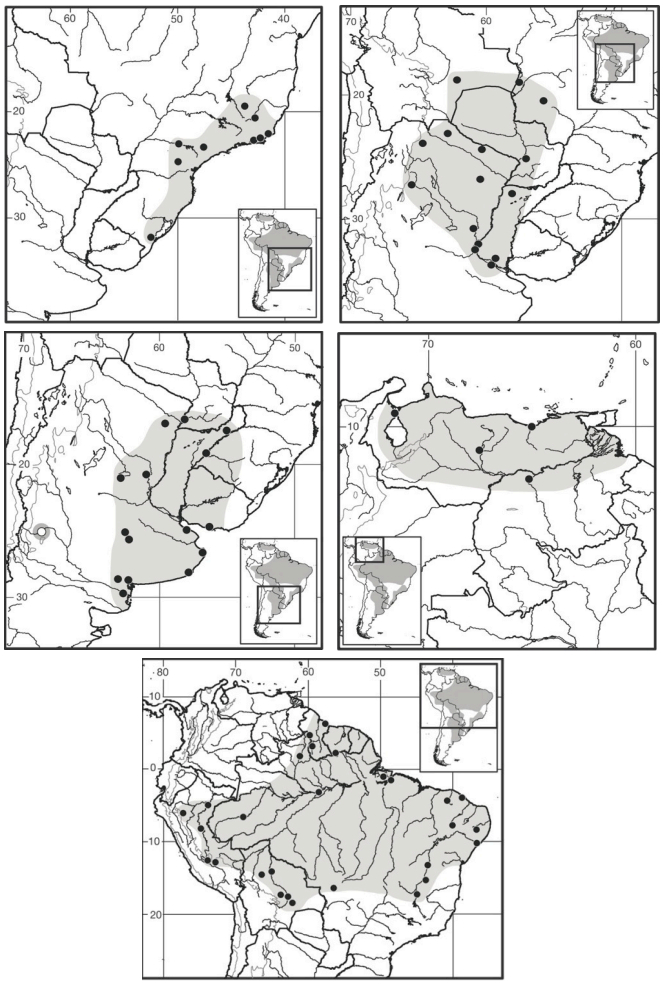


Figure 2. Known collection localities of the current valid species of genus *Holochilus*. First map shows *H. brasiliensis*, the second map shows distribution points of *H. chacarius*, in the third map are presented the geographic range of *H. vulpinus* (black dots) and *H. lagigliai* (white dot), the fourth map shows data from *H. venezuelae*, and the last one presents the distribution of *H. sciureus*. Maps were reproduced from Gonçalves et al., (2015).

Besides the broad current geographic distribution, *Holochilus* also has a wide temporal distribution, with considerable fossil records in South America, ranging from the middle Pleistocene until recently. In this fossil records (see PARDIÑAS; D'ELÍA; ORTIZ, 2002) there are samples appointed as *H. brasiliensis*, located in east-central Argentina and Brazil, samples denominated as *H. chacarius*, located in northern Argentina and southeastern Bolivia and also the recent publish data records for *H. lagigliai* in Mendoza (Argentina; FERNÁNDEZ et al. 2016).

1.2.3. Taxonomic History

Desmarest (1819) described the first specimen allegedly associated with the genus *Holochilus* under the name of *Mus brasiliensis*, based on a specimen collected by Auguste F. C. P. de Saint-Hilaire, which was donated to the Musee d'Histoire Naturelle de Paris. However, Desmarest in

the description of the species credited the name to E. Geoffroy. Next year, Desmarest (1820) suggests the similarity between *Mus brasiliensis* Desmarest 1819 and *Mus angouya* Fischer, 1814 (currently *Sooretamys angouya*) and treats them as synonyms.

Later, Brants (1827) described *Mus vulpinus*, crediting the name to Lichtenstein, which later also made a description of this species in the "Einleitung zum siebenten Heft" (1830), and gave to Brants the credit of naming the species.

Subsequently Brandt (1835) proposed the name *Holochilus* and placed it as subgenus of *Mus*, describing within this genus the species *Mus (Holochilus) leucogaster* and *Mus (Holochilus) anguya*. The later was applied for the Rat Angouya of Felix de Azara (1802) and *Mus angouya* Fischer, 1814, that like Desmarest, 1820 (see MUSSER et al., 1998), he believed to be a *Holochilus*.

Waterhouse (1839) identified a rodent collected by C. Darwin at Bahia Blanca, Argentina, as *Mus brasiliensis*, thinking to be similar to the specimen described by Desmarest in 1819; he stated that he had examined what he believed to be the original specimen of *brasiliensis* in Paris, and was careful to point out that *Mus angouya* Fischer 1814 (=*Sooretamys*), misinterpreted as *Mus brasiliensis* Desmarest 1819 by some authors, was a very different animal.

Wagner in 1842 raised *Holochilus* to the generic category and described a new species, *H. sciureus*. Furthermore he assumed that *Mus brasiliensis* (DESMAREST, 1819) would belong to *Holochilus*. In 1843, the same author decided to replace the name *Mus (Holochilus) anguya* Brandt (1835) by a new name, *H. canellinus*, because he believed that *Mus (Holochilus) anguya* Brandt 1835 did not represent the same entity as the *Mus angouya* Fischer (1814). Moreover, Wagner treated again *Holochilus* as a complete genus, listing in this new concept of the genus, the species *H. brasiliensis* Desmarest (1819), *H. leucogaster* Brandt (1835), *H. canellinus* Wagner 1843, *H. sciureus* Wagner 1842 and *H. vulpinus* Brants (1827).

Further up *Holochilus* was considered subgenus of *Hesperomys* by Burmesiter (1954) who described a new species, *H. robustus*, and also included in this taxon, the species *H. squamipes* Brants, (1827; in reality nowadays a species of *Nectomys*), *H. physodes* Brants (1827; synonym of *Euryoryzomys russatus*) and *H. vulpinus* Brants (1827).

Brandt (1855) in his classic monograph of rodent classification explained that he extracted the skulls from the specimens described by him in 1835 [*Mus (Holochilus) leucogaster* and *Mus (Holochilus) anguya*] and came to the conclusion that in fact those specimens were histricognats; recognizing its own error in relation to the comparison with the *Mus angouya* of Desmarest, Brandt proposed the name *H. langsdorffii* for the taxon he previously called *Mus* (*Holochilus) "anguya",* and classified *Holochilus* in the Spalacopoides family of the suborder Hystricomorphi. For the other names of myomorph rodent associated with *Holochilus* so far (*brasiliensis, vulpinus* and *sciureus*) Brandt proposed a new genus *Holochilomys*, which he placed in the Myoides family and suborder Myomorphi. However this action was neglected by most of his contemporary mammalogists. Only Peters (1861) used the name *Holochilomys* (*Holochilus*, BRANDT 1835) as Brandt suggested, but

without any bibliographic reference to the name. Palmer (1904) suggested that *Holochilomys* Peters (which is actually Brandt) must be an amendment of *Holochilus* Brandt 1835. According to Voss and Abramson (1999) *Holochilomys* seems to be a *nomen nudum*, which was not formally described and also not mentioned for decades, being the last reference of this name in the mammalogy literature in Cabrera (1961), who listed *Holochilomys* Peters, 1861 (which is actually BRANDT, 1855) as a junior synonym for *Holochilus*.

Within that time interval Lilljeborg in 1866 raised again *Holochilus* to the generic category. Thomas (1882, 1884) considered *Nectomys* as a subgenus of *Holochilus* and later, Thomas (1897) described a new species called *Holochilus nanus*, adding that *Holochilus* and *Nectomys* would be different genera. Still in this same work he attributes the name *H. darwini* to the specimens from Bahia Blanca, which were identified tentatively by Waterhouse (1839) under the name *H. brasiliensis* Desmarest (1819; see above).

Trouessart (1898) moved the genus *Megalomys* to *Holochilus*, as a subgenus of the later and suppressed Brandt's *M. anguya* in favor of Wagner's H. *canellinus*. According to Tate (1932), Trouessart's decision probably stems from the idea that *Mus (Holochilus) anguya* of Brandt was pre-occupied by *Mus angouya* of Desmarest (Tate, 1932). However, since the homonym definition rule does not apply to *ou* or *u*, *Mus angouya* Desmarest can not be considered a concern for the name *Mus anguya* Brandt, and *Oryzomys angouya* (=*Sooretamys angouya;* see CHIQUITO; D'ELÍA; PERCEQUILLO, 2014) and *Holochilus anguya* are valid names connoting perfectly distinct animals (TATE, 1932).

Posteriorly, Thomas (1901) described *H. guianae* for specimens distributed between the range distribution of *H. sciureus* and *H. nanus.* Next year Miller and Rehn (1902) designed *leucogaster* Brandt as the type of genus *Holochilus*. Allen in 1904 studying mammals collected in Colombia and Venezuela described *H. venezuelae*. Trouessart (1905) removed *Megalomys* of *Holochilus*. In 1906, Thomas publishes a paper with notes about some rodents and describes two new species *H. chacarius*, from Concepcion (Paraguay) and *H. balnearum* from Tucuman (Argentina) .

Nine years later, Osgood (1915) described a new species based on the material collected in Peru and Brazil, *H. amazonicus*. Again, Thomas in 1921, based on the material collected by Mr. Edmund Heller, also from an expedition in the region of Cuzco in Peru, described *H. incarum*. In 1937, Morrison-Scott describes a subspecies, *Holochilus sciureus berbicensis*, from Guyana.

Hershkovitz (1955), in the only revisionary study of *Holochilus* put together all the species previously described in a single species, *H. brasiliensis*, recognizing 10 subspecies (table 1), and described a new taxon, *H. magnus* for specimens from eastern Uruguay. Cabrera (1961) considers the exact same taxonomic arrangement proposed by Hershkovitz (1955). Posteriorly Voss and Carleton (1993) synonymized *H. magnus* of Herskovitz with *Hesperomys molitor* of Winge (1887) and allocated this species in a new genus, *Lundomys*, with *H. molitor* as species-type and only valid species. Steppan (1996) described a fossil species inside *Holochilus*, that he called H. *primigenus*,

from Tarija Basin in Bolivia, further on time Machado et al., (2013) proposed a new generic name for this species, *Reigomys*.

More recently, Voss and Abramson (1999) with the propose of conserve the name *Holochilus,* elected *H. sciureus* as a type species for the genus *Holochilus*, and moved *H. leucogaster* (the former type species designated by MILLER; REHN, 1902 that was known to be a hystricomorphous and referable to *Proechimys*) to the genus *Proechimys.*

Musser and Carleton (2005) consider three valid species: *H. brasiliensis* Desmarest, 1819 [including *anguya* (BRANDT, 1835); *canellinus* Wagner, 1843; *darwini* Thomas, 1897; *leucogaster* (BRANDT, 1835); *russatus* (WAGNER, 1848); and *vulpinus* (BRANTS, 1827)], *H. chacarius* [including *balnearum* THOMAS, 1906] and *H. sciureus* [including *amazonicus* OSGOOD, 1915; *berbericensis* MORRISON-SCOTT, 1937; *guianae* THOMAS, 1901; *incarum* THOMAS, 1920; *multannus* AMEGHINO, 1889; *nanus* THOMAS, 1897; and *venezuela* J. A. ALLEN, 1904].

Pardiñas et al., (2013) based on material collected by Humberto La Giglia, director of the Museo Municipal de Historia Natural de San Rafael (MHNSR, San Rafael, Mendoza, Argentina) described the most recent species for the genus, *Holochilus lagigliai*, from the Mendoza region.

Gonçalves et al., (2015) based in the extensive cytogenetic variability recognized six species within *Holochilus*: *H. brasiliensis, H. chacarius, H. sciureus, H. venezuelae, H. vulpinus,* and *H. lagigliai*. In the same year D'Elía et al. (2015) based in molecular markers also recovered six species-level lineages, *H. brasiliensis, H. vulpinus, H. chacarius, H. sciureus* and two forms that they left unnamed.

Thus, according to the data presented above the genus *Holochilus* currently presents six formal namely species (*brasiliensis* DESMAREST, (1819); *sciureus* WAGNER, 1842; *chacarius* THOMAS, 1906; *H. vulpinus* (BRANTS, 1827); *H. venezuelae* ALLEN, 1904*;* and *lagigliai* PARDIÑAS et al., 2013) and other eight nominal taxa related to the genus (*multannus* AMEGHINO, 1889; *nanus* THOMAS, 1897; *darwini* THOMAS, 1897; *guianae* THOMAS, 1901; *balnearum* THOMAS, 1906; *amazonicus* OSGOOD, 1915; *incarum* THOMAS,1920; *berbericensis* MORRISON-SCOTT, 1937).

As noted above, *Holochilus* has a taxonomic history full of nomenclatural uncertainties (table 1), and this fact allied to controversial phylogenetic hypotheses, results in species poorly defined within the genus. There is no systematic study covering both morphological and molecular data, and the only study that includes DNA data (D'ELÍA et al., 2015) had a reduced sampling of nuclear DNA sequences and fails to support the monophyly of the genus using the mitochondrial gene Cytb.

Regarding the phylogenetic relationships within the genus *Holochilus*, there are few studies investigating the relationships between species. This is probably because the taxonomy of this group is not well established. In all reconstructions involving only living species, the genus appears to be monophyletic; however in the phylogeny proposed by Machado et al., (2013), that included also the

extinct species (*H. primigenius*), the genus appears as paraphyletic, with *H. primigenius* appearing as sister group of *Carletonomys cailoi,* and the living species (*H. brasiliensis*, *H. chacarius* and *H. sciureus*) appear as sister group of *Noronhomys vespuccii*. Finally, the phylogeny published in D'Elía et al., (2015) included sequences of 1 mitochondrial and 3 nuclear loci, and the monophyly of *Holochilus* was not recovered on the mitochondrial loci. Although the concatenated matrix shows the presence of two main groups, the *brasiliensis* species group formed by *H. brasiliensis* and *H. vulpinus,* and the *sciureus* species group composed by *H. chacarius*, *H. sciureus*, and 2 currently unnamed forms.

Table 1. Main taxonomic arrangements proposed by different authors for the taxa of the species group of the genus *Holochilus*, with species and subspecies. Underlined names do not belong to *Holochilus* anymore. Right-aligned names are species names and left-aligned names are subspecies names.

Tate 1932	Gyldenstolpe 1932	Hershkovitz 1955	Cabrera 1961	Musser & Carleton 1993	Musser & Carleton 2005	Gonçalves et al. 2015
amazonicus	*amazonicus*	*brasiliensis*	*brasiliensis*	*brasiliensis*	*brasiliensis*	*brasiliensis*
anguya	*balnearum*	*amazonicus*	*amazonicus*	*chacarius*	*chacarius*	*chacarius*
balnearum	*chacarius*	*balnearum*	*balnearum*	*magnus*	*sciureus*	*sciureus*
brasiliensis	*darwini*	*berbicensis*	*brasiliensis*	*sciureus*		*venezuelae*
chacarius	*guianae*	*brasiliensis*	*darwini*			*vulpinus*
darwinii	*incarum*	*guianae*	*guianae*			*lagigliai*
guianae	*nanus*	*incarum*	*incarum*			
incarum	*physodes*	*leucogaster*	*leucogaster*			
leucogaster	*leucogaster*	*nanus*	*venezuelae*			
nanus	*russatus*	*venezuelae*	*vulpinus*			
physodes	*sciureus*	*vulpinus*	*magnus*			
russatus	*simpsoni*	*magnus*				
sciureus	*venezuelae*					
simpsoni	*vulpinus*					
venezuelae						
vulpinus						

All this data suggest that the diversity within the genus *Holochilus* is underestimated, being the morphological and molecular variation in this genus still not enough explored, as well the sister relationship within this taxa and with the others members clade D of Oryzomyini tribe. Therefore, the following chapters of this thesis is an attempt to clarify these issues, providing evidences about the morphological and genomic variation of *Holochilus*, both in an inter and intraspecific approach (Chapters 1, 2 and 3); additionally we present a model-based species delimitation integrating genomic and morphometric data, with a coalescent phylogenetic reconstruction for the species of *Holochilus*

(Chapter 3), as well as for the relationship between *Holochilus* and the other members clade D of Oryzomyini tribe (Chapter 4), which contemplates the diversity of species existing in these taxa as recognized here and suggestively indicated in previous works. Besides the description of all these patterns, all the chapters also investigates the processes responsible for the diversification in *Holochilus* or in the members of clade D, addressing some aspects of temporal and spatial differentiation that could lead to the present observed patterns.

In that context, the main goals for this thesis will comprehend the following issues: *(i)* evaluate and describe non-geographic variation related with age and sex in rodents of genus *Holochilus*; and establish what is role of the environment in shaping this kind of variation; *(ii)* quantitatively assess potential genomic parallels in three species of genus *Holochilus* from different biomes of South America to evaluate whether geographic structuring of genomic variation reflects shared life history traits versus biome specific associations; *(iii)* perform analysis of species delimitation in a model-based statistical approach to recognize single evolutionary entities inside genus *Holochilus*, gathering information from morphometric and genomic data; *(iv) g*enerate a genomic phylogeny for the genus *Holochilus*, including hypotheses about the origin and evolution of species; *(v)* test the monophyly, and access the internal relationship among members of clade D of tribe Oryzomyini; and *(vi)* establish biogeographical hypothesis on the evolution of genera of clade D of tribe Oryzomyini, and on the understandings of the origins and evolution of open/forest dwellers.

REFERENCES

ALLEN, J. A. New mammals from Venezuela and Colombia. **Bulletin of the American Museum of Natural History**, v. 20, p. 327–335, 1904.

AMEGHINO, F. Contribución al conocimiento de los mamíferos fósiles de la República Argentina. **Actas de la Academia Nacional de Ciencias en Córdoba**, v.6, p. 1–1027, 1889.

AZARA, F. Apuntamientos para la historia natural de los quadrupedos del Paraguay, y Rio de la Plata. En la imprenta de la viuda de Ibarra, Madri, 1802.

BEMMELS, J. B; TITLE, P.O.; ORTEGO, J.; KNOWLES, L.L. Tests of species-specific models reveal the importance of drought in postglacial range shifts of a Mediterranean-climate tree: insights from iDDC modelling and ABC model selection. **Molecular Ecology,** v. 25, p. 4889 – 4906, 2016.

BONVICINO, C.R. A new species of *Oryzomys* (Rodentia, Sigmodontinae) of the *subflavus* group from the Cerrado of Central Brazil. **Mammalian Biology**, v. 68, p. 78–90, 2003.

BRANDT, J.F. Mammalium rodentium exoticorum novorum vel minus rite cognitorum. Musei Academici Zoologici descriptiones et icones - **Memóires de L'Académie Impériale des Sciences de Saint-Pétersbourg**, v. 6, p.358–442, 1835.

BRANDT J.F. Untersuchungen über die craniologischen entwickelungsstufen und die davon herzuleitenden verwandtschaften und classificationen der nager der jetztwelt, mit besonderer beziehung auf die gattung Castor. **Memóires de L'Académie Impériale des Sciences de Saint-Pétersbour**g, v.7, p. 125–336, 1855.

BRANTS, A. Het Geslacht der Muizen door Linnaeus opgesteld, Volgens de Tegenswoordige toestand der Wetenschap in Familien, Geslachten en Soorten verdeeld. **Gedrukt ter Akademische Boekdrukkery**, xii + 190 pp, 1 plate, 1827.

BRAUN, J. K. Systematic relationships of the tribe Phyllotini (Muridae: Sigmodontinae) of South America. **Oklahoma Museum of Natural History**, **Special Publication**, p. 1-50, 1993.

BURMEISTER, H. Systematische Uebersicht der Thiere Brasiliens : welche während einer Reise durch die Provinzen von Rio de Janeiro und Minas geraës gesammlt oder beobachtet wurden von Hermann Burmeister. Saugethiere (Mammalia). Berlin: George Reimer, 1954. p. 342.

CABRERA, A. Catalogo de los mamiferos de America del Sur. **Revista del Museo Argentino de Ciencias Naturales "Bernadino Rivadavia",** v. 4, p. 309–732, 1961.

CAMARGO, A.; SITES Jr., J. Species Delimitation: A Decade After the Renaissance. In: PAVLINOV, Y.I. (Ed.) **The Species Problem - Ongoing Issues**. INTECH Open Access Publisher, 2013. p 225–247.

CAMARGO, A.; AVILA, L. J.; MORANDO, M.; SITES Jr., J.W. Accuracy and precision of species trees: effects of locus, individual and base pair sampling on inference of species trees in lizards of the *Liolaemus darwinii* Group (Squamata, Liolaemidae). **Systematic Biology**, v. 61, p. 272–288, 2012.

CARLETON, M.D.; MUSSER, G.G. Systematic studies of oryzomyine rodents (Muridae, Sigmodontinae): a synopsis of *Microryzomys*. **Bulletin of the American Museum of Natural History**, v.191, p. 1-83, 1989.

CARLETON, M. D.; OLSON, S.L. Amerigo Vespucci and the rat of Fernando de Noronha: a new genus and species of Rodentia (Muridae: Sigmodontinae) from a volcanic island off Brazil´s continental shelf. **American Museum Novitates**, v. 3256, p. 1-59, 1999.

CARSTENS, B.C.; PELLETIER, T.A.; REID, N.M.; SALTER, J.D. How to fail at species delimitation. **Molecular Ecology**, v. 22, p. 4369 – 4383, 2013.

CHIQUITO, E.A.; D'ELÍA, G.; PERCEQUILLO, A.R. Taxonomic review of genus *Sooretamys* Weksler, Percequillo & Voss (Rodentia: Cricetidae: Sigmodontinae): an integrative approach. **Zoological Journal of Linnean Society**, v. 94, p. 365 – 378, 2014.

D'ELÍA, G. Phylogenetics of Sigmodontinae (Rodentia, Muroidea, Cricetidae), with special reference to the akodont group, and with additional comments on historical biogeography. **Cladistics**, v. 19, p. 307–323, 2003.

D'ELÍA, G.; PARDIÑAS, U.F.J. Subfamily Sigmodontinae Wagner, 1843. In: PATTON J.L; PARDIÑAS U.F.J.; D'ELÍA G. (Ed.) **Mammals of South America**, Vol. 2, Rodents. University of Chicago Press, 2015. p. 63–73.

D'ELÍA, G.; GONZALEZ, E.M.; PARDIÑAS, U.F.J. Phylogenetic analysis of sigmodontine rodents (Muroidea), with special reference to the akodont genus Deltamys. **Mammalian Biology**, v. 68, p 351–364, 2003.

D'ELÍA, G.; PARDIÑAS, U.F.J; MYERS, P. An introduction to the genus *Bibimys* (Rodentia: Sigmodontinae): phylogenetic position and alpha taxonomy. In: LACEY E.; MYERS P. (Ed.) **Mammalian Diversification: From Chromosomes to Phylogeography**. University of California Publications in Zoology, 2005. p. 211–246.

D'ELÍA, G.; OJEDA, A.A.; ONDACA, F. M; GALLARDO, M.H. New data of the long-clawed mouse *Pearsonomys annectens* (Cricetidae, Sigmodontinae) and additional comments on the distinctiveness of *Pearsonomys*. **Mammalian Biology**, v. 71, p. 39–51, 2006a.

D'ELÍA, G.; LUNA, L; GONZÁLEZ, E.M.; PATTERSON, B.D. On the Sigmodontinae radiation (Rodentia, Cricetidae): An appraisal of the phylogenetic position of *Rhagomys*. **Molecular Phylogenetics and Evolution**, v. 38, p 558–564, 2006b.

D'ELÍA, G.; PARDIÑAS, U.U.F.J.; TETA, P.; PATTON, J.L.J.; ELÍA, G.D. Definition and diagnosis of a new tribe of sigmodontine rodents (Cricetidae: Sigmodontinae), and a revised classification of the subfamily. **Gayana**, v. 71, p. 187–194, 2007.

D'ELÍA, G.; HANSON, J.D.; MAULDIN, M.R.; TETA, P.; PARDIÑAS, U.F.J. Molecular systematics of South American marsh rats of the genus *Holochilus* (Muroidea, Cricetidae, Sigmodontinae). **Journal of Mammalogy**, v. 43, p. 1–14, 2015.

de QUEIROZ, K. Different species problems and their solution. **BioEssays,** v. 26, p. 67–70, 2005a.

de QUEIROZ, K. Ernst Mayr and the modern concept of species. **Proceedings of the National Academy of Sciences**, v. 102, p. 6600–6607, 2005b.

DESMAREST, A.G. Le Rat du Brésil, *Mus brasiliensis*, Geoffr. In: _____. Nouveau dictionnaire d'histoire naturelle, appliquèe aux art, principalement à l'agriculture et à l'economie rurale et domestique, à la medicine, etc. Par une société de naturalists et d'agriculteurs, tome XXIV. Nouv. éd. Paris: Deterville, 1819. p. 62.

DESMAREST, A.G. **Mammalogie ou description des espèces de mammifères**. Veuve Agasse, Imprimeur-Libraire, Paris, 1820. p. 555.

DOMINGOS, F.M.C.B.; COLLI. G.R.; LEMMON, A.; LEMMON M.E.; BEHEREGARAY, L.B. In the shadows: Phylogenomics and coalescent species delimitation unveil cryptic diversity in a Cerrado endemic lizard (Squamata: *Tropidurus*). **Molecular Phylogenetics and Evolution,** v. 107, p. 455–465, 2017.

DOMÍNGUEZ-BELLO, M.G; ROBINSON, M.D. A Comparison of Digestive Adaptations in Two Neotropical Cricetid Rodents (Holochilus venezuelae and Zygodontomys microtinus). **Physiological and Biochemical Zoology**, v. 64, p. 1542–1551, 1991.

DONOGHUE, M.J. A critique of the biological species concept and recommendations for a phylogenetic alternative. **Bryologist**, v. 88, p. 172–181, 1985.

EATON, D.A.R; REE, R.H. Inferring phylogeny and introgression using RADseq data: an example from flowering plants (Pedicularis: Orobanchaceae). **Systematic biology**, v. 62, p. 689–706, 2013.

EDWARDS, D.L.; KNOWLES, L.L. Species detection and individual assignment in species delimitation: can integrative data increase efficacy? **Proceedings of the Royal Society of London B**, v. 281, p. 1–7, 2014.

EMMONS, L. H.; PATTON, J.L. A new species of *Oryzomys* (Rodentia: Muridae) from Eastern Bolivia. **American Museum Novitates**,v. 3478, p. 1–26. 2005.

FERNÁNDEZ, F.J.; TORRES, J; TAMMONE, M.; LÓPEZ, J.M.; PARDIÑAS, U.F.J. New data on the endemic cricetid rodent *Holochilus lagigliai* from central-western Argentina: fossil record and potential distribution. **Mammalia**, in press, 2017.

FISCHER, G. **Zoognosia Tabulis Synopticis illustrata**. N.S. Vsevolozky, Moscow, 1814. p. 465.

FORMOSO, A.E.; SAUTHIER, D.E.U.; PARDIÑAS, U.F.J. Mammalia, Rodentia, Sigmodontinae, *Holochilus brasiliensis* (Desmarest, 1819): Distribution extention. **Check list**, v. 6, p. 195–197, 2010.

FROST, D. R.; KLUGE, A. G.; HILLIS, D. M. Species in contemporary herpetology: comments on phylogenetic inference and taxonomy. **Herpetological Review,** v. 23, p. 46–54, 1992.

GANDINI, A.M. Anatomia comparada de estruturas do esqueleto pós-craniano dos roedores semiaquáticos *Holochilus brasiliensis* e *Nectomys squamipes* (Cricetidae, Sigmodontinae). 2014. 47 p. Monografia – Universidade Luterana do Brasil, Rio Grande do Sul, 2014.

GOMÉZ-LAVERDE, M.; ANDERSON, R.P.; GARCÍA, L.P. Integrated Systematic reevaluation of the Amazonia genus *Scolomys* (Rodentia: Sigmodontinae). **Mammalian Biology**, v. 69, p. 119–140, 2004.

GONÇALVES, P. R.; TETA, P.; BONVICINO, C. Genus *Holochilus* Brandt, 1835. In: PATTON J.L; PARDIÑAS U.F.J.; D'ELÍA G. (Ed.) **Mammals of South America**, Vol. 2, Rodents. University of Chicago Press, 2015. p. 325–335.

HARRINGTON R.C.; NEAR T.J. Phylogenetic and coalescent strategies of species delimitation in Snubnose Darters (Percidae: *Etheostoma*). **Systematic Biology**, v. 61, p. 63–79, 2012.

HERSHKOVITZ, P. South American marsh rats genus *Holochilus*, with a summary of sigmodont rodents. **Fieldiana,** Zoology, v. 37, p. 639–673, 1955.

HERSHKOVITZ, P. Evolution of Neotropical cricetine rodents (Muridae) with special reference to the phyllotine group. **Fieldiana,** Zoology, v. 46, p. 1–524, 1962.

HERSHKOVITZ, P. South American swamp and fossorial rats of the scapteromyine group (Cricetinae, Muridae), with comments on the glans penis in murid taxonomy. **Zeitschrift für Säugetierkunde**, v. 31, p. 81–149, 1966.

HERSHKOVITZ, P. The recent mammals of the Neotropical region: a zoogeographic and ecologic review. In: KEAST, A.; ERK, F.C.; GLASS, B. (Ed.) **Evolution, mammals, and southern continents.** State University of New York Press, 1972. p. 311–431.

HIPP, A. L.; EATON D.A.R.; CAVENDER-BARES, J.; FITZEK, E.; NIPPER, R.; MANOS, P.S. A framework phylogeny of the American oak clade based on sequenced RAD data. **PLoS ONE**, v. 9, e93975, 2014.

HOOPER, E.T.; MUSSER, G.G. Notes on classification of the rodent genus *Peromyscus*. **Occasional Papers of the Museum of Zoology, University of Michigan**, v. 635, p.1–13, 1964.

HUANG, J-P; KNOWLES, L.L. The species versus subspecies conundrum: quantitative delimitation from integrating multiple data types within a single Bayesian approach in Hercules beetles. **Systematic Biology**, v. 65, p. 685–699, 2016.

KAWAZOE, U; PINTO, A.C.M. Importância epidemiológica de alguns animais silvestres na esquistossomose mansônica. **Revista de Saúde Pública**, v. 17, p. 345–346, 1983

KNOWLES, L.L.; CARSTENS, B.C. Delimiting species without monophyletic gene trees. Systematic Biology, v. 56, p. 887– 895, 2007.

LANGGUTH, A.; BONVICINO, C.R. The *Oryzomys subflavus* group, with description of two new species (Rodentia, Muridae, Sigmodontinae). **Arquivos do Museu Nacional**, v. 60, p. 285–294, 2002.

LICHTENSTEIN, H. Darstellung neuer oder wenig bekannter Säugethiere in Abbildungen und Beschreibun: von fünf und sechzig Arten auf funfzig colorirten Steindrucktafeln nach den Originalen des Zoologischen Museums der Universität zu, 1830.

LILLJEBORG, W. Systematisk oversigt af de gnagande diiggdjuren, Glires. Kongl. Akad. Boktryekeriet, Uppsala., 1866. p. 59.

LIRA, M.G.S.; MIRANDA, G.S.; RODRIGUES, J.G.M.; NOGUEIRA, R.A.; GOMES, G.C.C.; CANTANHÊDE, L.G.; SILVA-SOUZA, N. Aspectos Biológicos de *Holochilus* sp., hospedeiro natural da esquistossomose. **Ciência Animal Brasileira**, v. 17, p. 143–153, 2016.

MACHADO, L.F.; LEITE, Y.L.R.; CHRISTOFF, A.U.; GIUGLIANO, L.G. Phylogeny and biogeography of tetralophodont rodents of the tribe Oryzomyini (Cricetidae: Sigmodontinae). **Zoologica Scripta**, v. 43, p. 119–130, 2013.

MARTINO, A.M.G.; AGUILERA. M. Food habits of *Holochilus venezuelae* in rice fields. **Mammalia,** v. 53, p. 545–561, 1989.

MASSATTI, R.; KNOWLES, L.L. Contrasting support for alternative models of genomic variation based on microhabitat preference: species-specific effects of climate change in alpine sedges. **Molecular Ecology,** v. 25, p. 3974–3986, 2016.

MASSOIA, E. Las ratas nutrias argentinas del genero Holochilus descriptas como *Mus braziliensis* por Waterhouse (Mammalia - Rodentia - Cricetidae). **Revista de Investigaciones Agropecuarias**, v. 8, p. 141–148, 1971.

MASSOIA, E. Ataques graves de *Hochilus* y otros roedores a cultivos de cana de azúcar. **Instituto Nacional Tecnología. Agropecuaria (I.N.T.A.),** p. 1–12, 1974.

MASSOIA, E. Mammalia. In: RINGUELET, R. (Ed.) **Fauna de agua dulce de la Republica Argentina**. Fundación Editorial Ciencia y Cultura, 1976. p. 1–128.

MASSOIA, E. El estado sistemático y zoogeografia de *Mus brasiliensis* Desmarest y *Holochilus sciureus* Wagner (Mammalia-Rodentoa-Cricetidae). **Physys**, v. 39, p. 31–34, 1981.

MAYDEN, R.L. A hierarchy of species concepts: the denouement in the saga of the species problem. In: CLARIDGE M.F.; DAWAH H.A.; WILSON M.R. (Ed.) **Species: The Units of Biodiversity**. Chapman & Hall, 1997. P. 381–424.

MAYR, E; ASHLOCK, P.D. **Principles of Systematic Zoology**. Mcgraw-Hill College, New York, 1991. p. 475.

MILLER G.S.J.; REHN J.A.G. Systematic results of the study of North American land mammals to the close of the year 1900. **Proceedings of the Boston Society of Natural History**, v. 30, p. 1–352, 1902.

MORRISON-SCOTT, T.C.S. An apparently new form of Cricetinae from Bristish Guiana. **Annals and Magazine of Natural History**, v. 20, p. 535–538, 1937.

MUSSER, G.G.; CARLETON, M.D. Superfamily Muroidea. In: WILSON, D.E.; REEDER, D.A. (Ed.) **Mammal species of the world. A taxonomic and geographic reference.** 3rd ed. The Johns Hopkins University Press, Baltimore, 2005. p. 894–1531.

MUSSER, G.G.; CARLETON, M.D.; BROTHERS, E.; GARDNER, A.L. Systematic studies of Oryzomyine rodents (Muridae, Sigmodontinae): diagnoses and distributions of species formerly assigned to *Oryzomys* "*capito*". **Bulletin of the American Museum of Natural History**, v. 236, p. 1-376, 1998.

NAZARENO, A.G.; BEMMELS, J.B.; DICK, C.W.; LOHMANN, L.G. Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species. **Molecular Ecology Resources**, in press, 2017.

OLAVE, M.; SOLA, E.; KNOWLES, L.L. Upstream analyses create problems with DNA-based Bayesian species delimitation. **Systematic Biology**, v. 63, p. 263–271, 2014.

OLAVE, M; AVILA, L.J.; SITES Jr, J.W.; MORANDO, M. Hidden diversity within the lizard genus *Liolaemus*: Genetic vs morphological divergence in the *L. rothi* complex (Squamata: Liolaeminae). **Molecular Phylogenetics and Evolution**, v. 107, p. 56–63, 2017.

OSGOOD, W. H. New mammals from Brazil and Peru. **Field Museum of Natural History, Zoological Series**, v. 10, p. 187–198, 1915.

PACHECO, V. **Phylogenetic analyses of the Thomasomyini (Muroidea: Sigmodontinae) based on morphological data.** 2003. 398 p. Thesis (Ph.D. in Philosophy) - The City University of New York, New York, 2003.

PALMER T.S. Index generum mammalium: a list of the genera and families of mammals. North American Fauna. Washington,Govt. Print. Off., 1904. p. 984.

PARADA, A.; D'ELÍA, G.; PALMA, R.E. The influence of ecological and geographical context in the radiation of Neotropical sigmodontine rodents. BMC Evolutionary Biology, v. 15, p. 1–17, 2015.

PARDIÑAS, U.F.J.; D'ELÍA, G.; ORTIZ, P.E. Sigmodontinos fósiles (Rodentia, Muroidea, Sigmodontinae) de América del Sur: estado actual de su conocimiento y prospective. Mastozoología Neotropical, v. 9, p. 209–252, 2002.

PARDIÑAS, U.F.J.; TETA, P.; VOGLINO, D.; FERNÁNDEZ, F.J. Enlarging rodent diversity in west–central Argentina: a new species of the genus Holochilus. Journal of Mammalogy, v. 94, p. 231–240, 2013.

PARDIÑAS, U.F.J.; TETA, P.; SALAZAR-BRAVO, J. A new tribe of Sigmodontinae Rodents (Cricetidae). **Mastozoología Neotropical**, v. 22, p.171–186, 2015.

PATTON, J.L.; HAFNER, M.S. Biosystematics of the native rodents of the Galapagos Archipelago, Ecuador. In*:* BOWMAN, R.I.; BERSON, M.; LEVITON, A.E. (Eds) **Patterns of evolution in Galapagos organisms**. San Francisco: AAAS Pacific Division, 1983. p. 539-568.

PATTON, J.L.; da SILVA, M.N.F.; MALCOLM, J.R. Mammals of the Rio Juruá and the evolutionary and ecological diversification of Amazonia. **Bulletin of the American Museum of Natural History**, v. 244, p. 1–306, 2000.

PERCEQUILLO, A.R; HINGST-ZAHER, E.; BONVICINO, C.R. Systematic review of genus Cerradomys Weksler, Percequillo and Voss, 2006 (Rodentia: Cricetidae: Sigmodontinae: Oryzomyini), with description of two new species from eastern Brazil. **American Museum Novitates,** v. 3622, p. 1–46, 2008.

PERCEQUILLO, A.R.; WEKSLER, M.; COSTA, L.P. A new genus and species of rodent from the Brazilian Atlantic Forest (Rodentia: Cricetidae: Sigmodontinae: Oryzomyini), with comments on oryzomyine biogeography. **Zoological Journal of the Linnean Society**, v. 161, p. 357–390, 2011.

PETERS, W. **Uber einige merkwurdige Nagethiere (Spalacomys indicus, Mes tomentosus und Mus squamipes) des Konigl. Zoologischen Museums**. Abhandlungen der Koniglichen Akademie der Wissenschaften zu Berlin, 1861. p. 139–156.

PETERSON, B. K.; WEBER, J.N.; KAY, E.H.; FISHER, H.S.; HOEKSTRA, H.E. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. **PloS One**, v. 7, e37135, 2012.

PINE, R.H.; TIMM, R.M.; WEKSLER, M. A newly recognized clade of trans-Andean Oryzomyini (Rodentia: Cricetidae), with description of a new genus. **Journal of Mammalogy,** v. 93, p. 851–870, 2012.

PRADO, J.R.; PERCEQUILLO, A.R. On the geographic distribution of the genera of Tribe Oryzomyini on South America, with some comments on the patterns of diversity and Reig´s Areas of Original Differentiation. **Arquivos de Zoologia**, v. 44, p. 1–120, 2013.

PRADO, J.R.; PERCEQUILLO, A.R. Systematic Studies of the Genus Aegialomys Weksler et al., 2006 (Rodentia: Cricetidae: Sigmodontinae): Geographic Variation, Species Delimitation, and Biogeography. **Journal of Mammalian Evolution**, in press, 2017.

REIG, O.A. Diversity patterns and differentiation of high Andean rodents. In: VUILLEUMIER, F.; MONASTERIOS, M. (Ed.). **High altitude tropical biogeography**. Oxford University Press, 1986. p. 404–439.

ROSEN, D.E. Fishes from the uplands and intermontane basins of Guatemala: revisionary studies and comparative geography. **Bulletin of the American Museum of Natural History**, v. 162, p. 267–376, 1979.

SALAZAR-BRAVO, J.; PARDIÑAS, U.F.J.; ZEBALLOS, H.; TETA, P. Description of a New Tribe of Sigmodontine Rodents (Cricetidae: Sigmodontinae) with an Updated Summary of Valid Tribes and Their Generic Contents. **Occasional Papers Museum of Texas Tech University**, V. 338, P. 1–24, 2016.

SAMUELS, J.X.; VAN VALKENBURGH, B. Skeletal Indicators of Locomotor Adaptations in Living and Extinct Rodents. **Journal of Morphology**, v. 269, p. 1387–1411, 2008.

SAUTHIER W.O.U.; ABBA A.M.; SAUTHIER D.E.U. Nests of *Oligoryzomys* sp. an *Holochilus brasiliensis* (Rodentia, Cricetidae) eastern Entre Rios Province, Argentina. **Mastozoología Neotropical**, v. 17, p. 207–211, 2010.

SHAFER, A.B.A.; PEART, C.R.; TUSSO, S.; MAAYAN, I.; BRELSFORD, A.; WHEAT, C.W.; WOLF, J.B.W. Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. **Methods in Ecology and Evolution**, 2016.

SILVA, W.O.; PIECZARKA, J.C.; ROSSI, R.V.; SCHNEIDER, H.; SAMPAIO, I.; MIRANDA, C.L.; SILVA, C.R.; CARDOSO, E.M.; NAGAMACHI, C.Y. Diversity and Karyotypic Evolution in the Genus *Neacomys* (Rodentia, Sigmodontinae). **Cytogenetic Genome Research**, v. 146, p. 296–305, 2015.

SIMPSON, G.G. **Princípios de Taxonomia Animal**. Fundação Calouste Gulbenkian, 1962. p. 254.

SIMPSON, G.G.; ROE, A.; LEWONTIN, R.C. **Quantitative Zoology**. Dover Publications, 2003. p. 454.

SMITH, M.F.; PATTON, J.L. Diversification of South American muroid rodents: evidence from mitochondrial DNA sequence data for the Akodontine tribe. **Biological Journal Linnean Society**, v. 50, p. 149–177, 1993.

SMITH, M.F.; PATTON, J.L. Phylogenetic relationships and the radiation of sigmodontine rodents in South America: evidence from cytochrome b. **Jounal of Mammalian Evolution**, v. 6, p. 89–128, 1999.

STEPPAN, S.J. Revision of tribe Phyllotini (Rodentia: Sigmodontinae), with a phylogenetic hypothesis for the Sigmodontinae. **Fieldiana**, Zoology, v. 80, p. 1–112, 1995.

STEPPAN, S.J. A new species of *Holochilus* (Rodentia: Sigmodontinae) from the middle Pleistocene of Bolívia and its phylogenetic significance. **Journal of Vertebrate Paleontology**, v. 16, p. 522–530, 1996.

TAKAHASHI, T.; NAGATA, N.; SOTA, T. Application of RAD-based phylogenetics to complex relationships among variously related taxa in a species flock**. Molecular Phylogenetics and Evolution**, v. 80, p. 137–144, 2014.

TATE, G.H.H. The Taxonomic History of the South and Central American cricetid rodents of the genus Oryzomys- part 2: Subgenera Oligoryzomys, Thallomyscus, and Melanomys. **American Museum Novitates**, p. 1–18, 1932.

TAVARES, W.C.; PESSÔA, L.M.; GONÇALVES, P.R. New species of *Cerradomys* from coastal sandy plains of southeastern Brazil (Cricetidae: Sigmodontinae). **Journal of Mammalogy**, v. 92, p. 645–658, 2011.

THOMAS, O. On a collection of rodents from North Peru. **Proceedings of the Zoological Society of London**, v. 50, p. 98–111, 1882.

THOMAS, O. On a collection of Muridae from Central Peru. **Proceedings of the Zoological Society of London**, p. 447–458, 1884.

THOMAS, O. Description of four new South American mammals. **Annals and Magazine of Natural History**, v. 20, p. 218-221, 1897.

THOMAS, O. New Mammals from Peru and Bolivia, with a List of Those Recorded from the Inambari River, upper Madre de Dios. **Annals and Magazine of Natural History**, v.7, p. 178–90, 1901.

THOMAS, O. Notes on South-American Rodents. **Annals and Magazine of Natural History,** v. 18, p. 442–448, 1906.

THOMAS, O. Report on the Mammalia Collected by Mr. Edmund Heller During the Peruvian Expedition of 1915 Under the Auspices of the Yale University and the National Geographie Society. **Proceedings of the United States National Museum,** v. 68, p. 217–49, 1921.

TROUESSART, E.-L. **Catalogus mammalium tam viventium quam fossilium. Fasciculus III. Rodentia II**. Berolini:R. Friedländer & Sohn, 1898. p. 1 +vi 664.

TROUESSART, E.-L. **Catalogus mammalium tam viventium quam fossilium. Quinquennale supplementum anno 1904. Fasciculus II. Rodentia II**. Berolini:R. Friedländer & Sohn, 1905. p. vii+ 929.

TURVEY, S.T.; WEKSLER, M.; MORRIS, E.L.; NOKKERT, M. Taxonomy, phylogeny, and diversity of the extinct Lesser Antillean rice rats (Sigmodontinae: Oryzomyini), with description of a new genus and species. **Zoological Journal of the Linnean Society**, v. 160, p. 748–772, 2010.

TWIGG, G.I. Notes on *Holochilus sciureus* in British Guiana. **Journal of Mammalogy**, v. 43, p. 369–374, 1962.

UTURUNCO, A.C.J.; PACHECO, V.R. Redescripción de *Eremoryzomys polius* (Rodentia: Cricetidae) y descripción de una nueva especie del género Eremoryzomys. **Mastozoología Neotropical**, v. 23, p. 483–503, 2016.

VOSS, R.S. An introduction to the neotropical muroid rodent genus *Zygodontomys*. **Bulletin of the American Museum of Natural History**, v. 210, p. 1–113, 1991.

VOSS, R.S. A revision of the brazilian Muroid rodent genus *Delomys* with remarks on "Thomasomyine" characters. **American Museum Novitates**, v. 3073, p. 1–44, 1993.

VOSS, R.S.; CARLETON, M.D. A new genus for *Hesperomys molitor* Winge and *Holochilus magnus* Hershkovitz (Mammalia, Muridae) with an analysis of its phylogenetic relationships. **American Museum Novitates**, v. 3085, p. 1–39, 1993.

VOSS, R.S.; ABRAMSON, N.I. Holochilus Brandt, 1835, Proechimys J.a. Allen, 1899 And Trinomys Thomas, 1921 (Mammalia, Rodentia): Proposed Conservation By The Designation Of H. Sciureus Wagner, 1842 As The Type Species Of Holochilus. **Bulletin of Zoological Nomenclature**, v. 56, p. 255–261, 1999.

WAGNER, C.E; KELLER, I.; WITTWER, S.; SELZ, O.M.; MWAIKO, S.; GREUTER, L.; SIVASUNDAR, A.; SEEHAUSEN, O. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. **Molecular Ecology**, v. 22, p. 787–798, 2013.

WAGNER, J. A. Beschreibung einiger oder minder bekannter Nager. **Archives fur Naturgeschichte,** v. 8, p.1–33, 1842.

WAGNER, J. A**. Die Säugthiere in Abbildungen nach der Natur mit Beschreibungen von Dr. Johann Christian Daniel von Schreber. Supplementband.** Dritter Abtheilung: Die Beutelthiere und Nager (erster Abschnitt). In der Expedition des Schreber'schen Säugthier- und des Esper'schen Schmetterlingswerkes, und in Commission der Voss'schen Buchhandlung in Leipzig, 1843. p. 614.

WATERHOUSE, G. R. **The zoology of the voyage of H. M. S.Beagle, Part II. Mammalia.** Edited and superintended by Charles Darwin. London, Smith, Elder & Co, 1839. p. 97.

WEKSLER, M. Phylogeny of Neotropical oryzomyine rodents (Muridae: Sigmodontinae) based on the nuclear IRBP exon. **Moleculer Phylogenetics and Evolution**, v. 28, p. 331–349, 2003.

WEKSLER, M. Phylogenetic relationships of oryzomyine rodents (Muroidea: Sigmodontinae): separate and combined analyses of morphological and molecular data. **Bulletin of the American Museum of Natural History**, v. 296, p. 1–149, 2006.

WEKSLER, M. Tribe Oryzomyini Vorontsov, 1959. In: PATTON J.L; PARDIÑAS U.F.J.; D'ELÍA G. (Ed.) **Mammals of South America**, Vol. 2, Rodents. University of Chicago Press, 2015. p. 63–73.

WEKSLER, M.; PERCEQUILLO, A.R.; VOSS, R.S. Ten new genera of Oryzomyine Rodents (Cricetidae; Sigmodontinae). **American Museum Novitates**, v. 3537, p. 1–29, 2006.

WILEY, E.O.; LIEBERMAN, B.S. **Phylogenetics: Theory and Practice of Phylogenetic Systematics, 2nd Edition.** Wiley-Blackwell, 2011. p. 432.

WILEY, E.; MAYDEN, R. The evolutionary species concept. In: WHEELER, Q.; MEIER, R. (Ed.) **Species concepts and phylogenetic theory: A debate** Columbia University Press, 2000. p. 70–92.

WILSON, D.E.; REEDER, D.M. **Mammal species of the world: a taxonomic and geographic reference. 3rd ed**. Johns Hopkins University Press, 2005. p. 2142.

WINGE, H. **Jordfunde og nulevende Gnavere (Rodentia) fra Lagoa Santa. Minas Geraes, Brasilien: med udsigt over gnavernes indbyrdes slaegtskab.** Kjøbenhavn: F. Dreyer, 1887. p. 178.

2. ONTOGENETIC AND SEXUAL VARIATION IN SOUTH AMERICAN MARSH RATS (SIGMODONTINAE*:* *HOLOCHILUS*) AND THE ROLE OF ENVIRONMENT IN SHAPING THE NON-GEOGRAPHIC VARIATION

ABSTRACT

Ontogenetic and sexual variation has been documented in rodents, and it is imperative to control this variation prior to geographic and taxonomic approaches. An uncommon approach on such non-geographic variation is the evaluation of the effects of environment in the ontogenetic trajectories by comparing natural populations with a population grown under controlled conditions. Here we report analyses of craniometrical variation within wild-collected populations and a captive colony of the genus *Holochilus* to detect if environmental and taxonomic effects can change the patterns of non-geographic variation. We expect that populations from the same locality but grown in different environment conditions exhibits different rates of morphological change, while populations from different species but both grown in a non-controlled condition exhibits similar rate of change. Sexual dimorphism exhibited small degree of variation among populations. The greater ontogenetic variation is found in the younger age classes, but oldest individuals also show larger degree of differentiation. There are also great differences in the ontogenetic trajectories among samples, where individuals from the captive population exhibited the lower degree of variation between all age classes. The differences in the degree of non-geographic variation might be related to short-term responses to habitat fluctuations or to long-term evolutionary changes.

Keywords: Morphology; Morphometric; Rodent; Skull; Diet; Oryzomyini; Ontogenetic trajectory

2.1. Introduction

One of the primary tasks of the taxonomist is to unmask variation such as sexual dimorphism, age, seasonal effects, and other forms of individual polymorphism that can mislead the delimitation of species limits. This is not an easy task, as proved by the long lists of synomyms in many groups (Mayr & Ashlock, 1991), that do not represent valid species, whose type specimens represent age, sex or individual variants of former described species, making this data available for important comparisons among different species.

Ontogenetic and sexual variation in morphometric traits has been documented in rodents and, in practical terms, this variation has been proved useful to evaluate differences among individuals of distinct age classes and to establish if individuals of two sexes are different in size and shape, and if those age and sex categories should be pooled or not for all the subsequent data analysis. Dental and cranial dimensions have been broadly used in age and sexual dimorphism studies with rodents (Myers & Carleton, 1981; Patton & Rogers 1883; Musser & Williams 1985; Carleton & Musser 1989; Voss, 1991; Brandt & Pessôa 1994; Chimimba & Dippenaar 1994; Carmadella, Pessôa & Oliveira 1998; Oliveira, Strauss & Reis, 1998; Monteiro, Lessa & Abe, 1999; Hingst-Zaher, Marcus & Cerqueira,

2000; Mullin, Pillay & Taylor, 2001; Percequillo, Hingst–Zaher & Bonvicino, 2008; Fernandes *et al*., 2009; Prado & Percequillo, 2011, Abreu-Júnior *et al*., 2012) and represent valuable markers for non-geographic, geographic, taxonomic variation, and also to understand evolutionary trajectories, since the skull supports an elaborate trophic apparatus, the brain, and most of the sensorial organs, and might exhibit morphological responses to environmental pressures (Voss, Marcus & Escalante, 1990; Herring, 1993).

Most studies of intrapopulacional variation were conducted on wild-collected specimens housed in museum collections. Although they provide a lot of relevant insights about patterns of age and sex variation in natural populations, it would be crucial also to evaluate how wild species bred under controlled conditions would develop and if they can provide inferences about the genetic basis of morphological evolution and the effects of natural selection (Voss, Marcus & Escalante, 1990).

Regarding this, rodents of genus *Holochilus* provide an unusual opportunity to estimate the effects of environment into the ontogenetic trajectories by comparing the variation of a natural population and one that was grown under controlled conditions. Species of the genus *Holochilus* (Brandt, 1835) are oryzomines of large body size, specialized in semi-aquatic habitats, mainly inhabiting marshes, swamps, grasslands, wetlands and other open areas of South America, with a predominantly herbivorous diet (Massoia, 1971, 1976; Formoso, Sauthier & Pardiñas, 2010). The diet consist basically of leaves and other items of plant origin (Hershkovitz, 1955), and rodents of this genus frequently attack sugar cane, corn and rice fields, causing great damages to this crops (Twigg, 1962; Martino & Aguilera, 1989). Holochilus presents several herbivorous specializations in its anatomy, such as the gut morphology, adapted for an herbivorous diet (Domínguez-Bello & Robinson 1991), the gall bladder losses (which can be attributed to the low fat content of plant diets and presents a high metabolic rate; Domínguez-Bello & Robinson, 1991), the reduction or loss of the mesoloph/id (loph localized between the paracone and metacone in the upper molars and between the metaconid and entoconid in the lower molars, and connected to the median mure; Hershkovitz, 1962), and the lamination of molar cusps and a reduction in the number of molar folds (condition designate as tetralophodont, Hershkovitz, 1962).

Regarding to reproduction, adults individuals starts the reproductive period after about three months of life, the fertile period of females lasts about four days and the gestation lasts on average 21 days (Lira *et al*., 2016), with approximately 2 to 10 individuals per litter (Twigg, 1962; Lira *et al*., 2016). In addition, results observed for breeding of *Holochilus* in captivity seem to indicate that in a favorable environment, with sufficient space for survival and abundant food, there is no time frame for its reproduction (Kawazoe & Pinto, 1983).

There are genetic (e.g. ancestral traits) and non-genetic factors (e.g. differences in environment conditions) that can contribute to the variation related with age and sex. Describe how this variation occurs in different taxonomic levels (among populations of the same species or among

sibling species) is the first step to understanding the genetic forces (such as natural selection) that may be acting to shape this type of morphological variation (Voss, Marcus & Escalante, 1990).

Here we report analyses of craniometrical variation (qualitative changes on skull morphology and quantitative comparisons between cranial-dental measurements that are the standard in the study of the non-geographic variation in rodents) within a wild-collected population and a captive colony with parental individuals from the same locality in eastern Paraguay, and a third sample from a different species of genus *Holochilus* from northeastern Brazil. We compare the results of age and sexual variation to detect possible environmental influence on these traits and to evaluate the effects of the taxonomy on the patterns of non-geographic variation. We compare populations from three completely different environment: from a temperate/arid climatic zone in northeastern part of Brazil (Alagoas_Wild), with vegetation characterized by xerophile/perennial forest; from the region of Humid Chaco in Paraguay (Par_Wild), with the vegetation characterized by inundated areas covered by grasslands; and from the same locality in Paraguay (Par_Captive), but grown in controlled conditions. We expect that populations from the same locality but experiencing growth in different environment conditions (Par_Wild X Par_Captive) will exhibit a different rates of morphological change, and that population from different species but both grown in a non-controlled condition (Alagoas_Wild X Par_Wild) exhibits a similar rate of morphological change. This is our hypothesis, as we believe that samples from wild will present a response to selection, when compared with a population under a controlled condition and in the absence of selective pressure.

2.2. Material and Methods

2.2.1. Samples

We examined and measured 302 specimens of the genus *Holochilus* obtained from a wild population in the State of Alagoas (Brazil) housed at the Museu Nacional, Universidade Federal do Rio de Janeiro (Rio de Janeiro, Brazil), collected by the Serviço Nacional de Peste (freely translated as National Plague Service; see Oliveira & Franco (2005) for more details about this collection) from 1952 to 1955. The specimens were sampled at three localities, namely “Anadia”, “Viçosa” and “Quebrangulo”, approximately 80km far from each other, and we pooled them together in order to form a more robust sample named Alagoas_Wild. The captive-bred sample of the genus *Holochilus* consists of 154 individuals descendent from a former natural population trapped in two localities, distanced by approximately 300km from each other, namely “24 km NW (by air) of Villa Hayes, and west bank of Rio Paraguay, Estancia la Golondrina”, and “4 km NW Puerto Fonciere”, both in Presidente Hayes Departament (Paraguay), deposited in the University of Michigan Museum of Zoology (Ann Arbor, USA); we named this sample as Par_Captive. We also present data from 36 individuals collected at the same localities as the Captive parental individuals, the sample called

Par_Wild. Here it is important to specify that samples from Alagoas and Paraguay belong to different species. See Appendix A for a complete list of examined material with information about the collected localities.

2.2.2. Age Classes

We classified all the specimens in age classes according to the eruption and wear of upper molars, as well as characteristics of the pelage and degree of ossification of the basisphenoid-basioccipital (=sphenoccipital) suture, following criteria defined by Voss (1991), Oliveira, Strauss & Reis (1998) and Prado & Percequillo (2011). Based on these criteria, we recognized five age classes, described as follows (Fig. 3):

Age class 1: First and second molars with almost no apparent wear, main cusps high, small exposure of dentine and with the labial flexus open and conspicuous; third molar is usually non-erupted or newly erupted, with the main cusps still closed, without dentine exposition, and anterior lamina frequently isolated, configuring the paraflexus confluent with the hypoflexus (Fig. 3). Basisphenoid-basioccipital suture frequently open. Roots of all molars not visible in occlusal view. Pelage sparse, thin, and wavy, composed by abundant viliforms and few setiforms and aristiforms; pelage with dorsal color predominantly grayish-brown and ventral color grayish (Alagoas_Wild: 11 individuals; Par_Captive: 6 individuals).

Age class 2: First and second molars also with minor wear; some flexi closed, forming fossets; third molar already showing small degree of wear, with exhibition of dentine and connection of anterior lamina to posterior lamina, isolating the paraflexus from the hypoflexus (Fig. 3). Basisphenoid-basioccipital suture varies among frequent open and almost closed. Anterior root of first molar moderately exposed, more visible in occlusal view. Pelage sparse, thin, and wavy, composed mostly by viliforms, but with more setiforms and aristiforms; pelage with dorsal color predominantly grayish-brown with a light darker-brown middle band and, ventral color grayish with orange spots (Alagoas_Wild: 106 individuals; Par_Captive: 44 individuals).

Age class 3: First and second molar with medium wear, cusps conspicuously eroded and with large dentin exposure; labial flexi more closed (especially metaflexus, forming a large metafosset) and posterofesset completely obliterated in M1 and M2; third molar exhibit marked wear (Fig. 3). Basisphenoid-basioccipital suture varies among almost closed and closed. Anterior root of first molar more exposed, medial lingual root of first, second and third molars moderately exposed in occlusal view. Pelage dense and lax, with abundant viliforms and very abundant setiforms, and more aristiforms (but not as abundant as setiforms); pelage with dorsal coloration generally brownish but exhibiting a darker-brown middle band and, ventral color grayish with orange spots (Alagoas_Wild: 108 individuals; Par_Captive: 52 individuals; Par_Wild: 9 individuals).

Age class 4: First and second molar with heavy wear, cusps become flat and there is a massive exposure of dentine; all labial flexi almost closed. Third molar appears quite flat, with major exposure of dentine (Fig. 3). Basisphenoid-basioccipital suture closed. Anterior root of first molar greatly exposed, medial lingual root of first, second and third molars exposed in occlusal view. Pelage dense and lax, with abundant viliforms and very abundant setiforms, and more aristiforms (but not as abundant as setiforms); pelage with dorsal coloration generally brownish but exhibiting a darker-brown middle band and, ventral color grayish with orange spots (Alagoas_Wild: 65 individuals; Par_Captive: 36 individuals; Par_Wild: 10 individuals).

Age class 5: Three molars are completely worn, with large exposition of dentine (Fig. 3); most flexi and fossets not visible, completely eroded. Basisphenoid-basioccipital suture closed. Anterior root of first molar greatly exposed, as well as the medial lingual root of first, second and third molars, and also the posterior root of third molar in occlusal view. Pelage dense and lax, with abundant viliforms and very abundant setiforms, and more aristiforms (but not as abundant as setiforms); pelage with dorsal coloration generally brownish but exhibiting a darker-brown middle band and, ventral color grayish with orange spots (Alagoas_Wild: 13 individuals; Par_Captive: 16 individuals; Par_Wild: 17 individuals).

2.2.3. Morphometric Data

Cranial and dental measurements were taken with a digital caliper (accurate to 0.01 millimeter) under a stereomicroscope, including: length of the upper molar series (LM); breadth of M1 (BM1); length of incisive foramen (LIF); breadth of incisive foramen (BIF); breadth of the incisor tips (BIT); breadth of palate (BP); length of nasal (LN); breadth of nasal (BN); least interorbital breadth (LIB); breadth of braincase (BB); breadth of zygomatic plate (BZP); depth of incisor (DI); breadth of the occipital condyles (BOC); length of palatal bridge (LPB); orbital fossa breadth (OFB); breadth of rostrum (BR); length of interparietal (LI); breadth of interparietal (BI); bullar breadth (BUB); lambdoidal breadth (LB), condyle-zygomatic length (CZL). Appendix B shows an illustration of all morphometric variables and its descriptions.

2.2.4. Morphological Data

In order to document qualitative changes between age classes and sexes, we evaluated in Alagoas_Wild and Par_Wild populations characters of the skins, such as structure and texture of the pelage, tail and ear pilosity, length of the mystacial vibrissae, and number and length of scales on the tail, and general coloration on the dorsal, lateral, and ventral sides. We also analyzed in detail cranial regions such as the rostrum, zygomatic arches, secondary palate, interorbital region, braincase, cranial

foramina, pterygoid region, mandible, molars and molar roots. Anatomical terminology follow Reig (1977) for the molars, Wahlert (1985) for the cranial foramina, and Carleton (1980), Voss (1988), and Weksler (2006) for general cranial characters.

2.2.5. Ontogenetic and sexual variation analyses

In order to access the non-geographic variation we performed traditional statistical tests (see Prado & Percequillo, 2011). We used the morphometric variables to calculate descriptive statistics for all age classes, and applied the Student's *t* test to check for sexual dimorphism in each population. We also calculated the ratio of the differences between mean values of males and females for all variables in order to estimate the dimension of the differences related to sex (Abreu-Júnior *et al*., 2012). We then employed analysis of variance (ANOVA and MANOVA) and Tukey's *post hoc* test to check for age variation with males and females pooled or not pooled (see results).

We selected 14 qualitative morphological traits out of several morphological characters that exhibited some degree of ontogenetic variation and we calculated the frequencies of the character states for each age class, plotting them in pie–chart graphics for visual identification of sharp discontinuities.

2.2.6. Comparisons in the ontogenetic development across populations

We performed morphometric comparisons amongst age classes employing four different Canonical Discriminant Analysis (CDA). We firstly performed a CDA using all individuals and the log10-transformed variables, considering each age from each of the three sampled population as a different factor, resulting in 13 different factors (e.g. five age classes in Alagoas_Wild sample, five age classes in Par_Captive sample, and three age classes in Par_Wild sample). We then constructed a Dice Leraas diagram with the scores of the first discriminant function (that accounted for the higher percentage of variation, see results) by age class and sampled population plotting them in the same graph for better visualization of the differences. Next we performed a second CDA using all individuals and the log10-transformed variables, but this time we pooled the three populations in each age class, resulting in 5 different factors (e.g. age class one in Alagoas sample was considered the same factor as age class one in the others two samples), and the same proceedings of the first analysis to construct the Dice Leraas diagram were applied to build the graph.

Subsequently, we performed again the same two CDAs (following the same workflow of the first two), but now instead of employ the log10-transformed variables we used the residuals of a linear regression with species and sex as the factors (lm (VARIABLES ~ SPECIES + SEX)), aiming to remove the effect of these factors. The motivation for these two last CDAs was the fact that specimens

from Alagoas and the two populations trapped in Paraguay (Captive and Wild) belong to different species (see Chapter 3 of this document), and this might include an extra source of variation in the analyses, so we wanted to remove from the CDA all the differences that were related with sex, geography or taxonomy, and keep only differences related with age class variation.

Finally, we performed a PCA analyses (with the log10-transformed data) and used the scores of the first principal components in different Linear Regressions, in order to compare ontogenetic trajectories across populations by analyzing the correspondence of the regression lines inclinations. The first component of a principal component analysis (PCA) derives from the multivariate generalization of allometry, and offers us the possibility to compare ontogenetic trajectories (Klingenberg & Froese, 1991; Wilson & Sánchez-Villagra, 2010). We first performed two Linear Regressions in Alagoas_Wild and Par_Captive populations (using age classes 1, 2, 3, 4 and 5) and plotted the two regressions lines with first PCA scores; we then performed three others Linear Regressions in each of the three populations and using age classes 3, 4 and 5, and plotted the them with first PCA scores.

2.2.7. Association between morphological traits and environment

The potential impact of environmental factors on morphological differentiation between the two wild populations (Par_Wild and Alagoas_Wild) was tested using Distance-based redundancy analysis (dbRDA; Legendre & Anderson 1999). We tested for the relationship between individual pairwise morphological distances and corresponding climatic variables. The environmental variables (Hijmans *et al*., 2005) were extracted at each sampled locality, using the function *extract* of the Raster package (Hijmans *et al*., 2016) in R software (R Core Team 2014), and include: Annual Mean Temperature (1), Mean Diurnal Range (2), Isothermality (3), Temperature Seasonality (4), Maximum Temperature of Warmest Month (5), Minimum Temperature of Coldest Month (6), Temperature Annual Range (7), Mean Temperature of Wettest Quarter (8), Mean Temperature of Driest Quarter (9), Mean Temperature of Warmest Quarter (10), Mean Temperature of Coldest Quarter (11), Annual Precipitation (12), Precipitation of Wettest Month (13), Precipitation of Driest Month (14), Precipitation Seasonality (15), Precipitation of Wettest Quarter (16), Precipitation of Driest Quarter (17), Precipitation of Warmest Quarter (18), and Precipitation of Coldest Quarter (19). Following a PCA analysis was performed with the bioclimatic data and the scores of the First Principal Component (which represented 79.74% of the variance) were used in the dbRDA analyzes, conditioned or not on geographic distances (i.e., removing the effect of geographic distance separating individuals). Mahalanobis geographic distances among all individuals were calculated using the function *pairwise.mahalanobis* of the HDMD package (McFerrin, 2013) in R software (R Core Team 2014), and the geographic distance matrix was calculated with *earth.dist* function of the Fossil package

(Vavrek, 2015), and transformed into continuous rectangular vectors via principal coordinates analyses using the *pcnm* function of the Vegan package (Oksanen *et al*., 2012).

Additionally to the dbRDA analyzes the correlation between the environmental and morphometric variables matrices was computed using the function *matcor* of the CCA package (González & Déjean, 2012). The dbRDA and correlations analyses were performed between individuals of the two wild populations (Alagoas_Wild and Par_Wild) and for age classes 3, 4 and 5 (ages separated and together). The significance levels of the model generated by dbRDA analyzes were recovered by ANOVA.

2.3. Results

2.3.1. Description of Non-Geographic Variation: Age and Sex

Morphometric analysis

Sexual dimorphism analyses performed on the Alagoas_Wild sample showed great significant sexual variation in age classes 3 and 4 (Fig. 4; Appendix C, Table C.1). Only variables related to molars, interorbital, and interparietal regions (CSM, LM1, LIB and IL) did not show significant differences. However, the difference between the means for each sex in each variable and the proportion of the difference calculated for all age classes exhibited small variation, and opposite to the student *t* test, the classes 1 and 5 presented the higher range of variation. Age class 1 exhibited the percentage of difference that range from 0.06% to 12.37% in males and 0.06% to 11.01% in females, and the length of palate (LPB) was the variable responsible for the higher proportion of difference. Age class 2 presented the lower range of variation, males varying between 0.05% and 2.63% and females between 0.05% and 2.70%, and the variables related with the nasal (LN, BN) showed the higher proportion of difference. In age class 3 the proportion of the difference in females varies between 1.23% and 5.17% and from 1.22% to 4.92% in males, and the orbital fossa breadth (OFB) presented the higher proportion of difference. In males of age class 4 the proportion varies from 0.12% to 4.70% and from 0.12% and 4.69% in females individuals, and the Lambdoidal breadth (LB) and orbital fossa breadth (OFB) exhibited the higher proportion of variation. Finally, the range of the proportion of variation in males of age class 5 goes from 0.02% to 9.87%, and in females it varies from 0.02% to 8.98%, with the interorbital length (IL) being the variable with higher proportion of the difference (Appendix C, Table C.2).

Individuals from the Par_Captive sample presented the lowest number of variables with sexual dimorphism, and were all found in age classes 3, 4 and 5 (Fig. 4; Appendix C, Table C.3). The range of the percentage of variation was also low and homogeneous across age classes. In classes 2, 3 and 4 the range of variation goes from 0.04% to 6.16%, and the class 5 exhibited the larger range, with

values between 0.05% and 7.76%; variables related with incisive foramen (BFI; age class 2), nasal (BN; age classes 3 and 4), and palate (LPB; age class 5) exhibited the higher proportion of variation (Appendix C, Table C.4). Par_Wild sample showed significant differences only in few variables, especially in age class 5, where variables related with incisive foramen, interorbital and overall size of the skull exhibited higher differences (note that this population only has individuals from age classes 3, 4 and 5; Fig. 4; Appendix C, Table C.5). However, this population exhibited the greater range in the proportion of the difference; age class 3 follow the patterns of the others population varying between 0.68% and 8.30%, but the age class 4 and 5 presented a range of variation that goes from 0.01% to 25.81%. Variables related with the braincase (BB), interorbital (IL) and nasal (BN) showed the higher proportion of variation in age class 3, 4 and 5, respectively (Appendix C, Table C.6).

In addition, most variables that exhibited differences between both sexes were heterogeneous among samples and age classes: on the sample Alagoas_Wild, 10 variables were dimorphic on both age classes 3 and 4, but only one of these is also dimorphic in age class 5; on Par_Captive, no variable are consistently distinct among age classes; and on Par_Wild, just one variable is consistently distinct on age classes 3 and 4. Regarding age classes, individuals from class 3, only 3 variables are consistently different among pairs if the three samples (BM1, Par_Captive and Par_Wild; BN, Alagoas_Wild and Par_Captice; and BOC, Alagoas_Wild and Par_Wild); on the class 4, only one variable is shared as dimorphic by two samples (Alagoas_Wild and Par_Wild); and on class 5, CZL is dimorphic in Par_Captive and Par_Wild.

The ANOVA performed with Alagoas_Wild sample showed significant difference amongst age classes for all variables. We can observe that age class 1 is the mostly different from the others, except by LIB (least interorbital breadth). Age class 2 also showed significant differences between other age classes for all variables, except by LM, BM1 and LIB that did not show differences in comparisons between any age classes. Regarding the age class 3, the females presented more differences with age class 4 than with age class 2, while males showed inverse pattern when compared to age classes 2 and 4 (e.g. females individuals exhibited lower mean values, bringing them closer to class 2, whereas males exhibited higher mean values, bringing them closer to class 4). Finally, age classes 4 and 5 exhibited few variables with significant differences between each other (Appendix C, Table C.7), however we can also notice that female individuals presented more differences than males.

The Par_Captive sample exhibited few variables with significant difference, only related with age classes 1 and 2 (Appendix C, Table C.8). The statistical analyses performed on the Par_Wild sample with age classes 3 to 5 showed only two variables (BFI and PB) with significant differences between ages 3 and 4 (Appendix C, Table C.9).

Regarding the variation in the multivariate approach (MANOVA), both age and sex factors were shown to be statistically different in the Alagoas_Wild sample, although the interactions between them were not (Table 2). The Par_Captive sample did not show any significant variation (Table 2), and Par_Wild presented multivariate differences only regarding to sex (Table 2). In Appendix C

(Tables C.10, C.11, and C.12) we present a complete list of descriptive statistics for all age classes and variables for each population separately.

Morphologic analysis

Considering the external traits, there was variation in the color patterns among ages (as presented above, in the description of age classes), and the number of caudal scales (per cm) decreased with the growth of the individuals (Fig. 5). In general, the morphologic variation among age classes in skulls of specimens from Alagoas_Wild and Par_Wild samples is related with the development cranial crests (supraorbital, temporal, lambdoidal, occipital, and postorbital), where younger individuals presented smoother crests or none at all, while older ones frequently showed well-developed crests (Figs. 6, 7, 8, 9 and 10).

The changes in the shape of the skull of *Holochilus* through ontogenetic development is unparalleled in other members of tribe Oryzomyini. The rostral region grows in length, but also in height, with a conspicuous enlargement of the zygomatic notch, both laterally and antero-posteriorly, with the excavation of the medial wall of the notch (dorsal to the nasolacrimal capsules) and with the growth in width of the zygomatic plate and its zygomatic spine (that is variably present), respectively; there is also a noticeable development of the nasolacrimal capsules. Regarding the zygomasseteric complex (including the interorbital region) and the associated masticatory apparatus, there are deep changes: there is a projection of the molar series ventrally, accompanied by a growth in the height of the skull in this region (a consequence of the projection of supraorbital crests dorsally); the zygomatic arches become quite robust, deep and more projected laterally; and there is also a conspicuous increase on the depth of the dentary; younger individuals showed a shorter palate, with the mesopterygoid fossae reaching the molars, while individuals of older ages classes presented an intermediate length palate, which is characterized by the mesopterygoid fossae reaching the maxillary bones, but not the molars; the parapterygoid plates become wider and more deeply excavated. The braincase also experiments shape modifications, changing from a rounded, smooth and un-crested braincase towards a elongated, narrower and conspicuously crested structure, with supraorbital, temporal, postorbital, lambdoidal and occipital crests highly developed and acute. On the occipital region, the paraoccipital process became well developed in all individuals only after the age class 4.

The capsular process of the lower incisor alveolus was conspicuous only after the age class 3. We observed an increment of the number of fossets in the caudal portion of the molars, more precisely on the region of the posteroloph of M1 and M2, during the ontogenetic development.

2.3.2. Interspecific Variation: ontogenetic development

In the analysis depicted in figure 11A, its observable that all samples are quite distinct among themselves, with samples from Alagoas_Wild presenting mean score values increasing significantly from class 1 to 5; on the other hand, both samples from Paraguay exhibit an increase until the age class 4, with a pronounced decrease of the mean value score for age class 5. This first discriminant function (CDA; Fig 11A) was responsible for 75.23% of the variation, and the most important variables were CZL (positively), BOC and LN (negatively); the second function was responsible for 12.75%, and the variables BB (negatively) and CZL (positively). Scatterplots with the raw measurements among the most important variables in this first CDA (CZL, BOC and LN) corroborates that in fact the range of variation in Alagoas_Wild population is larger than in the others, and individuals from age classes 1 and 2 are smaller in size than specimens from the Paraguayan populations (Appendix C, Figs. C.1 and C.2).

Other analytical approaches (Figs. 11B, 11C and 11D) revealed similar patterns of variation although with less differences among mean score values of samples, with slight increment, stabilization or even decrease in size from class 4 towards class 5. On the second CDA (Fig. 11B), the first discriminant function was responsible for 71.88% of the variation, and the most important variables were CZL and DI (positively), and LM and BR (negatively); the second function was responsible for 13.44%, and the variables BOC (negatively) and CZL (positively) were the most important. The third CDA (Fig. 11C) revealed that the first discriminant function was responsible was responsible for 56.36% of the variation, and the most important variables were DI (negatively), BR (positively); the second function was responsible for 14.41%, and the variables BM1 (negatively) and DI (positively) were the most important. The last CDA (Fig. 11D) showed that the first discriminant function was responsible for 69.11% of the variation, and the most important variables were DI (positively), BR (negatively); the second function was responsible for 14.26%, and the variables positively BOC (negatively) and BIF (positively) were the most important.

The age variation across population (conspicuous or not) observed in the analyses above is reflected in the ontogenetic trajectory of the three populations (accessed by the tendency of the line of the linear regression with the first principal components analysis; Fig. 12). Figure 12A shows the two linear regressions performed with all age classes in the populations Alagoas_Wild and Par_Captive. Most of the individuals from Alagoas population presented higher PCA scores than the Par_Captive population, and the regression line of the Captive sample was conspicuous more parallel with X-axis than the regression line of Alagoas_Wild population, but both lines presented the same decreasing pattern of PCA scores, although this pattern is more inclined in Alagoas_Wild sample, meaning that this sample exhibited a range of differentiation much higher than the Par_Captive sample. However, in Fig 12B its observable that the regression line of Par_Captive population is even more paralleled with X-axis, that is the size of age classes 3, 4 and 5 almost do not change during the growth trajectory in

this population; Alagoas_Wild population presented the same growth pattern as the analysis above, and the Par_Wild sample showed an intermediate degree of inclination.

2.3.3. Morphological traits and environment

The First Principal Component Analysis with the bioclimatic variables represented 79.74% of the variance observed between environments. The most important variables were most related with differences in temperature, such as Isothermality (Bio3), Temperature Seasonality (Bio4), Temperature Annual Range (Bio7), and also Precipitation Seasonality (Bio 15).

Despite environmental differences, no significant association between environment and morphological differentiation was detected with dbRDA analyses (age classes together or separately), even when conditioned on the geographic distances between populations (i.e., controlling for the effects of geographic isolation; Table 3), which leads us to think that environment is not a good predictor of morphological differentiation in age classes 3, 4 and 5. The cross-correlation analysis between the environmental and morphological variables performed with both wild populations together also did not show great correlation between them in age classes 3 and 4 (Figs. 13A and 13B). Age class 4 presented some higher correlations between variables PB and LIB with all bioclimatic variables, except by bio1 and bio12. Age class 5 (Fig. 13C) exhibited higher correlations in at least five morphometric variables (LM, BFI, LN, LIB, BB, and BUB), and almost all bioclimatic variables. We also performed the same cross-correlation analyzes in the two wild populations separately. The analysis in the Alagoas_Wild population showed high correlation between the morphometric and bioclimatic variables only in age class 5, and the higher correlations were observed in the variables BFI, BIT, DI and BR (all related with the frontal part of the skull and incisive) and all bioclimatic variables (Appendix D, Fig D.1). In the Par_Wild population we observe that all age classes exhibit some variables with great correlations. In age class 3 the variables BIT, PB, OFB, BR and LI presented the higher correlations between the morphometric variables and all bioclimatic variables (Appendix D, Fig D.2). In age class 4 the variables LFI and BIT presented the higher correlations (Appendix D, Fig D.2). In age class 5 the variables BN, LIB and BB showed the higher correlations (Appendix D, Fig D.2).

2.4. Discussion

2.4.1. Sexual Variation

Compared to other sources of variation, the sexual dimorphism is often considered non-significant font of variation in rodent literature (Carleton & Musser, 1989; Musser & Williams, 1985;

Voss, 1991; Percequillo, Hingst–Zaher & Bonvicino, 2008; Abreu-Júnior *et al*., 2012), although some studies have shown that significant differences between sexes can also be found in this group (Myers & Carleton, 1981; Brandt & Pessôa, 1994; Carmadella, Pessôa & Oliveira 1998; Prado & Percequillo, 2011). These distinct interpretations reflect the biological differences among different species and populations, may also reflect the statistical methods to approach the variation, as well as the sample distribution, both geographically and across age classes.

In this study, we found significant sexual dimorphism in the student *t* test and MANOVA mostly in age classes 3 and 4 of Alagoas_Wild population. However, when we analyze the proportion of this variation we can observe that the percentage are less than 10% in most of the variables and in all populations (except for the variables PB, BN and IL in age class 4 of Par_Wild population that shows proportions between 10.03% and 25.81%, and variables LFI, BFI, BIT and BN in age class 5 of Par_Wild population that shows proportions between 10.41% and 13.23%).

Few studies have compared the differences between sexes through this approach for rodents, although it allows us to better understand the magnitude of the variation beyond the raw values of the means. Goldman (1918), found similar ratio values between males and females, and considered it correspondent with the range of individual variation, and therefore not related to true sexual dimorphism. Goldman (1918) also explained that it is difficult to access precisely this kind of information analyzing small series, which can be the reason we found higher proportions values in the Par_Wild population, which is the one with lower sample number.

Although we assume from the data above that the sexual dimorphism is overall negligible in the genus *Holochilus*, which was expected for an oryzomyine rodent (see Abreu-Júnior *et al*., 2012 for a more comprehensive study of dimorphism in the tribe Oryzomyini), some degree of sexual variation is presented in our samples, always linked with older age classes (3, 4 and 5). Such pattern is also expected, as it is during the adult life that the individuals experience higher levels of competition for food and shelter, and even reproductive habits, such as mating systems, which can influence the degree of sexual dimorphism. However we do not have enough information about the ecology of this genus to hypothesize about these possible explanations. Our findings are in the opposite direction of the results presented by Voss, Marcus & Escalante (1990), where tests for sex effects in captive population of the genus *Zygodontomys* revealed significant dimorphism in most cranial measurements, while natural population did not present such pattern. They believe that it is because they controlled all the sex comparisons not only for the usual condition such as food, temperature and age, but also for family (parental), and (within age classes) for natal litter size, isolating the sexual dimorphism by design.

2.4.2. Age Variation

Ecological adaptations can explain many of the patterns and mechanisms leading to morphological variation, both among species, populations and stages of life. The significant differences among age classes in the multivariate and univariate analyses is more conspicuous in age classes 1 and 2, even though the sample from Alagoas_wild exhibited lower degree of similarity even among age classes 3, 4 and 5.

For most of mammals, the bite force has a central role in shaping cranial morphological variation and evolution, and the determined growth in this group implies that the performance of the regions related with the bite apparatus must change during ontogenesis until reaches its adult size and morphology (Santana & Miller, 2016). Such allometry could explain lower variation in age classes 3, 4 and 5, when the masticatory apparatus of rodents are almost fully developed.

Furthermore, variables responsible for the most part of the variation in the discriminant analyses (CZL, BOC, LN, DI, BR), and that presented higher variation in younger classes, are related to skull size, nasal, rostrum, braincase and incisors, i.e., regions directly related to the bite apparatus. Populations from Paraguay (Wild and Captive), although we miss younger samples, reflected such pattern of ontogenetic variation in the univariate and multivariate analyses, presenting almost none variable with significant differences in the ANOVA (in age class 3, 4 and 5). Fig 11A shows a increase of values of the discriminant scores from age class 3 to 4, and a posterior decrease of these values from age 4 to age 5 in the Paraguayan populations, although the magnitude of these changes differ between these two populations; the variation between age classes is less conspicuous in the Wild population (note that this pattern was only recovered in the analysis with log10-transformed and 13 age classes factors on the. 11A). On the other hand Alagoas_Wild population did not exhibit the decrease of discriminant scores in age class 5.

This pattern also can be related with the bite force apparatus, as Chazeau *et al*. (2013) in a study with captive mouse lemurs found that older animals generally bite with more force, but the oldest age group that they analyzed showed a decreased of the bite-force capacity. Voss, Marcus & Escalante (1990) in a study with captive population of the rodent *Zygontomys* also reveals that in older animals variables related with the neurocranium and molars are proportionately smaller as a consequence of prolonged growth in the facial skeleton.

The overall skull size, facial skeleton, some parts of the neurocranium (posterior part of braincase), palate and incisors variables in the Par_Captive population of *Holochilus* grew continuously until the early adult stege; variables related braincase and orbit ceased growth earlier in age class 1 (Appendix C, Table C.8). This pattern corresponds partially to the pattern found by Voss, Marcus & Escalante (1990) in a captive population of *Zygodontomys*. According to the authors, the variables related with facial skeleton and incisors on this rodent have an indeterminate and continue

growth until age 5 variables related with the neurocranium, however, completes growth earlier in postnatal life, and the braincase and interorbit increase only slightly after molar eruption.

In the Alagoas_Wild population the molars and the interorbital region completes growth earlier in age class 1; palate, interparietal and the posterior part of neurocranium, grew continuously until the early adult phase (class 3), and the other variables have an indeterminate and continue growth throughout all age classes (Appendix C, Table C.7). We also notice that most of the morphological differences found in this study are related with younger age classes (1 and 2), except by: i) the length of the palate, that after age class 1 stabilizes; ii) the basisphenoid-basioccipital suture and paraoccipital process and coloration that show some variation in age class 3. In the Par_Wild population, as we have data only from the age classes 3, 4 and 5, we can observe that all variables stabilizes their growth during the older age classes.

The ontogenetic trajectories accessed by linear regressions exhibited different patterns of ontogenetic development differing in magnitude and direction of the variation. The ontogenetic line of the Par_Wild and Par_Captive samples were almost parallel (Fig. 12B), but the direction is slightly different, as well as the magnitude, where individuals from the captive population presented lower signal of age variation between the analyzed classes and its relationship is generally more linear and isometric. The direction and rate of the Alagoas population ontogenetic trajectory is reflected in the strongest line inclination observed in the data. The ontogenetic trajectories of Alagoas_Wild and Par_Wild populations showed similar direction but strong different magnitudes.

The most noteworthy variation among age classes considering the Alagoas_Wild and Par_Captive samples is on age class 1 and on a minor extent on age class 2 (Fig 12A): when these two classes are removed, there is more similarity between the two samples (and also with the Par_Wild sample; Fig. 12B); individuals of age class 5 are virtually equal on the three samples. Individuals of age class 1 from the Par_Captive sample are larger and more robust than specimens from Alagoas_Wild (Appendix C, Tables C.10 and C.11), and this phenomenon could have some explanations: i) in captivity, with less stressful conditions and constant food supply, females gave birth to larger offspring than females from the wild, that are under more critical conditions; ii) in captivity, molars were eroded more slowly than in the wild, and individuals grow without molar wear; iii) the species from Paraguay is larger than the species from Alagoas, and females gave birth to larger offspring. Differences in the number of offspring could give us some hint about differences in size of these litters, although our captive data shows that crossbreeding in the captive population generates 1-10 litters, with a modal number of 4 litter per cross, and data from the literature (Twigg, 1962; Lira *et al*., 2016) shows that the number of litters also can varies from 2-10 in natural populations.

Regardless the geographic origin of the samples, there is a general trend in the qualitative changes in the shape of the skull in *Holochilus*. As previously discussed, the more noticeable changes are associated to the masticatory apparatus, with the enlargement and increase in the robustness of the zygomassateric structures, and more uncommonly, with the increase in the height of the skull on this

region, with the projection of the molar series ventrally (on younger individuals, the molar series is on the same level of the palate, while in older specimens, the molar series is few millimeters ventral to the palate). The molar series is in a plane more distant from the cranium-mandibular articulation (glenoid fossa), favoring a more effective grinding action of all molars simultaneously. As the molar crowns in *Holochilus* are only moderately hypsodont, the alternative was the projection of the molar series ventrally; alternatively, this projection could also exist to compensate the molar wear (in addition to the exposition of molar roots to compensate the molar oclusal wear). These changes are probably accompanied by modifications on the zygomasseteric (superficial, lateral and medial) and temporal muscles associated to this region that will likely be longer and more robust (wide in cross section). *Holochilus* shares with other grazing herbivorous mammals these adaptations on the skull for herbivory, suggesting that this is the most herbivorous lineage on the oryzomyine radiation. The tribe Phyllotini abridges the most herbivorous lineages amongst the sigmodontines, but the solutions to attend this alimentary preference were slightly different (see Steppan, 1995), with most genera exhibiting highly hypsodont molars, convex skulls (*Galenomys*) and distinct zygomasseteric apparatuses, with high and wide zygomatic plates and zygomatic arches well projected ventrally (*Galenomys* and *Reithrodon*).

2.4.3. Environmental variation and skull morphology

Non-geographic differences among populations and species may be resultant from genetic or environmental factors (and even the interaction between both) affecting several developmental mechanisms. As a consequence, in studies of this kind of variation it is essential to take into account ecological aspects to predict how these animals may have competed during environmental and climatic variations. Moreover, it is crucial to understand if the environmental differences between the sampling sites can be a useful approach for unraveling the environmental factors that directly influence phenotypic traits (Binning, Chapman & Dumont, 2010).

We have analyzed data from different species and environments, and the bioclimatic associations tests showed that the current differences among the climatic conditions of the sampled localities is not enough to explain the variance in the ontogenetic development observed in the two wild populations (Alagoas_Wild and Par_Wild). This fact might be related to the lack of samples from age classes 1 and 2 from Par_Wild, as these two classes were the most variable. Although, as both environments differ considerably, as there is some correlation between morphological distance and bioclimatic variables (specially in age class 5, see Fig. 13), and as the morphometric differences between these two population are conspicuous (see CDAs analyzes), the environment influence might be related to the divergence time that these two wild populations are experiencing in these completely different eco-regions. However, we do observe differences both in age variation between age classes

(in the CDAs analyzes) and ontogenetic trajectory between wild populations, which let us to think that the genetic pool is also an important force driving the differences in the ontogenetic development (as described above).

The variation observed in the degree of correlation between morphometric and environmental variables across age classes 3, 4 and 5 when we analyzed both population together and when we analyzed each population separately is conspicuous. The dataset containing both populations (Fig. 13) shows an increase of the number of variables highly correlated from age 3 to 5. The variables that exhibited highly correlation in age 5 are related with teeth, incisive foramen, interorbital region, braincase and bulla. However when we examine the correlation only in Alagoas_Wild population we do not observe high correlation in age class 3 and 4, but only in 4 variables (related with incisive foramen, teeth and rostrum) in age class 5. In the Par_Wild population the three age classes presented variables with high correlation, and age class 3 presented the largest number of variables with high correlations. The results indicate that populations located in drier climatic region from northeastern Brazil exhibited less influence of environmental variables than the population located the humid Chaco region in Paraguay, suggesting that the Paraguayan population is more environmental plastic than the Alagoas_Wild population.

Cranial and dental morphology has been shown to display a significant relationship with distinct demands of food processing inherent to varying life habits, specially in herbivorous diet species (Samuels, 2009). The mechanical effort required to process tough plant materials has been linked with the development of a more massive jaw, larger cheek tooth area and a deeper skull and rostrum to provide an increased area for muscle attachment, thus improving efficiency and power during mastication (Samuels, 2009). Plasticity among cranial dental traits related to food processing was recovered by Oliveira, Strauss & Reis (1998) in *Necromys lasiurus*. They also recovered that processes of molar wear, suture ossification, and molar-root exposure are less integrated in samples collected during the dry seasons, than in samples collected in the wet season. That is, in the dry season the suture ossification and root exposure occurs slowly than molar wear, and the molar wear is not highly correlated with age, which means that the effect of environment in the dry samples is higher. The development of the basisphenoid-basioccipital suture ossification in *Holochilus* did not differ among *Holochilus* populations, but the molar wear in Alagoas_Wild population seems to happen fast and in a higher degree than in Par_Wild population.

Concluding, age classes classification can represent “biological age”, as described by Oliveira, Strauss & Reis (1998), and the identification of these indexes can be extrapolated for the entire genus, they are subjected to taxonomic and environment variation, although genetic responses seems to have a stronger effect in age variation across populations than differences in bioclimatic conditions.

Understanding the patterns of non-geographic variation in rodents allowed us to increase the knowledge about how this variation is structured in this large and diverse group, and also to

comprehend how rodents can respond to spatial differences in the environment both in microevolutionary (between populations) and macroevolutionary scale (between species). Future studies with the genus *Holochilus* might aim to use quantitative genetic approaches to estimate the relative contributions of genetic and environmental factors in the variation observed in the current study, as well as explain differences related with the ontogenetic trajectory between population, by comparing data from captive and wild populations generating by a phenotypic P (collected in the wild without known genealogy) and a genotypic G (constructed from specimens grown in colony) matrices.

Our results show that different species and populations present different degrees of sexual dimorphism and ontogenetic variation, which might be related to short-term responses to habitat fluctuations or to long-term evolutionary changes driven by natural selection or genetic drift. Our results confirm our hypothesis that species from the wild exhibited similar rates of morphological change probably driven by selective forces, although both samples from Paraguay (wild and captive) are more similar in size than Alagoas sample as a response to a common and shared evolutionary history.

ACKNOWLEDGEMENTS

This work was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP; Proc: 2014/22444-0, 2012/24099-3, 2009/16009-1). We are also grateful to the curators and researchers of the Museu Nacional da Universidade Federal do Rio de Janeiro (João Alves de Oliveira) and University of Michigan Museum of Zoology (Priscilla Tucker and Cody Thompson) for allowing us to access the collections and collect the data.

REFERENCES

Abreu Junior EF, Brennand PGG, Chiquito EA, Jorge-Rodrigues CR, Libardi GS, Prado JR, Percequillo AR. 2012. Dimorfismo sexual na tribo Oryzomyini. In: Freitas TRO, Vieira EM, ed. *Mamíferos do Brasil: genética, sistemática, ecologia e conservação. 1ed.* Rio de Janeiro: Sociedade Brasileira de Mastozoologia, 115–134.

Binning SA, Chapman LJ, Dumont J. 2010. Feeding and breathing: trait correlations in an African cichlid fish. *Journal of Zoology* **282**:140–149.

Brandt RS, Pêssoa LM. 1994. Intrapopulacional variability in cranial characteres of Oryzomys subflavus (Wagner, 1842) (Rodentia: Cicretidae), in northeastern Brazil. *Zoologischer Anzeiger* **233**:45–55.

Brandt JF. 1835. Mammalium rodentium exoticorum novorum vel minus rite cognitorum. Musei Academici Zoologici descriptiones et icons. *Memóires de L'Académie Impériale des Sciences de Saint-Pétersbourg* **6:**358–442.

Carleton MD. 1980. Phylogenetic relationships of neotomine-peromyscine rodents (Muroidea) and a reappraisal of the dichotomy within New World Cricetinae. *Miscellaneous Publications Museum of Zoology*, *University of Michigan* **157**:1–146.

Carleton MD, Musser GG. **1989.** Systematic studies of oryzomyine rodents (Muridae, Sigmodontinae): a synopsis of *Microryzomys*. *Bulletin of the American Museum of Natural History* **191**:1–83.

Carmadella AR, Pessôa LM, Oliveira JA. 1998. Sexual dimorphism and age variability in cranial characters of *Oryzomys subflavus* (Wagner, 1842) (Rodentia: Sigmodontinae) from northeastern Brazil. *Bonner Zoologische Beiträge* **48**:9–18.

Chazeau C, Marchal J, Hackert R, Perret M, Herrel A. 2013. Proximate determinants of bite force capacity in the mouse lemur. *Journal of Zoology* **290:** 42–48.

Chimimba CT, Dippenaar NJ. 1994. Non-geographic variation in *Aethomys chrysophilus* (De Winton, 1897) and *A. namaquensis* (A. Smith, 1834) (Rodentia: Muridae) from southern Africa. *South African Journal of Zoology* **29**:107–117.

Domínguez-Bello MG, Robinson MD. 1991. A Comparison of Digestive Adaptations in Two Neotropical Cricetid Rodents (*Holochilus venezuelae* and *Zygodontomys microtinus*). *Physiological and Biochemical Zoology* **64:** 1542–1551.

Fernandes FA, Fornel R, Cordeiro-Estrela P, Freitas TRO. 2009. Intra- and interspecific skull variation in two sister species of the subterranean rodent genus *Ctenomys* (Rodentia, Ctenomyidae): coupling geometric morphometrics and chromosomal polymorphism. *Zoological Journal of the Linnean Society* **155**:220–237.

Formoso AE, Sauthier DEU, Pardiñas UFJ. 2010. Mammalia, Rodentia, Sigmodontinae, *Holochilus brasiliensis* (Desmarest, 1819): Distribution extention. *Check list* **6**:195–197.

Goldman EE. 1918. The rice rats of North America (genus *Oryzomys*). *North American Fauna* **43**: 1–100.

González I, Déjean S. **2012.** *CCA: Canonical correlation analysis*. R Package Version 1.2. Available at: http://CRAN.R-project.org/package=CCA

Herring SW.1993. Epigenetic and functional influences in skull growth. In: Hanken J, Hall BK, ed. *The skull.* Chicago: The University of Chicago Press, 153–206.

Hershkovitz P. 1955. South American marsh rats genus *Holochilus*, with a summary of sigmodont rodents. *Fieldiana Zoology* 37: 639–673.

Hershkovitz P. 1962. Evolution of Neotropical cricetine rodents (Muridae) with special reference to the phyllotine group. *Fieldiana Zoology* 46: 1–524.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**:1965–1978.

Hijmans RJ, van Etten J, Cheng J, Mattiuzzi M, Sumner M, Greenberg JA, Lamigueiro OP, Bevan A, Racine EB, Shortridge A. 2016. *Raster: Geographic Data Analysis and Modeling*. R Package Version 2.5–8. Available at: http://cran.r-project.org/package=raster

Hingst-Zaher E, Marcus LF, Cerqueira R. 2000. Application of geometric morphometrics to the study of postnatal size and shape changes in the skull of *Calomys expulsus*. *Hystrix* **11**:99–113.

Kawazoe U, Pinto ACM. 1983. Importância epidemiológica de alguns animais silvestres na esquistossomose mansônica. *Revista de Saúde Pública* 17:345–346.

Klingenberg CP, Froese R. 1991. A multivariate com- parison of allometric growth patterns. *Systematic Zoology* **40**: 410–419.

Legendre P, Anderson MJ. 1999. Distance-based redundancy analysis: testing multi-species responses in multi-factorial ecological experiments. *Ecological Monographs* **69**:1–24.

Libardi GS. 2013. Variação não-geográfica em Necromys lasiurus (Lund, 1840) (Cricetidae: Sigmodontinae) no Brasil. Unpublished Master Dissertation, Universidade de São Paulo.

Lira MGS, Miranda GS, Rodrigues JGM, Nogueira RA, Gomes GCC, Cantanhêde LG, Silva–Souza N. 2016. Aspectos Biológicos de *Holochilus* sp., hospedeiro natural da esquistossomose. *Ciência Animal Brasileira* **17**:143–153.

Martino AMG, Aguilera M. 1989. Food habits of *Holochilus venezuelae* in rice fields. *Mammalia* **53**:545–561.

Massoia E. 1971. Las ratas nutrias argentinas del genero *Holochilus* descriptas como *Mus braziliensis* por Waterhouse (Mammalia–Rodentia–Cricetidae). *Revista de investigaciones Agropecuarias* **8**:141–148.

Massoia, E. 1976. Mammalia. In: Ringuelet R, ed. *Fauna de agua dulce de la Republica Argentina*. Fundación Editorial Ciencia y Cultura, 1–128.

Mayr E, Ashlock PD. 1991. *Principles of Systematic Zoology*. New York: Mcgraw-Hill College.

McFerrin L. 2013. *HDMD: Statistical Analysis Tools for High Dimension Molecular Data (HDMD).* R Package Version 1.2. Available at: https://cran.r-project.org/web/packages/HDMD.

Monteiro LR, Lessa LG, Abe AS. 1999. Ontogenetic variation in skull shape of *Thrichomys apereoides* (Rodentia: Echimyidae). *Journal of Mammalogy* **80**:102–111.

Mullin SK, Pillay N, Taylor PJ. 2001. Non-geographic morphometric variation in the water rat *Dasymys incomtus* (Rodentia: Muridae) in southern Africa. *Durban Museum Novitates* **26**:38–44.

Musser GG, Williams MM. 1985. Systematic studies of oryzomyine rodents (Muridae): definitions of *Oryzomys villosus* and *Oryzomys talamancae*. *American Museum Novitates* **2810:**1–22.

Myers P, Carleton MD. 1981. The species of *Oryzomys* (*Oligoryzomys*) in Paraguay and the identity of Azara's "Rat sixième ou Rat à Tarse Noir ". *Miscellaneous Publications, Museum of Zoology, University of Michigan* **161**:1–41.

Oksanen J, Kindt R, Legendre P, O'Hara RB. 2012. *Vegan: community ecology package*. R package version 2.0–5. Available at: http://cran.rproject.org/web/packages/vegan/index.html

Oliveira JA, Franco SM. 2005. A coleção de mamíferos do Serviço Nacional de Peste no Museu Nacional, Rio de Janeiro, Brasil. *Arquivos do Museu Nacional* **63**:13–20.

Oliveira JA, Strauss RE, Reis SF. 1998. Assessing relative age and age structure in natural populations of *Bolomys lasiurus* (Rodentia: Sigmodontinae) in Northeastern Brazil. *Journal of Mammalogy* **79**:1170–1183.

Patton JL, Rogers MA. 1983. Systematic implications of non-geographic variation in the spiny rat genus *Proechimys* (Echimyidae). *Zeitschrift für Säugetierkunde* **48**:363–370.

Percequillo AR, Hingst–Zaher E, Bonvicino CR. 2008. Systematic review of genus Cerradomys Weksler, Percequillo and Voss, 2006 (Rodentia: Cricetidae: Sigmodontinae: Oryzomyini), with description of two new species from Eastern Brazil. *American Museum Novitates* **3622**: 1–46.

Prado JR; Percequillo AR. 2011. Ontogenetic and sexual variation in cranial characters of *Aegialomys xanthaeolus* (Thomas, 1894) (Cricetidae: Sigmodontinae) from Ecuador and Peru. *Papéis Avulsos de Zoologia* **51**:155–177.

R Development Core Team. 2014. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available at: http:// www.R-project.org

Reig OA. 1977. A proposed unified nomenclature for the enamelled components of the molar teeth of the Cricetidae (Rodentia). *Journal of Zoology* **181**:227–241.

Samuels JX. 2009. Cranial morphology and dietary habits of rodents. *Zoological Journal of the Linnean Society* **156**:864–888.

Santana SE, Miller KE, 2016. Extreme Postnatal Scaling in Bat Feeding Performance: A View of Ecomorphology from Ontogenetic and Macroevolutionary Perspectives. *Integrative & Comparative Biology* **56**:459–468.

Steppan, S.J. 1995. Revision of the tribe Phyllotini (Rodentia: Sigmodontinae), with a phylogenetic hypothesis for the Sigmodontinae. Fieldiana Zoology, **80**:1–112.

Twigg GI. 1962. Notes on *Holochilus sciureus* in British Guiana. *Journal of Mammalogy* **43**: 369–374.

Vavrek MJ. 2015. *fossil: Palaeoecological and Palaeogeographical Analysis Tools*. R package version 0.3.7. Available at: http://matthewvavrek.com/programs-and-code/fossil/.

Voss RS. 1988. Systematics and ecology of ichthyomyine rodents (Muroidea): patterns of morphological evolution in a small adaptive radiation. *Bulletin of the American Museum of Natural History* **188**:259–493.

Voss RS. 1991. An introduction to the Neotropical muroid rodent genus Zygodontomys. *Bulletin of the American Museum of Natural History* **210**:1–113.

Voss RS, Marcus LF, Escalante P.P. 1990. Morphological evolution in muroid rodents. I. Conservative patterns of craniometric covariance and their ontogenetic basis in the Neotropical genus *Zygodontomys*. *Evolution* **44:**1568–1587.

Wahlert J.H. 1985. Cranial foramina of rodents. In: Luckett WP, Hartenberger JL, ed. *Evolutionary relationships among rodents: a multidisciplinary analysis*. New York: Plenum Press, 311–332.

Weksler M. 2006. Phylogenetic relationships of oryzomyine rodents (Muroidea: Sigmodontinae): separate and combined analyses of morphological and molecular data. *Bulletin of the American Museum of Natural History* **296**:1–149.

Wilson LAB, Sánchez-Villagra MR. 2010. Diversity trends and their ontogenetic basis: an exploration of allometric disparity in rodents. *Proceedings of The Royal Society B* **277**:1221–1234.

Figures

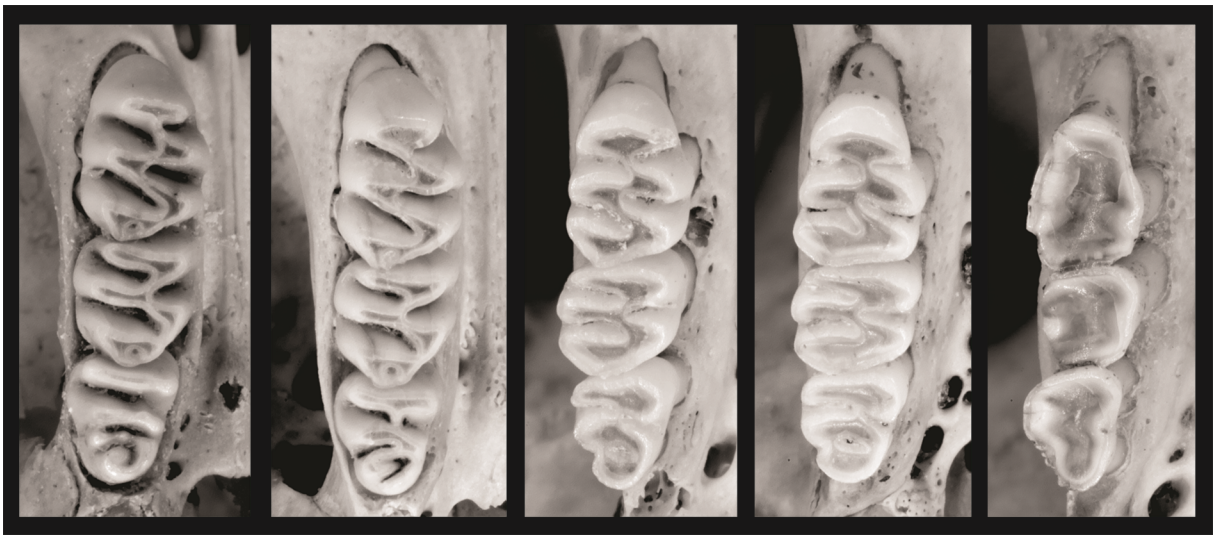


Figure 3. Examples of the upper molars wear and configuration used to identify the five age classes. From the left to right is the age class 1 to age class 5 in the Alagoas_Wild sample.

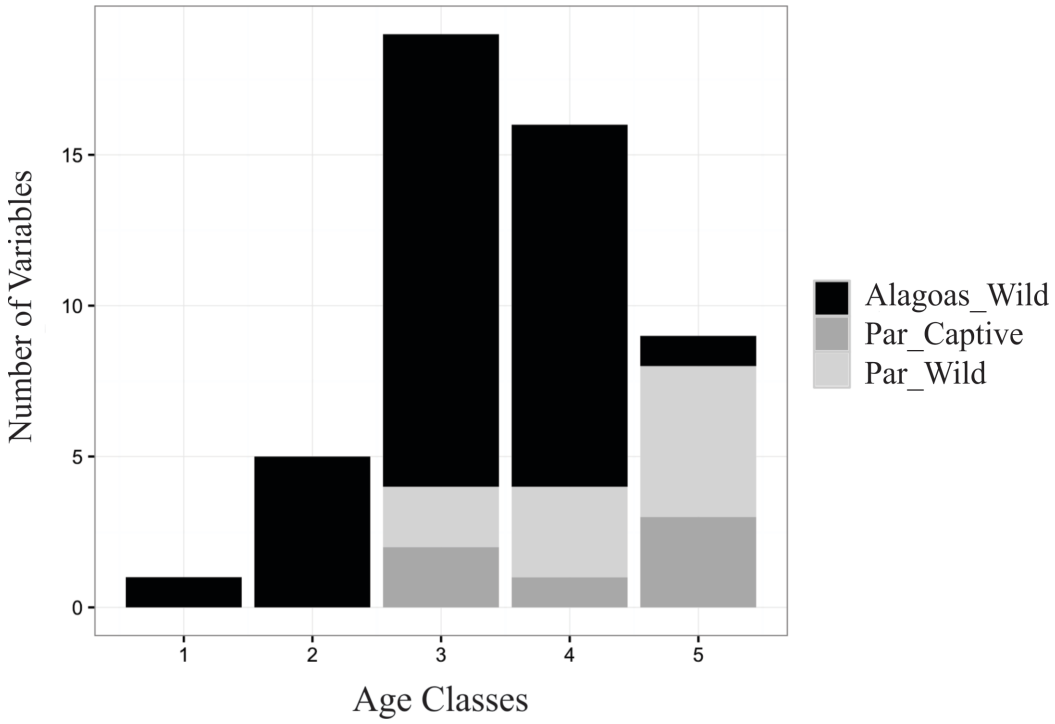


Figure 4. Number of morphometric variables that show significant statistic sex variation in the Student *t* test (y axis) between age classes and across populations. For age classes 1 and 2 there was no specimens from Par_Wild and Par_Captive.

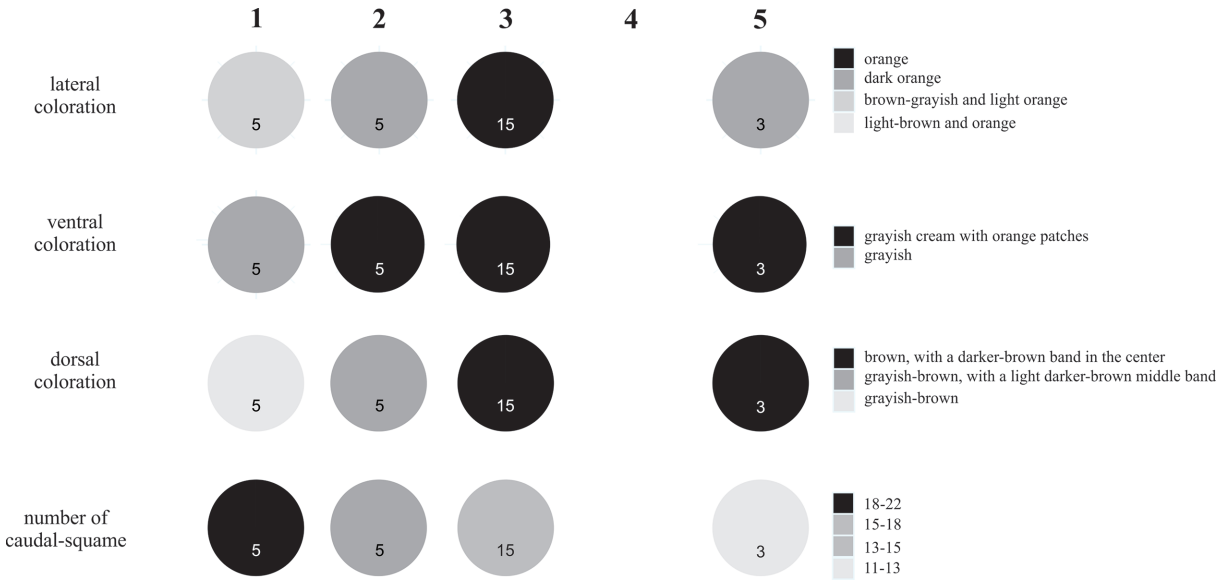


Figure 5. Pie charts describing the morphological variation in skin traits between age classes in the sample Alagoas_Wild. Numbers in the circles represent the numbers of individuals. Note that we do not have data for individuals from age class 4.

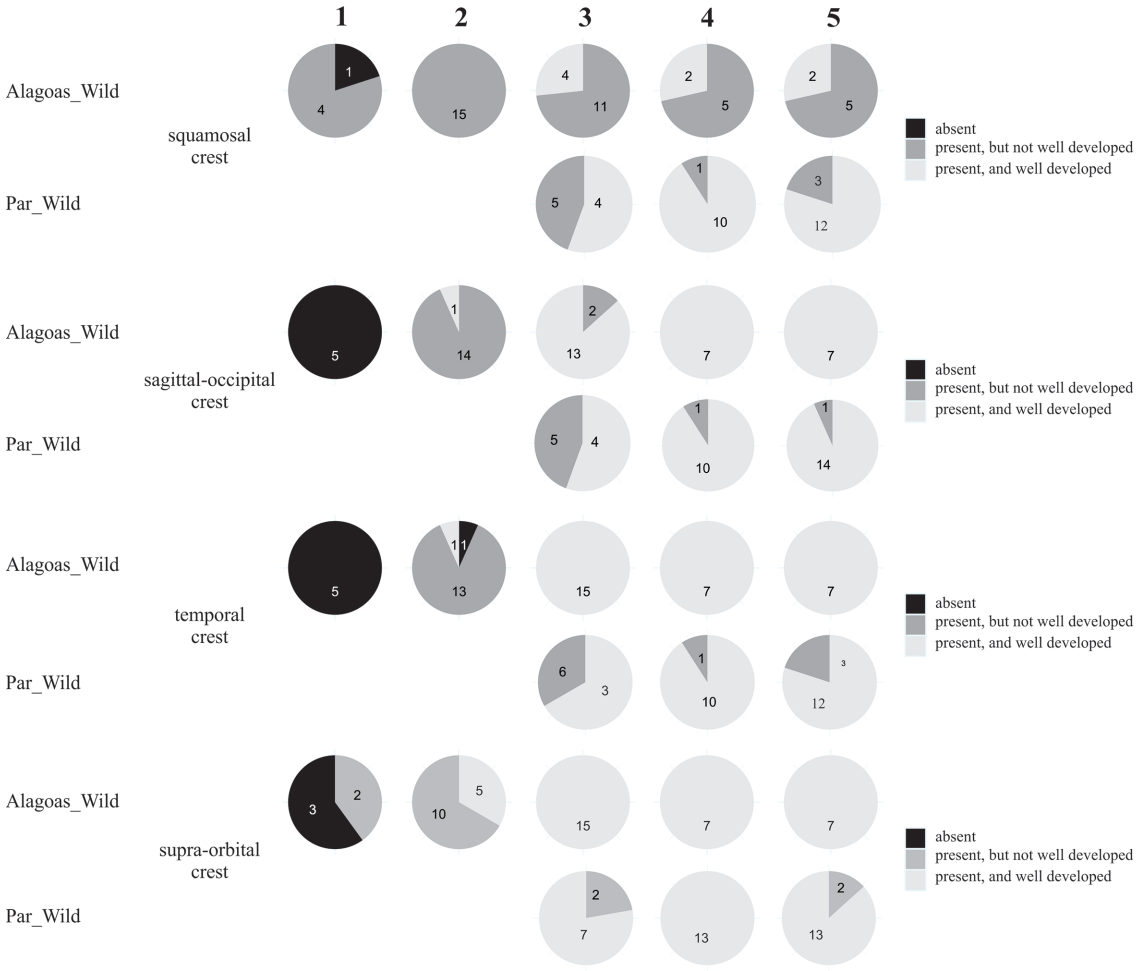


Figure 6. Pie charts describing the morphological variation in cranial traits between age classes in the samples Alagoas_Wild and Par_Wild. Numbers in the circles represent the numbers of individuals. Note that we do not have data for individuals from age classes 1 and 2 in Par_Wild population.

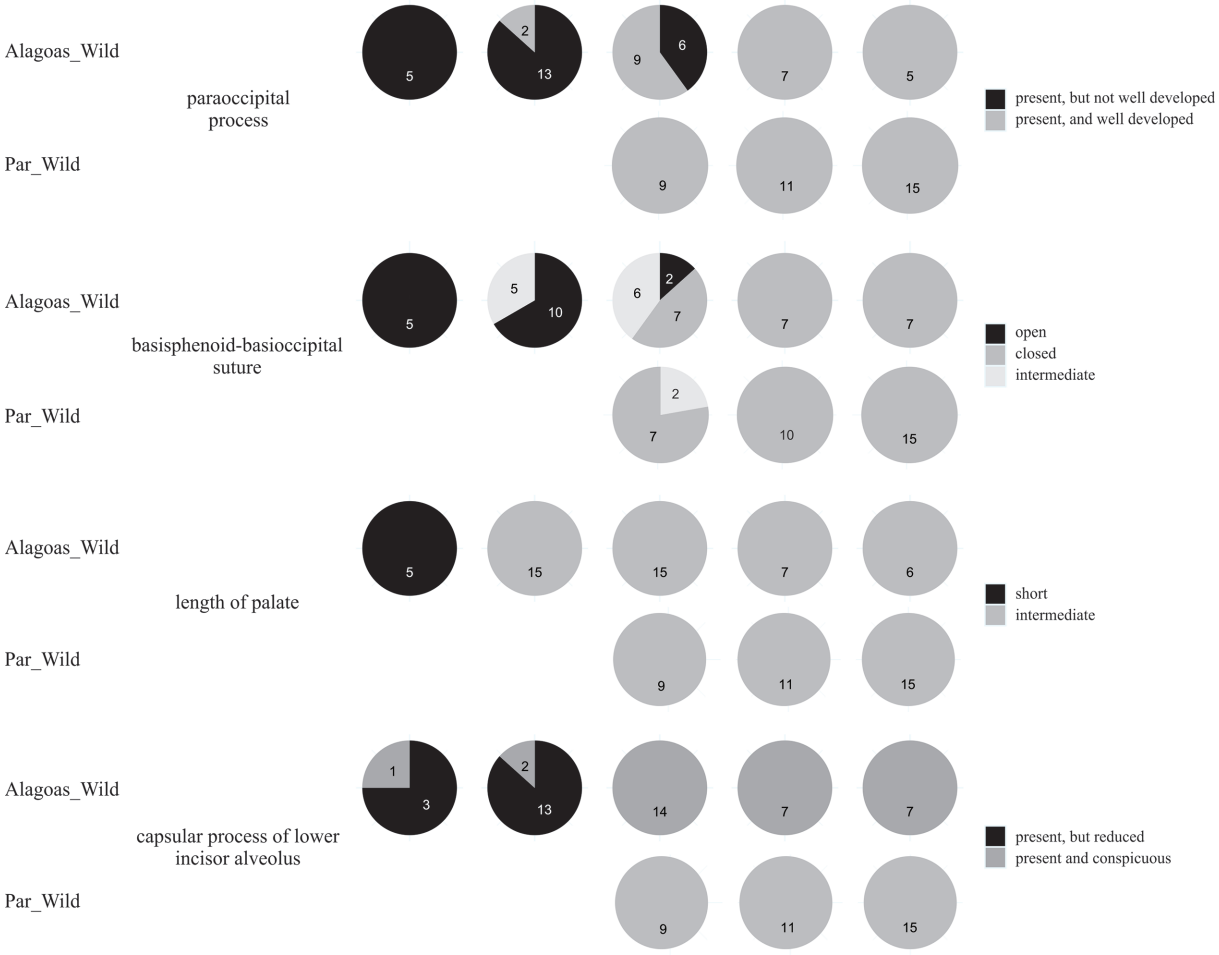


Figure 7. Pie charts describing the morphological variation in cranial traits between age classes in the samples Alagoas_Wild and Par_Wild. Numbers in the circles represent the numbers of individuals. Note that we do not have data for individuals from age classes 1 and 2 in Par_Wild population.

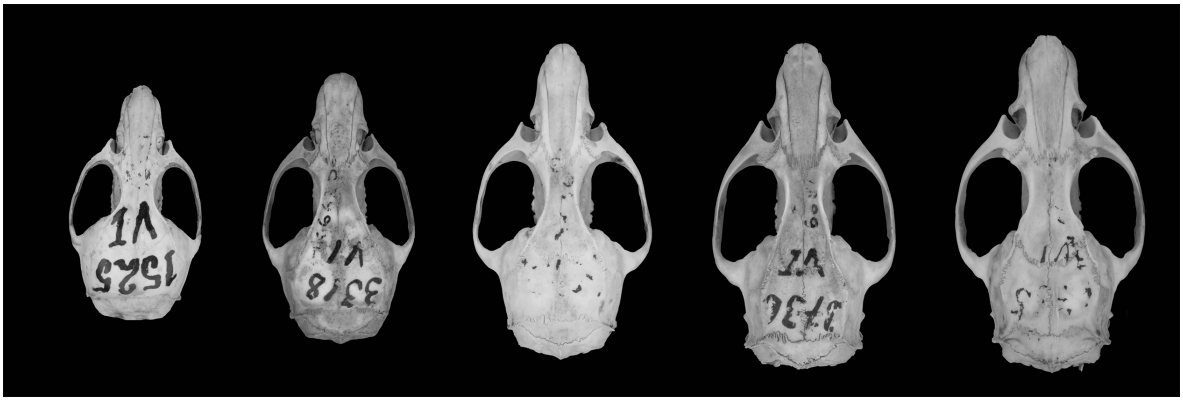


Figure 8. Dorsal view of the skull illustrating the ontogenetic development in the Alagoas_Wild sample; from the left to right are the age classes 1, 2, 3, 4 and 5.

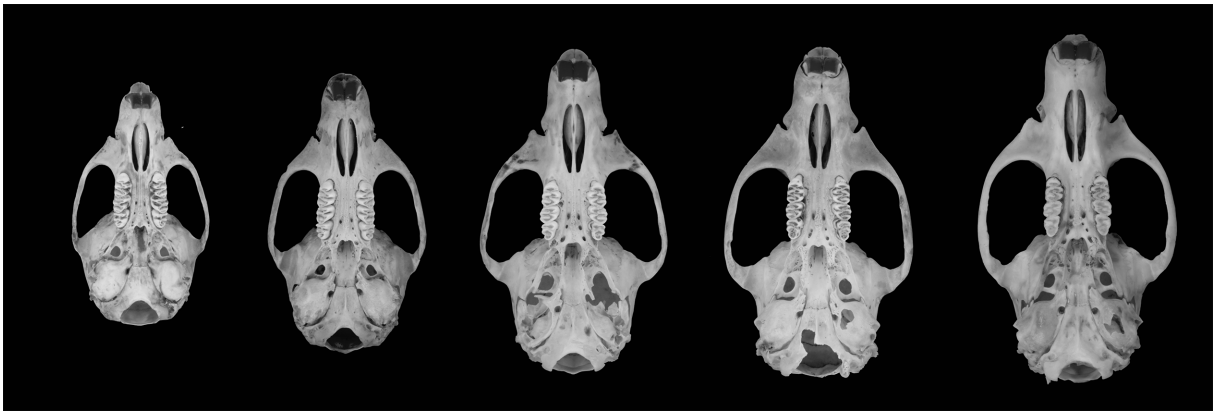


Figure 9. Ventral view of the skull illustrating the ontogenetic development in the Alagoas_Wild sample; from the left to right are the age classes 1, 2, 3, 4 and 5.

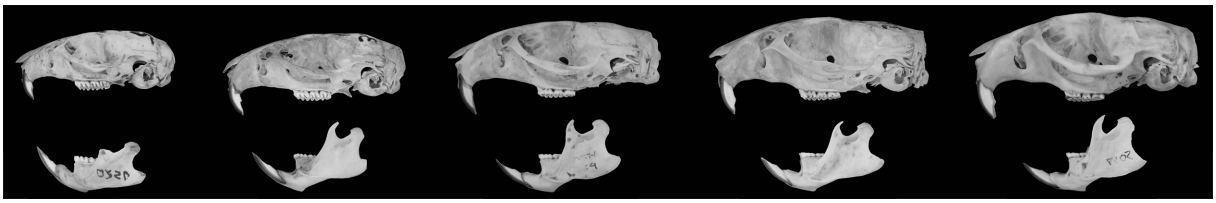


Figure 10. Lateral view of the skull and the mandible illustrating the ontogenetic development in the Alagoas_Wild sample; from the left to right are the age classes 1, 2, 3, 4 and 5.

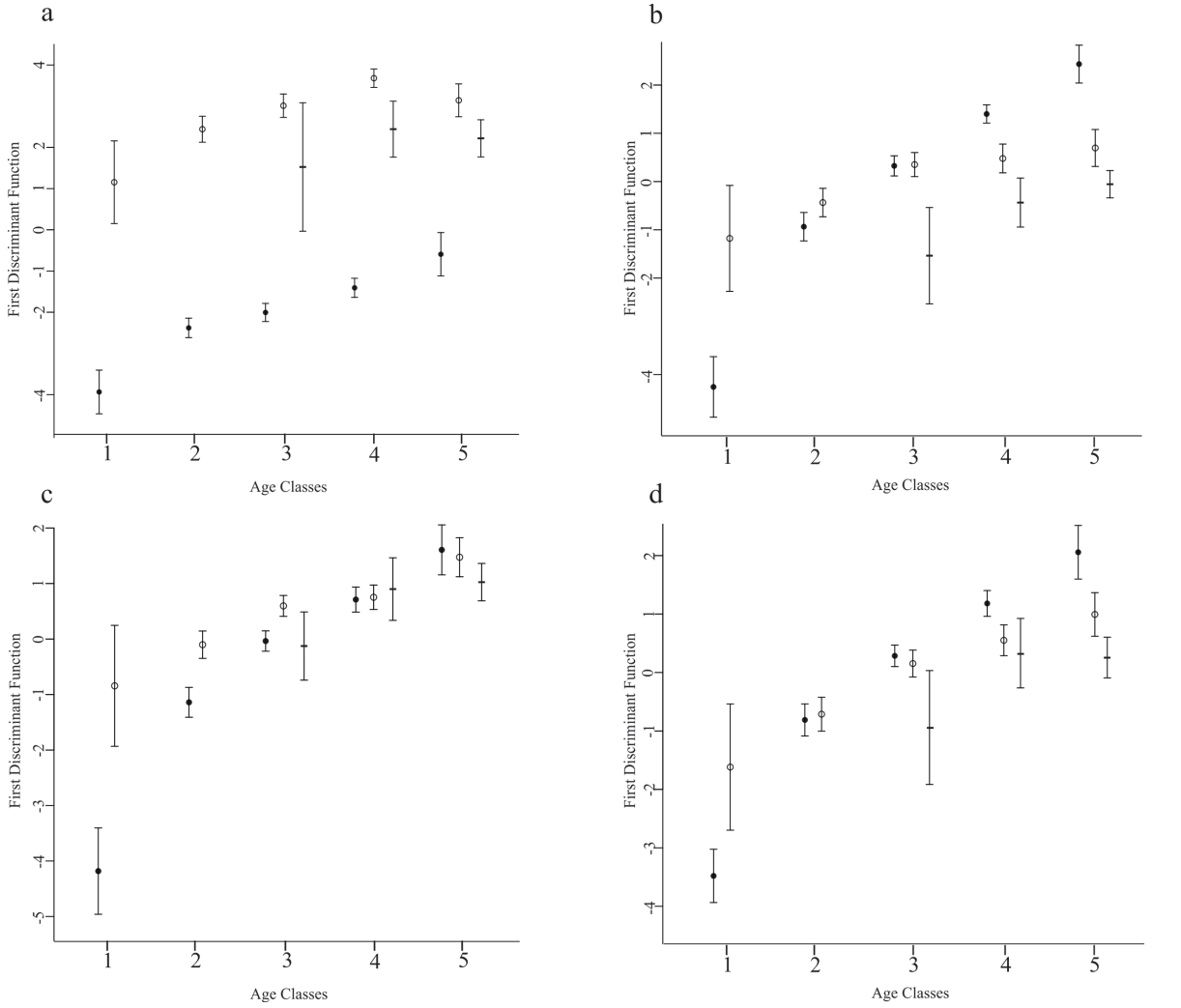


Figure 11. Dice Leraas diagram showing the mean and standard error of the First Canonical Discriminant scores between age classes and across populations. A- shows the Dice Leraas diagram for the CDA with age classes as 13 factors and log10-transformed variables; B- shows the Dice Leraas diagram for the CDA with age classes as 5 factors and log10-transformed variables; C- shows the Dice Leraas diagram for the CDA with age classes as 13 factors and the residuals of a linear regression removing the effect of sex and species variation; D- shows the Dice Leraas diagram for the CDA with age classes as 5 factors and the residuals of a linear regression removing the effect of sex and species variation. Empty circles represents sample from Par_Captive population, full circle represent sample from Alagoas_Wild population, and the black trace represent the scores from Par_Wild sample.

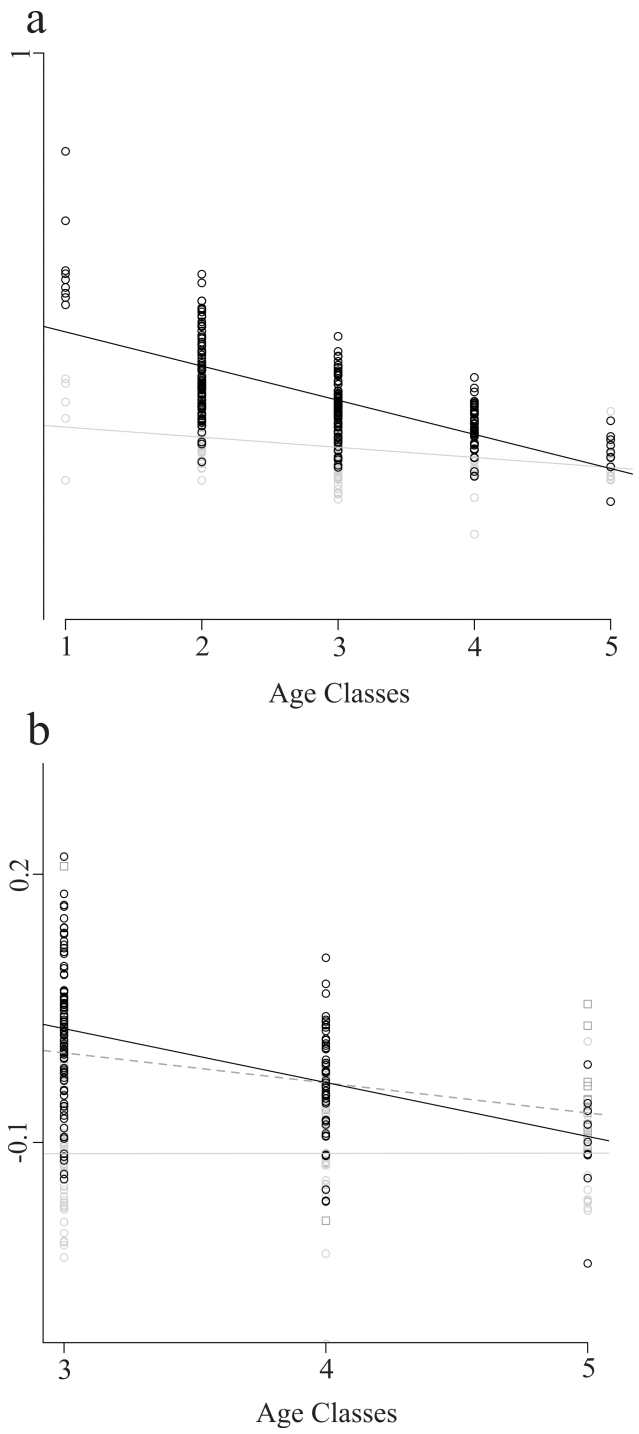


Figure 12. Linear Regression Analysis performed with the scores of the First Principal Component Analysis showing differences in the ontogenetic trajectories across populations. A- shows the Linear Regression between the 5 age classes of the Alagoas_Wild and Par_Captive populations; black line shows the ontogenetic trajectory of the Alagoas_Wild population, light grey line exhibits the trajectory of the Par_Captive sample; black dots represents the PC1 scores of the Alagoas_Wild population, grey dots represents the PC1 scores of the Par_Captive sample. B- shows the Linear Regression between the age classes 3, 4 and 5 of the Alagoas_Wild, Par_Captive and Par_Wild populations; black line shows the ontogenetic trajectory of the Alagoas_Wild population, light grey line exhibits the trajectory of the Par_Captive sample, and dashed grey line shows the ontogenetic trajectory of the Par_Wild sample; black dots represents the PC1 scores of the Alagoas_Wild population, grey dots represents the PC1 scores of the Par_Captive sample, and grey square represents the PC1 scores of the Par_Wild sample.

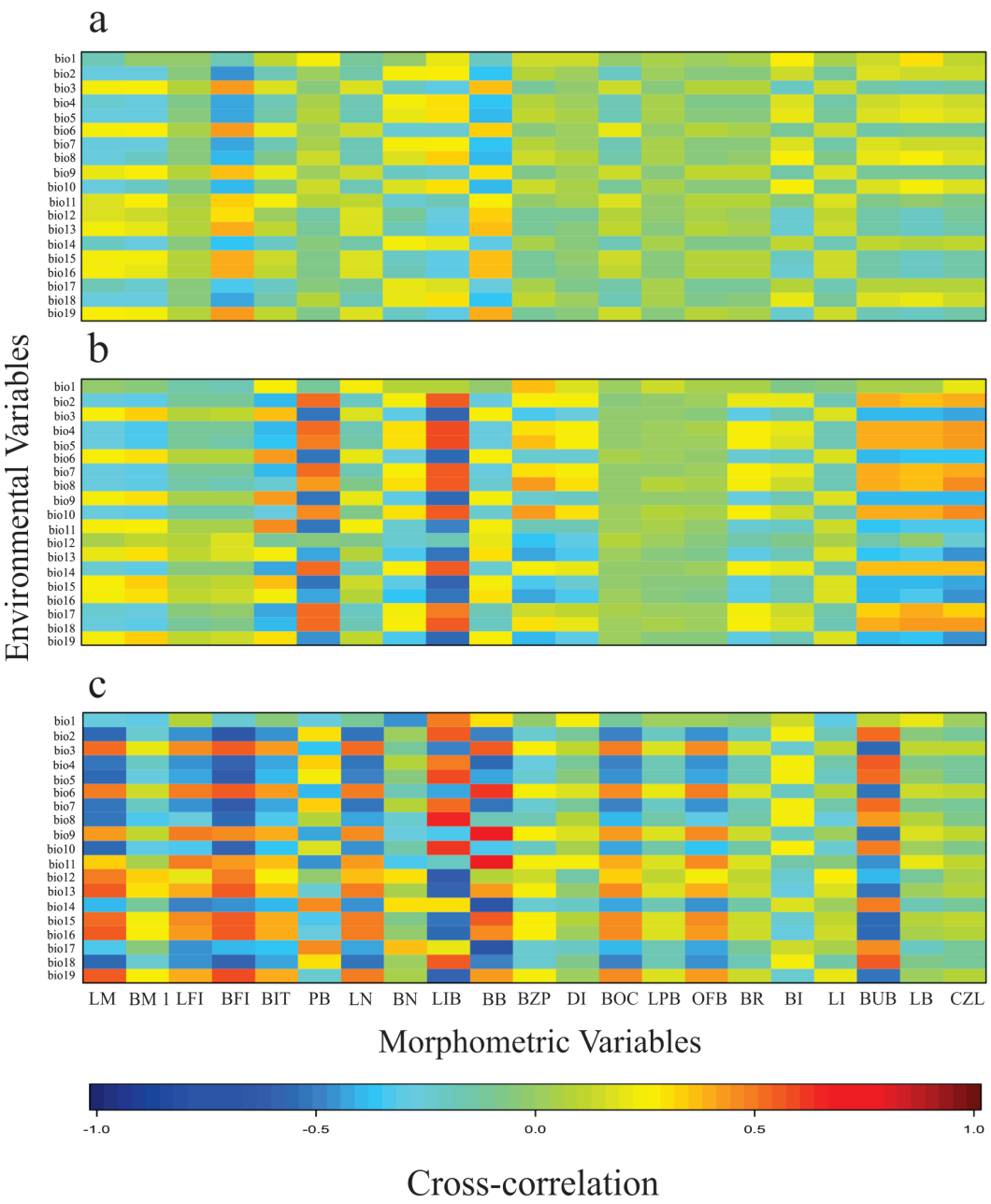


Figure 13. Correlation between morphometric and bioclimatic variables in the wild collected individuals (Alagoas)Wild and Par_Wild populations). A – shows correlation in age class 3. B – shows correlation in age class 4. C – shows correlation in age class 5.

Tables

Table 2. Results of the multivariate analysis of variance (MANOVA) performed with the fixed effects of age, sex, and iteration between them, for the three sampled populations.

		DF	Wilks	Approximate *F* value	Num DF	Den DF	P value
Alagoas_Wild	Age	1	0.42966	13.5903	21	215	**0.000**
	Sex	1	0.73642	3.6645	21	215	**0.000**
	Age*Sex	1	0.89801	1.1628	21	215	0.286
	Residuals	235					
Par_Wild	Age	1	0.14267	2.28927	21	8	0.114
	Sex	1	0.12614	2.63915	21	8	0.080
	Age*Sex	1	0.35263	0.69937	21	8	0.758
	Residuals	28					
Par_Captive	Age	1	0.52064	5.2174	21	119	**0.000**
	Sex	1	0.83041	1.1573	21	119	0.301
	Age*Sex	1	0.84156	1.0668	21	119	0.392
	Residuals	139					

Table 3. Tests of association between morphometric distances with environmental differences (PC1) and/or geographic distance among populations using distance-based redundancy analysis (dbRDA). Results are given for each geographic and environmental variable separately (marginal tests), as conditioned on the effects of geographic distance (conditional tests). F-statistics and p-values are presented.

		Marginal tests		**Conditional tests**	
	Variable	*F*	p-value	*F*	p-value
Total	Dist	-24.4	0.91		
	PC1	0.19	0.55	0	1
Age class3	Dist	-2.3677	0.77		
	PC1	0.57	0.50	0	1
Age class4	Dist	2.2223	0.16		
	PC1	-0.3405	0.83	0	1
Age class5	Dist	0	1		
	PC1	-0.406	0.75	0	1

APPENDIX A. Examined material

Details about the examined material separated by sampled population, with information about locality, sex, age class and catalog number are provided in this appendix, Appendix A.

ALAGOAS: BRAZIL: Alagoas: Anadia: Engenho Ferreiros: **F:** Age Class 3: MN 66994. Fazenda Bananas 1: **M:** Age Class 4: MN 66979, MN 66980. Fazenda Cajueiro: **F:** Age Class 4: MN 66846. Fazenda Vale Verde: **M:** Age Class 4: MN 66981. Sítio Igrejinha: **M:** Age Class 4: MN 66985. Sítio Mocós 1: **F:** Age Class 4: MN 66922, MN 66988. **M:** Age Class 4: MN 66987. Quebrangulo: Engenho Brejo da Folha: **F:** Age Class 2: MN 66694. **M:** Age Class 3: MN 66893. Age Class 4: MN 66894. Engenho Carangueija: **F:** Age Class 4: MN 66982. **M:** Age Class 2: MN 66983. Age Class 3: MN 66920. Age Class 4: MN 66984. Engenho Gitó: **F:** Age Class 3: MN 66919. Engenho Juliana: **M:** Age Class 2: MN 66847. Engenho Pedra Talhada: **F:** Age Class 2: MN 66925. **M:** Age Class 2: MN 66921. Engenho Riachão 1: **M:** Age Class 3: MN 66692. Engenho Riachão 2: **M:** Age Class 3: MN 66886. Fazenda Água Branca A: **M:** Age Class 2: MN 66818. Fazenda Bento de Barros: **M:** Age Class 2: MN 66689. Fazenda Guaribas: **M:** Age Class 4: MN 66986. Fazenda Passagem: **F:** Age Class 4: MN 66680. **M:** Age Class 2: MN 66679, MN 66681. Age Class 3: MN 66677, MN 66678. Age Class 5: MN 66676. Fazenda Periperi: **F:** Age Class 1: MN 66684. Age Class 2: MN 66685. Age Class 3: MN 66686. **M:** Age Class 2: MN 66687, MN 66688. Fazenda Pirauás: **F:** Age Class 1: MN 66696. Age Class 2: MN 66695, MN 66907, MN 66932. Age Class 5: MN 66905. **M:** Age Class 2: MN 66906, MN 66933. Age Class 3: MN 66904, MN 66939, MN 66941. Fazenda Poço da Serra: **M:** Age Class 3: MN 66691. Fazenda Riachão 2: **F:** Age Class 3: MN 66881. Fazenda Santa Cruz 2: **M:** Age Class 2: MN 66690. Fazenda Santa Terezinha 1: **F:** Age Class 3: MN 66990, MN 66991. **M:** Age Class 3: MN 66989. Fazenda Torresópolis: **M:** Age Class 4: MN 66682. Povoado Dois Braços: **F:** Age Class 2: MN 66951. Age Class 4: MN 66964. **M:** Age Class 3: MN 66948. Sítio Barro Preto: **F:** Age Class 4: MN 17454. Sítio Boqueirão: **F:** Age Class 4: MN 66978. Sítio Fernandes: **F:** Age Class 2: MN 67003. **M:** Age Class 2: MN 67002. Sítio Gabirú: **M:** Age Class 2: MN 66992. Sítio Gavião: **F:** Age Class 2: MN 67004. Sítio Goiabeira: **F:** Age Class 2: MN 17458, MN 66830, MN 66877, MN 66879. Age Class 3: MN 66827. **F:** Age Class 2: MN 17459, MN 66826. Age Class 3: MN 66828, MN 66878. Age Class 4: MN 66829. Sítio Gravatá: **F:** Age Class 1: MN 66923. Age Class 2: MN 66849. Age Class 4: MN 66693, MN 66852. **M:** Age Class 2: MN 66798, MN 66854 MN 66855, MN 66856. Age Class 3: MN 66853. Age Class 4: MN 66851. Age Class 5: MN 66797. Sítio Manivas: **F:** Age Class 4: MN 66683. Sítio Olho d'Água 1: **F:** Age Class 1: MN 66809. **M:** Age Class 1: MN 66861. Sítio Riachão: **M:** Age Class 3: MN 66977. Sítio Teixeira: **F:** Age Class 1: MN 66799. Age Class 2: MN 66815. **M:** Age Class 1: MN 66924. Age Class 2: MN 66814. Age Class 5: MN 66800. Viçosa: Engenho Biquara: **M:** Age Class 2: MN 67000. Engenho Cachoeira A: **F:** Age Class 3: MN 66768. **M:** Age Class 4: MN 66767. Engenho Itapicurú: **F:** Age Class 2: MN 66774, MN 66775. **M:** Age Class 2: MN 66850. Engenho Pedra Cascuda: **M:** Age Class 3: MN 66780. Engenho Retiro: **F:** Age Class 4: MN 66773. **M:** Age Class 3: MN 66771, MN 66772. Fazenda Aniceto: **F:** Age Class 5: MN 66712, **M:** Age Class 4: MN 66713. Fazenda Areias: **F:** Age Class 3: MN 66862, MN 66863. **M:** Age Class 2: MN 66808, MN 66810. Age Class 5: MN 66806. Fazenda Bananal: **M:** Age Class 2: MN 66999. Fazenda Cachoeira Grande: **F:** Age Class 1: MN 66746. **M:** Age Class 3: MN 66757, MN 66758. Fazenda Caldeirões: **F:** Age Class 2: MN 66927. Age Class 3: MN 66778, MN 66817. **M:** Age Class 3: MN 66816. Fazenda

Cambuim 1: **M:** Age Class 3: MN 66824. Fazenda Cambuim 2: **F:** Age Class 1: MN 66880. **M:** Age Class 2: MN 66884, MN 66885. Age Class 3: MN 66831, MN 66882, MN 66883. Fazenda Esperança: **M:** Age Class 3: MN 66857. Fazenda Lagoa: **F:** Age Class 3: MN 66844, MN 66918. **M:** Age Class 4: MN 66843. Fazenda Limoeirinho: **M:** Age Class 2: MN 66710. Fazenda Modelo São Luis: **F:** Age Class 2: MN 66802, MN 66803, MN 66804, MN 66805. Age Class 3: MN 66858. Age Class 3: MN 66777. **M:** Age Class 2: MN 66801. Fazenda Paraná: **F:** Age Class 2: MN 66840. Age Class 4: MN 66833. Age Class 5: MN 66832. **M:** Age Class 2: MN 66835, MN 66897, MN 66898. Age Class 3: MN 66834, MN 66836. Age Class 4: MN 66901. Age Class 5: MN 66895, MN 66896. Fazenda Paturi: **M:** Age Class 2: MN 66795. Fazenda Pedra Cascuda: **M:** Age Class 3: MN 66781. Age Class 4: MN 66782. Age Class 5: MN 66892. Fazenda Pedra de Fogo: **F:** Age Class 4: MN 66766. **M:** Age Class 3: MN 66761. Age Class 4: MN 66759, MN 66760, MN 66762, MN 66764, MN 66765. Fazenda Pindobinha: **F:** Age Class 3: MN 66735. **M:** Age Class 2: MN 66736. Age Class 3: MN 66734. Age Class 4: MN 66737. Fazenda Primavera II: **M:** Age Class 2: MN 66997. Age Class 3: MN 66998. Fazenda Riachão 2: **M:** Age Class 2: MN 66752, MN 66848. Age Class 3: MN 66741. Age Class 5: MN 66740. Fazenda São Braz: **M:** Age Class 2: MN 66776. Fazenda São José: **M:** Age Class 2: MN 66733, MN 66743. Fazenda São José B: **F:** Age Class 2: MN 66970. Age Class 3: MN 66842. Age Class 4: MN 66968, MN 66969. **M:** Age Class 2: MN 66917. Fazenda São Manoel: **F:** Age Class 2: MN 66756. Fazenda São Pedro: **F:** Age Class 3: MN 66738. **M:** Age Class 2: MN 66742. Age Class 3: MN 66739. Fazenda Torresópolis: **F:** Age Class 3: MN 66709, MN 66936. Age Class 4: MN 66935. **M:** Age Class 2: MN 66903. Age Class 3: MN 66934, MN 66937. Age Class 4: MN 66902. Sítio Alto do Céu: **F:** Age Class 2: MN 66926. **M:** Age Class 4: MN 66813, MN 66864. Sítio Alves: **F:** Age Class 2: MN 67005. **M:** Age Class 2: MN 66812. Age Class 3: MN 66811. Sítio Bananas: **F:** Age Class 2: MN 66726. Age Class 3: MN 66728, MN 66974, MN 66976. Age Class 4: MN 66973. **M:** Age Class 2: MN 66730, MN 66731, MN 66732. Age Class 3: MN 66727, MN 66845, MN 66975, MN 67006. Age Class 4: MN 66971, MN 66972. Sítio Boa Vista 4: **F:** Age Class 4: MN 66784. **M:** Age Class 2: MN 66839, MN 66899, MN 66900. Age Class 3: MN 66838. Age Class 4: MN 66785, MN 66837. Sítio Caboge: **F:** Age Class 2: MN 66908, MN 66911, MN 66915, MN 66943, MN 66957, MN 67008, MN 67009, MN 67012. Age Class 3: MN 66912, MN 66916, MN 66942, MN 66945, MN 66946, MN 66956, MN 66958, MN 66959, MN 66961, MN 66963, MN 66966, MN 67011. Age Class 4: MN 66909. **M:** Age Class 1: MN 66788, MN 66789. Age Class 2: MN 66913, MN 66914, MN 66947, MN 66965, MN 67013. Age Class 3: MN 66787, MN 66910, MN 66949, MN 66950, MN 66954, MN 66955, MN 66960, MN 67010. Age Class 4: MN 66952, MN 66953. Age Class 5: MN 66962. Sítio Cambuim 1: **M:** Age Class 3: MN 66823. Sítio Goiabeira: **F:** Age Class 3: MN 66931. Sítio Itapicurú: **F:** Age Class 3: MN 66769. Sítio Pedra de Fogo dos Pereiras: **F:** Age Class 3: MN 66745, MN 66753, MN 66754. Age Class 4: MN 66796. **M:** Age Class 4: MN 66755. Sítio Pirauás: **F:** Age Class 2: MN 66967. Sítio Poço Dantas: **F:** Age Class 2: MN 66825, MN 66876, MN 66930. Age Class 3: MN 66865, MN 66870, MN 66871, MN 66928. Age Class 4: MN 66866. **M:** Age Class 2: MN 66822, MN 66872, MN 66873, MN 66929. Age Class 3: MN 66819, MN 66820, MN 66869, MN 66874. Age Class 4: MN 66867. Age Class 5: MN 66821. Sítio Rio Branco: **F:** Age Class 3: MN 66887. Age Class 4: MN 66888. Sítio São José B: **F:** Age Class 4: MN 66783. Sítio Sapucaia: **M:** Age Class 2: MN 66779. Sítio Tangil: **F:** Age Class 3: MN 66717. **M:** Age Class 2: MN 66715. Age Class 4: MN 66716, MN 66718. Sítio Tangil 2,Viçosa: **F:** Age Class 1: MN 66721. Age Class 2: MN 66720. Age Class 3: MN 66719. **M:** Age Class 2: MN 66722, MN 66725. Age Class 3: MN 66723, MN 66724. Age Class 4: MN 66889, MN 66891. Sítio Trevas: **F:** Age Class 4: MN 66938. Sítio Uruçuba: **M:**

Age Class 3: MN 66744. Sítio Velho: **F:** Age Class 4: MN 66996. Sítio Vila Maria Lia: **M:** Age Class 3: MN 66711. Age Class 4: MN 66714.

CAPTIVE: PARAGUAY: Presidente Hayes: Lab Stock from 24 km NW (by air) of Villa Hayes, Estancia La Golondrina. **F:** Age Class 1: UMMZ 166166, UMMZ 166244, UMMZ 166318, UMMZ 166319, UMMZ 166348. Age Class 2: UMMZ 166004, UMMZ 166009, UMMZ 166014, UMMZ 166020, UMMZ 166026, UMMZ 166027, UMMZ 166035, UMMZ 166160, UMMZ 166162, UMMZ 166163, UMMZ 166164, UMMZ 166165, UMMZ 166170, UMMZ 166262, UMMZ 166316, UMMZ 166317, UMMZ 166351, UMMZ 166367, UMMZ 166368, UMMZ 166687. Age Class 3: UMMZ 165996, UMMZ 165998, UMMZ 166001, UMMZ 166015, UMMZ 166021, UMMZ 166022, UMMZ 166023, UMMZ 166028, UMMZ 166036, UMMZ 166154, UMMZ 166159, UMMZ 166161, UMMZ 166241, UMMZ 166343, UMMZ 166354, UMMZ 166356, UMMZ 166357, UMMZ 166364, UMMZ 166366, UMMZ 166370, UMMZ 166446. Age Class 4: UMMZ 165995, UMMZ 165999, UMMZ 166006, UMMZ 166010, UMMZ 166019, UMMZ 166167, UMMZ 166336, UMMZ 166342, UMMZ 166344, UMMZ 166346, UMMZ 166359, UMMZ 166360, UMMZ 166361, UMMZ 166447. Age Class 5: UMMZ 166007, UMMZ 166011, UMMZ 166013, UMMZ 166152, UMMZ 166352, UMMZ 166355, UMMZ 166363. **M:** Age Class 1: UMMZ 166171. Age Class 2: UMMZ 166018, UMMZ 166030, UMMZ 166169, UMMZ 166172, UMMZ 166196, UMMZ 166242, UMMZ 166243, UMMZ 166245, UMMZ 166246, UMMZ 166263, UMMZ 166264, UMMZ 166265, UMMZ 166349, UMMZ 166353. Age Class 3: UMMZ 165994, UMMZ 165997, UMMZ 166000, UMMZ 166002, UMMZ 166003, UMMZ 166012, UMMZ 166016, UMMZ 166017, UMMZ 166024, UMMZ 166025, UMMZ 166029, UMMZ 166031, UMMZ 166153, UMMZ 166195, UMMZ 166337, UMMZ 166339, UMMZ 166347, UMMZ 166350, UMMZ 166358, UMMZ 166362. Age Class 4: UMMZ 166005, UMMZ 166008, UMMZ 166034, UMMZ 166037, UMMZ 166151, UMMZ 166158, UMMZ 166168, UMMZ 166240, UMMZ 166338, UMMZ 166340, UMMZ 166345, UMMZ 166371, UMMZ 176704. Age Class 5: UMMZ 166032, UMMZ 166033, UMMZ 166156, UMMZ 166157, UMMZ 166173, UMMZ 166320, UMMZ 166341, UMMZ 166448. Lab Stock from west bank of Rio Paraguay, 4 km NW of Puerto Fonciere. **F:** Age Class 2: UMMZ 166274, UMMZ 166406, UMMZ 166180. Age Class 3: UMMZ 166273, UMMZ 166399, UMMZ 166402, UMMZ 166404, UMMZ 166429, UMMZ 166181. Age Class 4: UMMZ 166272, UMMZ 166398, UMMZ 166400, UMMZ 166401, UMMZ 166403, UMMZ 166430. **M:** Age Class 2: UMMZ 166312, UMMZ 166275, UMMZ 166431, UMMZ 166434, UMMZ 166701, UMMZ 166177, UMMZ 166179. Age Class 3: UMMZ 166311, UMMZ 166396, UMMZ 166432, UMMZ 166433, UMMZ 166182. Age Class 4: UMMZ 166270, UMMZ 166271, UMMZ 166405. Age Class 5: UMMZ 166397.

WILD: PARAGUAY: Presidente Hayes: 24 km NW (by air) of Villa Hayes, Estancia La Golondrina. **F:** Age Class 3: UMMZ 166325, UMMZ 166238. Age Class 4: UMMZ 166327, UMMZ 166328, UMMZ 165992, UMMZ 166239. Age Class 5: UMMZ 166330, UMMZ 166686, UMMZ 166237. **M:** Age Class 3: UMMZ 166198, UMMZ 166199, UMMZ 166203. Age Class 4: UMMZ 166333, UMMZ 165990, UMMZ 166150, UMMZ 166234, UMMZ 166235. Age Class 5: UMMZ 166326, UMMZ 166331, UMMZ 166332, UMMZ 166334, UMMZ 166335, UMMZ 165991, UMMZ 165993, UMMZ 166149, UMMZ 166236. West bank of Rio Paraguay, 4 km NW of Puerto Fonciere. **F:** Age Class 3: UMMZ 166388, UMMZ 166175. Age

Class 4: UMMZ 166176. Age Class 5: UMMZ 166260, UMMZ 166387. **M:** Age Class 3: UMMZ 166386, UMMZ 166691. Age Class 5: UMMZ 166385, UMMZ 166389, UMMZ 166690.

APPENDIX B. Morphometric Variables

Descriptions of the morphometric variables used in the present study, as well as a figure showing the variables drawn in a rodent skull are provided in this appendix, Appendix B.

Length of the upper molar series (LM): measured from the anterior surface of the first upper molar to the posterior surface of the third upper molar, at the crown of the molars;

Breadth of M1 (BM1): greatest breadth of the first upper molar measure of the base crown, the height of the protocone;

Length of incisive foramen (LIF): the greatest length measured from the anterior edge to posterior edge of incisive foramen;

Breadth of incisive foramen (BIF): the greatest internal breadth, measured on the lateral margins of the incisive foramen;

Breadth of the incisor tips (BIT): breadth measured between the side faces of the two upper incisors attached;

Breadth of palate (PB): measured in the external lateral portion of the maxillary, between the second and third molar;

Length of nasal (LN): measured from the anteriormost end of the nasal to the naso-frontal suture;

Breath of nasal (BN): breadth measured from the widemost part of the lateral of the nasal;

Least interorbital breadth (LIB): shortest distance through the frontals in the orbital fossa;

Breadth of braincase (BB): greater width of the braincase, measured posterior to the squamosal root of the zygomatic arch;

Breadth of zygomatic plate (BZP): the shortest distance between the anterior and posterior margin of the inferior zygomatic root or zygomatic plate;

Depth of incisor (DI): corresponds to the depth of the incisor measured from the anterior to the posterior face, near the base of the incisor;

Breadth of the occipital condyles (BOC): greater breadth through lateral sides of the occipital condyles;

Length of palatal bridge (LPB): measured from the posterior margin of incisive foramen to the anterior margin of mesopterygoid fossa;

Orbital fossa breath (OFB): greatest breadth of the orbital fossa;

Breadth of rostrum (BR): measured across the rostrum, at the posterior extremity of the upper edge of the infraorbital foramen;

Length of interparietal (LI): greatest length (anteroposterior) of interparietal bone;

Breadth of interparietal (BI): greatest breadth of interparietal bone;

Bullar breadth (BUB): measured from the petrosal suture with the basioccipital to dorsal process of ectotympanic;

Lambdoidal breadth (LB): measure between lambdoidal ridges;

Condylo-zygomatic length (CZL): shortest distance between the posteriormost point of occipital condyle and the posteriormost point of the upper edge of the zygomatic notch.

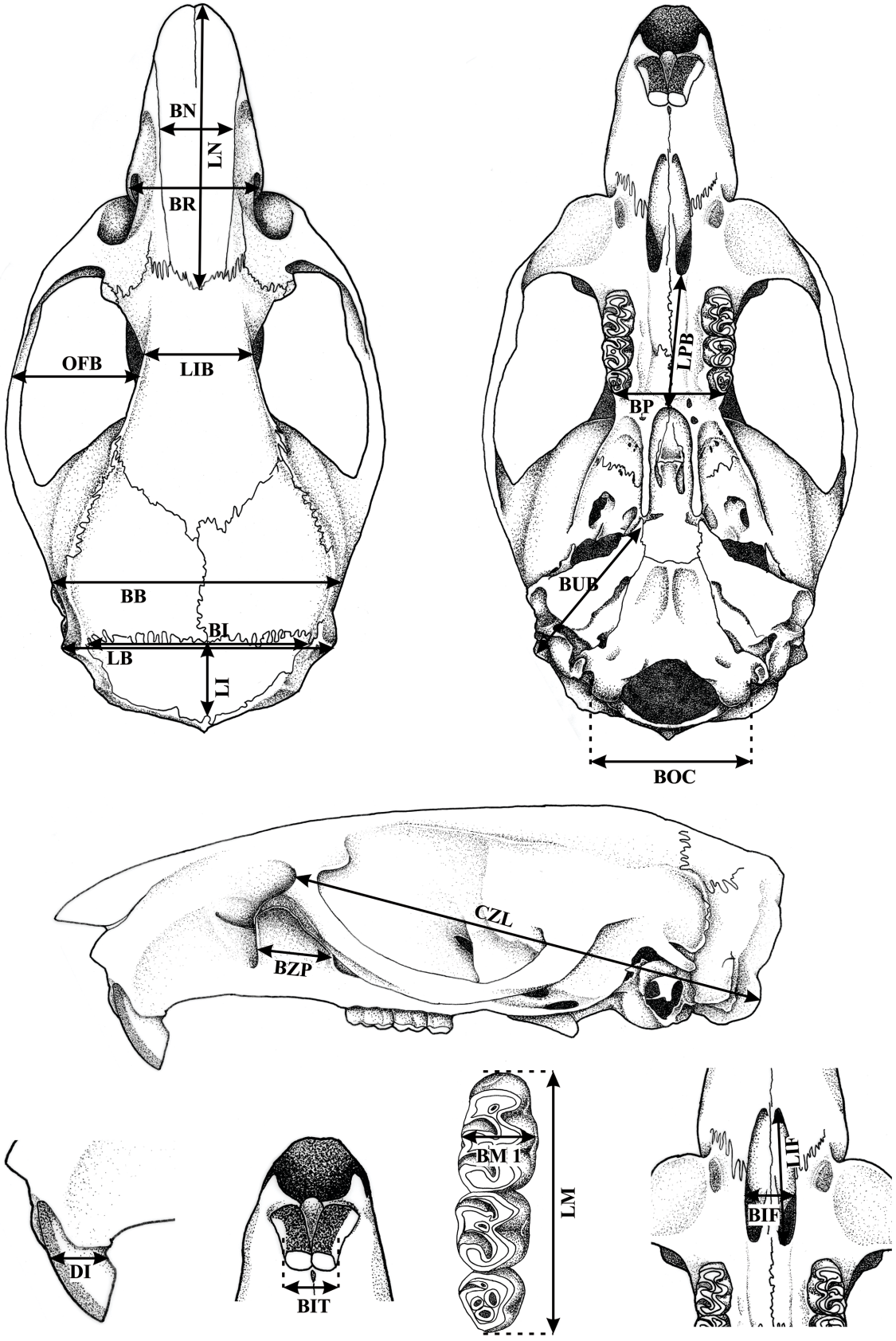


Figure B.1. Representation of the cranial-dental variables extract from the *Holochilus* specimens (illustration of a *Euryoryzomys russatus* skull made by Gustavo S. Libardi).

APPENDIX C. Non-Geographic Variation tests and Description of the samples

Results of the sexual (student *t* test and proportion of difference) and ontogenetic tests (ANOVA and Tukey), as well as the descriptive statistics (sample size, mean, standard deviation, minimum and maximum) are provided in this appendix, Appendix C.

Table C.1. Values of the student *t* test performed on the sample Alagoas_Wild to test sexual dimorphism within each age class. Values in bold represent significant p-values under 0.05. See main text to the description of the variables initials.

Var	Age Classes				
	1	2	3	4	5
LM	-0.1279	**-2.6403**	-1.9809	-1.0948	1.6107
BM1	-1.4377	**-2.0162**	-1.5234	0.3314	0.4562
LFI	-0.0078	-0.9212	**-4.0809**	**-2.1049**	0.6508
BFI	-0.7783	-0.0456	-1.2201	**-2.3766**	-0.2440
PIB	-0.7344	-0.4327	-1.9745	**-2.4208**	-0.9066
PB	0.3861	**-2.0359**	**-2.7161**	0.3489	0.7055
LN	-2.1497	-1.4611	**-3.9523**	**-3.6122**	0.0075
BN	0.5459	-1.4976	**-2.3419**	**-2.0205**	0.194
LIB	-1.3313	-1.0471	-1.9779	-1.6466	0.2607
BCB	-0.6840	-1.949	**-4.4952**	**-2.5291**	1.9289
BZP	0.1707	-0.4384	**-2.8697**	-1.6423	0.1677
DI	-0.7673	0.0262	**-2.3839**	-0.0616	-1.9024
CB	-0.0456	**-2.0342**	**-2.6004**	**-4.8806**	1.1009
LPB	0.8073	-1.5106	**-3.2482**	**-5.0447**	0.3075
OFB	0.3167	-1.2609	**-4.3587**	**-4.2775**	0.8187
BR	-0.1703	-1.0303	**-3.1324**	-1.588	-0.5198
IB	2.2486	-1.3211	**-2.7834**	**-2.0828**	**2.6167**
IL	0.8854	0.5551	-0.5604	0.4459	0.8402
BB	-0.8220	-1.5025	**-5.1386**	-0.0975	0.6400
LB	**-2.6137**	**-2.3013**	**-4.7519**	**-5.2161**	0.0088
CZL	0.2573	-1.5364	**-4.5544**	**-5.1294**	-1.6453

Table C.2. Difference between the mean in percentage calculated for each sexes and variable in each age class in the sample Alagoas_Wild. %M represent the percentage in males and %F represent the percentage in females. See main text to the description of the variables initials.

Var	Age Classes									
	1		2		3		4		5	
	%M	%F	%M	%F	%M	%F	%M	%F	%M	%F
LM	1.10	1.11	1.83	1.87	1.22	1.23	1.10	1.11	1.90	1.86
BM1	6.31	6.74	1.57	1.60	1.22	1.23	0.30	0.30	0.99	0.98
LFI	0.06	0.06	1.47	1.50	4.53	4.74	2.08	2.12	1.27	1.25
BFI	3.47	3.60	0.07	0.07	1.27	1.28	3.35	3.47	1.32	1.34
PIB	4.66	4.89	0.93	0.94	2.97	3.06	3.32	3.43	2.49	2.55
PB	1.18	1.17	1.57	1.59	1.89	1.92	0.43	0.43	2.28	2.23
LN	7.94	8.63	2.63	2.70	4.49	4.70	4.57	4.79	0.02	0.02
BN	2.00	1.96	2.47	2.53	3.50	3.62	3.81	3.96	1.11	1.10
LIB	4.03	4.20	1.14	1.15	2.39	2.44	2.15	2.20	2.00	1.96
BCB	1.41	1.43	1.39	1.40	2.37	2.42	1.69	1.72	2.53	2.47
BZP	1.56	1.54	0.91	0.92	3.98	4.14	1.86	1.90	1.07	1.06
DI	5.49	5.81	0.05	0.05	2.55	2.62	0.27	0.27	4.55	4.77
CB	0.27	0.27	1.80	1.84	1.99	2.03	3.80	3.95	4.17	4.01
LPB	12.37	11.01	2.28	2.33	3.33	3.44	4.62	4.84	1.70	1.67
OFB	2.15	2.11	2.18	2.23	4.92	5.17	4.70	4.94	1.14	1.13
BR	0.77	0.77	1.36	1.38	3.49	3.61	2.13	2.18	1.79	1.82
IB	6.88	6.44	1.48	1.50	3.74	3.89	3.88	4.03	5.42	5.14
IL	9.73	8.87	1.54	1.51	1.34	1.36	1.71	1.68	9.87	8.98
BB	3.07	3.17	1.70	1.72	5.03	5.29	0.12	0.12	1.69	1.66
LB	3.99	4.15	2.13	2.18	3.79	3.94	4.48	4.69	0.02	0.02
CZL	1.63	1.61	1.95	1.99	3.63	3.76	3.52	3.65	2.20	2.25

Table C.3. Values of the student *t* test performed on the sample Par_Captive to test sexual dimorphism within each age class. Values in bold represent significant p-values under 0.05. See main text to the description of the variables initials.

Var	Age Classes			
	2	**3**	**4**	**5**
LM	-1.2372	-1.9781	0.2981	**-2.9651**
BM1	-0.7981	**-2.3264**	0.2933	-1.2283
LFI	0.2235	0.5080	1.4618	-1.0865
BFI	-2.0334	-1.0727	-0.1987	-0.1591
PIB	-1.9032	1.1094	0.1062	-1.2021
PB	-0.6543	-1.7992	0.1580	-0.0255
LN	0.0968	-0.2764	0.6947	-1.9562
BN	-0.4062	**-2.1325**	1.5803	-1.6396
LIB	-0.5925	0.4369	0.9281	0.6482
BCB	-0.5692	-0.3604	1.0799	-0.4719
BZP	0.3994	-0.0209	**2.2546**	-1.7039
DI	0.2083	0.5701	1.6907	-0.6220
CB	0.7594	0.8927	1.0276	-0.9610
LPB	0.2104	-0.3262	0.1223	**-2.5349**
OFB	0.3929	-0.4619	-0.8060	-0.4911
BR	-0.9949	-0.2595	1.3393	-1.688
IB	0.1255	0.0695	-0.2251	-0.0565
IL	0.5529	-0.4770	0.6321	0.5751
BB	0.6970	-0.5322	0.2485	-0.5339
LB	-0.5373	-1.027	0.3338	0.1802
CZL	0.8032	-0.2146	1.7156	**-2.6618**

Table C.4. Difference between the mean in percentage calculated for each sexes and variable in each age class in the sample Par_Captive. %M represent the percentage in males and %F represent the percentage in females. See main text to the description of the variables initials.

Var	Age Classes							
	2		3		4		5	
	%M	%F	%M	%F	%M	%F	%M	%F
LM	1.27	1.29	2.01	2.05	0.40	0.39	6.30	6.72
BM1	0.88	0.89	2.43	2.49	0.43	0.43	3.55	3.69
LFI	0.43	0.43	0.76	0.75	3.12	3.03	3.64	3.77
BFI	5.80	6.16	3.03	3.12	0.84	0.85	0.64	0.65
PIB	4.14	4.32	1.98	1.94	0.28	0.28	3.64	3.78
PB	0.74	0.75	2.77	2.85	0.36	0.35	0.05	0.05
LN	0.19	0.19	0.38	0.38	1.26	1.25	4.78	5.02
BN	1.05	1.07	4.65	4.87	5.81	5.49	6.61	7.08
LIB	1.33	1.35	0.85	0.84	2.01	1.97	2.02	1.98
BCB	0.72	0.73	0.34	0.34	1.18	1.16	0.81	0.82
BZP	0.94	0.93	0.04	0.04	4.80	4.58	4.76	5.00
DI	0.40	0.40	1.17	1.16	3.55	3.43	1.58	1.61
CB	1.08	1.07	1.00	0.99	1.25	1.24	2.17	2.22
LPB	0.37	0.37	0.62	0.62	0.21	0.21	7.20	7.76
OFB	0.79	0.78	0.70	0.71	1.38	1.40	1.39	1.41
BR	2.53	2.60	0.58	0.58	3.90	3.75	5.23	5.52
IB	0.31	0.31	0.13	0.13	0.59	0.60	0.22	0.22
IL	3.42	3.31	2.34	2.39	3.71	3.58	5.33	5.06
BB	1.28	1.26	0.91	0.92	0.48	0.48	1.81	1.85
LB	0.83	0.84	1.26	1.27	0.63	0.62	0.26	0.25
CZL	1.10	1.09	0.20	0.20	1.95	1.91	4.36	4.56

Table C.5. Values of the student *t* test performed on the sample Par_Wild to test sexual dimorphism within each age class. Values in bold represent significant p-values under 0.05. See main text to the description of the variables initials.

Var	Age Classes		
	3	4	5
LM	-0.3387	1.2018	-0.7250
BM1	**-2.4645**	1.4637	-1.0116
LFI	0.4669	-1.4043	**-0.65123**
BFI	0.3687	-2.447	**-2.9986**
PIB	0.2179	0.4489	-2.3489
PB	0.2866	**-2.7168**	-0.3574
LN	0.5143	-0.6968	-1.0474
BN	0.7483	-2.1545	-2.2049
LIB	0.4076	-0.2026	**-2.8731**
BCB	-1.342	-1.5748	-0.5067
BZP	-0.1146	-0.1119	-0.5815
DI	1.2596	-0.2739	**-2.5853**
CB	**-2.5421**	**-2.8146**	-0.8286
LPB	0.4995	-2.2587	-0.0700
OFB	-0.6239	-1.1351	-0.9871
BR	0.3596	-0.3753	-1.5338
IB	1.357	-2.8902	-0.6463
IL	0.5592	**-1.9145**	-0.3177
BB	1.4333	-0.0996	-0.1914
LB	0.2251	-1.2256	-2.2091
CZL	0.3336	-1.5174	**-2.7826**

Table C.6. Difference between the mean in percentage calculated for each sexes and variable in each age class in the sample Par_Wild. %M represent the percentage in males and %F represent the percentage in females. See main text to the description of the variables initials.

Var	Age Classes					
	3		4		5	
	%M	%F	%M	%F	%M	%F
LM	0.68	0.69	1.89	1.86	1.88	1.91
BM1	4.00	4.17	1.83	1.79	3.56	3.69
LFI	2.07	2.03	3.79	3.94	10.41	11.62
BFI	2.66	2.59	8.57	9.38	10.64	11.90
PIB	1.53	1.50	1.33	1.31	9.76	10.82
PB	0.78	0.77	9.11	10.03	1.27	1.29
LN	3.61	3.49	2.46	2.53	2.42	2.49
BN	2.95	2.86	9.75	10.80	11.68	13.23
LIB	2.11	2.06	0.54	0.55	5.94	6.31
BCB	3.88	4.04	3.24	3.35	0.87	0.87
BZP	0.80	0.81	0.47	0.47	2.36	2.42
DI	7.11	6.64	0.82	0.83	7.69	8.33
CB	4.07	4.24	3.67	3.81	1.81	1.84
LPB	2.14	2.10	4.59	4.81	0.26	0.26
OFB	3.59	3.72	3.14	3.24	2.89	2.98
BR	1.85	1.82	2.17	2.22	7.23	7.80
IB	8.30	7.66	8.54	9.34	0.01	0.01
IL	7.00	6.54	20.51	25.81	2.91	3.00
BB	8.04	7.44	0.68	0.68	0.42	0.42
LB	1.17	1.15	3.09	3.19	3.52	3.64
CZL	1.40	1.38	2.45	2.51	4.33	4.53

Table C.7. Results of the Analysis of Variance (ANOVA) performed with sample Alagoas_Wild, followed by the results of the Tukey test. Values in bold and the * represent significant p-values under 0.05. See main text to the description of the variables initials. For the ages 3, 4 and 5, males and females were analyzed separately.

Var	ANOVA (F)	Tukey																		
		1/2	1/3F	1/3M	1/4F	1/4M	1/5	2/3F	2/3M	2/4F	2/4M	2/5	3F/4F	3M/4F	3F/4M	3M/4M	3F/5	3M/5	4F/5	4M/5
LM	5.843 *	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.999	0.575	0.999	0.527	0.985	0.999	0.973	0.635	0.999	0.984	0.999	0.999	0.999
BM1	2.981*	0.076	0.089	**0.005**	**0.011**	**0.016**	0.581	0.999	0.503	0.696	0.836	0.997	0.860	0.999	0.950	0.999	0.994	0.712	0.732	0.839
LFI	51.14 *	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.007**	**0.000**	**0.000**	**0.000**	**0.000**	0.122	0.999	**0.000**	0.745	**0.000**	**0.004**	**0.007**	0.142
BFI	12.82 *	**0.045**	**0.003**	**0.000**	**0.001**	**0.000**	**0.000**	0.517	**0.020**	0.265	**0.000**	**0.000**	0.993	0.999	**0.020**	0.178	**0.000**	**0.001**	**0.004**	0.306
PIB	50.38*	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.009**	0.345	**0.000**	**0.000**	**0.000**	**0.004**	0.448	0.989
PB	15.45*	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.923	**0.000**	**0.000**	**0.000**	**0.009**	**0.007**	0.646	**0.008**	0.812	0.115	0.915	0.999	0.999
LN	42.73*	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.114	**0.000**	**0.000**	**0.000**	**0.000**	0.155	0.999	**0.000**	**0.021**	**0.000**	0.232	0.799	0.999
BN	7.656*	**0.012**	**0.034**	**0.000**	**0.003**	**0.000**	**0.000**	0.999	0.185	0.876	**0.000**	0.089	0.856	0.998	**0.002**	0.459	0.096	0.824	0.680	1.000
LIB	2.681*	0.225	0.347	**0.015**	0.505	**0.039**	0.650	0.999	0.297	0.999	0.685	0.999	0.999	0.505	0.717	0.999	0.999	0.809	1.000	0.913
BCB	24.01*	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.994	**0.000**	**0.024**	**0.000**	**0.000**	0.204	0.989	**0.000**	0.416	**0.001**	0.618	0.386	0.999
BZP	51.18*	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.019**	**0.000**	**0.000**	**0.000**	**0.000**	**0.013**	0.807	**0.000**	0.063	**0.000**	**0.038**	0.506	0.915
DI	37.33*	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.036**	**0.000**	**0.000**	**0.000**	**0.000**	**0.026**	0.577	**0.006**	0.389	**0.000**	**0.000**	**0.046**	**0.032**
CB	18.18*	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.933	**0.000**	0.999	**0.000**	0.154	0.975	**0.038**	**0.010**	0.863	0.574	0.999	0.263	0.996
LPB	35.61*	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.247	**0.000**	**0.000**	**0.000**	**0.000**	0.156	0.999	**0.000**	**0.001**	**0.006**	0.385	0.721	0.992
OFB	50.88*	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.101	**0.000**	**0.000**	**0.000**	**0.000**	0.066	0.999	**0.000**	**0.009**	**0.000**	**0.000**	**0.001**	0.589
BR	13.63*	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.999	**0.004**	0.341	**0.000**	**0.006**	0.596	0.996	**0.004**	0.895	**0.020**	0.739	0.552	0.993
IB	10.39*	**0.012**	**0.007**	**0.000**	**0.002**	**0.000**	**0.000**	0.996	**0.000**	0.792	**0.000**	**0.026**	0.982	0.625	**0.003**	0.924	0.118	0.986	0.506	1.000
IL	6.047*	0.715	**0.034**	**0.009**	**0.049**	0.075	**0.004**	**0.031**	**0.001**	0.111	0.157	**0.008**	0.999	0.999	0.999	0.973	0.710	0.869	0.811	0.564
BB	17.57*	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.935	**0.000**	**0.007**	**0.000**	**0.049**	**0.001**	0.999	**0.000**	0.999	**0.014**	0.999	0.999	0.999
LB	40.64*	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.367	**0.000**	**0.000**	**0.000**	**0.000**	0.537	0.486	**0.000**	**0.023**	**0.000**	**0.024**	**0.000**	0.955
CZL	64.67*	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.023**	**0.000**	**0.000**	**0.000**	**0.000**	**0.023**	0.999	**0.000**	**0.001**	**0.000**	**0.000**	**0.000**	0.288

Table C.8. Results of the Analysis of Variance (ANOVA) performed with sample Par_Captive, followed by the results of the Tuckey test. Values in bold and the * represent significant p-values under 0.05. See main text to the description of the variables initials.

Var	ANOVA (F)	Tuckey									
		1/2	1/3	1/4	1/5	2/3	2/4	2/5	3/4	3/5	4/5
LM	1.547	0.994	0.613	0.946	0.778	0.199	0.956	0.689	0.678	0.999	0.945
BM1	1.022	0.925	0.999	0.994	0.999	0.338	0.929	0.687	0.883	1.000	0.962
LFI	3.171	0.898	0.196	0.593	0.152	0.065	0.798	0.112	0.677	0.979	0.563
BFI	3.33 *	0.999	0.887	0.877	0.151	0.515	0.557	**0.005**	0.999	0.108	0.168
PIB	4.073*	0.707	0.059	0.178	0.116	**0.015**	0.265	0.193	0.894	0.999	0.979
PB	3.61*	0.999	0.564	0.593	0.723	**0.012**	**0.031**	0.221	0.999	0.999	0.999
LN	5.023*	0.685	0.126	**0.028**	**0.030**	0.174	**0.010**	**0.036**	0.690	0.662	0.997
BN	0.758	0.944	0.723	0.842	0.635	0.839	0.980	0.772	0.994	0.991	0.954
LIB	1.676	0.968	0.531	0.647	0.480	0.315	0.592	0.422	0.998	0.995	0.976
BCB	3.521*	**0.047**	**0.003**	**0.024**	**0.026**	0.442	0.983	0.941	0.834	0.992	0.997
BZP	5.324*	0.971	0.102	0.267	0.154	**0.001**	0.052	**0.041**	0.920	0.999	0.956
DI	12.41*	0.327	**0.002**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.818	0.404	0.905
CB	3.875*	0.433	**0.017**	0.125	0.206	**0.026**	0.645	0.865	0.627	0.781	0.999
LPB	10.84*	0.321	**0.011**	**0.000**	**0.000**	**0.042**	**0.000**	**0.000**	0.137	0.114	0.971
OFB	6.402*	**0.029**	**0.000**	**0.001**	**0.000**	0.163	0.366	**0.033**	0.999	0.663	0.589
BR	1.66	0.944	0.395	0.645	0.652	0.212	0.725	0.796	0.950	0.995	0.999
IB	0.91	0.976	0.692	0.812	0.975	0.572	0.855	0.999	0.996	0.874	0.967
IL	0.44	0.999	0.994	0.999	0.997	0.765	0.999	0.940	0.904	0.999	0.977
BB	2.405	0.835	0.367	0.211	0.948	0.492	0.197	0.996	0.950	0.557	0.291
LB	5.116*	0.802	**0.030**	0.128	0.387	**0.001**	0.093	0.709	0.842	0.592	0.967
CZL	15.07*	**0.020**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.860	0.305	0.793

Table C.9. Results of the Analysis of Variance (ANOVA) performed with sample Par_Wild, followed by the results of the Tuckey test. Values in bold and the * represent significant p-values under 0.05. See main text to the description of the variables initials.

Var	ANOVA F	Tuckey		
		3/4	3/5	4/5
LM	0.515	0.981	0.763	0.621
BM1	0.376	0.999	0.772	0.742
LFI	2.769	0.634	0.074	0.376
BFI	3.407 *	**0.045**	0.101	0.765
PIB	1.868	0.300	0.164	0.965
PB	3.297 *	**0.045**	0.597	0.163
LN	1.052	0.522	0.348	0.976
BN	0.888	0.606	0.979	0.407
LIB	1.32	0.956	0.503	0.306
BCB	1.86	0.479	0.146	0.795
BZP	0.596	0.559	0.941	0.673
DI	2.893	0.100	0.092	0.988
CB	1.469	0.405	0.988	0.240
LPB	1.162	0.592	0.294	0.899
OFB	1.707	0.478	0.171	0.843
BR	1.838	0.156	0.667	0.410
IB	0.2	0.952	0.806	0.950
IL	0.254	0.996	0.860	0.800
BB	2.402	0.335	0.088	0.818
LB	0.546	0.571	0.915	0.730
CZL	2.448	0.157	0.118	0.997

Table C.10. Descriptive Statistics of the sample Alagoas_Wild. Sample size, mean ± standard deviation, minimum and maximum value for each variable in each age class. See main text to the description of the variables initials.

	1	2	3	4	5
LM	11 6.53±0.82 5.17-7.24	105 6.97±0.25 6.24-7.60	108 7.01±0.22 6.50-7.62	63 7.03±0.26 6.40-7.66	13 7.04±0.19 6.64-7.35
BM1	11 2.06±0.18 1.56-2.26	105 2.15±0.09 1.93-2.43	108 2.16±0.09 2.00-2.70	63 2.17±0.09 2.00-2.44	13 2.14±0.10 1.92-2.27
LFI	11 5.68±0.69 4.11-6.31	106 7.15±0.57 5.45-8.50	108 7.67±0.50 6.46-8.90	65 7.90±0.38 6.92-9.04	13 8.40±0.40 7.86-9.29
BFI	11 2.42±0.16 2.1-2.65	106 2.58±0.17 2.18-3.14	108 2.65±0.14 2.32-3.08	65 2.71±0.15 2.37-3.12	13 2.87±0.25 2.53-3.37
PIB	11 2.37±0.25 1.79-2.6	103 3.15±0.34 2.44-4.02	103 3.42±0.27 2.51-4.09	64 3.68±0.21 3.29-4.15	12 3.81±0.25 3.39-4.18
PB	11 6.62±0.32 6.14-7.20	106 7.05± 0.28 6.41-7.62	108 7.18±0.26 6.36-7.82	65 7.33±0.29 6.31-8.01	13 7.34±0.35 6.83-7.88
LN	9 11.27±1.01 8.98-12.39	104 14.14±1.25 10.71-16.66	104 15.01±0.97 12.37-18.36	62 15.70±0.82 12.42-17.20	12 16.07±0.40 15.45-16.90
BN	11 3.02±0.20 2.79-3.42	106 3.32±0.27 2.72-4.07	108 3.37±0.27 2.79-4.04	65 3.47±0.30 2.77-4.22	13 3.54±0.28 3.13-3.94
LIB	11 4.39±0.26 3.91-4.67	106 4.61±0.26 4.00-5.47	108 4.66±0.29 4.08-5.45	65 4.65±0.28 3.97-5.26	13 4.59±0.46 3.81-5.29
BCB	11 13.73±0.50 12.88-14.44	106 14.70±0.54 13.71-16.37	108 14.97±0.45 13.74-16.08	65 15.20±0.47 14.15-16.14	13 15.36±0.39 14.66-16.05
BZP	11 3.07±0.47 2.09-3.57	106 4.31±0.44 3.22-5.41	108 4.64±0.35 3.92-5.57	65 4.90±0.25 4.11-5.53	12 5.08±0.31 4.50-5.49
DI	11 1.70±0.18 1.32-1.95	105 2.11±0.19 1.64-2.55	105 2.23±0.12 1.96-2.55	65 2.32±0.14 1.95-2.81	13 2.48±0.16 2.77-2.12
CB	9 7.06±0.42 6.28-7.53	94 7.77±0.33 6.96-8.73	101 7.93±0.31 7.10-8.68	61 7.95±0.29 7.29-8.53	12 8.01±0.33 7.71-8.77
LPB	11 6.76±1.19 3.5-7.77	105 8.18±0.66 6.15-10.25	106 8.59±0.49 7.25-10.05	64 9.02±0.40 8.05-9.84	13 9.07±0.72 7.10-10.05
OFB	11 6.22±0.61 5.11-6.83	106 7.75±0.67 6.12-9.22	108 8.24±0.54 7.05-9.63	65 8.65±0.48 7.69-10.06	13 9.14±0.33 8.45-9.53
BR	11 4.35±0.33 3.75-4.94	106 4.89±0.33 4.16-5.67	108 5.00±0.30 4.21-5.79	65 5.10±0.32 4.46-5.93	13 5.22±0.31 4.67-5.90
IB	11 9.64±0.50 8.64-10.32	106 10.40±0.56 8.71-11.58	108 10.70±0.75 7.77-12.51	65 10.86±0.84 8.63-12.62	13 11.04±0.66 10.16-11.95
IL	11 2.39±0.38 1.74-2.98	106 2.56±0.35 1.73-3.47	108 2.77±0.34 2.04-3.44	65 2.74±0.33 1.91-3.79	13 2.93±0.39 2.21-3.74
BB	11 4.43±0.22 4.14-4.91	102 4.97±0.28 4.20-5.78	103 5.05±0.28 4.40-5.69	58 5.18±0.29 4.66-5.75	12 5.21±0.24 4.85-5.61
LB	11 11.98±0.36 11.27-12.75	106 13.23±0.65 11.88-14.74	108 13.74±0.65 12.20-16.18	65 14.09±0.64 12.89-15.52	13 14.56±0.53 13.76-15.29
CZL	9 20.97±1.60 18.12-22.42	103 25.27±1.63 20.69-28.83	107 26.59±1.25 23.73-29.67	64 27.63±0.99 25.61-30.17	13 28.98±0.71 27.77-30.35

Table C.11. Descriptive Statistics of the sample Par_Captive. Sample size, mean ± standard deviation, minimum and maximum value for each variable in each age class. See main text to the description of the variables initials.

	1	2	3	4	5
LM	6 6.53± 0.20 6.78- 7.14	44 6.82±0.23 6.41-7.24	52 6.94±0.26 6.18-7.51	36 6.87±0.27 6.33-7.51	16 6.93±0.37 6.27-7.78
BM1	6 2.06± 0.04 2.02- 2.16	44 2.03±0.07 1.87-2.2	52 2.07±0.08 1.78-2.2	36 2.05±0.09 1.86-2.24	16 2.07±0.12 1.82-2.34
LFI	6 7.60± 0.55 6.78- 8.21	44 7.78±0.49 6.77-9.08	51 8.04±0.42 7.27-9.05	34 7.90±0.48 6.87-8.7	16 8.12±0.55 6.94-8.99
BFI	6 2.43±0.31 2.11-2.84	44 2.45±0.24 2.06-3.08	51 2.53±0.26 1.73-3.12	34 2.54±0.31 1.76-3.28	16 2.72±0.22 2.4-3.09
PIB	5 3.33± 0.40 2.98- 4.01	44 3.48±0.26 2.92-4.08	46 3.65±0.21 3.09-4.21	33 3.60±0.26 3.21-4.19	16 3.64±0.23 3.25-3.96
PB	6 7.25±0.19 6.94- 7.51	44 7.24± 0.27 6.83-7.83	52 7.50±0.43 6.52-8.44	36 7.50±0.50 6.63-8.82	16 7.48±0.29 6.93-8.02
LN	6 14.78±1.01 13.2-16.01	44 15.27±0.97 13.26-17.49	52 15.67±0.77 13.78-17.26	36 15.91±0.85 14.06-18.03	16 16.00±0.81 14.15-17.14
BN	6 3.48±0.41 3.05-4.26	44 3.59±0.30 2.95-4.28	52 3.66±0.30 2.68-4.21	36 3.63±0.38 2.72-4.41	16 3.70±0.33 3.26-4.53
LIB	6 4.97±0.47 4.39-5.71	44 5.07±0.37 4.28-5.77	52 5.20±0.35 4.62-6.1	36 5.18±0.32 4.49-5.74	16 5.24±0.31 4.74-5.67
BCB	6 14.49±0.38 14.05-14.98	44 15.15±0.62 13.72-16.52	52 15.34±0.52 14.3-16.61	36 15.21±0.49 14.03-16.28	16 15.27±0.50 14.52-16.23
BZP	6 4.89±0.18 4.7-5.16	44 4.98±0.38 4.37-6.08	52 5.27±0.36 4.59-6.27	36 5.20±0.35 4.63-6.32	16 5.28±0.28 4.68-5.72
DI	5 2.16±0.18 1.96-2.46	44 2.30±0.14 2.04-2.67	47 2.43±0.16 2.03-3.03	34 2.47±0.15 2.19-2.75	16 2.51±0.12 2.32-2.83
CB	6 7.58±0.35 7.29-8.15	44 7.83±0.35 7.03-8.57	52 8.03±0.32 7.51-8.88	36 7.93±0.28 7.54-8.72	16 7.92±0.34 7.32-8.45
LPB	6 8.60±0.62 8.24-9.82	44 9.08±0.51 7.89-10.03	52 9.42±0.65 7.8-10.89	36 9.72±0.47 8.73-10.91	16 9.82±0.70 8.79-11.11
OFB	6 7.98±0.68 7.42-9.29	44 8.63±0.56 7.5-9.75	52 8.87±0.48 7.55-9.77	36 8.84±0.45 7.94-10.02	16 9.06±0.46 8.26-10.11
BR	6 5.12±0.49 4.63-6	44 5.26±0.44 4.49-6.56	52 5.46±0.43 4.6-6.62	36 5.39±0.45 4.73-6.37	16 5.41±0.37 4.82-6.25
IB	6 11.04±1.34 8.45-12.16	44 11.26±0.92 8.58-12.88	52 11.53±0.76 9.68-13.08	36 11.46±0.92 9.59-13.84	16 11.29±0.80 9.54-12.75
IL	6 2.68±0.26 2.28-3.07	44 2.64±0.52 1.62-3.78	52 2.76±0.48 1.71-3.73	36 2.67±0.47 1.97-4.36	15 2.75±0.41 2.07-3.66
BB	6 5.67±0.46 5.32-6.6	44 5.83±0.35 5.18-6.45	52 5.96±0.36 4.83-6.67	36 6.01±0.36 5.21-6.61	16 5.80±0.38 5.23-6.31
LB	6 14.31±0.84 13.64-15.93	44 14.66±0.74 13.15-16.33	52 15.24±0.68 13.84-16.64	36 15.08±0.84 13.59-17.01	16 14.94±0.44 14.38-16.13
CZL	6 26.90±1.18 25.57-28.73	44 28.39±1.27 25.58-30.53	52 29.34±0.99 27.06-31.44	36 29.58±1.02 27.54-32.58	16 29.95±1.14 27.05-31.57

Table C.12. Descriptive Statistics of the sample Par_Wild. Sample size, mean ± standard deviation, minimum and maximum value for each variable in each age class. See main text to the description of the variables initials.

	3	**4**	**5**
LM	9 6.81±0.19 6.52-7.09	10 6.83±0.17 6.57-7.1	17 6.74±0.30 6.18-7.32
BM1	9 2.1±0.06 1.98-2.2	10 2.10±0.04 2.02-2.2	17 2.07±0.11 1.84-2.25
LFI	8 7.55±0.44 6.76-8.07	10 7.77±0.35 7.15-8.22	17 8.03±0.57 6.3-8.8
BFI	8 2.38±0.22 2-2.6	10 2.65±0.19 2.43-2.95	17 2.59±0.25 2.18-3.07
PIB	8 3.3±0.30 2.94-3.78	10 3.49±0.15 3.3-3.73	15 3.51±0.30 3.01-3.97
PB	9 7.31±0.29 6.93-7.81	10 7.85±0.57 7.24-8.78	17 7.50±0.48 6.77-8.5
LN	9 14.78±1.55 12.69-17.56	10 15.31±0.84 14.18-16.27	16 15.40±0.80 14.41-17.45
BN	9 3.60±0.20 3.15-3.88	10 3.74±0.33 3.29-4.44	16 3.58±0.35 2.91-4.34
LIB	9 5.10±0.38 4.44-5.69	10 5.14±0.20 4.78-5.5	17 4.96±0.28 4.55-5.58
BCB	9 14.57±0.71 13.9-16.14	10 14.87±0.53 13.48-15.34	17 15.01±0.47 14.37-15.72
BZP	9 2.25±0.18 4.23-6.01	10 5.11±0.31 4.61-5.57	17 4.97±0.33 4.4-5.61
DI	8 2.43±0.16 1.9-2.46	10 2.41±0.10 2.25-2.61	15 2.40±0.16 2.1-2.79
CB	9 7.76±0.24 7.39-8.16	10 7.92±0.22 7.58-8.33	17 7.75±0.28 7.17-8.39
LPB	9 8.81±0.59 7.68-9.85	10 9.03±0.35 8.46-9.67	17 9.11±0.48 8.08-9.94
OFB	9 8.36±0.76 7.66-10.11	10 8.64±0.39 9.15-9.33	17 8.76±0.43 7.71-9.46
BR	9 5.03±0.36 4.43-5.6	10 5.37±0.47 4.77-6.32	17 5.17±0.35 4.33-5.71
IB	9 11.17±0.98 9.64-13.32	10 11.27±0.74 10.31-12.87	17 11.36±0.54 10.41-12.47
IL	9 2.60±0.40 1.88-3.33	10 2.59±0.55 1.74-3.36	17 2.7±0.35 2.08-3.49
BB	9 5.31±0.49 4.26-5.81	10 5.60±0.56 4.76-6.78	17 5.71±0.33 5.03-6.37
LB	9 14.45±1.03 13.27-16.07	10 14.78±0.61 13.78-15.91	17 14.57±0.53 13.71-15.44
CZL	9 27.72±1.83 24.52-31.38	10 28.75±0.79 27.25-29.77	17 28.72±0.94 26.68-30.6

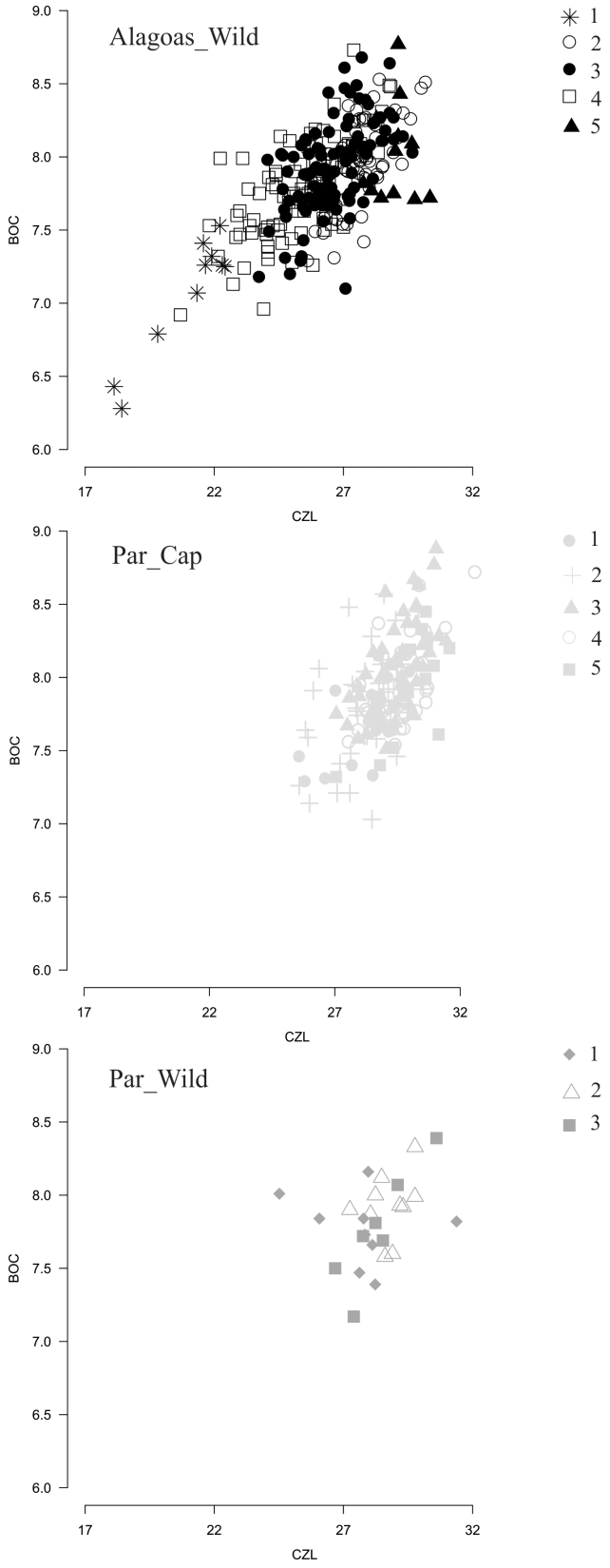


Figure C.1. Scatterplots showing the relation between the raw measurements of the CZL and BOC variables. This variables were the most important in the first CDA performed

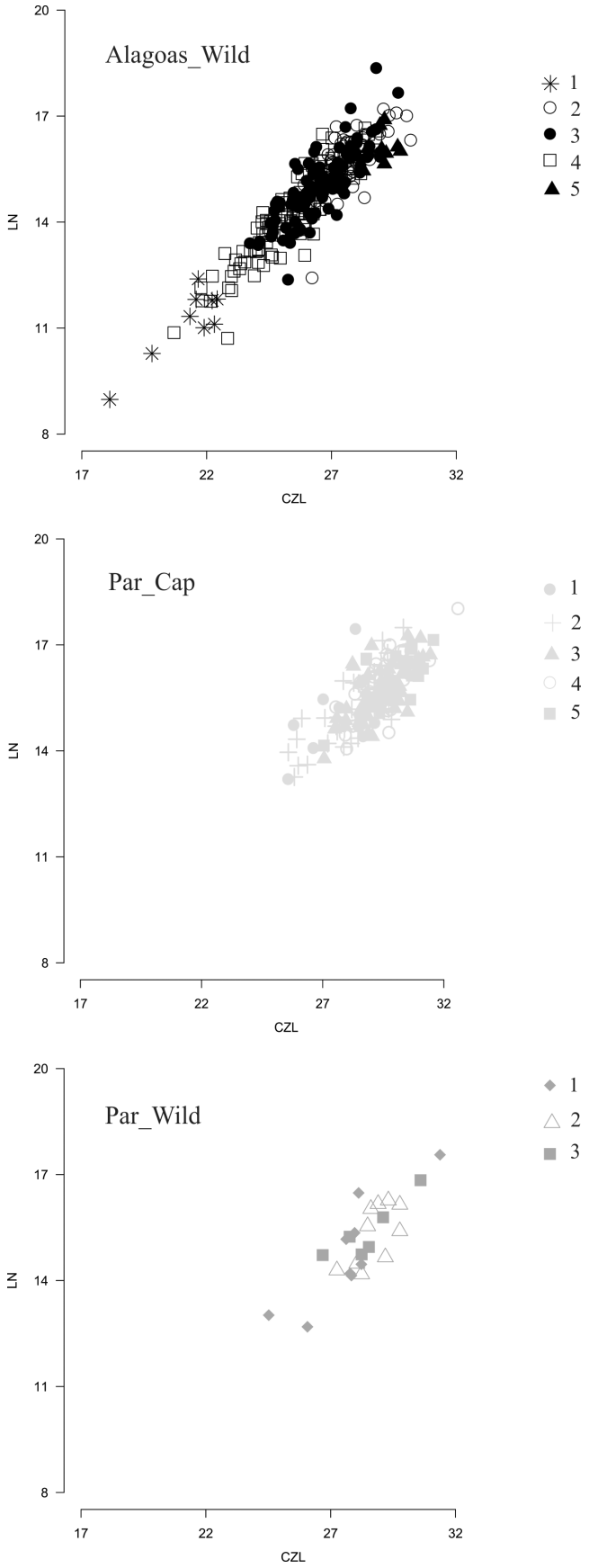


Figure C.2. Scatterplots showing the relation between the raw measurements of the CZL and LN variables. This variables were the most important in the first CDA performed.

APPENDIX D. Environmental and Morphometric Correlations tests

Results of correlations analyses between environmental and morphometric variables separately by age classes 3, 4 and 5 and by population locality, Alagoas_Wild (Fig. D.1) and Par_Wild (Fig D.2), are provided in this appendix, Appendix D.

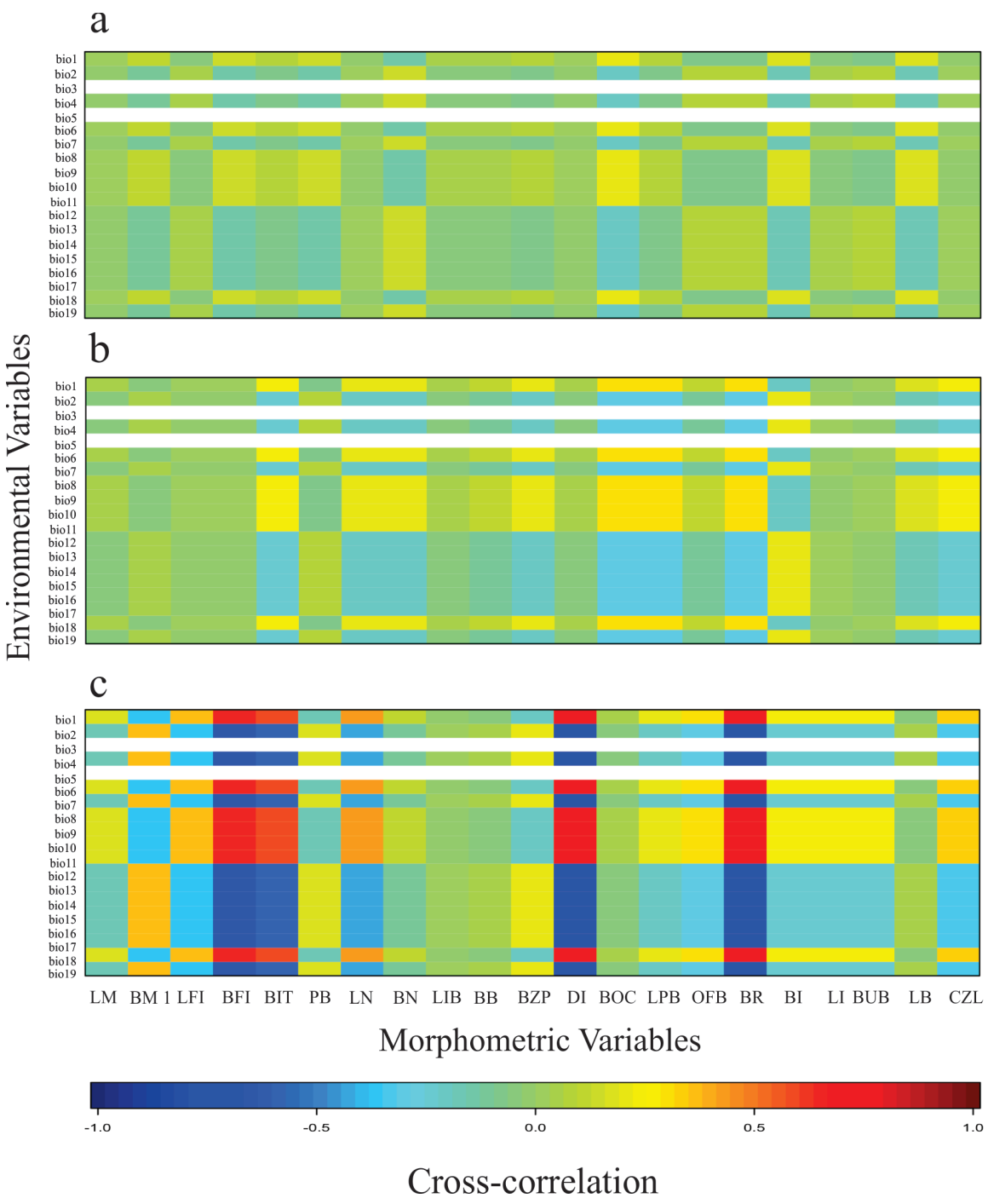


Figure D.1. Correlation between morphometric and bioclimatic variables in the wild collected individuals of Alagoas_Wild population. A – shows correlation in age class 3. B – shows correlation in age class 4. C – shows correlation in age class 5.

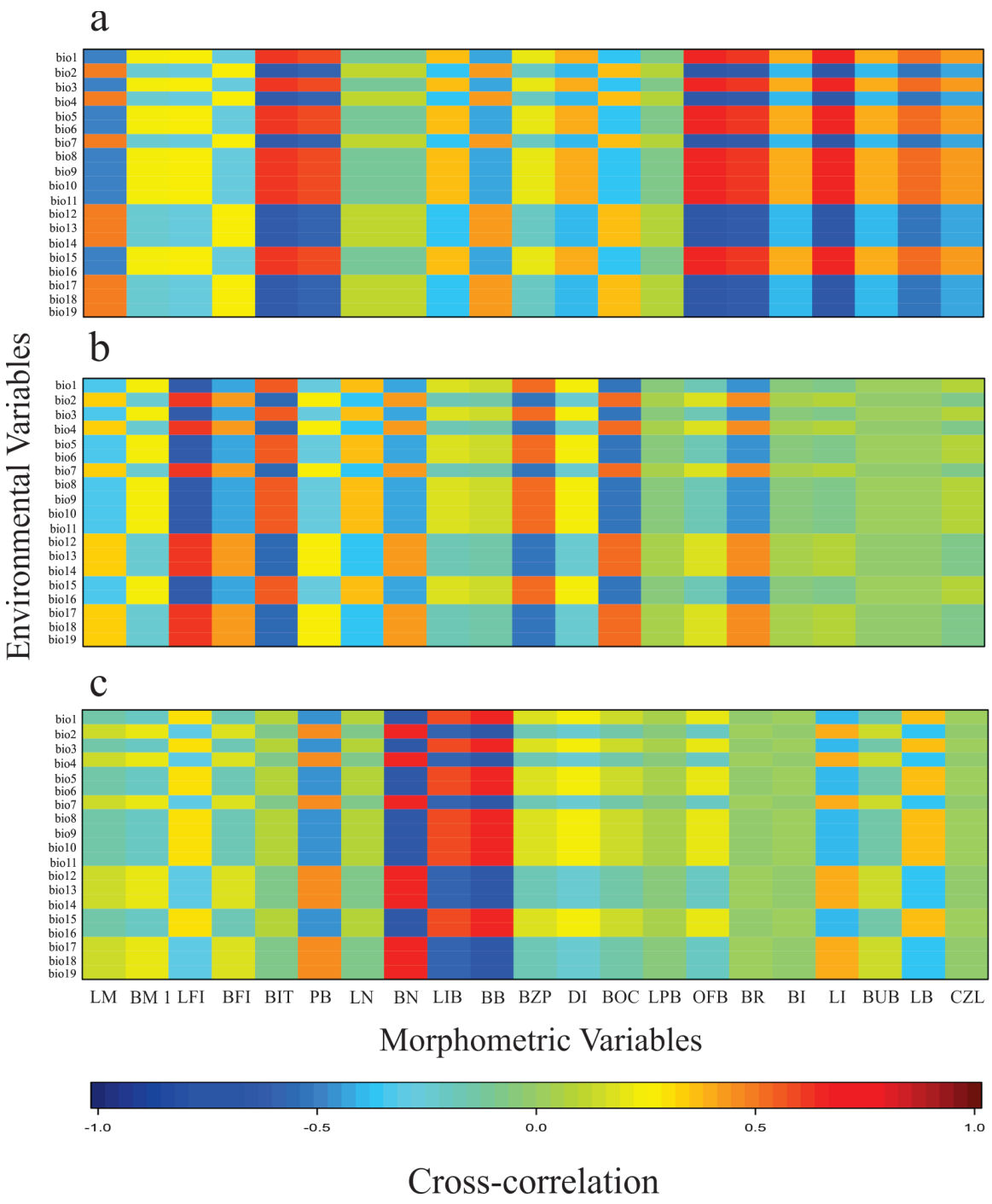


Figure D.2. Correlation between morphometric and bioclimatic variables in the wild collected individuals of Par_Wild population. A – shows correlation in age class 3. B – shows correlation in age class 4. C – shows correlation in age class 5.

3. GENOMIC SIGNATURES OF SPECIES' SIMILAR LIFE HISTORY TRAITS VERSUS BIOME SPECIFIC ASSOCIATIONS: QUANTITATIVE COMPARISON OF THREE SOUTH AMERICAN MARSH RAT SPECIES (RODENTIA: *HOLOCHILUS*)

ABSTRACT

The unevenness and the distribution through different biomes of the wetlands landscapes highly contribute to the rich biodiversity of these marshes. Given the current and past (historical stability) ecological differences in these habitats, species distributed across them may carry biome specific signatures in their patterns of genomic variation. In addition, despite the differences in the biomes where wetlands are inserted, it may be that patterns of genomic variation are predominated by life history of the constituent taxa themselves, rather than being dictated by abiotic factors of their environments. Here, we accessed the degree to which genomic structure varies among taxa and across the landscape given differences in the biomes they inhabit, regions of historical habitat stability, as well as differences in the current climatic variation to evaluate whether structure of genomic variation reflects shared life history traits versus biome specific associations, in three species of the South American marsh rat (*Holochilus*). A significant association between genomic variation and geography was confirmed using a series of complementary approaches, however, the degree to which geography predicts patterns of genomic variation differs among the species and across biomes. Ecological niche models of the LGM, Holocene and the present generated to compare the degree of stability of the distributions of the species across biomes revealed that biomes with larger areas of stability also presented more genomic structure, suggesting that historical dimension impacted population isolation/connectivity. Finally, the biomes the taxa inhabit not only differ geographically and environmentally (based on past climatic conditions), but also show significant association between the environmental space and the genetic variation that is not related with geography. Our results shows that genomic differences are observed across the taxa, despite similar natural histories, suggesting that the biomes play an important role in structuring genomic variation.

Keywords: South America; Climate change; Isolation by distance; Mammal; Next generation sequencing; Phylogeography; RADseq

3.1. Introduction

Combining results of distribution models and patterns of genetic structure to explore congruencies between scenarios of landscapes dynamics through time as well the studies of codistributed species responses to regional changes through comparative phylogeography are current in the literature of South America (Carnaval et al. 2009, Thomé et al. 2010, Leite et al. 2016, Prates et al. 2016, Thomé and Carstens, 2016, among others). However, despite the high genetic structure and biodiversity presented in tropical regions (Moritz et al 2000), comparative phylogeographic approaches in South America usually focus on abiotic features of specific region or biome (mostly in Atlantic Forest) and expand the taxonomic sampling, instead of compare the levels of contribution of abiotic factors (e.g., geography, geological, or climatic history) in different landscapes, controlling the

biotic factor (e.g., taxa with similar ecological or life history traits). Combining a multitaxon comparative framework with consideration of life-history traits can contribute to the understanding of historical and contemporary of abiotic and biotic factors that structure genetic variation (Papadopoulou and Knowles 2016) it is an important aspect to be adressed.

Here we test for a biome specific genomic signature versus a common signature of shared species-specific traits in wetland taxa. We focus specifically on South American marsh rats (genus *Holochilus*, Brandt 1835), and on three species that are distributed across three different wetland landscapes (Fig.14). Marsh rats are specialized semi-aquatic, large bodied rodents with a predominantly herbivorous diet (Hershkovitz 1955; Massoia 1971, 1976; Gonçalves et al. 2015). As *Holochilus* is a grass specialist, our concept of wetland includes these open landscapes, seasonally flooded grasslands. Species of this genus are widely distributed throughout the lowlands and highlands of South America, up to 2000 m (Prado and Percequillo 2013; Gonçalves et al. 2015). For this study, we chose three broadly distributed species: two that occur in two major forest-based patches of wetlands that cover Cis-Andean South America, and one that occurs through the large contiguous wetlands of the Chaco/Pantanal biome (Fig. 14). Of the two species that inhabit the patches of wetlands in forests, *Holochilus sciureus* is distributed along rivers throughout the Amazon biome (i.e., across Venezuela, Guyana, Suriname, Peru, northern Bolivia and Brazil), whereas *Holochilus vulpinus* is distributed throughout the patches of wetlands in the Atlantic Forest and in the Pampa biomes (i.e., occurring mainly in the border of Rio Parana and Rio Uruguay river basins of southern Brazil, Uruguay and Argentina, and eastern Paraguay). Unlike the forest-associated wetland taxa, *Holochilus chacarius* is distributed throughout the contiguous large wetland areas of the Chaco/Pantanal biomes (that extend into Brazil, Paraguay and northeastern Argentina).

The vastness of South American wetlands across the landscapes no doubt contributes to the rich biodiversity of these marshes. However, the wetlands are not uniform. These systems oscillate between aquatic and terrestrial phases, and the fluctuations can vary in depth, length, timing, and frequency (Hoorn et al. 2010; Junk 2014), producing high differentiation in these environments across biomes, which is reflected in the large number of names associated with these wet landscapes, such as the *igapó*, *varzea* and *campinaranas* in the Amazon basin, the savannas in Pantanal, and *veredas* in Atlantic forest and some open areas (Junk 2014; Wittmann 2012).

Given the differences in the wetland habitats across South America, the species distributed across them may carry biome specific signatures in their patterns of genomic variation. For example, differences in the continuity of habitat (e.g., the Chaco/Pantanal versus the Amazon and Atlantic Forest) could impact the relative connectivity of populations, producing more or less geographic genomic structure in the constituent taxa of the respective biomes. Likewise, levels of gene flow may vary as a function of differences in the relative ephemerality of habitat (e.g., contrasts between the narrow grass patches along the Rio Amazonas, which become dislodged and float downstream, versus the stable seasonally flooded grasslands of along Rio Paraná and the Pantanal and Chaco biomes). In

addition to the ecological differences of these contemporary biomes, patterns of genomic variation may be impacted by differences in the historical stability of the wetlands across biomes (Bush et al. 2007; Quattrocchio et al. 2008; Whitney et al. 2011). This includes not only shifts in the distribution of wetlands in the more distant past (e.g., large-scale shifts in habitats associated with Miocene-Pliocene), but also the recent and local shifts in habitat associated with climatic changes during the Pleistocene. For example, studies have shown how the Atlantic Forest biome was reduced in size during the last glacial maximum (LGM), introducing differences in the historic stability of the forest over the years (Carnaval and Moritz 2008). The impact of this stability on patterns of genetic variation has been documented in terrestrial taxa (Thomé et al. 2010; Leite et al. 2016), with shifts in the biome itself, in some cases, affecting multiple species in a similar manner (Carnaval et al. 2009).

Despite the differences in the biomes where wetlands are distributed, it may be that patterns of genomic variation are predominated by life history of the constituent taxa themselves, rather than being dictated by abiotic factors of their environments (see Papadopoulou and Knowles 2015, 2016). Wetland taxa share a number of characters related with specialization for aquatic life and herbivorous diet. These include not only morphological traits (e.g., among mammals a fusiform body shape, webbed feet, valvular ears, robust bones, elongate hind feet, loss of molars folds; Hershkovitz 1962; Stein 1988; Domínguez-Bello and Robinson 1991; Samuels and Van Valkenburgh 2008), but also behavioural traits (e.g., how mammals seek shelter and forage, as well as where they mate and disperse; Eisenberg 1981). Consequently, patterns of genomic variation may reflect these shared life history traits, rather than differences in the wetlands themselves. For example, the general degree of structuring of genomic variation across space may be more similar than expected based on differences in the current or historical environments across regions because of general constraints imposed by an aquatic life style (e.g., dispersal is constrained to water) and/or shared properties of the taxa themselves (Avise 2000).

We approach the question of biome specific genomic signature versus a common signature of shared species-specific traits in the wetland taxa by collecting genomic data using next generation sequencing (specifically, ddRADSeq; Peterson et al. 2012), and an analytical technique that provides not only a quantitative assessment of the similarity of genomic structure across species, but also provides spatially explicit information on how patterns of genomic variation are structured across the landscape (e.g., Procrustes and SpaceMix analysis of genomic and geographic structure). We couple this latter information with characterizations of the historical distributional stability of each species inferred from ecological niche models (ENMs) for the present and the past (Alvarado-Serrano and Knowles 2014), as well as variation in the contemporary environment based on bioclimatic variables (see Lanier et al. 2015). That is, we consider the degree to which genomic structure varies among taxa and across the landscape given differences in the biomes they inhabit (i.e., forest versus open wetlands), regions of historical habitat stability within each biome, as well as differences in the climatic variation within biomes.

3.2. Methods

3.2.1. DNA extraction, amplification, and sequencing

Genomic DNA data was generated for 68 individuals (26 individuals of *H. chacarius*; 24 individuals of *H. sciureus*; and 18 individuals of *H. vulpinus*) from 21 populations (9 populations of *H. chacarius*; 6 populations of *H. sciureus*; and 6 populations of *H. vulpinus*) that cover the entire range of the species distribution (Fig. 15; Appendix E Table E.1). Tissues were requested from museums (see Appendix F for full details of holdings). DNA was extracted from liver, muscle or skin of each individual using the Qiagen DNeasy Blood and Tissue Kit.

Four reduced representation libraries, were built using the Double Digest RADseq method (Peterson et al. 2012) and are briefly described here (see Appendix E for details). DNA was double digested with the restriction enzymes EcoR1 and MseI, unique barcodes (10 bp) and Illumina adapter sequences were ligated to the digested fragments, and the fragments were amplified by PCR. The four libraries were sequenced in four lanes of Illumina *HiSeq2000* to generate 150 base pair, single-end reads, which produced 270 million raw reads (see Appendix E for details).

3.2.2. Processing of Illumina Data

Raw sequence reads for each species were processed separately in Stacks v.1.35 (Catchen et al. 2013). The reads were demultiplexed and filtered using the *process radtags* script. Only reads with Phred score >10, unambiguous barcodes, and individuals with more than 500,000 reads were retained (additional details about data processing are given in Appendix E). A *de novo* assembly of filtered reads with a minimum coverage depth (m = 6) were used to identify putative loci with the USTACKS program. A catalog of consensus loci was constructed with the CSTACKS program from the USTACKS output files for each species, if the distance between individuals for a given locus ≤ 2. Alleles were called for loci using the SSTACKS program, and output files run in the POPULATIONS program to identify loci across populations, with the SNP data exported in Variant Call Format (vcf).

Two datasets with different levels of missing data were generated and used depending on the requirements of the respective analyses. Specifically, (1) a dataset that was not filtered based on missing data across individuals (beyond the initial filtering conducted prior to running the Populations program discussed above) was used to calculate genetic diversity statistics, and (2) a dataset with one randomly chosen SNP, and a maximum of 20% missing data per SNP was used in all other analyses (hereafter referred to as the unlinked-SNP dataset). The former dataset consists of 359,728 SNPs and 26 individuals of *H. chacarius*, 357,050 SNPs and 24 individuals of *H. sciureus* and 160,879 SNPs and 18 individuals of *H. vulpinus* (see Appendix E Table E.2). For the latter dataset of unlinked-SNPs, loci were selected using the toolset PLINK v.1.07 (Purcell et al. 2007), and consists of of 32,210 SNPs

and 24 individuals of *H. chacarius*, 17,513 SNPs and 21 individuals of *H. sciureus* and 8,035 SNPs and 15 individuals of *H. vulpinus* (see Appendix E Table E.2).

3.2.3. Population genetic summary statistics

Population genetic diversity statistics were calculated for all loci present in at least two populations and include: nucleotide diversity (π), major allele frequency, observed heterozygosity (H_O), and Wright's inbreeding coefficient (F_{IS}) at each locus. Population-level summaries of genetic diversity for each species were also characterized (i.e., average π, H_O and F_{IS}), and a one-way ANOVA was used to test for significant differences in genetic diversity between species (analyses conducted in R; R Core Team 2014).

3.2.4. Environmental niche models

Environmental niche models (ENMs) were generated from bioclimatic variables for the present, Holocene (6 kya) and the LGM (21 kya) with Maxent 3.3.3k (Phillips et al. 2006). Georeferenced occurrence data used in the modelling were representative of the entire ranges of the three species; vetted data were obtained by direct examination of specimens of museum specimens (listed in Appendix F) or from taxon-specific bibliographic sources (*i.e.* revisionary papers; e.g., Hershkovitz 1955; Pardiñas and Teta 2011; Pardiñas et al. 2013; D'Elia et al. 2015).

We used 19 bioclimatic informative variables to model current distributions (Bioclim), Holocene (MPI-ESM-P; Bioclim) and LGM distributions (MPI-ESM-P; Bioclim; Hijmans et al., 2005) for each species, and the occurrence data were rarefied using SDMToolBox at a resolution of 10 km to reduce spatial autocorrelation. To avoid overfitting of the distributional models, and considering the semi-aquatic habit of *Holochilus*, the geographical extent of the environmental layers used as the mask in the ENMs corresponded to the respective river basins where each of the species are broadly distributed with an additional 100km buffer.

We conducted a Principal Component Analysis to choose a subset of the most important variables using the *prcomp* function in R software (R Core Team 2014), and selected the variables with < 0.7 of correlation. In order to achieve a balance between goodness-of-fit and model complexity we used the *ENMevaluate* function from ENMeval package (Muscarella et al. 2014). Specifically, we test models over combinations of regularization parameters from 0.5 to 3 in intervals of 0.5 and combinations of features parameters choose according the sample size and Maxent instructions (Auto, Linear, Quadratic, Hinge, Linear + Quadratic and Linear + Quadratic + Hinge). The regularization and features parameters were chosen based on the best combination of values between Akaike Information Criterion (AIC; Warren and Seifert 2011) and AUC. For each model we used only the significant

Maxent's threshold to obtain binary predictions or each data set, climatically stable areas were inferred based on the intersection of the binary predictions under current and past climates to infer climatically stable areas. Each model parameter class was replicated 10 times for cross-validation. Additional details about the ENMs, including projected distributions for each time period, are given in Appendix G.

3.2.5. Association between genomic variation and geographic and/or environmental differences

The association between genomic variation and geography, as well as the environment, was assessed using a series of complementary approaches. Specifically, (i) Mantel tests and distance-based redundancy analysis (dbRDA), (ii) Procrustes analysis, and (iii) analysis of allele frequency covariance structure, each of which are described below. After conditioning on geography (i.e., controlling for the effects of geography), the association between genomic variation and environmental differences among populations were tested. This complement of approaches was selected because the tests differ in their respective assumptions, and hence potential to capture different aspects of how geography might structure genomic variation.

Mantel tests and dbRDA: The correlation between pairwise F_{ST}-values and Euclidean geographic distances among populations, as well as associations between genetic distance and environmental resistance, was examined using a Mantel and partial Mantel tests from the R package VEGAN (Oksanen et al. 2013), and a dbRDA (Legendre and Anderson 1999). For the partial Mantel test an environmental resistance matrix was generated based on raster maps obtained with the ENMs analysis among populations using CIRCUITSCAPE v4.0 (Shah and McRae 2008). dbRDA analysis was used to test for the relationship between pairwise genetic distances and corresponding climatic variables (represented by the First Principal Component of the same 19 climate layers used in the ENMs), and removing the effect of geographic distance separating populations (see He et al. 2013 for further details).

Because of consistency of the results between the Mantel tests with dbRDA, only the results form the Mantel are presented and discussed. For the Mantel tests, a sequential population drop out procedure, in which the test was repeated excluding one population at time, was also conducted to confirm that the results were robust.

Procrustes analyses: The association between genomic variation and geography, while retaining information about the relative positions of populations across space (i.e., both latitudinally and longitudinally) under an isolation by distance was also assessed through a Procrustes transformation approach, which performs a rotation of matrices to maximize the similarity between genes and geography (Wang et al. 2010). Specifically, the sum of squared Euclidean distances

between a PCA map of genetic variation and geographic coordinates was maximized using a rotation of the matrices, while preserving relative pairwise distances among points within the genetic and geographical maps. This analysis was performed using the *procrustes* and *protest* functions in Vegan package (Oksanen et al. 2013). The strength of the association is measured by t_0, which ranges from 0 to1, with low t_0-values indicating deviations between individuals in a genomic PC map from predictions based on geography. The significance of the statistic association between the first two PCs of genomic variation and the geographical coordinates of the populations for each species was evaluated based on 10,000 permutations, where geographical locations were randomly permuted across the different sample localities.

To evaluate the robustness of the results, the Procrustes analyses were repeated excluding one population at a time. We then compare the PCA coordinates from each reduced data sets with the original geographical data sets to identify the maximum extent to which the association between genes and geography (t'') might change as different populations were excluded (Wang et al. 2012, Knowles et al. 2016). In addition, we computed a similarity score (t') between PCA coordinates for the SNP data before and after removing population to assess whether any populations had a disproportionate effect on the relationship between genes and geography (Wang et al. 2012, Knowles et al. 2016).

Covariance structure of allele frequencies: In addition, spatial patterns of genomic variation and potential admixture were examined based on an allele frequency covariance matrix using the program SpaceMix (Bradburd et al. 2016). Following the developer's recommendations, 10 "fast" independent chains were run for 5-10^6 MCMC iterations, without conditioning populations on their locations and with no admixture, followed by a "long" run of for 10^8 iterations, with parameters sampled every 10^2 iterations, in which population locations were initiated at the origin (i.e., inferred from the "fast" runs), and all other parameters were drawn randomly from their priors at the start of each chain.

3.2.6. Bioclimatic characterization of biomes

Variation in the current environment across the landscape and biomes was quantified with a Principal Component Analysis of the 19 bioclimatic variables (from Bioclim). This information is presented graphically for PC1, PC2 and PC3 that were generated using the *rasterPCA* function from dismo package in R, with input variables rescaled from 0 to 1 (i.e., so that the PCs are not sensitive to differences in the units, and hence numerical range, used in measurements of environmental values, such as temperature degrees versus millimetres of precipitation). The *ggRGB* function from RStoolbox package in R was used to construct the bioclimatic map, with the red layer corresponding PC1, green layer corresponding PC2 and the blue layer corresponding to PC3.

To compare the environmental dispersion of sequenced populations relative to the total environmental space for each species (see Lanier et al. 2015), we extracted the environmental data (the first two Principal Component of the PCA performed with the 19 bioclimatic variables) from the occurrence points in each species used in the Maxent computations, and compared the overall environmental dispersion of these points to those of the sequenced populations. In addition, to examine whether the environment might make a significant contribution to patterns of genomic variation after controlling for the effects of geography. Specifically, environmental variables extracted from the principal components data for the sampled localities were used in a partial Mantel test and the dbRDA (e.g., He et al. 2013). We also used the residuals of the Procrustes analysis between genes and geography based on all populations as input in another Procrustes analysis pair with the extracted principal components data of environmental variables for the sampled localities of sequenced individuals.

3.3. Results

3.3.1. Population genetic summary statistics

Values of population genetic summary statistics were broadly overlapping among taxa (Appendix H Table H.1 Fig. H.1), except for a statistically significant difference in observed heterozygosity, H_{obs} (ANOVA: p=0.0001). There was also a relatively similar range of genetic diversity among populations in each of the species, with the exception of the South_Mamore population (S_MAM) in *H. sciureus* and Estancia Yacare population (YAC) in *H. chacarius*, which both show substantially higher inbreeding coefficients, F_{is} (Appendix H Table H.1 Fig. H.1).

3.3.2. Geographic structure of genomic variation

The correlation between pairwise F_{ST}-values and the Euclidean geographic distances among populations accessed by Mantel, partial Mantel tests and dbRDA was not significant in either species (Appendix H Table H.2 Table H.3). The association between environment and genomic variation not conditioned by geography was significant only in *H. chacarius*; re-analysis with sequential removal of individual populations confirmed the Mantel results are robust (i.e., the lack of significance does not a general sensitivity to the inclusion of any single population). This contrasts with a significant association between geography and genomic variation assessed by the Procrustes analysis. When retaining the geographic information about latitudinal and longitudinal position of populations, each species showed a statistically significant association between geography and genomic variation, although there are differences in the strength of the association across species, and hence across

biomes (Figs. 16 and 17). Specifically, species from the patchily distributed wetlands scattered throughout the forested biomes (*H. sciureus* and *H. vulpinus*) showed the strongest association between genes and geography, with a t_0 = 0.83 and 0.73, respectively, compared with a t_0 = 0.50 in the open wetland taxa (*H. chacarius*). Note that this association, and the differences among taxa in the strength of this association, is generally robust based on repeating the test with individual population sequentially excluded (Fig. 17). For example, the shift from the t_0-value (i.e., t_0-*t")* was typically less than 0.05 in all species, whereas the difference among species in t_0-values is as great as 0.23. Only the exclusion of the PAR population in *H. vulpinus* and the BOL in *H. chacarius* (see Fig. 17) results in a fairly large increase in the strength of the association between genomic variation and geography (i.e., an increase over t_0 of 0.07 and 0.10, respectively; Fig. 17, Appendix H Table H.4). The differences among species in the strength of the association between genomic variation and geography also does not appear to be attributable to the undue influence of any single population on the PCA of genomic variation (Appendix H Table H.4). Specifically, the association between the multivariate genomic space with all populations relative to one in which individual populations were sequentially removed is very high (i.e., *t'*-values >0.97 in all but three cases, and the lowest association was still 0.76; Appendix H Table H.4).

Regarding to the geographic distribution of genomic variation within each species, several patterns are reinforced from the results of different analyses. For example, when PC1 from a genomic PCA is projected onto the geographic distribution of sampled individuals (Fig. 18), it highlights how the similarity in multidimensional genomic space among individuals is very much related to where an individual occurs geographically, confirming the results from the Procrustes analyses (e.g., Figs. 16 and 17). Likewise, this geographic structure of genomic variation is also evident from analyses of the spatial covariance of alleles (i.e., the SpaceMix results; Appendix H Fig. H.2). Although there is a general correspondence with the PCA, suggesting that PCAs (and therefore the Procrustes analysis) are insensitive to biases that can be introduced from the sampling distribution of individuals (see Novembre and Stephens 2008), for some individuals of *H. sciureus* the position in multivariate genomic space differs somewhat. Specifically, we note that the S_MAM and N_MAM populations in *H. sciureus* are in an intermediate position relative to other populations in the geogenetic map of Spacemix (Appendix H Fig. H.2), unlike their position along the axes of PC1 and PC2 in the PCA (Fig. 18). This suggests that the sampling gap in *H. sciureus* in which central Amazonian samples were not available (see Fig. 15) may introduce some bias in the position of these individuals in PC space (see Bradburd et al. 2016). However, it is also noteworthy that inclusion of these populations do not strongly affect the results Procrustes analysis in this species, which is based on PCA (i.e., their inclusion neither produces large changes in the strength of the association between genes and geography, nor does it result in a significant distortion of the genetic PC map, based on results from the sequential population drop-out procedure; Fig. 17 and Appendix H Fig. H.2).

3.3.3. ENMs and bioclimatic characterization of biomes

The environmental characteristics of the biomes the three species inhabit clearly differ (Fig. 19). In particular, the environmental space for the area inhabited by *H. vulpinus* differs the most from the other regions. There is some resemblance between the environmental space occupied by *H. vulpinus* and the north-western region of *H. sciureus*, based on a PCA of the bioclimatic variables across the entire area encompassed by all three species' ranges. However, a large region (the central green area) with different environmental characteristics separates these two areas (see the bluish areas are separated by a large green area; Fig. 19).

The ENMs suggest stable habitat in most of the distributional range of *H. vulpinus*, based on comparisons of projected habitat suitability during the LGM and the present for the species (Fig. 16). In contrast, the past and current projections for *H. chacarius* reveal a small area of predicted stability for this species. The ENMs suggests the southwestern region of *H. sciureus'* distribution was a large area of environmental stability area, suggesting that this region could have acted like a refuge for this species given the instability across the rest of the species' range (Fig. 16).

The sampled populations of each species are environmentally representative of the species environmental range (Fig. 20). That is, the populations that were characterized genomically span the entire dispersion of environmental values in the PCA based on the larger geographic sampling of individuals used in the ENMs (Fig. 20).

3.3.4. Association of genomic variation and environmental conditions

In *H. vulpinus*, when we consider the deviations of individuals in genomic space from expectations based on where the individual was collected (i.e., the length of the lines connecting a sampled population to a sequenced individual; Fig. 16), the deviations primarily vary along a latitudinal axis. Moreover, the individuals tend to occupy the central area of the species distribution, with the exception of one population in the southeast (i.e., BAR; Fig. 15) whose individuals are deviated to the south, and the position of individuals in the PC genomic space show a strong correspondence with areas of stability (Fig. 16).

In contrast, the direction of departures based on where geographically individuals of *H. sciureus* were sampled tends to follow a longitudinal axis (Fig. 16). Nevertheless, this species also shows a general correspondence between the position of individuals in genomic space and areas of projected stability like *H. vulpinus*, with the exception of one of the three individuals sampled from the northeast population of GUI (Fig. 15). The deviations of individuals from N_MAM and S_MAM are short and convergent to the same area, the most suitable area of stability through time for this species.

Lastly, in *H. chacarius* the displacement of individuals from expected positions based on geography in the genomic PC map shows a consistent clustering of many individuals in the central part of the species range, even though the sampled populations correspond to areas of projected stability (Fig. 16). Only the most northern populations remain somewhat distinct from this general cluster (i.e., individuals from PNP and POC; Fig. 15).

Tests of a correlation between the residuals from the Procrustes analysis and the environmental distance based on principal components data of environmental variables for the sampled localities suggests that environmental differences contribute to some of the genomic differences observed among individuals in both *H. sciureus* and *H. chacarius*, but was not significant in *H. vulpinus* (Appendix H Table H.5). As with the tests of isolation by distance that did not retain the relative position of sampled populations in geographic space, neither the partial Mantel nor dbRDA detected significant contribution of environmental differences to genomic distances (Appendix H Table H.2 Table H.3).

3.4. Discussion

Our results show that although species present similar life histories traits, they exhibit distinct genomic structure across biomes, suggesting that the biome where the taxon inhabits played an important role on the genomic variation structure. Understanding the patterns of genomic structure in species as the specialized semi-aquatic genus *Holochilus* assisted us to contribute not only with the study of population genetics of *Holochilus* but also with the knowledge about current and past environmental dynamics in the South America biomes and their influence on the degree of connectivity between wetlands (i.e., landscape features that may facilitate or difficult dispersal), as well as the temporal patterns in the availability and behavior of the floods.

Our study highlights the importance of taking into consideration current and past biomes structure (i.e., distances between habitats, traversability of the surrounding landscape, connectivity, shelter, food, etc.), to understand patterns of genomic variation even when mobility capacity (i.e., morphological traits) and habitat requirements does not vary among species (similar life history traits).

3.4.1. Biome-specific associations

We expected that if the difference on environmental variables did not interfere in the diversity pattern of the three species of genus *Holochilus* – first because they share similar life history and second because they are all distributed in the South America wetlands – the species would present correlated of patterns of spatial genomic structure. Although, we found significant differences in

patterns across species, the degree of structure responded differently depending of the source of variation accessed (i.e. geography, current climatic conditions or historical habitat stability).

The higher association between genes and geography (*i.e.*, t_0 values from the Procrustes analyses) was found in the species distributed throughout the more forested habitat, *H. sciureus,* in Amazon biome; the second higher association was detected in *H. vulpinus*, a taxon that have part of its distribution associated with forested environments (southern of Atlantic forest biome) and part associated with open grassy area in Pampa biome; and the species that inhabits the most open environment (Chaco and Pantanal biomes), *H. chacarius*, showed the lower association. Eventhough Procrustes analyses presented significant statistical values neither Mantel nor dbRDA detected the same association, which could be explained by the greater statistical power of Procrustes analyses (Peres-Neto and Jackson 2001). The combination of features of wetlands patches (permeability of the matrix, number, size, proximity, contrast, borders) could explain these patterns of diversity variation (Neiff, 2001), by increasing or decreasing connectivity.

Differences in connectivity between the wetland patches in the three different species home range can explain why the importance of geography varies among taxa. The degree of connectivity among populations is related to grassland availability and seasonal flooding gradient and it varies consistently among biomes (Junk and Piedade 2010, Junk et al. 2014). In South America, large-river floodplains present water level oscillations that range from approximately 4-5 m (Rio Paraná and Paraguay) to up to 12 m (Amazon River with its major tributaries; Wittmann 2010). The extension and duration of the flood distribution also varies among biomes; Amazon has 20–25% of it territory is periodically flooded in until approximately 6 months of the year, while larger places in the Pantanal biome have flood permanence larger than 50% with extremes of 85% of the time (Junk and Piedade 2010; Assine et al., 2016), and the flooded landscapes in southern part of Atlantic Forest and Pampa biomes formed one of the mostly fragmented landscape components of the region (Wittmann et al. 2010).

Three patterns of wetland dynamics in the studied biomes have directly influenced the connectivity between populations of wetland inhabitants. Connectivity can be hampered in Amazon region because its presents the higher instability in the water level as well as landscapes of more fragmented wetlands during a greater period of the year; and also in the southern part of Atlantic Forest and Pampa region that show lower instability in the water level however more fragmented wetland landscapes. On the other hand connectivity can be facilitated in the Chaco and Pantanal regions, where the instability in the water level is lower, when compared with Amazon, and the wetland landscapes are larger and more connected for greater part of the year.

The Procrustes analyses also suggests the effect of environmental differences role for current environment in two species, *H. sciureus* and *H. chacarius*. However, the effects of environmental variables (current or past variables), become insignificant after controlling for the effect of geography, showing that the covariance in these landscapes can difficult to disentangle historical from

contemporary landscape effects (see He et al. 2013). In addition, the correspondence between local environmental differences and population genomic pattern can be more indicative of geographic proximity and common ancestry than differences in environmental variables (Lanier et al. 2015), which could explain why we found statistical significance in *H. chacarius,* an inhabitant of a large humid savanna, and not in *H. vulpinus* that occur in more heterogeneous type of landscapes.

Although geographic distance and current habitat similarity/dissimilarity among population localities might have influenced the degree of genomic differentiation among populations, with similar genomic patterns (e.g. well delimited genomic clusters and higher statistic association in Procrustes analyses) being found among populations associated with forested areas, (*H. sciureus* and *H. vulpinus*), biome-specific signatures seems to be more related with demographic process associated with past climate history inducing habitat shifts. In addition, species that show larger stability area as well as higher association between genes and geography (*H. sciureus* and *H. vulpinus*) also show the largest local deviations from geographic expectations (*i.e.*, the length of the lines in the Procrustes analyses, Fig. 16). It suggests that habitat stability is contributing mostly to regional structuring of genomic variation in the species associated with areas predominantly covered by forest (*H. sciureus* and *H. vulpinus*).

Species that occupies the southern part of Atlantic Forest and Pampas (*H. vulpinus*) exhibited the largest stable area, almost all its current distribution was recovered as stable along the time; this species, as described above, also exhibited great association between genes and geography but no association with local environmental variables, suggesting that these biomes along the years kept a homogeneously environmental variance necessary to provide the habitat conditions to the *H. vulpinus*.

In the Amazonian species (*H. sciureus*) one particular region (southwestern Amazon) was recovered as stable, together with smaller suitable areas dispersed throughout the biome.

This larger stable region might to have acted as a past refugia for *H. sciureus* and seems to have helped the genomic differentiation of the N_MAM and S_MAM populations from the others that are localized in the same current suitable area for the species, although does not explain the entire structure found in *H. sciureus* and the higher differentiation of the populations from the northern part of the species distribution.

The Pantanal and Chaco biomes almost did not favour the maintenance of stable areas for *H. chacarius.* In fact, the few regions of stability are localized in the southern part of the mask used in the ENMs, and some of these few regions are included on the current and past distributional range of *H. vulpinus*, suggesting the *H. chacarius* might never had inhabited this southern region of the mask. The lack of stability, together with the great difference between habitat suitability of the LGM and current, and the lower association between genes and geography may be a response to a recent and rapid demographic expansion, resulting in a more homogeneously genomic structure.

Thus, current and past features of the biomes matters in different magnitudes and are able to influence patterns of genomic variation even when comparing species that share the same habitats requirements.

3.4.2. Local versus broad scales signatures

Despite evidence for biome-specific signatures across broad spatial and temporal scales, there is evidence for a finer-scale common population structuring. We found correlated estimates of genetic diversity and Wright's inbreeding coefficient across species and among populations, and a differences regarding to the observed heterozygosity, with *H. sciureus* presenting the lowest values (Appendix H Table H.1 Fig. H.1).

Species traits similarities could explain some points. For example, as high-specialized semi-aquatic, herbivorous and open marshes inhabitants, these rodents have food and shelter always associated with grass and shrubs (as confirmed by the gut morphology and diet; Hershkovitz 1955, Domínguez-Bello and Robinson 1991), their nests are always close to water (Sauthier et al. 2010), and they present more adaptations to swim than to climb and run (Samuels and Van Valkenburgh 2008, Gandini 2014), revealing that the presence of specific type of vegetation and proximity of the water is an important component to the species survival. When we look to the current conditions of biomes studied here we observe that there is no local condition separating populations of *H. chacarius*, as the Pantanal and Chaco biomes are large humid savannas, and be a good swimmer and plant eater is a great advantage in these area. On the other hand, when we look to Amazon, Atlantic Forest and Pampas we can observe a greater environmental heterogeneity, with high suitable areas separated by larger unsuitable ones, which for a specialized organism would represent a limiting factor (even for small distances), and so the dispersion of these rodents would be dependent of the narrow grass patches along the seasonal flooded river systems, which become dislodged and float downstream as rafts during the rainy season when river water levels are high (Patton et al. 2001).

But even among the more forested biomes the similar patterns could be derived from different explanatory causes. For example, the maintained local genomic structure of *H. sciureus,* even without areas of stability in the north part of Amazon biome could be due to species-specific traits as the food and shelter specialization, combined with local differentiation in the composition of the wetlands landscapes in Amazon; whereas in the *H. vulpinus* distribution, species-specific traits more related with the semi-aquatic traits specialization, could be preventing these rodent to transpose larger dry areas, as the ones localized in the Pampa region. Although, unlike the Amazon forest where geographic records are really sparse and discontinuous throughout the range distribution, the distribution of *H. vulpinus* in the southern part of Atlantic Forest and even in the Pampa area is broad, showing that this species can persist even in more drier conditions.

That is, the results from both population and species level analyses suggest that these rodents also tend to diversify in a more local geographical scales, due to their high specialized traits, eventhough they are apparently capable long-distance dispersion over drier and great forested areas.

3.4.3. Hypotheses motivated by results

We applied different approaches to test the relationships between genomic variation, geography and environment, including one that permit graphic visualizations of deviations in genetic variation in relation to the geographical localities of sampled populations, and all these approaches provide us a useful framework for generating hypotheses.

The higher deviation in Procrustes analyses (regarding to magnitude and geographical orientation) belong to sampled populations from *H. sciureus*, in Amazon forest. This species also shows a more longitudinal deviation, with exception of the GUI population that shows both latitudinal and longitudinal departures. If we consider as a working hypothesis that the original population of *H. sciureus* is localized in the southwestern part of Amazon forest, we could further hypothesize that this region could had acted as refugia for *H. sciureus*, and with the increase of humid environments during the Holocene (increasing the suitable areas mostly longitudinally below Amazon River), the species experimented a demographic expansion towards the central and thus northern part of Amazon biome. The Rio Madeira basin (an area of stability for this species in our past ENM's), with its patches of open grasslands and wetlands (Eva et al., 2002), would be a main corridor for population dispersion through the narrow grass patches along the seasonal flooded river systems, which become dislodged and float downstream as rafts during the rainy season when river water levels are high (Patton et al., 2000). This postglacial expansion northwards could also be associated with founder effects and, as such, would promote relatively higher levels of genetic diversity in these populations.

The species distributed throughout the southern part of Atlantic Forest and Pampas, show a more latitudinal departures of genetic variation, which could be related to latitudinal orientation of the demographic expansion, but without faced huge shifts in the demographic history. The expansion of grasslands along the rivers in the Late Holocene, especially in the Rio de la Plata basin, probably towards their headwaters would favour the northwards dispersion of *H. vulpinus*. Although *Holochilus* is a genus mostly adapted to open and humid areas, the main grassland occupation during the LGM and the posterior increase of humidity during the Holocene could had enabled *H. vulpinus* to persist in these coastal areas without quite expansions or retractions on its geographic distributional range even in the driest periods. Considering the distinct drainage systems during the LGM (Thomaz et al., 2015) due to changes on sea level, it is possible that suitable wetlands occurred between Rio Uruguay and Lagoa dos Patos, playing a role on the gene flow between these coastal samples.

In the Pantanal and Chaco region the climatic data available, together with the results of the EMNs analyzes and the overlap in the genetic differentiation showed in the Procrustes analyses, supports the hyphotesis of a recent and rapid demographic expansion in *H. chacarius*. This demographic expansion probably started to happen during the mid-Holocene, where the enviroment began to present more suitable areas to the expansion of this species. According to the direction of the Procrustes deviations the expansion was predominantly northwards and westwards, but also southernly, along the Rio Paraguay, accompanying the expansion of wetlands. Therefore, while H. vulpinus may have expanded to the north, H. chacarius may also have spread southward along the moist zones of the Rio de la Plata basin, encountering each other on southern Paraguay, on the humid Chaco, where they present present parapatric distribution.

In these flood-dominated systems, being they inserted in major forested biomes or in extensive open areas, species of the genus *Holochilus* exhibited specific signatures in their patterns of genomic variation. Species from open habitats, with apparently major connectivity of populations, exhibited less structure genomic variation in the constituent taxa, while species from forest areas showed more structured genomic variation. These preliminary results suggest that open grassland specialists, such as the genus *Holochilus*, as well as the areas they inhabit exhibit complex biogeographic histories, as interesting as the histories of forest specialists groups, which is the main focus of the diversification studies being conducted in the Neotropical region.

Distinct diversification processes may not be easily intuited from the pattern of deviations visualized in the Procrustes plots, and other summaries analyzes, because they can leave similar signatures which difficult our power to distinguish them (see Knowles and Alvarado-Serrano, 2010; Brown and Knowles, 2012; He et al., 2013; Wang and Bradburd, 2014). Alternatively, they can be impacted by the timing of the events that cause deviations between genes and geography (e.g. Excoffier et al., 2009), meaning that the interpretation needs to take in account the uncertainty that would come with a single summary of genetic variation. Further studies including methodologies that integrate distributional, demographic and genomic (e.g. iDDC, He et al. (2013)) data, and the tempo of evolution, we will be able to provide a framework to test the hypotheses raised here and other possibilities that connect difference in population dynamics over space, time and species-specific traits.

Acknowledgements

This work was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP; Proc: 2014/22444-0, 2012/24099-3, 2009/16009-1). We are also grateful to the curators and researchers that generously granted us tissues samples for the next generation sequencing. We also would like to thanks Andrea Thomaz, Luciana Resende, Anna Papadopoulou and Melisa Olave for all

the help with bioinformatics processing of genomic data and giving helpful feedback on earlier drafts of the manuscript.

References

Alvarado-Serrano, D.F. and Knowles, L.L. 2014. Environmental niche models in phylogeographic studies: recent advances and precautions. – Mol. Ecol. Res. 14: 233–248.

Assine, M.L. et al. 2016. Avulsive Rivers in the Hydrology of the Pantanal Wetland. – In: Bergier, I and Assine, M.L. (eds), Dynamics of the Pantanal Wetland in South America. Springer, 83–110.

Avise, J.C. 2000. Phylogeography. The History and Formation of Species. – Harvard Univ. Press.

Bradburd, G.S. et al. 2016. A Spatial Framework for Understanding Population Structure and Admixture. – PLoS Genet. 12: e1005703.

Brown, J.L. and Knowles, L.L. 2012. Spatially explicit models of dynamic histories: examination of the genetic consequences of Pleistocene glaciation and recent climate change in the American Pika. – Mol. Ecol. 21: 3757–3775.

Brandt, J.F. 1835. Mammalium rodentium exoticorum novorum vel minus rite cognitorum. – Musei Academici Zoologici descriptiones et icons. Mem.Acad.Sci.St.Petersb. 3: 358-442.

Bush, M.B. et al. 2007. A regional study of Holocene climate change and human occupation in Peruvian Amazonia. – J. Biogeogr. 34: 1342–1356.

Carnaval, A.C. and Moritz, C. 2008. Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. – J. Biogeogr. 35: 1187–1201.

Carnaval, A.C. et al. 2009. Stability Predicts Genetic Diversity in the Brazilian Atlantic Forest Hotspot. – Science 323: 785–789.

Catchen, J. et al. 2013. Stacks: an analysis tool set for population genomics. – Mol. Ecol. 22: 3124–3140.

D'Elía, G. et al. 2015. Molecular systematics of South American marsh rats of the genus *Holochilus* (Muroidea, Cricetidae, Sigmodontinae). – J. Mammal. 43: 1–14.

Domínguez-Bello, M.G. and Robinson, M.D. 1991. A Comparison of Digestive Adaptations in Two Neotropical Cricetid Rodents (*Holochilus venezuelae* and *Zygodontomys microtinus*). – Physiol. Zool. 64: 1542–1551.

Eisenberg, J. F. 1981. The Mammalian Radiations. University of Chicago Press, Chicago.

Eva, H.D. et al. 2002. A Vegetation map of South America. Office for Official Publications of the European Commission.

Excoffier, L. et al. 2009 Detecting loci under selection in a hierarchically structured population. – Heredity. 103:285–298.

Gandini, A.M. 2014. Anatomia comparada de estruturas do esqueleto pós-craniano dos roedores semiaquáticos *Holochilus brasiliensis* e *Nectomys squamipes* (Cricetidae, Sigmodontinae).– Universidade Luterana do Brasil, Rio Grande do Sul.

Gonçalves, P.R. et al. 2015. Genus *Holochilus* Brandt, 1835. – In: Patton, J. et al. (eds), Mammals of South America. 2. Rodents. Univ. Chicago Press, pp. 325–335.

He, Q. et al. 2013. Integrative testing of how environments from the past to the present shape genetic structure across landscapes. – Evolution 67: 3386–3402.

Hershkovitz, P. 1955. South American marsh rats genus *Holochilus*, with a summary of sigmodont rodents. – Fieldiana Zool. 37: 639–673.

Hershkovitz, P. 1962. Evolution of Neotropical cricetine rodents (Muridae) with special reference to the phyllotine group. – Fieldiana Zool. 46: 1–524.

Hijmans, R.J. et al. 2005. Very high resolution interpolated climate surfaces for global land areas. – Intl. J. Clim. 25: 1965–1978.

Hoorn, C. et al. 2010. Amazonia Through Time: Andean Uplift, Climate Change, Landscape Evolution, and Biodiversity. – Science 330: 927–931.

Junk, W.J. and Piedade, M.T.F. 2010. An Introduction to South American Wetland Forests: Distribution, Definitions and General Characterization. – In: Junk, W.J. et al. (eds), Amazon floodplain forests: ecophysiology, biodiversity and sustainable management. Springer, pp. 3–25.

Junk, W.J. et al. 2014. Brazilian wetlands: their definition, delineation, and classification for research, sustainable management, and protection. – Aquatic Conserv: Mar. Freshw. Ecosyst. 24: 5–22.

Knowles, L.L. and Alvarado-Serrano, D.F. 2010. Exploring the population genetic consequences of the colonization process with spatio-temporally explicit models: insights from coupled ecological, demographic and genetic models in montane grasshoppers. – Mol. Ecol. 19: 3727–3745.

Knowles, L.L. et al. 2016. Quantifying the similarity between genes and geography across Alaska's alpine small mammals. – J. Biogeogr 43: 1464–1476.

Lanier, H.C. et al. 2015. Colonization from divergent ancestors: glaciation signatures on contemporary patterns of genetic variation in Collared Pikas (*Ochotona collaris*). – Mol. Ecol. 24: 3688–3705.

Legendre, P. and Anderson, M.J. 1999. Distance-Based Redundancy Analysis: Testing Multispecies Responses in Multifactorial Ecological Experiments. – Ecol Monogr. 69: 1–24.

Leite, Y.L.R. et al. 2016. Neotropical forest expansion during the last glacial period challenges refuge hypothesis. – Proc. Natl Acad. Sci. USA 113: 1008–1013.

Massoia, E. 1971. Las ratas nutrias argentinas del genero *Holochilus* descriptas como Mus braziliensis por Waterhouse (Mammalia–Rodentia–Cricetidae). – RIA 8: 141–148.

Massoia, E. 1976. Mammalia. – In: Ringuelet, R. (ed), Fauna de agua dulce de la Republica Argentina. Fundación Editorial Ciencia y Cultura, pp. 1–128.

Moritz, C. et al. 2000. Diversification of rainforest faunas: an integrated molecular approach. – Annu. Rev. Ecol. Syst. 31: 533–563.

Muscarella, R. et al. 2014. ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for MAXENT ecological niche models. – Methods Ecol. Evol. 5: 1198–1205.

Neiff, J.J. 2001. Diversity in some tropical wetland systems of South America. – In: Gopal, B. et al. (eds.) Biodiversity in wetlands: assessment, function and conservation. Backhuys Publishers, pp.157–186.

Novembre, J. and Stephens, M. 2008. Interpreting principal component analyses of spatial population genetic variation. – Nat. Genet. 40: 646–649.

Oksanen, J. et al. 2013. Package Vegan. Community ecology package. Version 2.0–10.

Papadopoulou, A. and Knowles, L.L. 2015. Genomic tests of the species-pump hypothesis: Recent island connectivity cycles drive population divergence but not speciation in Caribbean crickets across the Virgin Islands. – Evolution 69: 1501–1517.

Papadopoulou, A. and Knowles, L.L. 2016. Towards a paradigm shift in comparative phylogeography driven by trait-based hypotheses. – Proc. Natl Acad. Sci. USA 113: 8018-8024.

Pardiñas, U.F.J and Teta, P. 2011. Fossil history of the marsh rats of the genus *Holochilus* and *Lundomys* (Cricetidae, Sigmodontinae) in southern South America. – Estud. Geol. 67: 111–29.

Pardiñas, U.F.J. et al. 2013. Enlarging rodent diversity in west–central Argentina: a new species of the genus *Holochilus*. – J. Mammal. 94:231–240.

Patton, J.L. et al. 2000. Mammals of the Rio Juruá and the evolutionary and ecological diversification of Amazonia. – Bull. Am. Mus. Nat. Hist. 244:1–306.

Peterson, B. K. et al. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. – PloS One.7: e37135.

Phillips, S.J. et al. 2006. Maximum entropy modeling of species geographic distributions. – Ecol. Modell. 190: 231–259.

Prado, J.R. and Percequillo, A.R. 2013. On the geographic distribution of the genera of Tribe Oryzomyini on South America, with some comments on the patterns of diversity and Reig´s Areas of Original Differentiation. – Arq. Zool. 44: 1–120.

Prates, I. et al. 2016. Inferring responses to climate dynamics from historical demography in neotropical forest lizards. – Proc. Natl Acad. Sci. USA. 113:7978–7985.

Purcell, S.B. et al. 2007. PLINK: a toolset for whole-genome association and population-based linkage analysis. –Am. J. Hum. Genet. 81:559–575.

Quattrocchio, M.E. et. al. 2008. Landscape evolution and climate changes in the Late Pleistocene–Holocene, southern Pampa (Argentina): Evidence from palynology, mammals and sedimentology. – Quat. Int. 181: 123–138.

R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.r-project.org (accessed 15 October 2016)

Samuels, J.X. and Van Valkenburgh, B. 2008. Skeletal indicators of locomotor adaptations in living and extinct rodents. – J. Morphol 269:1387–1411.

Sauthier, W.O.U. et al. 2010. Nests of *Oligoryzomys* sp. an *Holochilus brasiliensis* (Rodentia, Cricetidae) eastern Entre Rios Province, Argentina. – Mastozool. Neotrop. 17: 207–211.

Shah,V.B., and McRae, B.H. 2008. Circuitscape: a tool for landscape ecology. Pp. 62–66. Proceedings of the 7th python in science conference (SciPy 2008). Pasadena, CA.

Stein, B.R. 1988. Morphology and allometry in several genera of semiaquatic rodents *Ondatra*, *Nectomys*, and *Oryzomys*. – J. Mammal 693:500–511.

Thomaz, A. et al. 2015. Testing the effect of palaeodrainages versus habitat stability on genetic divergence in riverine systems: Study of a Neotropical fish of the Brazilian coastal Atlantic Forest. – J. Biogeogr. 42: 2389–2401.

Thomé, M.T.C. and Carstens, B.C. 2016. Phylogeographic model selection leads to insight into the evolutionary history of four-eyed frogs. – Proc. Natl Acad. Sci. USA. 113:8010–8017.

Thomé, M.T. et al. 2010. Phylogeography of endemic toads and post-Pliocene persistence of the Brazilian Atlantic Forest. – Mol. Phylogenet. Evol. 55:1018–1031.

Wang, I.J. and Bradburd, G.S. 2014. Isolation by environment. – Mol. Ecol. 23: 5649–5662.

Wang, C. et al. 2010. Comparing spatial maps of human population-genetic variation using Procrustes analysis. – Stat. Appl. in Genet. Mol. Biol. 9: Article13.

Wang, C. et al. 2012. A quantitative comparison of the similarity between genes and geography in worldwide human populations. – PLoS Genet. 8: e1002886.

Warren, D.L. and Seifert, S.N. 2011. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. – Ecol. Appl. 21:335–342.

Whitney, B.S. et al. 2011. A 45 kyr palaeoclimate record from the lowland interior of tropical South America. – Palaeogeogr. Palaeoclimatol. Palaeoecol 307:177–192.

Wittmann, F. 2012. Tree species composition and diversity in brazilian freshwater floodplains. – In: Pagano, M.C. (ed), Mycorrhiza: Occurrence in Natural and Restored Environments. Nova Science Publishers, pp. 223–263.

Wittmann, F. et al. 2010. Phytogeography, species diversity, community structure and dynamics of central Amazonian floodplain forests. – In: Junk, W.J. et al. (eds), Amazon floodplain forests: ecophysiology, biodiversity and sustainable management. Ecological studies, vol 210. Springer, pp 61–102.

Figures

Figure 14. Examples of the open wetland biomes of the central, southern and northern region of South America, such as the (a) Poconé in Pantanal biome, (b) Miranda also in Pantanal biome, (c) Caçapava do Sul in Pampa biome (c) São José do Norte in Atlantic Forest biome, (d) lower Xingú River in Amazon biome, and (e) Japurá River also in Amazon biome, inhabited by (g) *Holochilus*.

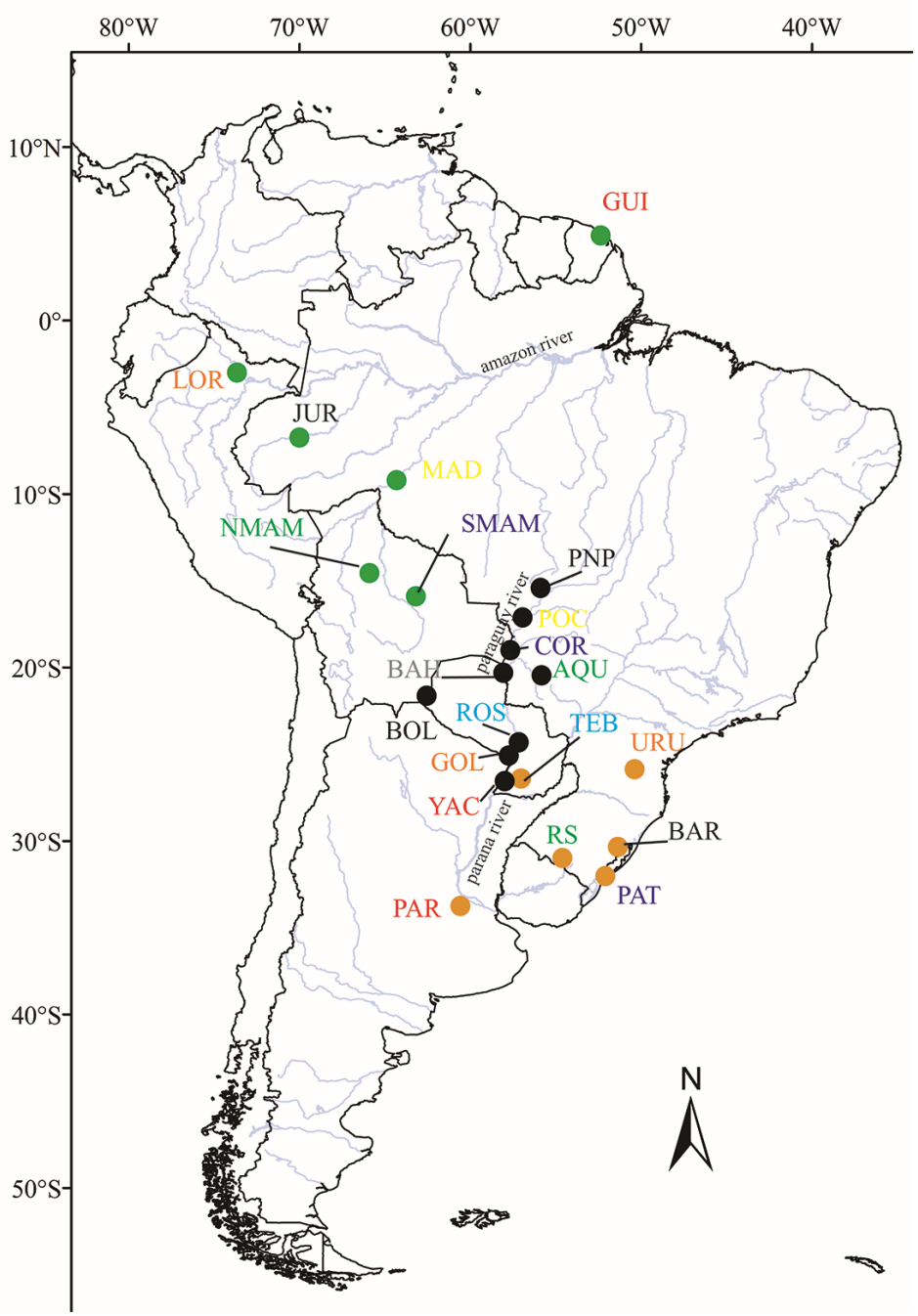


Figure 15. Sampled populations of the three widespread marsh rats in Brazil, with each species color-coded – that is, with *H. chacarius*, *H. sciureus*, and *H. vulpinus* shown in black, green, and orange, respectively. The sampling locations span each species' range, which are widespread and non-overlapping with taxa occupying fairly distinct biomes (see Fig. 14). Specifically, *H. sciureus* and *H. vulpinus* inhabit the wetlands of forested biomes; *H. sciureus* inhabits predominantly the wet grasslands along rivers in the Amazon biome, whereas *H. vulpinus*, occurs mainly through the Parana/Uruguay river basin, in the Atlantic Forest and part of Pampas biomes. In contrast, *H. chacarius*, is distributed throughout the relatively open wetland in the humid Chaco and Pantanal biomes.

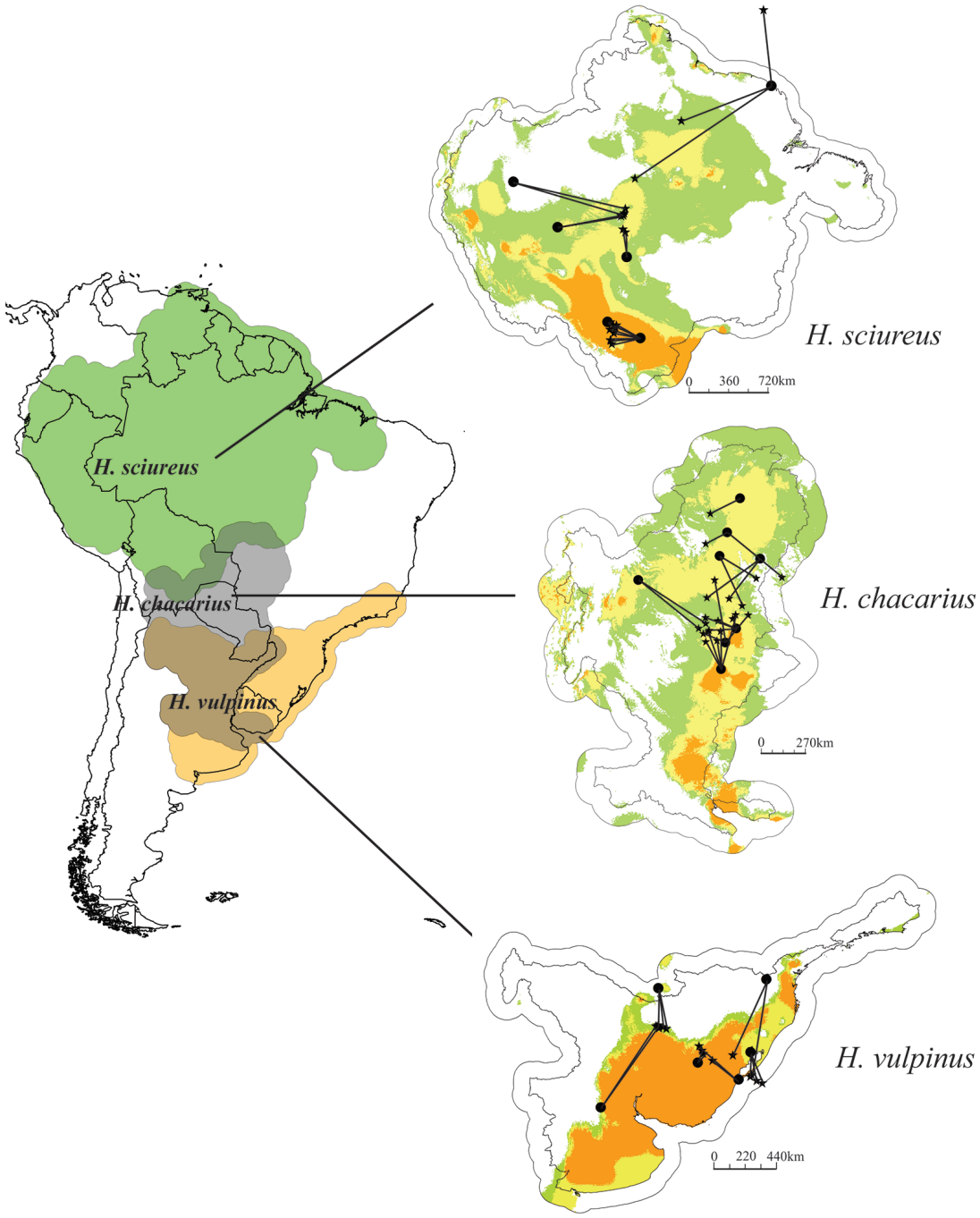


Figure 16. Plots of Procrustes-transformed PCA's of genomic variation with each individual mapped in the genomic PC-space (marked by stars) relative to the geographical location of sampled populations (marked by circles). The plots for each taxa are plotted on a map that highlights the relative stability of the region, where stability is defined as regions that have remained suitable overtime (see Appendix G, Fig. G.1 for distributional maps from ENMs for each geologic period). Stable area since LGM, 21 kya, are marked in orange, and yellow marks areas that have been stable since Holocene, 6 kya, relative to the unstable areas marked in green (i.e., areas that project a distribution for the present, but not the past). Note that the length of the line connecting individuals in the genomic PC-space to their geographical location represents the magnitude of the deviation from the expected pattern of genetic variation based on geography.

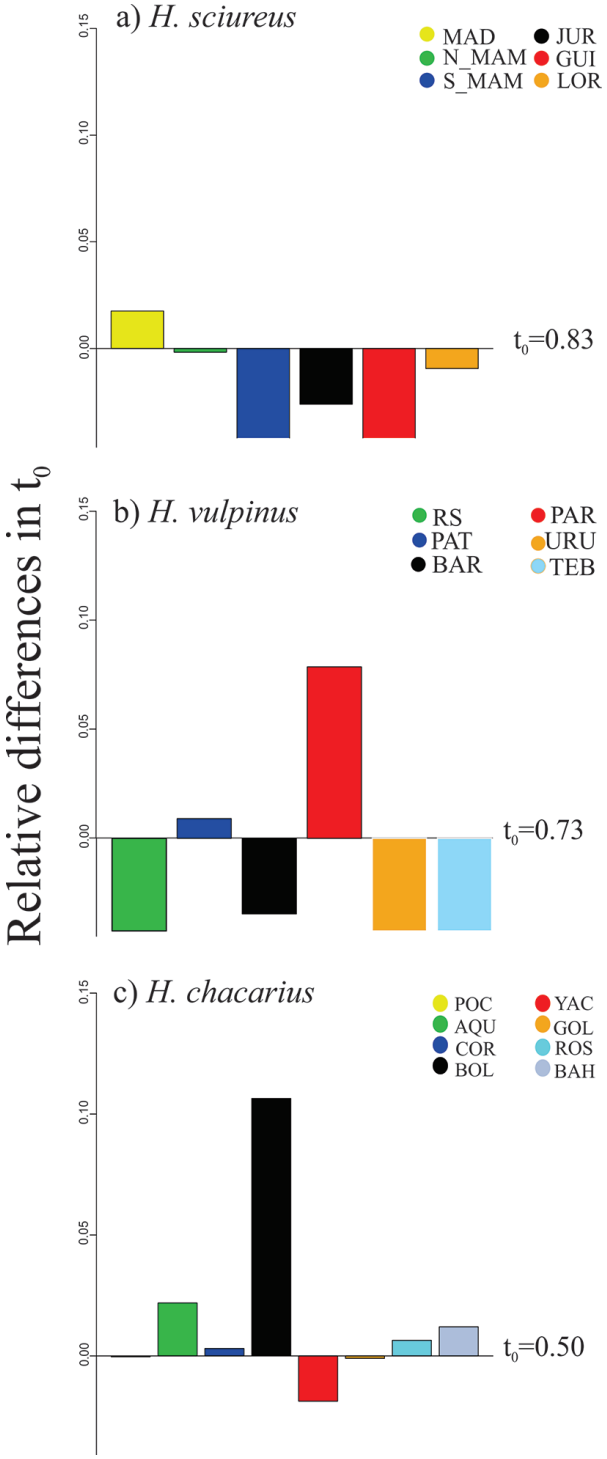


Figure 17. Comparison of the changes in the strength of the association between genes and geography with the exclusion of individual populations (i.e. *t"*) relative to when all populations are analyzed (i.e. t_0). Values for each species are standardized by t_0 (i.e. 0 on y-axis corresponds to t_0) such that positive values indicate a stronger association between genes and geography when a population is excluded, whereas negative values indicate a weaker association. Bar colours represent sampling populations following the coloured names of populations in Fig. 15.

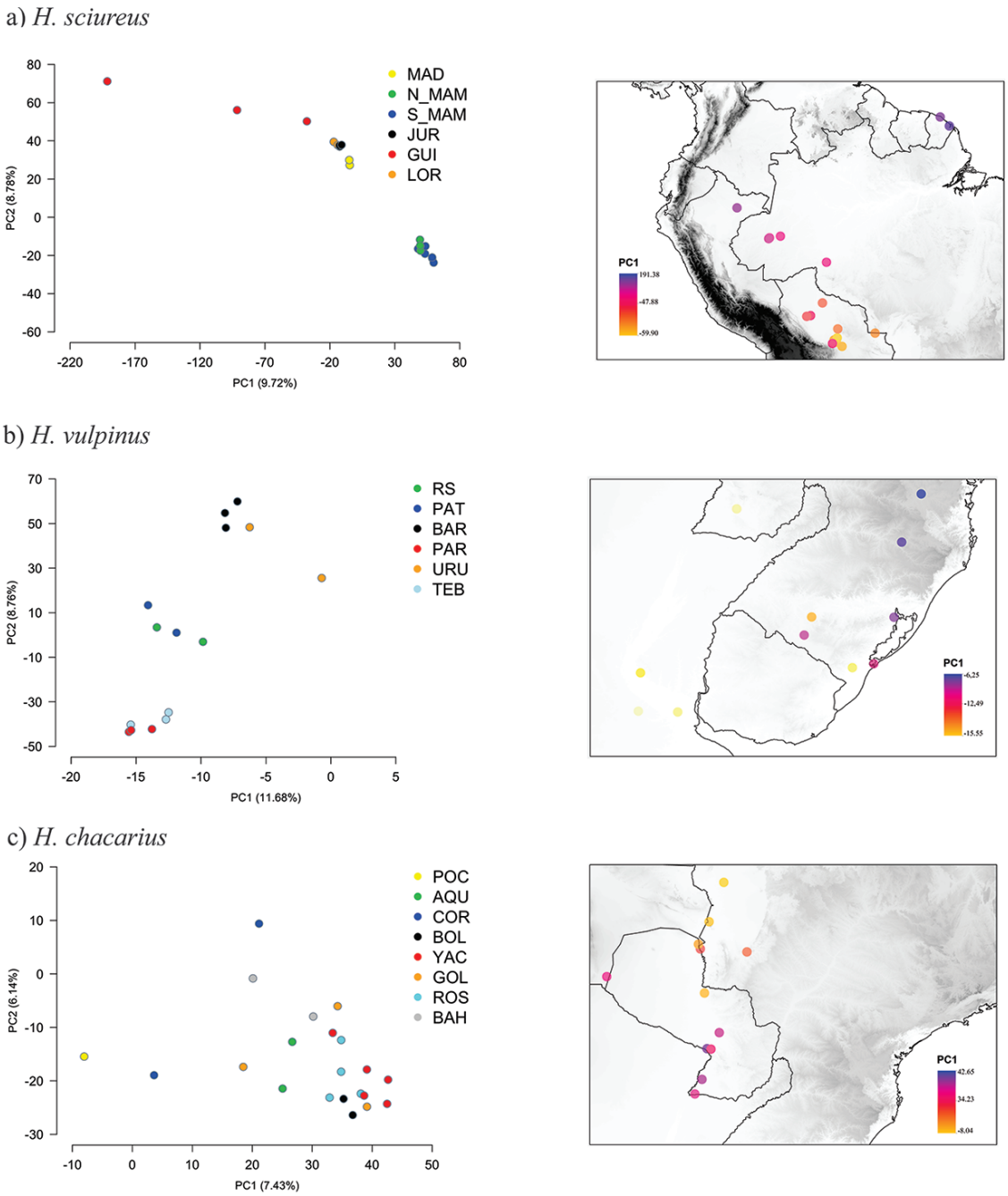


Figure 18. Distribution of individuals along PC1 of genomic variation based on the analysis of polymorphic SNPs; individuals are colour coded according to their population identities in the PC. For each PC analysis, the positions of individuals along P1 one are colour coded and shown on a map (i.e., different colors correspond to individuals with the greatest genomic difference along PC1).

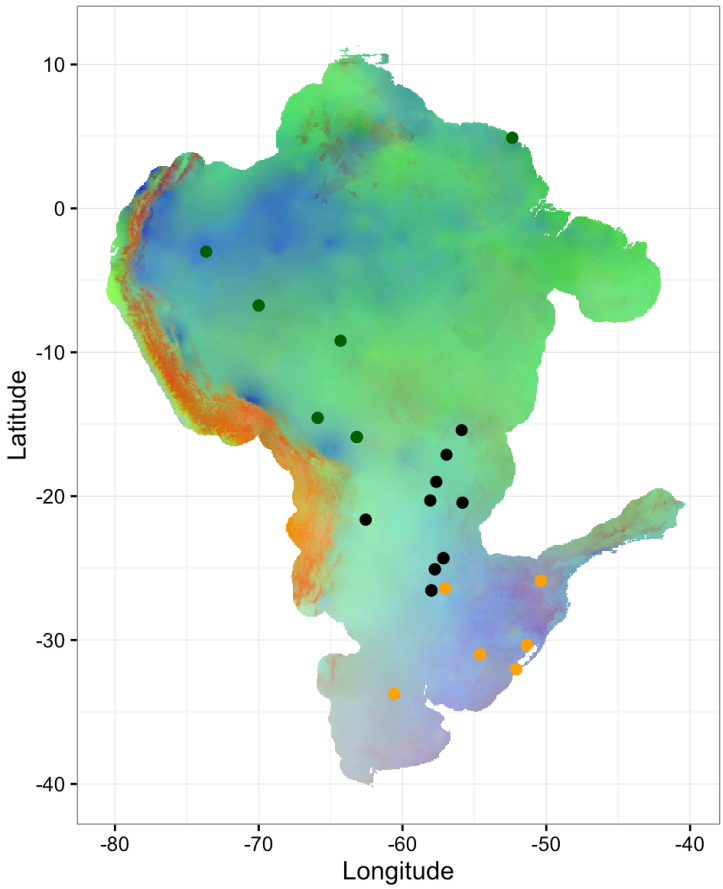


Figure 19. Map of the environmental variation across the region where the three species are distributed, where differences in colour depict geographic regions that differ the most from each other. Specifically, PC1, PC2 and PC3 of bioclimatic variables across the landscape were rescaled between 0 and 1, and the RGB color composite was calculated and plotted in the map with PC1 set as the red scale, PC2 as the green scale, and PC3 as the blue scale. Coloured dotes correspond to the populations presented in Fig 15.

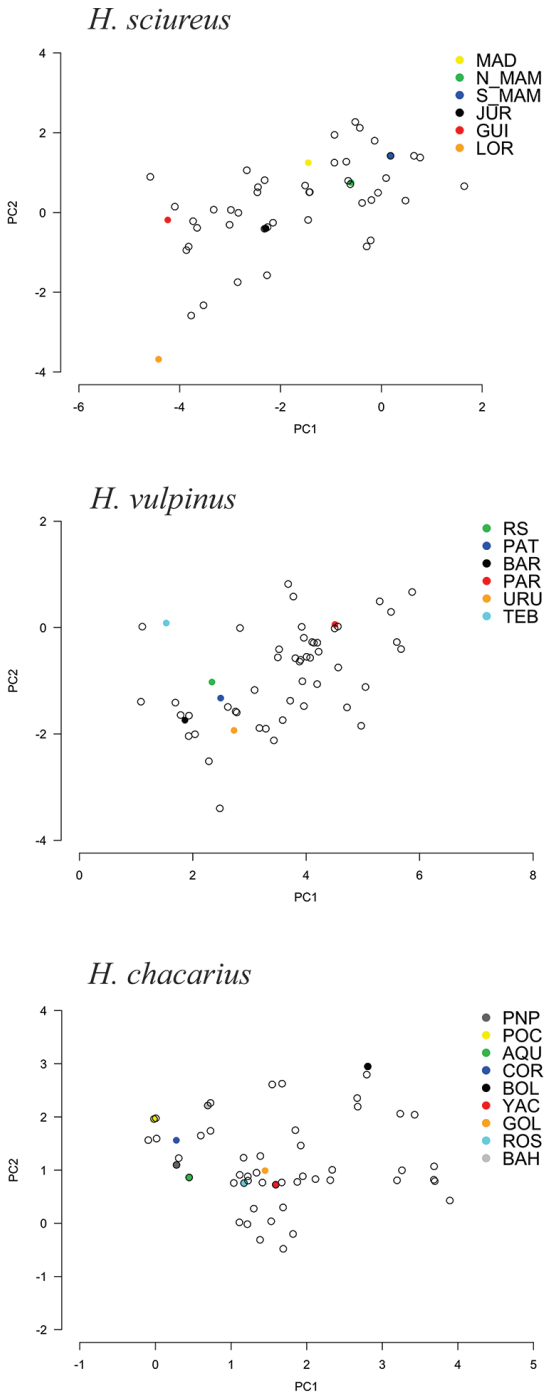


Figure 20. Dispersion in environmental space of the sampled populations used in our genetic analyses (marked as coloured dots) relative to the PC values for *Holochilus* sampling locations used in the ENMs. For *H. sciureus* PC1 is strongly positively correlated with Minimum Temperature of Coldest Month (Bio6) and explains most of the variation among populations (i.e., 55.49%), whereas PC2 explains relatively little variation among populations (19.66%). For *H. vulpinus* PC1 is strongly positively correlated with the Annual Precipitation (Bio12) and explains most of the variation among populations (48.56%), whereas PC2 explains 24.26% of the variation among populations. For *H. chacarius* PC1 is strongly positively correlated with the Minimum Temperature of Coldest Month (Bio6) and explains most of the variation among populations (54.4%), whereas PC2 explains 21.8% of variation among populations.

APPENDIX E. Summaries of geographical information and genomic data

Details about locality information of sequenced individuals (Table E.1) and protocols used to generate genomic data, as well as summaries of genomic data (Table E.2) are provided in this appendix, Appendix E.

Genomic data generation and processing. Genomic DNA was extracted from liver, muscle or skin of each individual using the Qiagen DNeasy Blood and Tissue Kit following the manufacturer's Animal Tissue Protocol. Four reduced representation libraries were built using the Double Digest RADseq method (see Peterson *et al.*, 2012). For each library, individuals were double digested with the restriction enzymes *Eco*RI and *Mse*I and uniquely tagged with a 10bp barcode. The digested products were pooled and size-selected using Pippin Prep (Sage Science, Beverly, MA, USA) and amplified by PCR with iProof™ High-Fidelity DNA Polymerase (BIO-RAD). DNA quantification and cleaning with Agencourt AMPure XP (Beckman Coulter, Indianapolis, IN, USA) was used for each step in the library construction procedure. Each genomic library was sequenced on an Illumina HiSeq2000 to generate 150bp single-end reads at The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Canada.

Raw sequence reads for each species were processed separately in Stacks v.1.35 (Catchen *et al.*, 2013). Specifically, the reads were demultiplexed and filtered using *process radtags,* with one mismatch in the adapter sequence (--adapter_mm) and a barcode distance of two (--barcode_dist). Individuals with less than 500,000 sequences were excluded. The USTACKS program was used to create a *de novo* assembly of reads with a minimum coverage depth (m = 6) into putative loci, with a maximum distance (in nucleotides) of M=2, and the *removal algorithm* (-r), the *deleveraging algorithm* (-d), and the model type equal bounded (--model_type) settings, and an error bound rate (ε) of 0.1 (--bound_high). A catalog of consensus loci among individuals was constructed with the CSTACKS program from the USTACKS output files, using all of the individuals, allowing for two mismatches between sample tags when building the catalog (-n 2); loci were identified using SSTACKS under default options. SSTACKS output files were loaded into the POPULATIONS module (with parameters: -r 0 -p 2 -m 6 --min_maf 0 --max_obs_het 0.5) and with the SNP data exported in Variant Call Format (vcf). SNP's from the five last base pairs in the 3' end of all loci were removed because increased number of SNPs, as were loci with high theta values, since these are suggestive of sequencing and assembly errors. All STACKS modules were run under parallel execution with 8 threads in the University of Michigan flux. A summary of the number of pre- and post-processing reads, as well as the number utilized by *Stacks*, is given in Table E.2 in this appendix.

Table E.1.Locality information for each sampled population and the number of individuals used in analyses of *H. sciureus* and *H. vulpinus* that are distributed in the wetlands of the forested areas, and *H. chacarius* from the open wetland area (see Fig. 15 for the population locations and names).

Species	Population	Locality	Latitude	Longitude	N of Individuals
H. sciureus	French Guiana	Kaw, Régina, French Guiana	4.48	-52.04	4
		Piste de l Anse Sinnamary, French Guiana	5.38	-52.95	
		RN1 pK-35, Macouria, French Guiana	4.91	-52.37	
	Jurua	Altamira, right bank Rio Juruá, Amazonas, Brazil	-6.58	-68.9	3
		Near Miranda. Left bank Rio Juruá, Amazonas, Brazil	-6.75	-70	
		Penedo, right bank Rio Juruá, Amazonas, Brazil	-6.83	-70.08	
	Loreto	Iquitos, Loreto, Peru	-3.74	-73.24	2
		Maynas, Loreto, Peru	-3	-73.66	
	Madeira	UHE Jirau, Porto Velho, Rondonia, Brazil	-9.264	-64.641	5
		UHE Santo Antônio, Porto Velho, Rondonia, Brazil	-8.801	-63.949	
	North_Mamore	Mamore, San Ramon, Beni, Bolivia	-13.28	-64.71	3
		Rio Matos, 6 KM E of Estacion Biologica de Beni, Beni, Bolivia	-14.63	-66.28	
		Rio Tijamuchi, Beni, Bolivia	-14.56	-65.9	
	South_Mamore	3.5 KM W Estacion Pailon, Beni, Bolivia	-17.65	-62.75	7
		Ayacucho, Beni, Bolivia	-17	-63.55	
		Estancia Cachuela Esperanza, Beni, Bolivia	-16.78	-63.23	
		Estancia San Marcos 6 km W of Ascension, Beni, Bolivia	-15.89	-63.18	
		San Miguel Rincon, Beni, Bolivia	-17.23	-63.32	
		San Rafael de Amboro, Beni, Bolivia	-17.35	-63.71	
H. vulpinus	Barra do Ribeiro	Barra do Ribeiro, Rio Grande do Sul, Brazil	-30.34	-51.36	3
	Lagoa dos Patos	Arroio Grande, Rio Grande do Sul, Brazil	-32.17	-52.86	2
		Rio Grande, Rio Grande do Sul, Brazil	-32.03	-52.09	
	North Uruguay River	Anita Garibaldi, Santa Catarina, Brazil	-27.68	-51.13	4
		UN-SIX Petrobrás, Parana, Brazil	-25.87	-50.38	
	Parana River	6 km S (by road) of Puerto Ibicuy, Entre Ríos,	-33.78	-59.16	3

		Argentina			
		Las Cuevas, 35 km (by air) SSE of Diamante in Rio Parana, Entre Ríos, Argentina	-32.36	-60.5	
		Pergamino, FARM INT, GRID X, Buenos Aires, Argentina	-33.75	-60.58	
	Rio Grande do Sul	Dom Pedrito, Rio Grande do Sul, Brazil	-30.99	-54.6	2
		São Gabriel, Rio Grande do Sul, Brazil	-30.33	-54.33	
	Tebicuary	5Km ENE Ayolas, Misiones, Paraguay	-27.4	-56.9	4
		610m S Hotel Centu Cue, Misiones, Paraguay	-26.43	-57.03	
		Margin of Rio Tebicuary, 1 Km upstream (E) Hotel Centu Cue, Misiones, Paraguay	-26.43	-57.03	
		Margin of Rio Tebicuary, 1.2 Km upstream (E) Hotel Centu Cue (opposite margin), Misiones, Paraguay	-26.41	-57.03	
H. chacarius	Aquidauana	Fazenda Rio Negro, Aquidauana, Mato Grosso do Sul, Brazil	-20.45	-55.83	2
	Bahia Negra	4 Km N (by air) of Bahia Negra, Estancia Dona Julia, W bank of Rio Paraguay, Alto Paraguay, Paraguay	-20.14	-58.15	3
		6 Km SE (by air) of Bahia Negra, W bank of Rio Paraguay along Riacho Ramos, Alto Paraguay, Paraguay	-20.3	-58.07	
		17 Km N (by air) of Bahia Negra, W bank Rio Negro, Estancia Immaculada Concepcion, Alto Paraguay, Paraguay	-20.08	-58.16	
	Corumba	Fazenda Alegria, Corumbá, Mato Grosso do Sul, Brazil	-19	-57.65	2
	Estancia Bolivar	Estancia Bolivar, Tarija, Bolivia	-21.63	-62.56	2
	Estancia Golondrina	24 Km NW (by air) of Villa Hayes, Estancia la Golondrina, Presidente Hayes, Paraguay	-25.08	-57.75	2
	Estancia Yacare	Estancia Yacaré, Ñeembucú, Paraguay	-26.55	-57.99	7
	PNPantanal	Parque Nacional do Pantanal, Mato Grosso do Sul, Brazil	-15.41	-55.89	2
	Pocone	Base de Pesquisas do Pantanal, CENAP/IBAMA, 110 Km SSW Poconé,	-17.12	-56.94	2

		Mato Grosso, Brazil			
	Rosario	Island in middle of Rio Paraguay, 10 Km (by air) NW of Rosario, San Pedro, Paraguay	-24.31	-57.16	4

Table E.2. Summary of genomic data collected in each population of *H. sciureus*, *H. vulpinus* and *H. chacarius*. The counts of raw number of reads from the Illumina run and the number of reads utilized after processing (i.e., after excluding reads with low quality scores and ambiguous barcodes), are shown, as are the number of loci (N_Loci[a]) before and after (N_Loci[b]) filtering for missing data, and the number of SNPs (N_SNPs) before filtering for missing data (the number of SPNs after the missing data filtering is the same as the N_Loci[b]).

Species	Population	N	Raw reads	Utilized reads	N_Loci[a]	N_Loci[b]	N_SNPs
H. sciureus	French Guiana	4	5,370,721	3,819,253	139,198	17,513	357,050
	Jurua	3	8,646,897	7,292,126			
	Loreto	2	9,582,246	7,784,946			
	Madeira	5	9,784,223	7,245,494			
	North_Mamore	3	16,058,85	14,014,829			
	South_Mamore	7	34,574,074	29,984,890			
H. vulpinus	Barra do Ribeiro	3	7,870,773	5,941,191	82,559	8,035	160,879
	Lagoa dos Patos	2	2,035,029	1,384,351			
	North Uruguay River	4	4,481,875	2,678,133			
	Parana River	3	20,156,007	17,538,909			
	Rio Grande do Sul	2	4,074,971	3,188,331			
	Tebicuary	4	14,391,985	12,214,919			
H. chacarius	Aquidauana	2	11,317,718	9,800,472	155,697	32,210	359,728
	Bahia Negra	3	16,461,054	13,956,686			
	Corumba	2	8,167,719	6,937,914			
	Estancia Bolivar	2	9,754,574	8,628,898			
	Estancia Golondrina	2	10,470,970	9,095,254			
	Estancia Yacare	7	53,455,260	46,007,223			
	PNPantanal	2	4,257,699	3,346,804			
	Pocone	2	5,226,472	4,247,099			
	Rosario	4	31,751,126	27,496,960			

APPENDIX F. Summaries of samples collected and scientific collections

Details about the scientific collections visited to gather information about the distributional records of *Holochilus* specimens, as well as details about vouchers of sequenced individuals are provided in this appendix, Appendix F.

List of visited scientific collections:

Musee National d'Histoire Naturelle (Paris, French)

Museo Argentino de Ciencias Naturales "Bernardino Rivadavia" (Buenos Aires, Argentina)

Museo de Historia Natural de San Rafael (Mendoza, Argentina)

Museo Nacional de Historia Natural (Montevideo, Uruguay)

Museu da Universidade Federal de Minas Gerais (Belo Horizonte, Brazil)

Museu de Zoologia da USP (São Paulo, Brazil)

Museu Nacional da Universidade Federal do Rio de Janeiro (Rio de Janeiro, Brazil)

Museum für Naturkunde (Berlin, Germany)

Museum of Southwestern Biology (Albuquerque, United States)

National Museum of Natural History (Washington DC, United States)

The Field Museum (Chicago, United States)

The Natural History Museum (London, England)

Universidade Federal do Espírito Santo (Vitória, Brazil)

University of Michigan Museum of Zoology (Ann Arbor, United States)

Zoologische Staatssammlung München (Munich, Germany)

List of specimens used in the molecular analyses: H. chacarius: **ARGENTINA**: Corrientes: 0.5 km n of Itati, Island in Rio Parana: UMMZ166514. **BOLIVIA:** Tarija: Estancia Bolivar: MSB235493. **BRASIL**: Mato Grosso: Base de Pesquisa do Pantanal, CENAP/IBAMA, 110 km SSW Poconé, Poconé: CTA1539, MVZ198015. Mato Grosso do Sul: Fazenda Alegria, Corumbá: LBCE5361, LBCE5409. Fazenda Rio Negro, Aquidauana: LBCE4432, LBCE4433. Parque Nacional do Pantanal: MZ35144, MZ35145. **PARAGUAY**: Ñeembucú: Estancia Yacaré: TTU108438, TTU108439, TTU108458, TTU108460, TTU108462. Estancia Yacare, 1.68 Km NW of Puesto San Fernando: UMMZ174829. Presidente Hayes: 24 km NW (by air) of Villa Hayes, Estancia la Golondrina: UMMZ165996, UMMZ166707. West bank of Rio Paraguay, 4 km NW of Puerto Fonciere: UMMZ166389. SAN PEDRO: Island in Rio Paraguay, 10 km (by air) NW of Rosario: UMMZ166250, UMMZ166375, UMMZ166391, UMMZ166409. ALTO PARAGUAY: 4 km N (by air) of Bahia Negra, Estancia Dona Julia, W bank of Rio Paraguay: UMMZ166255. 6 km SE (by air) of Bahia Negra, W Bank of Rio Paraguay along Riacho Ramos: UMMZ166256. 17 km N (by air) of Bahia Negra, W bank Rio Negro, Estancia Immaculada Concepcion:UMMZ166381.

H. sciureus: **BOLIVIA:** El Beni: Mamore, San Ramon: MSB208404. Rio Tijamuchi: MSB211440. Rio Matos, 6 KM E of Estacion Biologica de Beni: MSB211485. Santa cruz: San Miguel Rincon: MSB210547. Ayacucho: MSB236995. Estancia Cachuela Esperanza: MSB55296. 3.5 KM W Estacion Pailon: MSB55298.

San Rafael de Amboro: MSB55988. Estancia San Marcos 6 km w OS ascension: MSB99051. **BRASIL:** Amazonas: Altamira, right bank Rio Juruá: MVZ193736. Near Miranda, left bank Rio Juruá: MVZ190357. Penedo, right bank Rio Juruá: MVZ193734. Rondonia: UHE Jirau, Porto Velho: BRFMP10668, MJ742. UHE Santo Antônio, Porto Velho: MC164, MPP24, SA88. **FRENCH GUIANA**: Kaw: T4581, T4595. Macouria: RN1 pK-35: T6822. Sinnamary, Piste de l Anse: T6876. **PERU:** Loreto: Iquitos: TTU75634. Maynas: TTU98745.

H. vulpinus: **ARGENTINA**: Buenos Aires: Pergamino, FARM INT, GRID X: MSB204443. Entre Rios: 6 km S (by road) of Puerto Ibicuy: UMMZ166479. Las Cuevas, 35 km (by air) SSE of Diamante in flood plain of Rio Parana: UMMZ166524. **BRASIL:** Parana: UN-SIX Petrobrás: UFSC5167. Santa Catarina: Anita Garibaldi: UFSC5074. PCH Passos Maia, Rio Chapecó: UFSC5005, UFSC5006. Rio Grande do Sul: Arroio Grande: MCNU1946. Barra do Ribeiro: MCNU3424, MCNU3425, MCNU3426. Dom Pedrito: MCNU1943. Rio Grande: MCNU3427. São Gabriel: MCNU2342. **PARAGUAY:** Paraguari: UMMZ174824. Margin of Rio Tebicuary, 1 Km upstream (E) Hotel Centu Cue: UMMZ175059. Margin of Rio Tebicuary, 1.2 Km upstream (E) Hotel Centu Cue (opposite margin): Missiones: 5KM ENE AYOLAS: UMMZ125495. 610 m S Hotel Centu Cue: UMMZ174820.

APPENDIX G. Summaries of ENM settings and projections of current, Holocene and LGM distributions

The specific settings used in the ecological niche modeling (ENMs) are provided in Table G.1, as are maps of the projected distributions of *H. sciureus*, *H. vulpinus* and *H. chacarius* (Fig. G.1) for the present, holocene and last glacial maximum (LGM) in this appendix, Appendix G.

Table G.1. For each species, the number of distribution points and variables used in niche modeling are shown. Specifically, only the subset of bioclimatic variables with < 0.70 Pearson's r correlation were used in the ENMs. The bioclimatic variables include: Annual Mean Temperature (1), Mean Diurnal Range (2), Isothermality (3), Temperature Seasonality (4), Maximum Temperature of Warmest Month (5), Minimum Temperature of Coldest Month (6), Temperature Annual Range (7), Mean Temperature of Wettest Quarter (8), Mean Temperature of Driest Quarter (9), Mean Temperature of Warmest Quarter (10), Mean Temperature of Coldest Quarter (11), Annual Precipitation (12), Precipitation of Wettest Month (13), Precipitation of Driest Month (14), Precipitation Seasonality (15), Precipitation of Wettest Quarter (16), Precipitation of Driest Quarter (17), Precipitation of Warmest Quarter (18), and Precipitation of Coldest Quarter (19). See the Methods for more details. The Maxent's threshold used for each species was: 10 percentile training presence logistic threshold for *H. chacarius,* the Maximum test sensitivity plus specificity logistic threshold for *H. sciureus*, and 10 percentile training presence logistic threshold for *H. vulpinus*.

Species	Number of Points	Variables	Highest Contribution	Features	Regularization
H. sciureus	46	3,4,6,12,14,18,19	6	LGH	2.5
H. vulpinus	54	2,3,5,6,12,15	12	LQ	2.0
H. chacarius	52	3,6,12,14	6	LQ	0.5

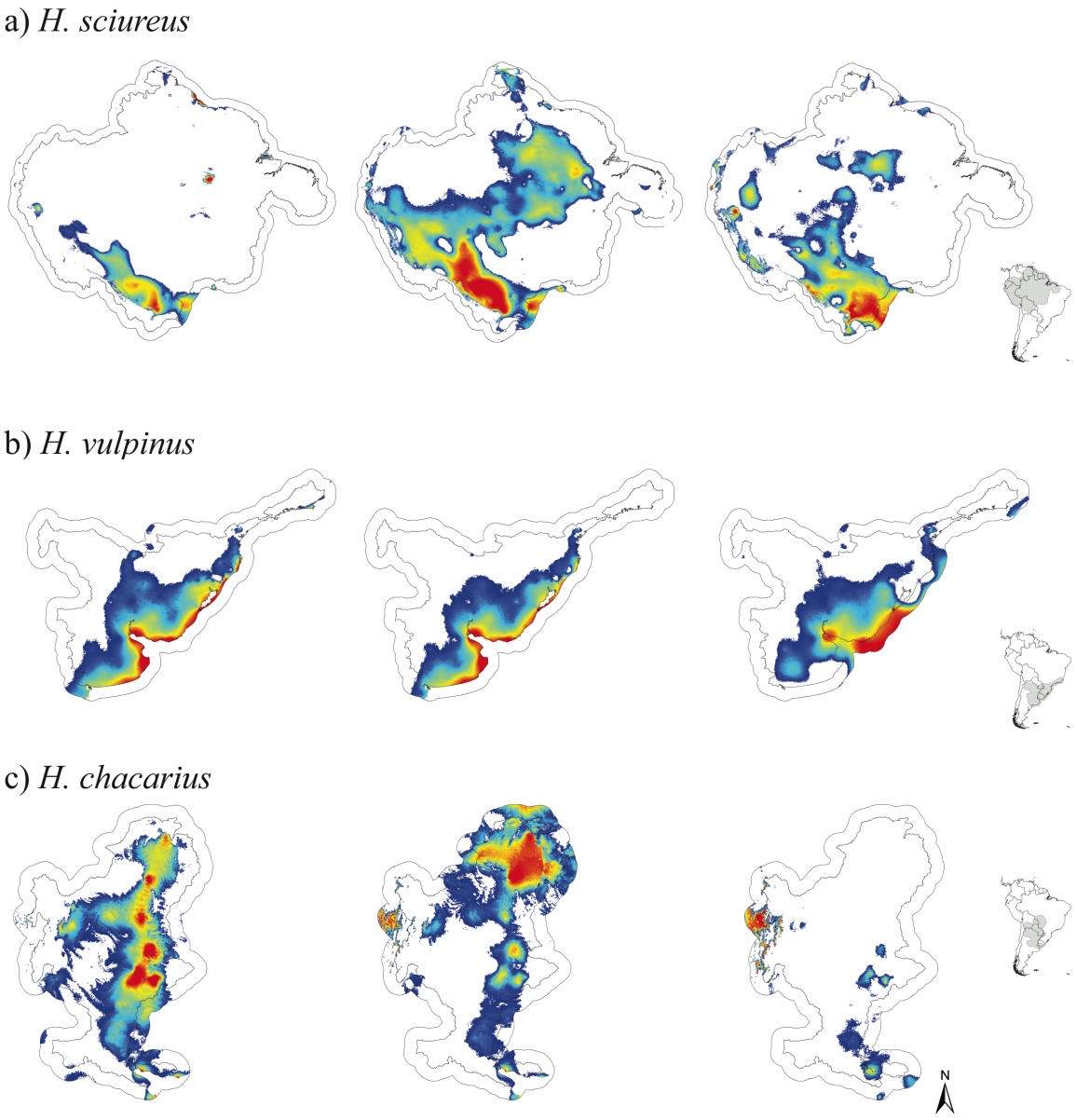


Figure G.1. Projections of the distributions for *H. sciureus*, *H. vulpinus* and *H. chacarius* for the present (left), Holocene, 6 mya (middle) and LGM, 21 mya (right); only the specific distribution areas are shown (i.e., see grey area of small inset for the position of each distribution in South America). Warmer colours indicate higher suitability of habitat, while cooler colours indicate unsuitable habitat; see main test for details about ENMs.

APPENDIX H. PC-maps of genetic variation and summaries of genetic variation

In addition to descriptions of genetic variation from population genetic summary statistics in each of the three taxa (see Table H.1 and Fig. H.1), tests of association between genetic distances and geographic distance and/or environmental differences among populations using Mantel and dbRDA (Table H.2 and H.3, respectively), Procrustes analyses showing the strength of the association between genes/geography based on sequential exclusion of populations (Table H.4), as well as results of Procrustes analyses showing the significance of the overall similarity in the association between bioclimatic variables and the residuals of the Procrustes analyses based on all populations (Table H.5), and inferred geogenetics maps from SpaceMix analyses (Fig. H.2) are provided in this Appendix H.

Table H.1. Summaries of genetic diversity (average observed heterozygosity (H_{obs}), average nucleotide diversity (π), and Wright's inbreeding coefficient (F_{IS})) per sampled population for each of the species, as well as sample sizes and percent of sites that were polymorphic within each population.

		N	**%poly**	**H_{obs}**	**π**	**F_{is}**
H. sciureus	French Guiana	4	0.38	0.0634	0.0963	0.0581
	Jurua	3	0.38	0.0708	0.1023	0.0528
	Loreto	2	0.36	0.092	0.1139	0.0328
	Madeira	5	0.48	0.0819	0.1204	0.0677
	North_Mamore	3	0.59	0.1192	0.1718	0.0867
	South_Mamore	7	1.05	0.1135	0.2207	0.2214
H. vulpinus	Barra do Ribeiro	3	0.46	0.1547	0.1794	0.0402
	Lagoa dos Patos	2	0.29	0.1325	0.1449	0.0186
	North Uruguay River	4	0.33	0.122	0.1581	0.0575
	Parana River	3	0.48	0.1511	0.1797	0.0484
	Rio Grande do Sul	2	0.40	0.1462	0.1658	0.0294
	Tebicuary	4	0.49	0.1496	0.1756	0.0438
H. chacarius	Aquidauana	2	0.40	0.1217	0.147	0.0379
	Bahia Negra	3	0.52	0.1323	0.1723	0.0649
	Corumba	2	0.38	0.1129	0.1346	0.0325
	Estancia Bolivar	2	0.18	0.0896	0.091	0.0021
	Estancia Golondrina	2	0.36	0.1134	0.1347	0.0319
	Estancia Yacare	7	0.84	0.1179	0.2003	0.1709
	PNPantanal	2	0.4	0.1214	0.1434	0.0331
	Pocone	2	0.38	0.1293	0.1458	0.0247
	Rosario	4	0.66	0.1296	0.1914	0.1097

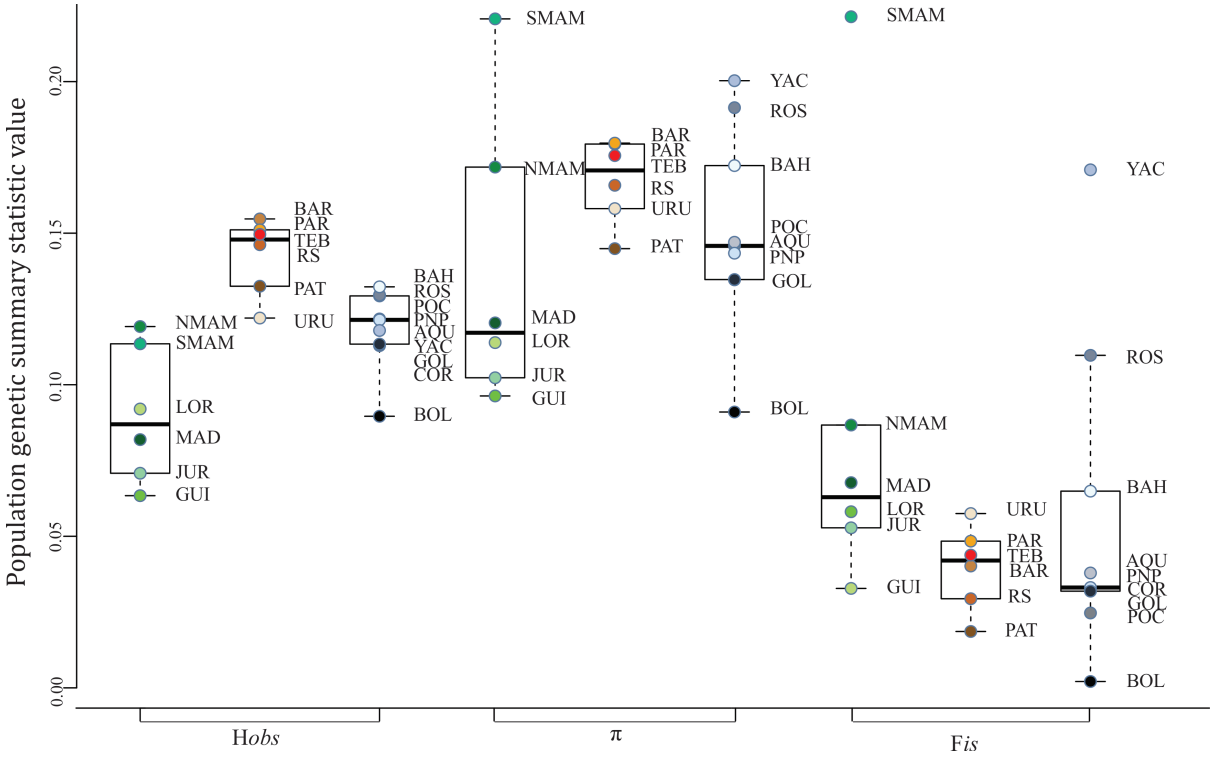


Figure H.1. Comparison of genetic diversity per sampled population across each species, with *H. sciureus, H. vulpinus* and *H. chacarius* shown in the box plots (calculated with *boxplot* function in R) for each summary statistic, respectively.

Table H.2. Results of Mantel and partial Mantel tests (with geography and ENM) between pairwise F_{ST} values and resistance matrices. Four resistance matrices are tested: the first is calculated from geography only, the second from current habitat suitability score calculated based on the ENMs, the third from Holocene habitat suitability and the last from the LGM suitability. Correlation coefficients (r) and p-values are presented; values in bold represent significant p-values.

	Mantel test		**Mantel test Pop Excluded**			**Partial Mantel test**		
	r	p-value	Pop	r	p-value		r	p-value
H. sciureus	0.352	0.141	GUI	0.137	0.466	Current	-0.3788	0.933
			JUR	0.482	0.116			
			LOR	0.447	0.116	Holocene	-0.3788	0.933
			MAD	0.317	0.233			
			N_MAM	0.207	0.433	LGM	-0.3788	0.93333
			S_MAM	0.786	**0.016**			
H. vulpinus	0.165	0.344	BAR	0.164	0.358	Current	-0.3412	0.730
			PAT	0.196	0.325			
			URU	-0.230	0.683	Holocene	-0.4495	0.9
			PAR	0.338	0.3			
			RS	0.013	0.433	LGM	0.09446	0.666
			TEB	0.329	0.291			
H. chacarius	0.0054	0.520	AQU	0.035	0.452	Current	-0.1849	0.598
			BAH	-0.003	0.511			
			COR	0.098	0.33			
			BOL	-0.185	0.774	Holocene	-0.01844	0.36
			GOL	-0.199	0.778			
			YAC	0.295	0.124			
			PNP	0.040	0.43	LGM	-0.01844	0.352
			POC	0.073	0.416			
			ROS	-0.054	0.546			

Table H.3. Tests of association between genetic distances with geographic distance and/or environmental differences (PC1) among populations using distance-based redundancy analysis (dbRDA). Results are given for each geographic and environmental variable separately (marginal tests), as conditioned on the effects of geographic distance (conditional tests). F-statistics, p-values, and the percentage of variance explained by each variable are presented; values in bold represent significant p-values.

		Marginal tests			**Conditional tests**		
	Variable	*F*	p-value	% Variance	*F*	p-value	% Variance
H. sciureus	Dist	1.129	0.283	62.88			
	PC1	0.901	0.566	18.38	1.459	0.194	22.48
H. vulpinus	Dist	1.440	0.087	68.37			
	PC1	0.705	0.856	14.98	1.006	0.487	15.86
H. chacarius	Dist	1.352	0.32	80.23			
	PC1	4.245	**0.00**	37.75	0.729	0.56	8.33

Table H.4. Evaluation of the robustness of the association between genes and geography across the taxa based on (i) a comparison of the changes in the association between genes and geography (t_0) with the exclusion of individual populations, reported as the *t"*-value (following the convention of Wang et al. 2010), as well as the angle of the PCA map that optimally minimizes the sum of squared Euclidean distance between the PCA map from the SNP data and the geographical map (i.e., θ, measured in degrees of rotation), and (ii) the association between the new PCA coordinates with a population excluded and the PCA coordinates before removing any populations, *t'*. High values of *t'* reflect robustness of the PCA maps, whereas a low value identifies a population with a disproportionate effect on the relative position of populations in genomic PCA space.

Species	Population		*t"*	θ	*t'*	θ
H. sciureus	French Guiana	GUI	0.68	-57.44	0.76	-56.01
	Jurua	JUR	0.80	59.35	0.99	3.42
	Loreto	LOR	0.82	61.46	0.99	2.98
	Madeira	MAD	0.85	64.76	0.99	0.86
	North_Mamore	N_MAM	0.83	68.52	0.99	-0.69
	South_Mamore	S_MAM	0.76	-54.47	0.97	-11.88
H. vulpinus	Barra do Ribeiro	BAR	0.67	56.87	0.98	-3.29
	Lagoa dos Patos	PAT	0.71	54.37	0.99	-1.03
	North Uruguay River	URU	0.76	50.59	0.99	-0.86
	Parana River	PAR	0.80	-42.51	0.98	-3.51
	Rio Grande do Sul	RS	0.71	-55.87	0.99	0.64
	Tebicuary	TEB	0.71	69.07	0.99	2.65
H. chacarius	Aquidauana	AQU	0.52	-48.00	0.99	-0.93
	Bahia Negra	BAH	0.51	42.59	0.99	-1.85
	Corumba	COR	0.50	-78.62	0.86	-7.83
	Estancia Bolivar	BOL	0.60	48.60	0.99	-4.77
	Estancia Golondrina	GOL	0.50	44.86	0.99	1.01
	Estancia Yacare	YAC	0.48	36.10	0.99	-9.38
	Pocone	POC	0.50	74.83	0.82	74.25
	Rosario	ROS	0.50	-71.78	0.99	0.84

Table H.5. Evaluation of the general association between the 19 bioclimatic variables and the residuals of from the gene and geography Procrustes analyses of all individuals in each of the taxa. In this case, higher values of t_{0E} indicate a very strong association between the environmental variables and the residual variation not explained by the effects of geography. Values with * represent a t_{0E} with significant *p-value* based on 10,000 permutations, where the residuals from the geographical and genetic Procrustes analysis were randomly permuted across the different 19 environmental variables.

	t_{0E}
H. sciureus	0.8108*
H. vulpinus	0.4069
H. chacarius	0.4515*

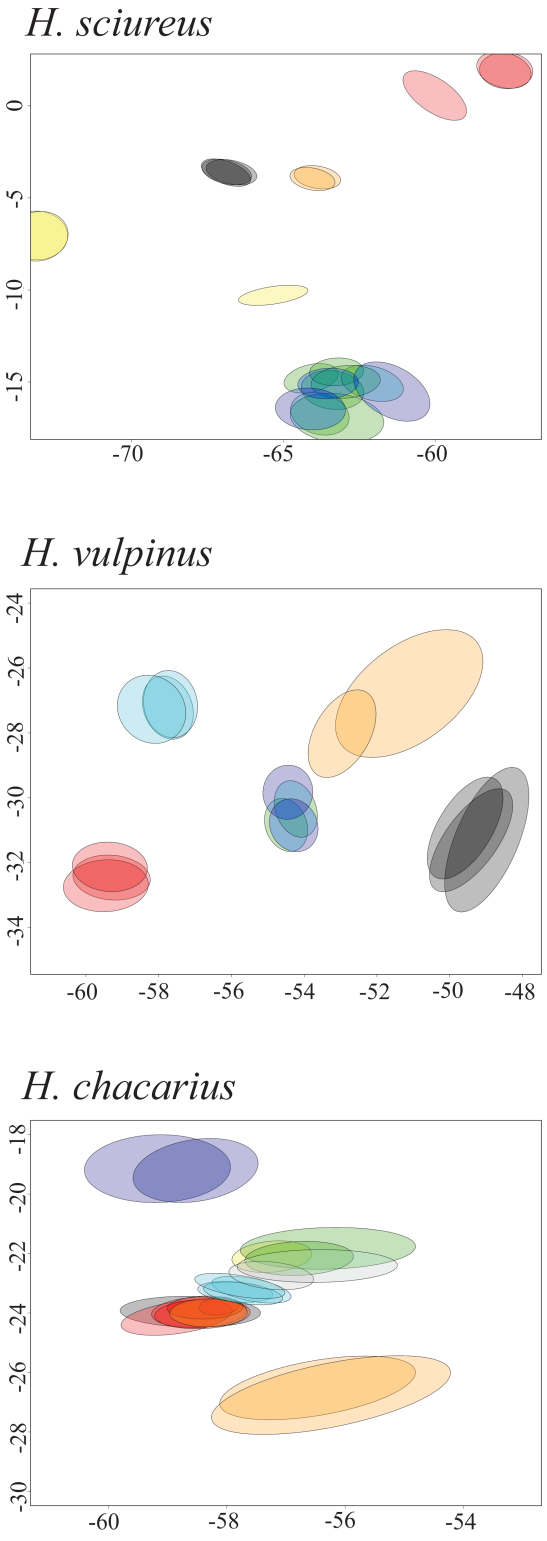


Figure H.2. Geogenetic Maps from SpaceMix analyses of *H. sciureus*, *H. vulpinus* and *H. chacarius*; analyses were conducted using geographic coordinates of populations as priors. Individuals are colour code according to their population identities (see Fig. 15).

4. COMBINING GENOMIC AND MORPHOLOGICAL DATA IN A MODEL-BASE APPROACH TO ASCERTAIN SPECIES-DELIMITATION IN THE SOUTH AMERICA MARSH RATS (*HOLOCHILUS,* SIGMODONTINAE, CRICETIDAE)

ABSTRACT

Since the description of the genus *Holochilus*, several names of the species group-taxa were attributed to it, with authors presenting distinct arrangements. This is due to the difficulty in study the pattern of morphological variation in the genus, and there is no past or current approach focused in combines the morphologic and genetic variation of the group in a broad geographic scale. Here, based on a comprehensive morphological and genomic sampling we estimated a coalescent-based phylogeny, and applied a model-based statistical framework for species delimitation that integrates phenotypic and genomic data in genus *Holochilus*. The results suggest that *Holochilus* is monophyletic and strongly supported, and reveal the potential for recognizing at least eight distinct species, with high support values, and with clear geographic structure. We discuss and highlight how RADseq data and a coalescent-based phylogeny can provide phylogenetic information for resolving relationships among species of genus *Holochilus*, supplying the resolution that was not achieved by Sanger sequencing data and other phylogenetic methods. We also demonstrate how the distributional and diversification patterns of *Holochilus* are correspondent with other vertebrates, and we present some conservations and taxonomic implications of the results provided here.

Keywords*:* RADSeq; iBPP; Species delimitation; Neotropics; Coalescence

4.1. Introduction

Since the description of the genus *Holochilus* in 1835, several names of the species group-taxa were attributed to this genus, with authors presenting distinct arrangements. This scenario lasted until Hershkovitz (1955) synonymized all described species of *Holochilus* in one single species, *H. brasiliensis*, recognizing 10 subspecies, and described another one, *H. magnus*. Later, Voss and Carleton (1993) assigned *H. magnus* under a new genus called *Lundomys*. Subsequent studies have revealed that the *brasiliensis* concept of Hershkovitz actually includes other species (Gardner and Patton, 1976; Massoia, 1980, 1981; Reig, 1986; Aguilera and Perez-Zapata, 1989; D'Elía et al., 2015; Gonçalves et al., 2015), which were validated based on limited studies in terms of geographical coverage, size of the samples, markers studied and methodological approaches, causing disagreement regarding the number of species of genus *Holochilus*. According to D'Elía et al. (2015) the disagreement is related mainly to the difficulty in study the morphological variation in the genus. However, no past or current approach evaluated the morphologic and genetic variation of the group in a broad geographic scale.

A summary of the accumulated knowledge on the genus was recently established by Gonçalves et al. (2015), who listed six living species, namely, *H. brasiliensis, H. chacarius, H. lagigliai, H. sciureus, H. venezuelae,* and *H. vulpinus*. *H. brasiliensis* occurs in southeastern and

southern Brazil, from Espírito Santo state to coastal Rio Grande do Sul state, including the eastern border of Minas Gerais, and west into the Humid Chaco Paraguayan; *H. chacarius* is distributed through the lowlands of the Paraguayan and northern Argentinean Chaco until the Brazilian boarder; *H. sciureus* occupies the lowlands of the northeastern, central and northern part of Brazil, Guyanas, Bolivia and Peru; *H. venezuelae* occurs from the lower Orinoco basin to Lago Maracaibo; *H. vulpinus* is distributed through Paraguay and east-central part of Argentina throughout Uruguay and Rio Grande do Sul state in Brazil; and *H. lagigliai* is known only from late Holocene records in the Argentinean Mendoza Province (Gonçalves et al. 2015; Fernández et al., 2017).

Regarding the phylogenetic relationships, only D'Elía et al. (2015) focused to elucidate relationships among species group taxa of *Holochilus*. Their results showed that the mitochondrial Cytb B reconstruction fails to support the monophyly of the genus. Additionally, their results also provided a larger number of polytomies in intra-specific level both in the mitochondrial and nuclear reconstructions. However, in the concatenated analysis the monophyly was recovered and the tree was resolved, revealing six species-level lineages. Their phylogeny shows two well-supported species groups, that they called the *brasiliensis* group (including *H. brasiliensis* and *H. vulpinus*) and the sciureus group (composed by *H. chacarius, H. sciureus,* and 2 currently unnamed species).

D'Elía et al (2015) results highlight the limitation in the resolution of traditional genetic markers to recover phylogenetic relationships (e.g., mtDNA or a few nuclear loci), and may also obstruct our skills to distinguish among alternative evolutionary scenarios. As a consequence, species limits can differ depending of the data and criteria used to delimit taxa (Edwards and Knowles 2014).

Along the years progress has been made in the species concept area (de Queiroz, 2011), however only recently the issue of species delimitation methods in practice coupled with theoretic philosophical concepts has growing in the literature (Camargo et al., 2012). The expansion of the use of high amount of loci has highlighting these issues, and bringing more support to decisions of species boundaries, making possible even recognize taxa under incomplete lineage sorting (Carstens and Knowles 2007). These vast amounts of genome data can also provide the high resolution for estimate fine-scale phylogeny structure (Wagner et al., 2013).

Moreover, it is been broadly discussed in the literature that a multifaceted, integrative approach to delimiting species should be conducted, a procedure that has being employed for some mammalogists since the 1980's (see comments on Chiquito et al., 2014). Although species delimitation from different data sources usually relies upon several sequential analyses of each separate data type (e.g., Chiquito et al. 2014; Prado and Percquillo, 2017), the ideal approach is to combine multiple data into a common framework to improve the accuracy of estimated species boundaries (Solis-Lemus et al., 2015).

In that sense, for the first time in the genus *Holochilus* and in the tribe Oryzomyini, we estimated a coalescent-based genomic phylogeny, and applied a model-based statistical framework for species delimitation integrating patterns of morphological quantitative divergence and genetic

distinction derived from the coalescent tree and (Solis-Lemus et al., 2015). Although quantitative methods for species delimitation are mainly applied with genetic data only, we concur that multiple data types can increase the precision of inferences (Dayrat, 2005), especially under divergence scenarios where expectations for species distinctiveness under the coalescent can be compromised (Solis-Lemus et al., 2015). Size and shape are informative markers on species delimitations (see Chapter 1), being largely employed on the systematics of sigmodontine rodents (see recent contributions by Chiquito et al., 2014; Prado et al, 2017), and their inclusion on our model based approach will fill the gap recognized by D'Elía et al. (2015), regarding the morphological variation on the genus.

Therefore, our goal is to establish hypothesis on the diversity of the genus, through an extensive sampling of previously unexplored geographical areas, incorporating genomic and morphological data in a model-base approach. Consequently, we aim to detect putative species assigning individuals to them, and estimate the species relationships and the probability that species are evolutionarily distinct in a coalescent-based framework.

4.2. Material and Methods

4.2.1. Samples and Collections

We analyzed 553 specimens of genus *Holochilus* from most of the currently known distribution of the genus (Gonçalves et al., 2015; Appendix I). For 466 specimens we studied only morphologic traits under a morphometric approach, while for 74 specimens we also generated genomic DNA data, and for 13 individuals we have both morphological and genomic data (Appendix I; Table I.1, Fig I.1). In addition to those specimens of the genus *Holochilus*, we also generated genomic data for seven species distributed in five genera, namely *Nectomys*, *Pseudoryzomys*, *Lundomys, Drymoreomys* and *Necromys*, employed as outgroup in phylogenetic analyses. We chose *Drymoreomys, Pseudoryzomys* and *Lundomys* because they belong to the same clade as *Holochilus* (sensu Weksler, 2006; Pine et al. 2012), and *Necromys* as an external group of the tribe Oryzomyini. These specimens were examined during visits to scientific collections and we present a complete list of the collections and specimens studied, as well as their collection localities in Appendix I.

4.2.2. Morphometric Data

Cranial and dental measurements were taken with a digital caliper (accurate to 0.01 millimetre) under a stereomicroscope, including: length of the upper molar series (LM); breadth of M1 (BM1); length of incisive foramen (LIF); breadth of incisive foramen (BIF); breadth of the incisor tips

(BIT); breadth of palate (BP); length of nasal (LN); breadth of nasal (BN); least interorbital breadth (LIB); breadth of braincase (BB); breadth of zygomatic plate (BZP); depth of incisor (DI); breadth of the occipital condyles (BOC); length of palatal bridge (LPB); length of interparietal (LI); breadth of interparietal (BI); lambdoidal breadth (LB), condyle-zygomatic length (CZL). Appendix J shows an illustration of all morphometric variables and its descriptions.

4.2.3. DNA extraction, amplification, and sequencing

DNA was extracted from liver, muscle or skin of each individual using the Qiagen DNeasy Blood and Tissue Kit. Four reduced representation libraries, were built using the Double Digest RADseq method (Peterson et al., 2012). DNA was double digested with the EcoR1 and MseI restriction enzymes, unique barcodes (10 bp) and Illumina adapter sequences were ligated to the digested fragments, and the fragments were amplified by PCR. Barcoded product from each individual were pooled to select fragment with size between 350 and 450 bp and the libraries were sequenced in the Illumina *HiSeq2000*, according to manufacturer's instructions to generate 150 base pairs, single-end reads. Twenty-nine *Holochilus* samples were discarded due to low sequencing coverage (<500,000 sequence reads), which reduced the number of samples to 94 (including seven outgroup samples).

4.2.4. Processing Illumina Data

Data was processed using *pyRAD* v3.0 (http://pyrad.googlecode.com) pipeline. First, sequences were identified allowing for one base mismatch in their sample-specific barcode. The restriction site and barcode were trimmed from each sequence, and bases with a FASTQ quality score below 20 were replaced with N. Sequences having more than >5 Ns were discarded. Then, for each sample, sequences were clustered using a threshold of 90%. Error-rate and heterozygosity were estimated from the base counts in each site across all clusters, and the averages were used to establish consensus sequences. Clusters with less coverage than a minimum depth of 6 were excluded in order to ensure accurate base calls. Consensus sequences from all samples were clustered by sequence similarity, with their input order randomized, using the same similarity threshold as the within-sample clustering (90%). The resulting clusters were then aligned with Muscle in the pyRAD pipeline. Any locus appearing heterozygous at the same site across more than 20 samples were discarded. The remaining clusters are treated as RAD loci, i.e., multiple alignments of putatively orthologous sequences, which were assembled into phylogenetic data matrices.

Following Huang (2016), a series of customized R script were applied to evaluate, visualize and filter even more the results. Sequences were chopped to keep the first 110bp in order to exclude

sites with high variation towards the end of the alignments. Next, for removing suspicious clusters of paralogous loci we excluded from the data set the loci with a maximum pairwise sequence divergence >15% (based on distances observed by D'Elía et al. (2015) for *Holochilus*) between taxa. Finally, all invariable loci and loci that contained samples from <4 species were also excluded. The summarized results before and after data filtering can be found in the Appendix K, Table K.1, Figs. K.1. K.2 and K.3. All customized R codes can be found in Huang (2016).

4.2.5. Species tree estimation

An R script was used to convert the edited loci file (as described in the previous topic) to a file that contained aligned unlinked SNP data from each locus. We randomly choose one parsimony-informative SNP from each locus to produce a concatenated data set (see Huang, 2016 for further details). The output was then manually edited, saved in *nexus* format and imported into PAUP* (ver. 4.0d147; Swofford, 2002).

We include a total of 94,091 loci (each SNP was considered as a independent locus) and 94 individuals for the species tree estimation using the coalescent-based program SVDQuartets (implemented in PAUP*), we chose the coalescent approach because it can produce better accuracy than traditional concatenated approaches (Heled and Drummond, 2010; Xi et al., 2014). We evaluate all possible quartets, selecting tree using the QFM quartet assembly, and we also performed bootstrapping with 1000 replicates to calculate branches support.

4.2.6. Morphological Variation

Assigning individuals to putative populations is an important step in species delimitation analysis (iBPP), which can potentially be prone to error (both in molecular and morphological approach; Solis-Lemus et al., 2015). There is no standard method for individual assignment, so in order to recover morphometric diagnosable clusters and use them in iBPP, we performed a series of Discriminant Function Analysis (DFA) in the entire sample available. Note that the purpose of these analyzes is only to assign morphometric cluster variation to discrete samples, that eventually will be hypothesized as putative taxa for subsequent analysis, and not finely describe and discuss aspects of morphological geographic variation within the genus. For this set of analyzes only specimens considered adults (see Chapter 1 of this document) were used and the sample was first assessed for univariate normality, according to the Kolmogorov-Smirnov test (KS) (Sokal and Rohlf, 1997). The outliers individuals were excluded from the final dataset and the variables were transformed to Logarithmic (base 10).

The first step was to group our total sample in fifty-four small clusters (Appendix L, Table L.1, column called cluster 1) to permit feasible comparisons, following the geographic proximity criterion (Musser, 1968; Vanzolini and Williams, 1970), and compare them throughout a first round of DFAs (Appendix L, Fig L.1). After the first round of DFAs, we identified fifteen different morphometric larger clusters (Appendix L, Table L.1, column called cluster 2) and these cluster were compared again throughout a second round of DFAs (Appendix L, Fig L.2 and L.3). The second round of DFAs identified eleven morphometric clusters that were used in the species delimitation process (Appendix L, Table L.1, column called Final cluster). All DFAs were performed with geography as the criterion of grouping variable; at his point we did not employ information from phylogenetic inference to avoid bias on our approaches.

The DFAs were performed with the 18 morphometric variables described above, except by the DFA that included the individual from Mendoza region, that was performed with 16 variables (excluding LN, BN and BR). We decided to include this cluster, even with only one individual and fewer morphometric variables available, because this region is known as the endemic distribution of the current recognized species *H. lagigliai* (Pardiñas et al., 2013).

4.2.7. Species delimitation

A multilocus, coalescent species delimitation analysis was conducted using the program iBPP (Solis-Lemus et al., 2015). This method uses a user-specified guide tree, integrating genes and traits. The groups recovered on morphological variation analysis together with the clades recovered by the coalescent species tree were considered to establish the putative species to be tested in the multi-coalescent species delimitation program iBPP (Solís-Lemus et al., 2015), and the species tree topology was used as the guide tree.

Note that the individual from Mendoza did not enter in these analyses because we do not have DNA data for it. The program was used to measure the posterior probability for the presence of independent evolving lineages within *Holochilus*, using both morphological and molecular data, separately and in combination, to explore both the phylogenetic signal of each and the concordance between them. iBPP is an expansion of the original program BPP that incorporates models of evolution for continuous quantitative traits under a Brownian motion process (Solis-Lemus et al. 2015). Both programs (BPP and iBPP) assume a known assignment of individuals to putative taxa and a guided tree of these lineages. The species' hypotheses are generated by collapsing one or more nodes in the guide tree (Rannala and Yang 2013; Solis-Lemus et al. 2015).

We tested ours datasets separately and in combination. We included a prior distribution of θ and τ = G(1, 1000). We used the auto finetune parameter. The MCMC chains were run for 50,000 generations, with a burnin of 10,000 generations and parameters sampled every five generations.

We first tested one model with eight independent lineages recovered in SVDQuartets analysis with high bootstrap support (considering samples from French Guiana as part of the clade D), and then we performed another set of species delimitation analysis considering the Guyana cluster as a independent lineage, and as a sister group of clade D (see results for futher explanations).

4.3. Results

4.3.1. Summary of RADseq data

A total of 390,318,064 raw reads were obtained with the Illumina sequencing and from this set of reads 345,929,421 passed the initial quality control (Appendix K, Table K.1). After filtering and clustering the reads were reduced to an average of 89,965 clusters with coverage greater than six, yielding a mean coverage depth of 30.4. Sequence clustering, alignment across samples and editing using customized R scripts (Appendix K, Figs K.1 and S3.2) provided a final data set containing 94,091 variable loci. The average number of loci per sample after filtering was 45,630, however the proportion of shared loci varied among taxa (Appendix K, Fig K.3).

4.3.2. Species Tree

The coalescent species tree showed eight strongly supported clades (Fig. 21) with two main divisions (A, B) and (C, D, E, F, G, and H), and the monophyly of *Holochilus* was highly supported. The first main division separates individuals from the south part of Brazil, Uruguay and northeastern Argentina (clade B) from the individuals from the southeast part of Brazil (clade A). Between the second main divisions a first split is observed separating specimens from the central part of Brazil (clade H) from the others, followed by a second split putting the clade G (specimens from Venezuela) apart from the others. Then, individuals from Paraguay and northern Argentina (cluster C) separated, followed by individuals from clade F (northeastern part of Brazil). Finally specimens from clades D and E (Amazonia and Bolivia, respectively) appear as two sister lineages. We also can observe that specimens from French Guiana (Kaw, Sinnamary and Macouria) are localized in two different small clades with great support, although one specimen from Loreto (Peru) share the same clade with Sinnamary and Macouria specimens. The sister relationship between all the eight lineages was highly supported by bootstrapping (100% bp support in all branching patterns, except by the split between C and (D, E and F).

4.3.3. Morphometric differences

The first round (Appendix L, Fig. L.1) compared samples geographically close. The Brazilian northeastern sample (yellow symbols) presented highly overlapping discriminant scores between all individuals, except by the individual from Juazeiro, Bahia. The DFA comparing individuals from Bahia and Minas Gerais (dark blue symbols) states also showed great overlapping of scores. The DFA comparing specimens only from the central part of Brazil (light blue symbols) demonstrated that these individuals are also similar regarding to the DFA scores. The analysis comparing individuals from the southeastern part of Brazil (olive symbols) exhibited great distinction between the specimens from the São Paulo and Rio de Janeiro from the others. The discriminant comparing the individuals from the southern region (brown symbols) reveled a degree of structure between the samples, but not enough for us to consider them as different groups. The analysis between the specimens from Paraguay and northern Argentina (black, gray, pink and dark red symbols) exhibited high overlap of the scores. The DFA comparing specimens from the southwestern Amazon biome showed distinction between specimens from Bolivia, Rio Peruate and the others samples (dark green, light green and green symbols, respectively). The analysis between the Venezuela and Colombia specimens (orange symbols) reveled great similarity between samples. Finally, the DFA comparing individuals from the northern part of Brazil and French Guyana (red and purple symbols, respectively) showed the separation of these two groups. According to these results we grouped the samples in larger clusters and compare them again through other set of DFA (Appendix L, L.2 and L.3). The DFA I presented great distinction between individuals from Paraguay/Argentina (black symbols), Belem (red), São Paulo (salmon) and Central (light blue) clusters; nevertheless, the sample from Silva Jardim, Rio de Janeiro (that belongs to São Paulo cluster) appeared apart from other São Paulo specimens, being more similar to specimens from Central cluster. DFA II showed great distinction between all clusters, but the Guyana cluster (purple) presented intermediate scores values in relation with the Bolivia and Amazon cluster (dark green and light green symbols); samples from Colombia and Venezuela (orange) appear as a distinct cloud of scores, along the second function. DFA III reveals the distinction between all morphometric clusters and, again, the similarity between the sample from Rio de Janeiro (that belongs to São Paulo cluster) and individuals from Minas Gerais cluster; samples from southern and southeastern regions (olive an brown) are distinct related to other samples and slightly distinct among themselves. The DFA IV shows a partial overlap between individuals from Paraguay, Amazon, Peruate and Bolivia clusters, but this last sample is slightly distinct. DFA V shows great distinction between all the clusters from northeastern Brazil (yellow, red, light blue and dark blue) and the individual from Juazeiro appears more similar to individuals from Minas Gerais cluster. DFA VI, comparing south central samples, showed the overlap of individuals from São Paulo (salmon) and Central (light blue) cluster, and a partial distinction between all the others clusters (black, olive and brown symbols, representing Paraguay, Southeast and South groups).

DFA VII, with samples from northern and southwestern Amazon basin, revealed discrimination among Guyana (purple), Amazon (dark green) and Peruate (gray) samples. Comparison DFA VIII shows the distinction between all clusters from the southern range of the genus (Paraguay, South and Mendoza). Finally, DFA IX shows again the similarity between the specimen from Silva Jardim, Rio de Janeiro sample and Minas Gerais cluster, and the noteworthy size distinction among so close geographic samples.

The rounds of Discriminant Function analyses described above (Appendix L, Fig. L.1, L.2 and L.3) allowed us to interpret the variation on the multivariate space in 11 morphometric discrete clusters (Fig 22). The first one is formed by samples from Venezuela and Colombia region (*Venezuela* cluster; orange). The second clusters individuals from western Amazon region (*Amazon* cluster; light green). The third groups specimens from Bolivia (*Bolivia* cluster; dark green). The fourth cluster comprehends individuals from Chaco and Pantanal region (*Paraguay* cluster; black). The fifth cluster group specimens from the northern part of Brazil, Guyana and French Guiana (*Guyana* cluster; purple). The sixth group includes samples from the northeastern part of Brazil (*Northeast* cluster; yellow). The seventh group clusters specimens from the southern part of Brazil and east of Argentina (*South* cluster; brown). The eighth comprehends the sample from Mendoza (*Mendoza* cluster; pink). The last three clusters were the most difficult to delimit. The cluster formed by individuals from Bahia, Minas Gerais, Rio de Janeiro and Juazeiro sample in Brazil was called *Minas Gerais* (dark blue). Initially, the individual from Juazeiro, Bahia (Appendix L, Fig. L.1) was compared with other samples from the Northeast cluster, and the individual from Rio de Janeiro was compared with samples of the São Paulo cluster (due to the geographical proximity). However, these specimens appeared grouped with individuals from Minas Gerais cluster in all previous DFAs that we performed (see Appendix L, Figs. L.2 and L.3). The *Central* (red) cluster comprehends specimens from Pará, Maranhão, Tocantins, Goiás and São Paulo states in Brazil. The sample from São Paulo locality (Appendix L, Fig. L.1; salmon dot on Figure 22), although geographically distant from the other samples of this cluster, presented great size and shape similarity with samples from Pará, Maranhão, and other samples from the Central group (Tocantins, Mato Grosso, Goiás and Lagoa da Confusão) and great morphometric differences when compared with other samples from São Paulo state; therefore, we pooled this sample along with the Central cluster. The *Southeast* (olive) cluster groups samples from the eastern part of São Paulo state, and it is morphometric very similar with samples from the South (brown) cluster, appearing in all DFAs partially overlapped with individuals of the South cluster, but as there is a sampling gap between these two groups, we decided to treat the Southeast sample as a distinct morphometric cluster. The last cluster is the one formed by the individual from Mendoza, and was called *Mendoza* cluster.

Subsequently, the DFA comparing the 11 morphometric clusters (Fig. 23) reveals a great overlap between all clusters; only individuals from the *South* and *Southeast* clusters appear slightly apart in the multivariate space, along the first DF. The first DFA counts for 41.98% of the variation,

and the most important variables in the first DFA is Least Interorbital Breadth (LIB), positively, and Breadth of Palate (BP), negatively. The second DFA is responsible for 24.12% of the variation, and the variables Length of Palatal Bridge (LPB), positively, and Breadth of the Occipital Condyles (BOC), negatively, are the most important to explain the variation.

The DFA with reduced number of variables in order to include the specimen from *Mendoza* cluster (Fig. 24) shows the same pattern of variation and the sample from Mendoza is morphometric similar to individuals from *South* and *Southeast* clusters. The first DFA counts for 44.43% of the variation, and the second DFA is responsible for 19.97% of the variation. The most important variables to explain the variation are the same as the previous DFA.

4.3.4. Species delimitation

For the species delimitation analyzes we first tested the independent lineages recovered by the species tree, plus the morphometric data identified during the DFAs as follow: species A in the iBPP analyzes comprehended individuals from *Minas Gerais* morphometric cluster and specimens from clade A of the species tree; species B included individuals from *South* morphometric cluster and individuals from clade B of the species tree; species C included samples from *Paraguay* morphometric cluster and individuals from clade C of the species tree; species D contained specimens from *Amazon* and *Guyana* morphometric clusters and specimens from clade D of the species tree; species E comprised the individuals from *Bolivia* morphometric cluster and specimens from clade E of the species tree; clade F comprised samples from the *Northeast* morphometric cluster and individual from clade F of the species tree; clade G comprehended samples from *Venezuela* morphometric cluster and individual from clade G of the species tree; and clade H encompassed individuals from *Central* morphometric cluster, and samples from clade H of the species tree.

Specimens from the Southeast cluster were not used during the iBPP runs to avoid error in the lineages assignment. These samples are not too apart from the specimens of the *South* cluster in the multivariate space, but they are geographically far away from each other, and unfortunately we do not have genomic data from samples from the eastern portion of São Paulo state to compare the datasets.

This first model tested in iBPP showed that irrespective of whether genomic data alone, morphometric data alone, or genomic and morphometric data combined are analyzed, the analyses are in agreement with eight distinct lineages, with all the lineages presenting Posterior Probability 1 (Fig. 25).

Then we tested a second model, with individuals from Guiana region as an independent species. The motivation for this analyzes is because the results from genomic and morphometric data are not in agreement regarding to the *Guyana* cluster; in multivariate space we can observe the

distinctiveness of this samples, but the species tree shows individuals from this region as part of the clade D (although they also form two French Guiana clades, highly supported, inside the clade D with another individual from Peru, see details above).

The second model tested showed that morphometric data alone support the specimens from Guiana region as a separated lineage, but genomic only or genomic and morphometric data combined collapsed individuals from Guiana region and samples from clade D into one species (Fig. 25).

4.4. Discussion

Our results arouse from an integrative approach to objectively identify evolutionary independent lineages by analyzing different data types under a common statistical framework (phenotypic data and molecular data). These results also show that RADseq data can provide phylogenetic information for resolving relationships among species of genus *Holochilus*, providing the resolution that was not achieved by Sanger sequencing data. Our results not only recovered more highly supported independent lineages (that were corroborated by morphological analyses), but also presented resolution among tips.

4.4.1. Species Delimitation

The guide tree plays an important role in the result of the species delimitation model (Leaché and Fugita, 2010). Any number of species tree inference could be used in the Bayesian species delimitation analyzes (Leaché and Fugita, 2010), here we used as guide tree the topology generated in SVDQuartets (Fig. 25). Additionally, we tested a second topology (Fig 25), based on morphological differences presented by specimens from the Guyana region, which could represent a different independent lineage. Because speciation is a transitional process influenced by both time and effective population size (Avise and Ball, 1990; Maddison, 1997; Weisrock et al., 2010) we decided to test this probability. Sometimes, for recently diverged sets of populations, reciprocal monophyly cannot be evident, and characteristic topological patterns can help to identify independently evolving lineages (Knowles and Carstens, 2007).

The results suggest that the genus *Holochilus* is monophyletic and strongly supported, and reveal the potential for recognizing at least eight distinct species with high support values in every iBPP run, and with clear geographic structure (Fig. 21 and 26). The first node of the phylogeny proposed here splits the genus in two main lineages, one distributed throughout the Uruguayan and part of Brazilian and Argentinean coast (coastal clade), while the other main cluster occupies the inland South America and the northeastern portion of Brazilian coast (inland clade). This study is not the first to suggest this arrangement, as originally proposed by Massoia (1981) and Voss and Carleton

(1993) with the recognition of groups *brasiliensis* and *sciureus* based on morphologic traits, and later recovered by the phylogenetic arrangement proposed by D'Elía et al. (2015). The geographical distribution and composition of the main clades recovered by D'Elía et al. (2015) are partially correspondent with the results provided in this study. Samples from the lineage that they called *brasiliensis* group are correspondent with samples from our clades A and B (our coastal clade). However they recalled a sample from Entre Rios (Argentina) as a member of the *sciureus* group (note that specimens considered by them in *sciureus* group, here we consider belonging to the inland clade), but our analysis recovered samples from the same region as a member of clade B (coastal clade). Further studies, gathering precisely morphological descriptions of samples that belong to these two main lineages, would be able to attest if the two clusters detected in the present study refer to the distinct morphological groups (*brasiliensis* and *sciureus* groups) described in Massoia (1981) and Voss and Carleton (1993).

The topological pattern recovered within the continental main cluster suggest that the diversification process started with the separation of the individuals from the central part of Brazil from the other regions; then, specimens from the Colombia and Venezuela diverged, followed by a southern diversification of specimens from the Chaco and Pantanal region. The clade composed by individuals from Amazon region is the most derived in *Holochilus* and present a sister relationship with individuals from the northeastern part of Brazil.

Previous studies (D'Elía et al., 2015: Fig. 4: pp. 9), based on smaller sample sizes and fewer loci, have proposed partially congruent phylogenetic arrangement in the concatenated analysis, however when considering the individuals gene tree the topologies are completely different and without resolution. The main differences between the species delimitation and its boundaries considered here in comparison with D'Elía et al., (2015) is the fact that for them samples from Santa Cruz (Bolivia) and Entre Rios (Argentina) are part of the species called by them as *H. sciureus*. In their concept *H. sciureus* comprehends all the specimens localized in Amazon biome, including individuals from Amazonas state in Brazil, Suriname, Peru, Bolivia and one individual from Entre Rios (Argentina). Our analyzes show that samples from that region are actually split in three different species. Individuals from Entre Rios (Argentina) belong to species B, individuals from the northern part of Amazon biome belong to species D, and specimens from Bolivia belong to species E. Additionally, D'Elía et al., (2015) also consider the sample from Colombia as a different new species, called by them as *H.* sp. 2, however here (although we do not present genomic data for specimens from Colombia, only from Venezuela), due to the morphometric similarity and geographic proximity of samples from Colombia and Venezuela we decide to assign specimens from both areas under the species G. Despite these differences, the fact that both types of data (Sanger and SNPs) and methods of inference (Bayesian analysis and SVDQuartets) recovered similar recently derived clades (although Sanger data sets presented lower support between clades), support the robustness of these relationships and suggest that both types of data are informative at this phylogenetic depth. However, it is important

to notice that D'Elía et al (2015: 8) stated that their decisions on species limits (and their geographic distributions; D'Elía et al., 2015: 3, fig. 1) were based on the cyt b tree, and when compared with analysis presented here, our species delimitations are quite distinct.

The incorporation of multiple data types (phenotypic and genetic) to delimit species boundaries also give us more confidence in the independent lineages designated in this study. Although genetic data alone would have been enough to support the species delimited here (Fig. 25), integrative species delimitation require the assignment of specimens taking in account ecological, behavioral, morphological and/or geographic patterns (even though here we only used genomic, morphological and geographic patterns; Edwards and Knowles, 2014) to putative taxa, and this initial assignment when based on genetic data could lead to mistakes that will compound downstream analyses (Olave et al., 2014). Moreover, using data from all possible axes along which speciation can proceed, species arising from different processes of speciation will be better adressed (Solis-Lemus et al., 2015).

Finally, in this study we present data for the recognition of eight species; however we do not include in the species delimitation analyzes individuals from Mendoza region, that nowadays correspond to a valid species within the genus, we only present morphometric data for one sample. As our morphometric analyses, even with fewer variables did show great differences between the holotype and its geographic neighbor's species (species B and C of this study), and even without including this sample in the iBPP models we recognized this sample also as a independent lineage in this study.

4.4.2. Biogeography and Dispersal corridors

In general the distribution of *Holochilus* species is congruent with the distribution of the major terrestrial South America biomes/ecoregions (see Morrone, 2010), being the allopatric speciation with niche conservatism one of the possible explanation to lead the differentiation in this group. This statement is supported by the fact that *Holochilus* species even when distributed through such distinct environments are always associated with of wet and open areas. However, the distribution of some species surpass the geographic limits of some biomes/ecoregions, as Amazon and Atlantic Forest, where we can observe more than one species.

Briefly, species A is distributed throughout the Brazilian Atlantic Forest, Cerrado and Caatinga biomes, extending from the southeastern Brazil until the right bank of Rio São Francisco (the sample from Juazeiro, Bahia). Samples from species B are distributed throughout the Atlantic Forest and Pampa biomes, including eastern Paraguay, central-eastern Argentina, Uruguay, and the south part of Brazil. The species C is distributed by the Chaco/Pantanal biomes, including the northern/northwestern part of Argentina, Paraguay and Mato Grosso do Sul state in Brazil. Species D

encompasses the species localized in the northern, western and southwestern part of Amazon biome, limited by the Rio Madeira. Species E refers to the species distributed by the southern part of Amazon biome, on the headwaters of rives Madeira-Mamoré-Guaporé. Species F encompasses the specimens localized in the northeastern part of Brazil, on the left bank of Rio São Francisco. Species G is defined here as the species distributed throughout the northern part of South America, in the open areas of Venezuelan and Colombian Llanos and coastal forests. Species H occupies the central part of Brazil and it is associated with the Cerrado biome and transitional areas between Cerrado and Amazon and Atlantic forests.

Some biogeographic pattern already recovered in the literature can be detected in the phylogeny proposed here, such as the existence of distinct lineages in Atlantic Forest. Here we detected three species occurring in this biome (A, B and F). Two of them (A and B) are sister taxa, and the break between them correspond to the phylogeographic breaks found in others mammals, for example *Euryoryzomys russatus* (Miranda et al., 2007), and *Desmodus rotundus* (Martins et al., 2007). Although there are several examples of the biodiversity breaks in Atlantic Forest, few studies have addressed its evolutionary origins (Costa and Leite, 2012). For long time studies has been justifying these breaks on the base of the refuges theory or taking the rivers barriers as the main diversification driver (Leite et al., 2016). However, nowadays recent studies (Thomaz et al., 2015; Leite et al., 2016) have been advocating that the emergence of the continental shelf during the LGM, must have been crucial in the biogeographic history of the Atlantic Forest, even more than rivers and refuges.

The third species occurring in this biome (species F) is in fact more closely related with the species that inhabits the Amazon biome than with the other species from Atlantic Forest. Despite the sister-group relationship between northern and southern Atlantic Forest clades be frequently found in some groups, this close relationship between taxa from Amazon and from northeastern Brazil is not uncommon (see Vivo, 1996), and has been detected since Costa (2003) in a paper about phylogeographic patterns of small mammals. Costa (2003) observed that some taxa from northern Atlantic Forest are more closely related to taxa distributed in Amazon region than to southern Atlantic Forest individuals. Batalha-Filho et al., (2013), also detected with birds the same pattern, proposing two distinct spatiotemporal pathways connecting the Atlantic and the Amazonian forests in the past. The first pathway refers to old connections between the two biomes during the Miocene along the southern margin of the modern Cerrado in Mato Grosso, or on the transition towards the Chaco and savannas of Bolivia and Paraguay. The second pathway proposed by Batalha-Filho et al., (2013), occurred through the Cerrado and Caatinga in northeastern Brazil during the Pliocene to Pleistocene. In this paper we do not present dates for the diversification process in *Holochilus*, but other studies (Machado et al., 2013; Parada et al., 2013; Chapter 4 of this document) suggested that the crown age of the genus possibly date somewhere between Late Pliocene and Early Pleistocene, which would make possible the use of the second pathway proposed by Batalha-Filho et al., (2013) as a dispersal route for individuals from Amazon biome and the northeastern part of Brazil.

Data from Brennand (2015) with the genus *Hylaeamys*, based on a genomic approach, also recovered a close sister relationship between samples from the northeastern Atlantic Forest, north to the São Francisco river, *H. oniscus*, to a species from the western Bolivian Amazonia, *H. acritus*. Interestingly, the other species of the genus *Hylaeamys* that inhabits the southern portion of the Atlantic Forest, *H. seuanezi*, also is sister to one species from western Bolivian, Brazilian, Peruvian, Ecuadorean and Colombian Amazonia, *H. perenensis*. This suggests multiples events of speciation between different portions of the Atlantic Forest and Amazon.

There is also a fourth species that occurs in the eastern portion of South America, marginally to the Atlantic Forest, species H. The distribution of this species ranges from the eastern Amazonia, throughout the Central Brazil, reaching the highlands of São Paulo, at the border of Serra do Mar. This distribution penetrates between the ranges of the sister species of the coastal clade (species A and B). This is a unusual distribution for oryzomyine rodents, although there is some similarity with the distribution of *Hylaeamys megacephalus* (see Percequillo, 2015) on the southern bank of Amazonas river. Species H is the basal lineage on the inland clade, sister to the ancestral of all other species: this suggests that the ancestral diversification of the genus was on the southeastern South America, with subsequent diversification towards the west, to the inland.

Another important biogeographic pattern found in our phylogeny that is often cited in the literature (see Wallace, 1852; Cracraft and Prum, 1988; Patton et al., 1997, 2000; Costa, 2003) is the division of Amazon biome into divergent phylogeographic areas. The diversity pattern found in *Holochilus* is congruent with the divergence between the southern and northern part of Amazon biome, that correspond to the SW-SE and NW-NE biogeographic regions presented in others studies as Wallace (1852), Cracraft and Prum (1988), Patton et al. (1997, 2000) and Costa (2003). Again several hypotheses have been raised to explain the differentiation among Amazon lineages, however most of these hypotheses can be referred in three main categories (Bonvicino and Weksler, 2012): climatic oscillations (such as refuges), geomorphic factors (such as paleofeatures, rivers, mountains), and ecological gradients.

Our knowledge about the biogeographic process leading the diversification in the Neotropics is still in its infancy, and a comprehensive biogeographic framework based on robust phylogenetic hypothesis, together with the assessment of historical demographic shifts based on coalescent methods will allow the identification of congruent patterns among organism, and as consequence a better understanding of the biogeographic history of South America continent.

4.4.3. Conservation and Taxonomic Implications

Hidden diversity in Neotropical fauna has been documented by several molecular studies (Patton et al. 2000), especially after the advance of statistical methods to recognize such lineages, but

have rarely been applied to genomic-scale datasets (Domingos et al., 2017). Our survey is one of the first attempt to integrate genomic and morphometric data in the study of rodent diversity in the Neotropical region, increasing the knowledge on the most diverse order in mammals. We revealed significant increase in the diversity of this South American marsh rats, bringing important implications for the conservation status in this group: formerly, it was recognized 6 species (Gonçalves et al., 2015; D'Elía et al., 2015) and we presently hypothesizes for the existence of 8 species (plus the species from Mendoza that we did not test); despite the difference on the number of species, our models redefined the boundaries of these species.

This study applied coalescent methods of species delimitation using a powerful integrative dataset of genomic and morphometric data. Other patterns of morphological variation, as binary, meristic, and geometric morphometric variables within *Holochilus* remain to be investigated and the patterns recovered here can change. The same is true for assessing the relative roles of geography and environment in the generation of diversity in the genus. The identification of the present evolutionary lineages can allow future species descriptions and the formulation of conservation strategies. For now, even though not all the species recovered here are already formally described, geographic information about the delimited lineages is already available and can be used for management purposes and conservation planning.

For example, the IUCN red list only recognizes three species within the genus *Holochilus*, namely *H. sciureus, H. chacarius*, and *H. brasiliensis*. *H. sciureus* is current listed as "Least concern", although their concept of *H. sciureus* is now split in five specific lineages (Fig. 26, lineages D, E, F, G and H) plus the individuals from Guyana region, that although were not recognized as a different species, present a highly morphological structure pattern. The recognition of these multiple species restricts the distribution and, consequently decreases the number of populations and individuals, resulting in species more rare than previously supposed. Also *H. sciureus* that was widespread through different biomes and regions became diferente species endemic to different biomes. For example, species D, E and G endemic to the Amazon biome, species F endemic of the Caatinga and northeastern Atlantic Forest biomes, and species H most distributed in the Cerrado biome. The same is true for *H. brasiliensis* that now comprehends three lineages, species A and B and part of species H. Species A is now distributed in Atlantic Forest and Cerrado biomes, species B occurs in Pampas and Atlantic forest biomes, and samples from species H occurs in part of Atlantic Forest biome.

Formal taxonomic descriptions with the correct assignment of names will take in consideration the occurrence of type specimens available and provide a completely diagnose for each species. For now we recognize nine species within *Holochilus*, for six of them there are specific names available in the literature: *Holochilus brasiliensis* (Desmarest, 1819) for lineage A; *Holochilus vulpinus* (Brants, 1827) for lineage B; *Holochilus chacarius* Thomas, 1906 for lineage C; *Holochilus sciureus* Wagner, 1842 for lineage D; *Holochilus venezuelae* Allen, 1904 for lineage G; and *Holochilus lagigliai* Pardiñas et al., 2013 for specimens of Mendoza region. Beside of this names we

call here the other species as: *Holochilus* sp bolivia for lineage E; *Holochilus* sp northeast for lineage F; and *Holochilus* sp central for lineage H.

We are confidente about the lineages diagnosed here, and for the nomenclatural balance we decided to consider the names as it is described above, but the assignment of the names problably will change when data from the type localities is crossed with the lineages geographic distribution. This is true specially in the case of the name *H. sciureus*, which is here applied to individuals from the Amazon biome, when the type locality is localized somewhere in the São Francisco river. Only a detailed description and evalution of the type specimens of genus *Holochilus* (which was not our goal in this study) will provide the correct allocation of these names.

Acknowledgements

This work was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP; Proc: 2014/22444-0, 2012/24099-3, 2009/16009-1). We are also grateful to the curators and researchers that generously granted us tissues samples for the next generation sequencing, and permitted the access to the mammal collections to the examination of specimens. We also would like to thanks Andrea Thomaz, Luciana Resende, Anna Papadopoulou, Melisa Olave, and Jen-Pan Huang for all the help with libraries preparation and bioinformatics processing of genomic data, and Edson F. de Abreu-Junior for reading the first draft of this manuscript.

References

Aguilera, M., Perez-Zapata, A., 1989. Cariologia de *Holochilus venezuelae* (Rodentia:Cricetidae). Acta Cie Venez. 40,198–207.

Allen, J. A., 1904. New mammals from Venezuela and Colombia. Bull. Am. Mus. Nat. Hist. 20, 327–335.

Avise, J.C., Ball, Jr., R.M., 1990. Principles of genealogical concordance in species concepts and biological taxonomy. Oxf Surv Evol Biol. 7, 45–67.

Batalha-Filho, H., Fjeldsa, J., Fabre, P-H., Miyaki, C.Y., 2013. Connections between the Atlantic and the Amazonian forest avifaunas represent distinct historical events. J Ornithol. 154, 41–50.

Bonvicino, C.R., Weksler, M., 2012. Speciation in Amazonia: patterns and predictions of a network of hypotheses, in: Patterson, B.D., Costa, E.P. (Eds.), Bones Clones and Biomes. University of Chicago Press, Chicago, pp. 283–306.

Brants, A., 1827. Het Geslacht der Muizen door Linnaeus opgesteld, Volgens de Tegenswoordige toestand der Wetenschap in Familien, Geslachten en Soorten verdeeld. Gedrukt ter Akademische Boekdrukkery, Berlin.

Brennand, P.G.G., 2015. The genus *Hylaeamys* Weksler, Percequillo and Voss, 2006: species definition and phylogenys of the forest clade of Oryzomyini tribe. PhD Thesis, Universidade de São Paulo.

Camargo, A., Sites Jr., J., 2013. Species Delimitation: A Decade After the Renaissance, in: Pavlinov, Y.I. (Eds.), The Species Problem - Ongoing Issues. INTECH Open Access Publisher, pp. 225–247.

Cartens, B.C., Knowles, L.L., 2007. Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from Melanoplus grasshoppers. Syst Biol. 56, 400–411.

Chiquito, E.A., D'Elía, G., Percequillo, A.R., 2014. Taxonomic review of genus *Sooretamys* Weksler, Percequillo & Voss (Rodentia: Cricetidae: Sigmodontinae): an integrative approach. Zool. J. Linn. Soc. 94, 365 – 378.

Costa, L.P., 2003. The historical bridge between the Amazon and the Atlantic forest of Brazil: a study of molecular phylogeography with small mammals. J Biogeogr. 30, 71–86.

Costa, L.P., Leite, Y.L.R., Fonseca, G.A.B., Fonseca, M.T., 2000. Biogeography of South American forest mammals: endemism and diversity in the Atlantic Forest. Biotropica. 32, 872–881.

Cracraft, J., Prum, R.O., 1988. Patterns and processes of diversification: speciation and historical congruence in some Neotropical birds. Evolution. 42, 603–620.

D'Elía, G., Hanson, J.D., Mauldin, M.R., Teta, P., Pardiñas, U.F.J., 2015. Molecular systematics of South American marsh rats of the genus *Holochilus* (Muroidea, Cricetidae, Sigmodontinae). J Mammal. 43, 1–14.

Dayrat, B., 2005. Towards integrative taxonomy. Biol J Linn Soc. 85, 407–415.

de Queiroz, K., 2011. Branches in the lines of descent: Charles Darwin and the evolution of the species concept. Biol J Linn Soc. 103, 19–35.

Desmarest, A.G., 1819. Le Rat du Brésil, *Mus brasiliensis*, Geoffr, in: Desmarest, A.G., (Eds.), Nouveau dictionnaire d'histoire naturelle, appliquèe aux art, principalement à l'agriculture et à l'economie rurale et domestique, à la medicine, etc. Par une société de naturalists et d'agriculteurs, tome XXIV. Nouv. éd. Paris: Deterville, pp. 62.

Domingos, F.M.C.B., Colli. G.R., Lemmon, A., Lemmon M.E., Beheregaray, L.B., 2017. In the shadows: Phylogenomics and coalescent species delimitation unveil cryptic diversity in a Cerrado endemic lizard (Squamata: *Tropidurus*). Mol. Phylogenet. Evol. 107, 455–465.

Edwards, D.L., Knowles, L.L., 2014. Species detection and individual assignment in species delimitation: can integrative data increase efficacy? Proc R Soc Lond B Biol Sci. 281, 1–7.

Fernández, F.J., Torres, J., Tammone, M., López, J.M., Pardiñas, U.F.J., 2017. New data on the endemic cricetid rodent *Holochilus lagigliai* from central-western Argentina: fossil record and potential distribution. Mammalia, in press.

Gardner, A.L., Patton J.L., 1976. Karyotypic variation in oryzomyine rodents (Cricetinae) with comments on chromosomal evolution in the Neotropical Cricetinae complex. Occas. pap. Mus. Zool. Univ. Mich. 49, 1–48.

Gonçalves, P.R., Teta, P., Bonvicino, C., 2015. Genus *Holochilus* Brandt, 1835, in: Patton J.L, Pardiñas U.F.J., D'Elía G. (Eds.), Mammals of South America, Vol. 2, Rodents. University of Chicago Press, Chicago, pp. 325–335.

Heled, J., Drummond, A.J., 2010. Bayesian inference of species trees from multilocus data. Mol. Biol. Evol. 27, 570–580.

Hershkovitz, P., 1955, South American marsh rats genus *Holochilus*, with a summary of sigmodont rodents. Fieldiana, Zoology. 37, 639–673.

Huang, J-P., 2016. Parapatric genetic introgression and phenotypic assimilation: testing conditions for introgression between Hercules beetles (Dynastes, Dynastinae). Mol Ecol. 25, 5513–5526.

Knowles, L.L., Carstens, B.C., 2007. Delimiting species without monophyletic gene trees. Syst. Biol. 56, 887–895.

Leaché, A.D., Fugita, M.K., 2010. Bayesian species delimitation in West African forest geckos (Hemidactylus fasciatus). Proceedings of the Royal Society B: Biological Sciences. 277, 3071– 3077.

Leite, Y.L.R., Costa, L.P., Loss, A.C., Rocha, R.G., Batalha-Filho, H., Bastos, A.C., Quaresma, V.S., Fagundes, V., Paresque, R., Passamani, M., Pardini, R., 2016. Neotropical forest expansion during the last glacial period challenges refuge hypothesis. Proceedings of the National Academy of Sciences. 113, 1008–1013.

Machado, L.F., Leite, Y.L.R., Christoff, A.U., Giugliano, L.G., 2013. Phylogeny and biogeography of tetralophodont rodents of the tribe Oryzomyini (Cricetidae: Sigmodontinae). Zool. Scr. 43, 119–130.

Maddison, W.P., 1997. Gene Trees In Species Trees. Syst. Biol. 46, 523–536.

Massoia, E., 1981. El estado sistemático y zoogeografia de *Mus brasiliensis* Desmarest y *Holochilus sciureus* Wagner (Mammalia-Rodentoa-Cricetidae). Physys. 39, 31–34.

Massoia, E., 1980. El estado sistem.tico de cuatro especies de cricetidos sudamericanos y comentarios sobre otras especies congenericas (Mammalia–Rodentia). Ameghiniana. 17, 280–287.

Miranda, G.B., de Andrades-Miranda, J., Oliveira, L.F.B., Langguth, A.R., Mattevi, M.S., 2007. Geographic patterns of genetic variation and conservation consequences in three South American rodents. Biochem. Genet. 45, 839–856.

Morrone, J.J., 2010. América do Sul e geografia da vida: Comparação de algumas propostas de regionalizaçãao, in: Carvalho, C.J.B., Almeida, E.A.B. (Eds.), Biogeografia da América do Sul: Padrões & Processos. Roca Ltda, São Paulo, pp. 14–40.

Musser, G.G., 1968. A systematic study of the Mexican and Guatemalan gray squirrel, Sciurus aureogaster, F. Cuvier (Rodentia: Sciuridae). Misc. publ. - Mus. Zool., Univ. Mich. 137, 1–112.

Olave, M., Sola, E., Knowles, L.L., 2014. Upstream analyses create problems with DNA-based Bayesian species delimitation. Syst. Biol. 63, 263–271.

Parada, A., Pardiñas, U.F.J., Salazar-Bravo, J., D'Elía, G., 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.

Pardiñas, U.F.J., Teta, P., Voglino, D., Fernández, F.J., 2013. Enlarging rodent diversity in west–central Argentina: a new species of the genus *Holochilus*. J Mammal. 94, 231–240.

Patton, J.L., Silva, M.N.F., Lara, M.C., Mustrangi, M.A., 1997. Diversity, differentiation, and the historical biogeography of nonvolant small mammals of the neotropical forests, in: Laurence, W.F., Bierrgaard Jr, R.O. (Eds.), Tropical forest remnants: ecology, management, and conservation of fragmented communities. University of Chicago Press, Chicago, pp. 455–465.

Patton, J.L., da Silva, M.N.F., Malcolm, J.R., 2000. Mammals of the Rio Juruá and the evolutionary and ecological diversification of Amazonia. Bull. Am. Mus. Nat. Hist. 244, 1–306.

Percequillo, A.R., 2015. Genus *Hylaeamys* Weksler, Percequillo and Voss, 2006, in: Patton J.L, Pardiñas U.F.J., D'Elía G. (Eds.), Mammals of South America, Vol. 2, Rodents. University of Chicago Press, Chicago, pp. 335–346.

Pine, R.H., Timm, R.M., Weksler, M. A., 2012. Newly recognized clade of trans-Andean Oryzomyini (Rodentia: Cricetidae), with description of a new genus. J Mammal. 93, 851–870.

Peterson, B. K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PloS One. 7, e37135.

Prado, J.R., Percequillo, A.R., 2013. On the geographic distribution of the genera of Tribe Oryzomyini on South America, with some comments on the patterns of diversity and Reig's Areas of Original Differentiation. Arq. Zool. São Paulo. 44, 1–120.

Prado, J.R. Percequillo, A.R., 2017. Systematic Studies of the Genus Aegialomys Weksler et al., 2006 (Rodentia: Cricetidae: Sigmodontinae): Geographic Variation, Species Delimitation, and Biogeography. J Mamm Evol., in press.

Rannala B., Yang Z., 2013. Improved reversible jump algorithms for Bayesian species delimitation. Genetics. 194, 245–25

Reig, O.A., 1986. Diversity patterns and differentiation of high Andean rodents, in: Vuilleumier, F., Monasterios, M. (Eds.), High altitude tropical biogeography. Oxford University Press, Oxford, pp. 404–439.

Sokal, R.R., Rohlf, F.J., 1997. Biometry. W.H. Freeman and Company, New York.

Solis-Lemus, C., Knowles, L.L., Ane, C., 2015. Bayesian species delimitation combining multiple genes and traits in a unified framework. Evolution. 69, 492–507.

Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Ver. 4. Sinauer, Sunderland, Massachusetts.

Thomas, O., 1906. Notes on South-American Rodents. Ann. Mag. Nat. Hist. 18, 442–448.

Thomaz, A., Malabarba, L.R., Bonatto, S.L, Knowles, L.L., 2015. Testing the effect of palaeodrainages versus habitat stability on genetic divergence in riverine systems: Study of a Neotropical fish of the Brazilian coastal Atlantic Forest. J. Biogeogr. 42, 2389–2401.

Vanzolini, P.E., Williams, E.E., 1970. South American anoles: the geographic differentiation and evolution of the *Anolis chrysolepis* species group (Sauria, Iguanidae). Arq. Zool. São Paulo. 19, 1–298.

Vivo, M., 1996. How many species of mammals are there in Brazil? Taxonomic practice and diversity evaluation, in: Bicudo, C.E.M., Menezes, N.A. (Eds.), Biodiversity in Brazil: a first approach - Conselho Nacional de Desenvolvimento Científico e Tecnológico, pp. 313–321.

Voss, R.S., Carleton, M.D., 1993. A new genus for *Hesperomys molitor* Winge and *Holochilus magnus* Hershkovitz (Mammalia, Muridae) with an analysis of its phylogenetic relationships. Am. Mus. Novit. 3085, 1–39.

Wagner, C.E., Keller, I., Wittwer, S., Selz, O.M., Mwaiko, S., Greuter, L., Sivasundar, A., Seehausen, O., 2013. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. Mol Ecol. 22, 787–798.

Wagner, J. A., 1842. Beschreibung einiger oder minder bekannter Nager. Arch.Naturgesch. 8, 1–33.

Wallace, A.R., 1952. On the monkeys of the Amazon. Proc. Zool. Soc. Lond. 20, 107–110.

Weisrock, D. W., Rasoloarison, R. M., Fiorentino, I., Ralison, J. M., Goodman, S. M., Kappeler, P. M., Yoder, A.D., 2010. Delimiting species without nuclear monophyly in Madagascar's mouse lemurs. PLoS One. 5, e9883.

Weksler, M., 2006. Phylogenetic relationships of oryzomyine rodents (Muroidea: Sigmodontinae): separate and combined analyses of morphological and molecular data. Bull. Am. Mus. Nat. Hist, 296, 1–149.

Xi, Z., Liu, L., Rest, J.S., Davis, C.C., 2014. Coalescent versus concatenation methods and the placement of Amborella as sister to Water Lilies. Syst. Biol. 63, 919– 932.

Figures

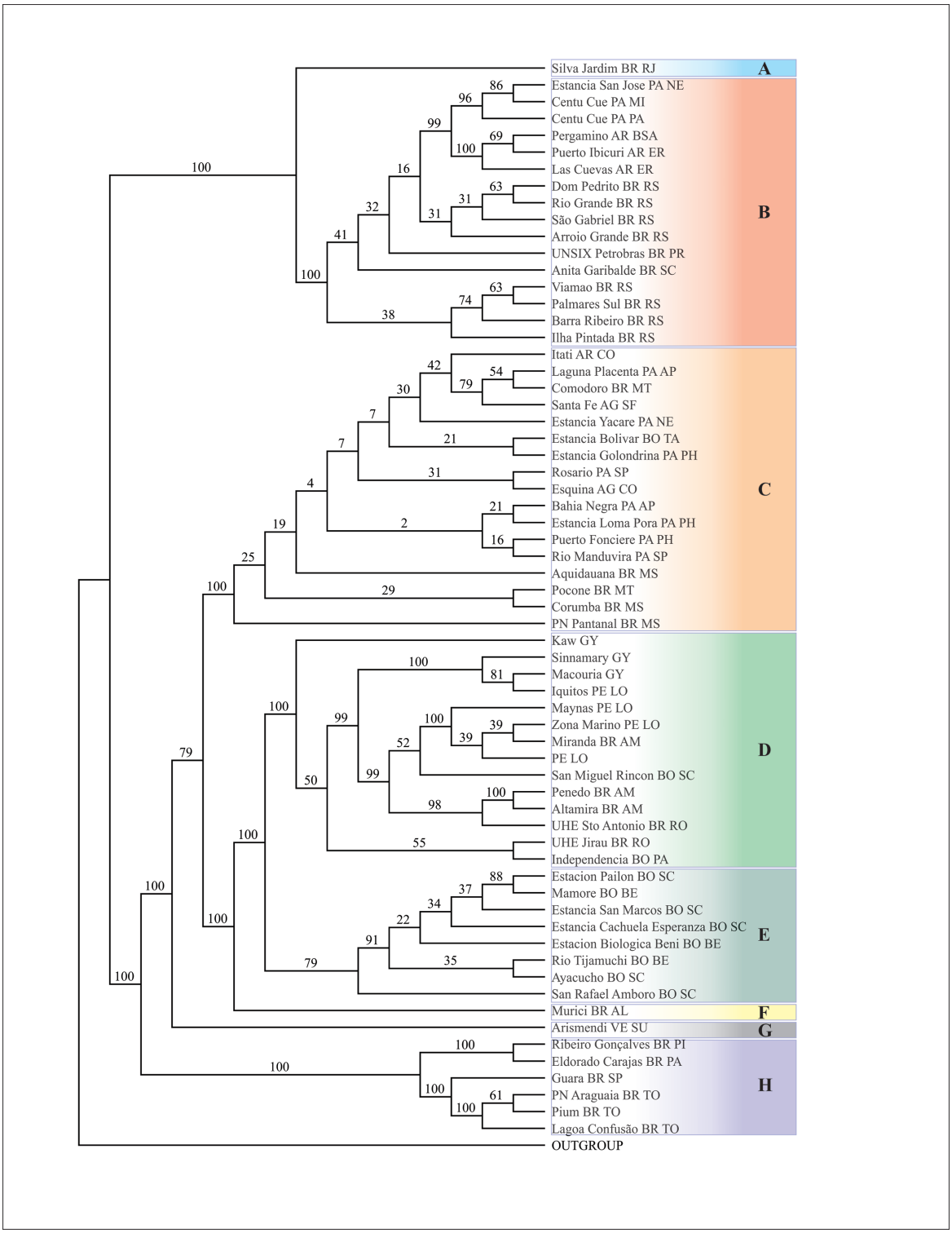


Figure 21. Species tree recovered for *Holochilus* populations as resolved by exhaustive quartet sampling in a data matrix containing 94,091 SNPs and 94 taxa using SVDquartets. Bootstrap values are shown in the nodes. Letters A, B, C, D, E, F, G and H define clades discussed in the text. Branch lengths are not meaningful.

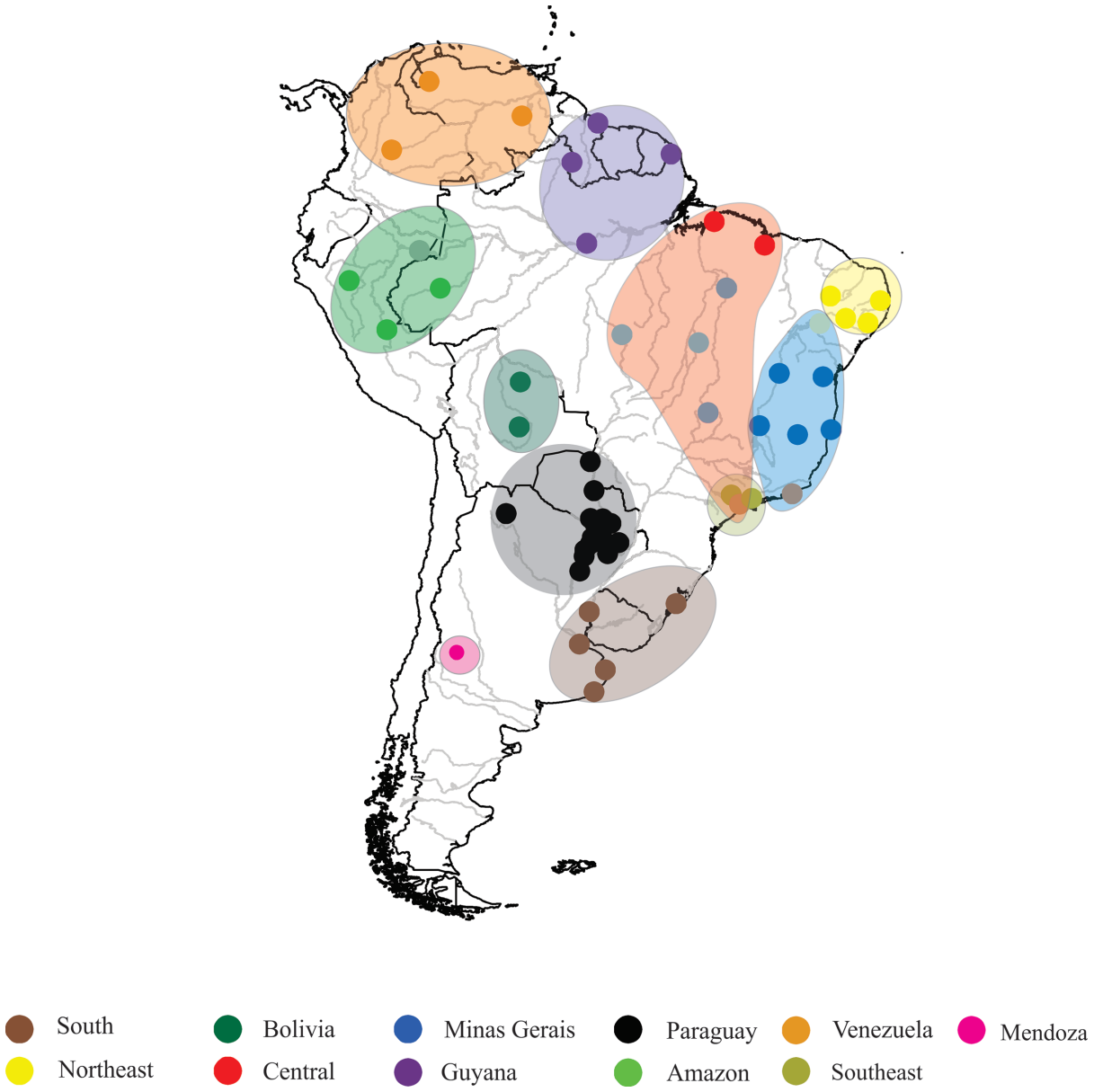


Figure 22. Map showing the geographic location of the morphometric diagnosticable clusters recovered after the rounds of Discriminant Analyses between *Holochilus* populations. Legend names in the figure define the clusters discussed in the text.

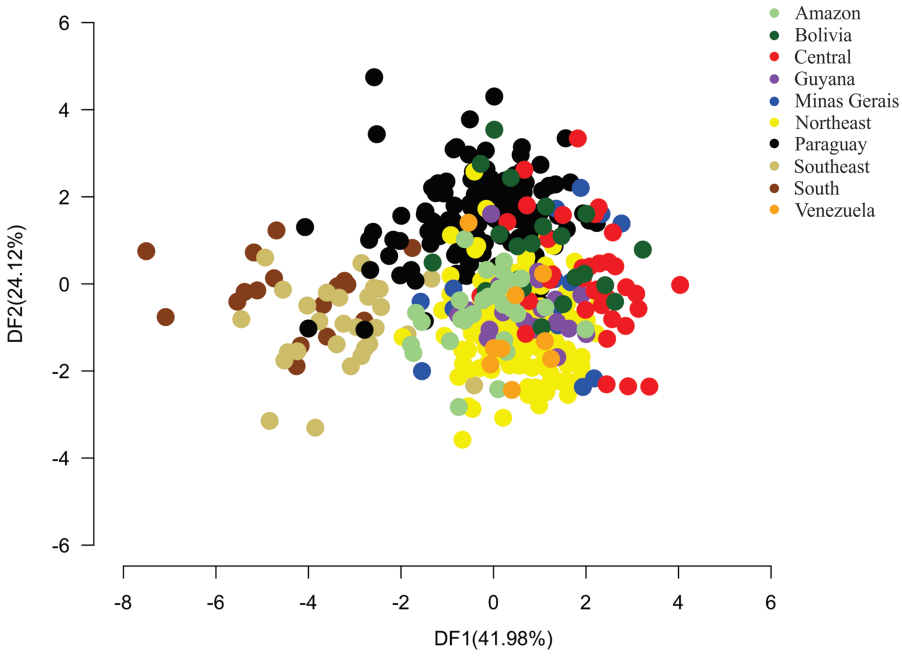


Figure 23. Scatterplot with the individual scores of the first two discriminant function (DF1 and DF2) of the Discriminant Function Analyses performed with the 10 morphometric clusters recognized in *Holochilus* samples.

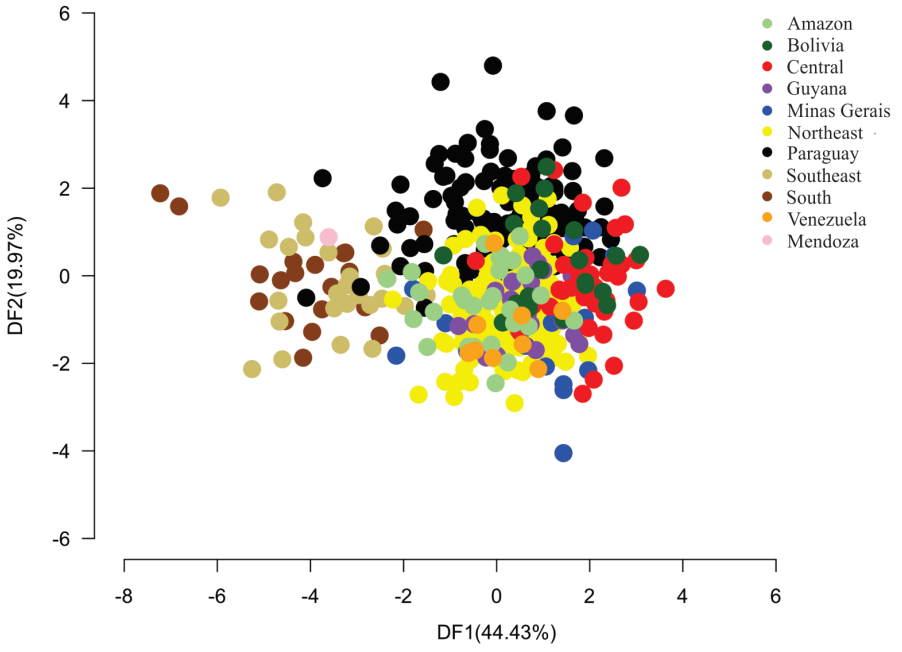


Figure 24. Scatterplot with the individual scores of the first two discriminant function (DF1 and DF2) of the Discriminant Function Analyses performed with the 10 morphometric clusters plus the individual from Mendoza recognized in *Holochilus* samples.

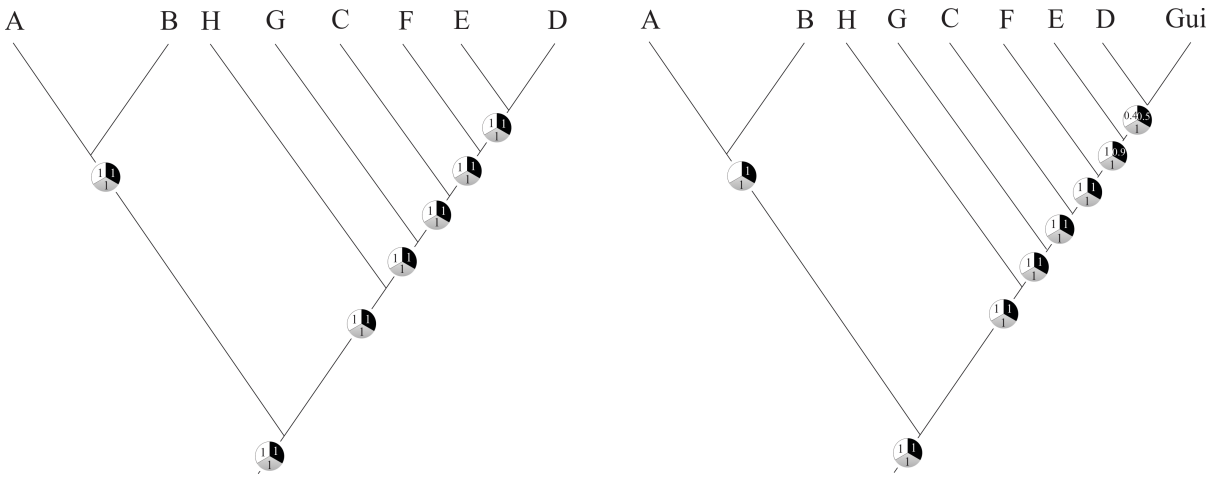


Figure 25. Phylogenetic relationship among *Holochilus* lineages showing the delimited species and their Posterior Probability estimated with iBPP on different runs. Circles with grey color represent morphometric data only; White color represent molecular data only; and black color represent morphometric and molecular data combined. Left tree represent the highly supported lineages recovered by SVDQuartets and the right tree correspond to the same topology but the individuals from Guyana region as a sister group of clade D.

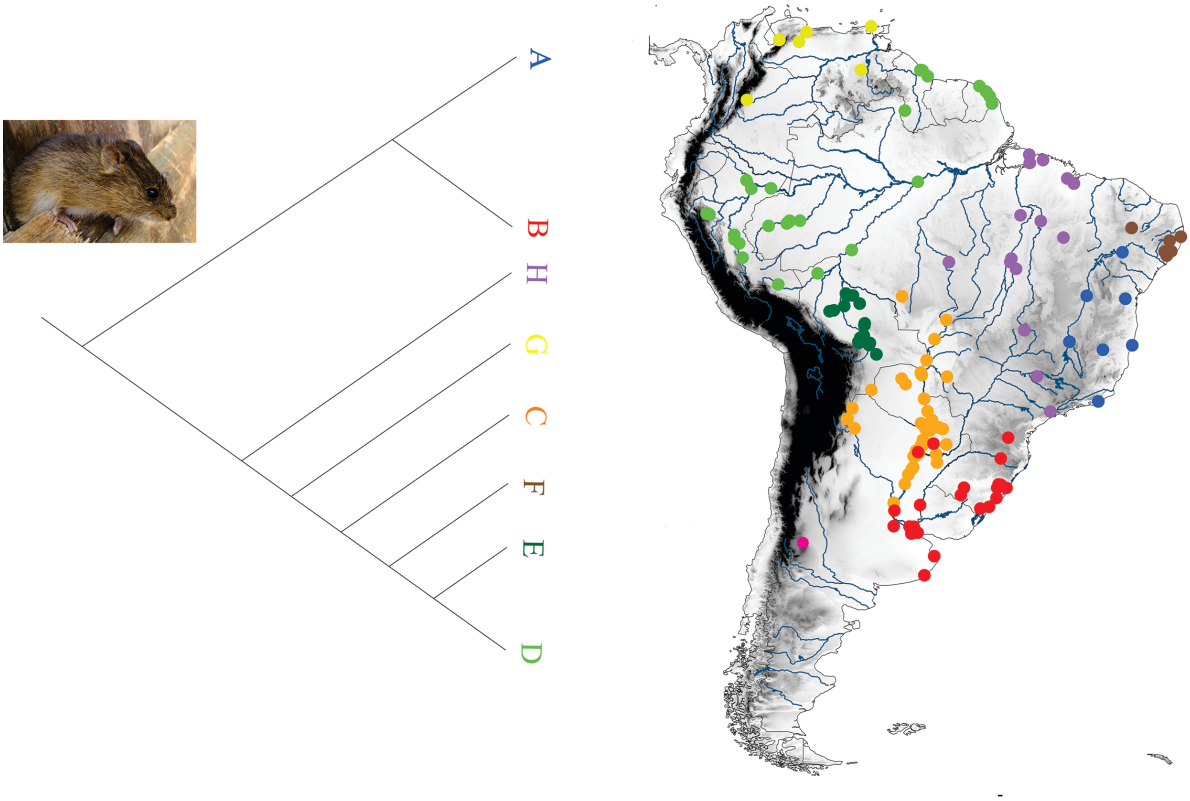


Figure 26. Geographic distribution of *Holochilus* species present in this report. *Holochilus brasiliensis* for lineage A; *Holochilus vulpinus* for lineage B; *Holochilus chacarius* for lineage C; *Holochilus sciureus* for lineage D; *Holochilus venezuelae* for lineage G; and *Holochilus lagigliai* for specimens of Mendoza region. Beside of this names we call here the other species as: *Holochilus* sp bolivia for lineage E; *Holochilus* sp northeast for lineage F; and *Holochilus* sp central for lineage H.

APPENDIX I. Summaries of samples collected

Details about the scientific collections visited or assessed throughout loans to gather morphological and molecular data of *Holochilus* specimens, as well as details about vouchers (locality and Catalog Number) of individuals sequenced (Table I.1) and morphologically analyzed, and a map showing the geographic distribution of the samples (Fig I.1) are provided in this appendix, Appendix I.

List of scientific collections assessed:

AMNH – American Museum of Natural History (New York, United States)

BMNH – The Natural History Museum (London, England)

FMNH – The Field Museum (Chicago, United States)

LSUMZ – Louisiana State University Museum of Zoology (Baton Rouge, United States)

MACN – Museo Argentino de Ciencias Naturales "Bernardino Rivadavia" (Buenos Aires, Argentina)

MCNU – Museu de Ciências Naturais da ULBRA

MHNSR – Museo de Historia Natural de San Rafael (Mendoza, Argentina)

MN – Museu Nacional da Universidade Federal do Rio de Janeiro (Rio de Janeiro, Brazil)

MNHN – Musee National d'Histoire Naturelle (Paris, French)

MNHN – Museo Nacional de Historia Natural (Montevideo, Uruguay)

MSB – Museum of Southwestern Biology (Albuquerque, United States)

MUFAL – Museu de História Natural da Universidade Federal do Alagoas (Maceió, Brazil)

MVZ – Museum of Vertebrate Zoology (Berkeley, United States)

MZUSP – Museu de Zoologia da USP (São Paulo, Brazil)

TTU – Museum of Texas Tech University (Lubbock, United States)

UFES – Mammal Collection of Universidade Federal do Espírito Santo (Vitória, Brazil)

UFMG – Mammal Collection of Universidade Federal de Minas Gerais (Belo Horizonte, Brazil)

UMMZ – University of Michigan Museum of Zoology (Ann Arbor, United States)

USNM – National Museum of Natural History (Washington DC, United States)

ZMB – Museum für Naturkunde (Berlin, Germany)

ZSM – Zoologische Staatssammlung München (Munich, Germany)

List of specimens used in the morphological analyses: **ARGENTINA**: Buenos Aires: Ajó: BMNH 1909.12.1.24, BMNH 1909.12.1.25, BMNH 1909.12.1.72a. Buenos Aires: MACN 17866. Buenos Aires, 35km N.O. de Ayacucho: MACN 15736. Delta I.N.T.A: MACN 18661, MACN 20333. Isla Ella, Delta del Parana: BMNH 1917.6.1.14, BMNH 1917.6.1.15. Otamendi, I.N.T.A Delta: MACN 20388, MACN 19335. Zelaya: MACN 5386. Chaco: Moreno Cue, Estancia Las Rosas, Depto Bermejo: MACN 22617. Rio de Oro, Depto Bermejo: MACN 14341. Corrientes: Estero Valenzuela, Depto Capital: MACN 22626, MACN 22628, MACN 22630, MACN 22637, MACN 22640, MACN 22645, MACN 22646, MACN 22648. Cañada del Pirayu, Ruta Nacional 12, Depto Capital: MACN 22627. San Cayetano, CAPRIM, Depto Capital: MACN 22634. Bella Vista, Depto Saladas: MACN 14050. Laguna Galarza, Depto Santo Tomé: MACN 22631. Goya: BMNH 1898.12.3.1, BMNH 1898.12.3.2. Loma Alta, Depto Mburucuyí: MACN 22656. Entre Rios: 6 km S (by road) of Puerto

Ibicuy: UMMZ 166268, UMMZ 166269, UMMZ 166395. Parque Nacional El Palmar: MACN 18601. Formosa: Kilometro 182, Riacho Pilaga, 10 Mi NW: USNM 236323, USNM 236324. Parque Nacional Rio Pilcamayo: MACN 20769. Jujuy: Villa Carolina, Rio Lavallen: BMNH 1920.1.7.10, BMNH 1920.1.7.16, BMNH 1920.1.7.8, BMNH 1920.1.7.12, BMNH 1920.1.7.19, BMNH 1920.1.7.21, BMNH 1920.1.7.14, BMNH 1920.1.7.9, BMNH 1920.1.7.15, BMNH 1920.1.7.18. Salta: Oran, Tabaca: MACN 17873, MACN 17877, MACN 17878, MACN 17882, MACN 17886, MACN 17889, MACN 16223, MACN 16226. **BOLIVIA:** El Beni: Itenez, Lago Victoria: USNM 390767, USNM 390768. Magdalena, La Granja: USNM 461031. Rio Tijamuchi: MSB55986. San Joaquin: FMNH 118819, USNM 461032, USNM 461034, USNM 364758, USNM 364759. Santa Cruz: 3.5 km W Estacion El Pailon: MSB55298. 6 km W (by road) of Ascension: MSB55987. 8 km SE Tita: MSB55299, MSB55300. Buenavista: BMNH 1926.1.5.12, BMNH 1926.1.5.11, BMNH 1926.12.4.50, BMNH 1928.2.9.35, BMNH 1926.12.4.51. San Rafael de Amboro: MSB55988. Uruma: MACN 13106. Warnes: USNM 390765. **BRAZIL:** Alagoas: Anadia, Engenho Ferreiros: MN 66994. Anadia, Fazenda Bananas I: MN 66979, MN 66980. Anadia, Fazenda Cajueiro: MN 66846. Anadia, Fazenda Vale Verde: MN 66981. Anadia, Sitio Igrejinha: MN 66985. Anadia, Sitio Mocós I: MN 66987, MN 66988, MN 66922. Quebrangulo, Engenho Carangueija: MN 66982, MN 66984. Quebrangulo, Engenho Gitó: MN 66919. Quebrangulo, Engenho Riachão II: MN 66881, MN 66886. Quebrangulo, Fazenda Guaribas: MN 66986. Quebrangulo, Fazenda Passagem: MN 66677, MN 66678, MN 66680. Quebrangulo, Fazenda Peri-Peri: MN 66686. Quebrangulo, Fazenda Pirauás: MN 66939, MN 66941, MN 66904, MN 66905. Quebrangulo, Fazenda Poço da Serra: MN 66691. Quebrangulo, Fazenda Santa Terezinha I: MN 66989, MN 66990, MN 66991. Quebrangulo, Povoado Dois Braços: MN 66948, MN 66964. Quebrangulo, Sitio Barro Preto: MN 17454. Quebrangulo, Sitio Boqueirão: MN 66978. Quebrangulo, Sitio Goiabeira: MN 66878, MN 66829. Quebrangulo, Sitio Gravatá: MN 66851, MN 66852, MN 66797, MN 66693. Quebrangulo, Sitio Riachão: MN 66977. Viçosa, Engenho Cachoeira A: MN 66767, MN 66768. Viçosa, Engenho Itapicurú: MN 66769. Viçosa, Engenho Retiro: MN 66771, MN 66773. Viçosa, Faz Aniceto: MN 66712, MN 66713. Viçosa, Faz Areias: MN 66862. Viçosa, Faz Cachoeira Grande: MN 66757, MN 66758. Viçosa, Fazenda Caldeirões: MN 66816, MN 66817, MN 66778. Viçosa, Fazenda Cambuim II: MN 66882, MN 66831. Viçosa, Fazenda Esperança: MN 66857. Viçosa, Fazenda Lagoa: MN 66918, MN 66843, MN 66844. Viçosa, Fazenda Paraná: MN 66895, MN 66896, MN 66901, MN 66832, MN 66833, MN 66834, MN 66836. Viçosa, Fazenda Pedra Cascuda: MN 66780, MN 66781, MN 66782. Viçosa, Fazenda Pedra de Fogo: MN 66766, MN 66759, MN 66760, MN 66764. Viçosa, Fazenda Pindobinha: MN 66734, MN 66735. Viçosa, Fazenda Primavera II: MN 66998. Viçosa, Fazenda Riachão II: MN 66740, MN 66741. Viçosa, Fazenda São José B: MN 66968, MN 66969, MN 66842, MN 66783. Viçosa, Fazenda São Pedro: MN 66738. Viçosa, Fazenda Torresópolis: MN 66937, MN 66902, MN 66934, MN 66935. Viçosa, Fazenda Areias: MN 66863. Viçosa, Sitio Alves: MN 66811. Viçosa, Sitio Bananas: MN 67006, MN 66971, MN 66972, MN 66973, MN 66974, MN 66975, MN 66976, MN 66845, MN 66727, MN 66728. Viçosa, Sitio Boa Vista IV: MN 66837, MN 66838, MN 66784, MN 66785, MN 66786. Viçosa, Sitio Caboge: MN 66942, MN 66944, MN 66945, MN 66946, MN 66950, MN 66952, MN 66953, MN 66954, MN 66955, MN 66956, MN 66958, MN 66960, MN 66962, MN 66963, MN 66966, MN 67010, MN 67011, MN 66909, MN 66910, MN 66912, MN 66787. Viçosa, Sitio Cambuim I: MN 66875, MN 66823, MN 66824. Sitio Pedra de Fogo dos Pereiras: MN 66745. Viçosa, Sitio Poço Dantas: MN 66865, MN 66869, MN 66870, MN 66858, MN 66871, MN 66874, MN 66928, MN 66819, MN 66820. Viçosa, Sitio Rio Branco: MN 66887, MN 66888. Viçosa, Sitio

Tangil II: MN 66891, MN 66719, MN 66723, MN 66724. Viçosa, Sitio Tangil: MN 66716, MN 66717, MN 66718. Viçosa, Sitio Urucuba: MN 66744. Sitio Velho: MN 66996. Viçosa, Sitio Vila Maria Leia: MN 66714. Amapá: Vila Velha do Caciporó, Oiapoque: MN 20611, MN 20612, MN 20617, MN 20637, MN 20638, MN 20640, MN 20627, MN 20628, MN 20630, MN 20632, MN 20633, MN 20634. Amazonas: Eirunepe: FMNH 136887. Igarape Grande: FMNH 136888. Ipixuna, Purus River: USNM 461416, USNM 461420. Itacoatiara, Amazon River: FMNH 20136. João Pessoa, Jurua Rive: FMNH 140815. Bahia: Bom Jesus da Lapa: MN4147, MN4148. Caravelas: MZUSP2692. Jaguaquara, Fazenda Trançadal: MN67014. Juazeiro, Fazenda do Horto: MN8339. Ceará: Crato: MN 1510, MN 1511, MN 1513, MN 1515, MN 1520, MN 1522, MN 1763, MN 7749, MN 7750, MN 1549, MN 1550, MN 7759, MN 7760. Goiás: Anápolis: MN4428, MN4361, MN4262, MN34181. Maranhão: São Bento: BMNH 1925.5.21.9, MNHN 1991/639, MNHN 1991/641, MNHN 1991/642, MNHN 1991/640, MNHN 1986/1016, MNHN 1986/1017. São Jeronimo: MSB279070, MSB279098, MSB279116, MSB279119. Mato Grosso: Ilha de Taiamã: MZUSP13463, MZUSP19537, MZUSP13462. Passo de Lontra: MZUSP 27430. Minas Gerais: Lagoa Santa: MAM 204, MN4088. Pirapora: MN4271, MN4207, MN4208, MN4205. Para: Belém: USNM545965, USNM545966, USNM545968, USNM545969, USNM545971, USNM545972, BMNH 1913.12.18.11. Belém, Utinga: USNM 394720. Belém, Utinga, Bacia do Água Preta: USNM 394726, USNM 394723, USNM 394724, USNM 394725. Peixe Boi: BMNH 1911.4.28.26. Soure, Ilha do Marajã, Foz do Rio Amazonas: BMNH 1897.4.1.2. Pernambuco: Caruaru, Serra dos Cavalos: MNHN 1967/1464. Saltinho, Rio Formoso: UFPB 2695. São Lourenço: BMNH 1903.10.1.38, BMNH 1903.10.1.37. Rio de Janeiro: Silva Jardim, Imbaú: MN61802. Rio Grande do Sul: Lagoa dos Patos: BMNH 1888.11.30.9. São Paulo: Americana, Fazenda Santo Angelo, Represa de Americana: MN 24177, MN 67106. Caçapava, Fazenda José Nane: MN 24174. Caçapava, Gleba José, Bairro Santa Luzia: MN 67108. Paulínia, Córrego Argentini: MN 67110. Paulínia, Fazenda Saltinho: MN 67109. Pindamonhangaba: MN 67111, MN 67112. Pindamonhangaba, Fazenda Boa Esperança: MN 67116. Pindamonhangaba, Fazenda Instituto Agronômico, MN 67234, MN 67235. São José dos Campos, Fazenda Honda: MN 67117. São José dos Campos, Fazenda Honda, Canal principal, left bank, Distrito Eugênio Mello: MN 67134. São Paulo: MNHN 1991/632, MNHN 1991/635, MNHN 1991/633, MNHN 1991/631, MNHN 1991/638. Taubaté, Fazenda Antonio Taino, Bairro barranco: MN 24173, MN 67131. Taubaté, Fazenda Antonio Tavares: MN 24175. Taubaté, Fazenda Kangutti, Bairro do Remédio: MN 24170. Taubaté, Gleba Paulo Japonês, Bairro barranco: MN 67124, MN 67125, MN 67126, MN 67127, MN 67128, MN 67129, MN 67132, MN 67133, MN 67136, MN 67137, MN 67139. Tocantins: Lagoa da Confusão, Fazenda Lago Verde: UFESMAM 1305, UFESMAM 1306. Palmeiras do Tocantins, Usina Hidreletrica Estreito: UFESMAM 2070. **COLOMBIA:** Meta: Villavicencio, Finca Buque: USNM 507261. **GUYANA:** Berbice: Blairmont: BMNH 1937.6.24.4. Demerara: Hyde Park: BMNH 1907.6.20.8. Nonpareil Plantation: FMNH 53925. Versailles: BMNH 1937.6.24.12. Upper Takutu-Upper Essequibo: Kanuku Uts: BMNH 1901.6.4.88, BMNH 1901.6.4.87. **FRENCH GUIANA:** Ile de Cayenne: MNHN 1983/358. **PARAGUAY:** Alto Paraguai: 4 km N (by air) of Bahia Negra, Estancia Dona Julia, W bank of Rio Paraguai: UMMZ 166255, UMMZ 166256, UMMZ 166258, UMMZ 166257, UMMZ 166382, UMMZ 166384, UMMZ 166383. Caagazu: 24 km NNW Carayao, Estancia San Ignacio: UMMZ 133970. Central: 17 km E (by road) Luque: UMMZ 124225, UMMZ 124226, UMMZ 124228, UMMZ 124227. Chaco: 50 km WNW Fortin Madrejon: UMMZ 125489, UMMZ 125493, UMMZ 125484, UMMZ 125485, UMMZ 125490, UMMZ 125491, UMMZ 125492. Cordillera: Compania Minas, Cue, Emboscada, Property of Fam Filipinni:

UMMZ 174968. Itapua: Estancia San Isidro, 5.56 Km NW of houses: UMMZ 174863. Misiones: 5 km ENE Ayolas: UMMZ 125495. Centu Cue: UMMZ 174879. Neembucu: Estancia Santa Teresa, 3.88 Km S of Puesto Anastacio: UMMZ 174842. Estancia Yacare, 1.68 Km NW of Puesto San Fernando: UMMZ 174829. Estancia Yacare, 3.59 Km NNE of Puesto San Fernando: UMMZ 174843. Margin of Rio Tebicuary, 1.2 Km upstream (E) Hotel Centu Cue (opposite margin): UMMZ 174824, UMMZ 174822. Presidente Hayes: 15.5 km NNW Chaco-I: UMMZ 125997, UMMZ 126000, UMMZ 126002, UMMZ 126004, UMMZ 126075, UMMZ 126077, UMMZ 125999, UMMZ 126001, UMMZ 126003, UMMZ 126078. 24 km NW (by air) of Villa Hayes, Estancia La Golondrina: UMMZ 166328, UMMZ 166325, UMMZ 166327, UMMZ 166330, UMMZ 166331, UMMZ 166333, UMMZ 166334, UMMZ 166335, UMMZ 165990, UMMZ 165991, UMMZ 166149, UMMZ 166150, UMMZ 166198, UMMZ 166199, UMMZ 166203, UMMZ 166234, UMMZ 166236, UMMZ 166237, UMMZ 166238, UMMZ 165992, UMMZ 165993, UMMZ 166235, UMMZ 166239. 24 km NW Villa Hayes: UMMZ 133976, UMMZ 133980, UMMZ 133971, UMMZ 133972, UMMZ 133974, UMMZ 133977, UMMZ 133979. West bank of Rio Paraguai, 4 km NW of Puerto Fonciere: UMMZ 166385, UMMZ 166389, UMMZ 166690, UMMZ 166260, UMMZ 166386, UMMZ 166387, UMMZ 166388, UMMZ 166175, UMMZ 166176. San Pedro: East bank of Rio Paraguai, 22 km (by air) SSW Rosario: UMMZ 166211, UMMZ 166210, UMMZ 166212, UMMZ 166213, UMMZ 166214, UMMZ 166215. Island in middle of Rio Paraguai, 10 km (by air) NW of Rosario: UMMZ 166247, UMMZ 166250, UMMZ 166253, UMMZ 166266, UMMZ 166689, UMMZ 166251, UMMZ 166252, UMMZ 166254, UMMZ 166249, UMMZ 166174. Island in Rio Paraguai, 10 km (by air) NW of Rosario: UMMZ 166373, UMMZ 166375, UMMZ 166377, UMMZ 166378, UMMZ 166380, UMMZ 166372, UMMZ 166374, UMMZ 166376. **PERU:** Loreto: Cumeria: BMNH 1928.5.2.170. Rio Amazonas, Boca Rio Peruate: FMNH 88916, FMNH 88919, FMNH 88915. San Jeronimo: BMNH 1928.5.2.165, BMNH 1928.5.2.163, BMNH 1928.5.2.164, BMNH 1928.5.2.166, BMNH 1928.5.2.162, BMNH 1928.5.2.169, BMNH 1928.5.2.168, BMNH 1928.5.2.167. Yarinacocha: FMNH 55472, FMNH 62093, FMNH 62089, FMNH 55476. Madre de Dios: Manu: FMNH 84297. San Martin: Moyobamba, Rio Mayo: BMNH 1924.12.12.25, BMNH 1924.12.12.26, BMNH 1924.12.12.28. Yurac Yacu: BMNH 1927.1.1.98. Ucayali: Masisea, Rio Ucayali: BMNH 1924.2.22.22, BMNH 1924.2.22.20. **VENEZUELA:** Apure: Nulita, 29 km SSW Santo Domingo, Selvas De San Camilo: USNM 442260. Bolivar: Hato San Jose, 20 km W La Paragua: USNM 406231. Carabobo: Montalban, 2 km SE Montalban, Potrerito: USNM 442443. Portuguesa: El Cruce, La Esperanza Farm: USNM 520678. Trujillo: Valera, 30 km NW Valera, Nr. El Dividive: USNM 372654, USNM 372655, USNM 372657, USNM 372664.

Table I.1. Sampled specimens used in next-generation sequencing, with species, voucher number, locality of origin and the information if the taxon was used or not in the morphological analyses.

Taxon	Voucher	Locality	Dpt/State/Prov	Country	Morpho data
Holochilus chacarius	UMMZ166514	0.5 km N of Itati, Island in Rio Parana	Corrientes	Argentina	no
Holochilus chacarius	UMMZ166518	0.5 km W of Esquina, Island in Rio Parana	Corrientes	Argentina	no
Holochilus chacarius	UMMZ166526	12 km (by road) E of Santa Fe on Island in Rio Parana	Santa Fe	Argentina	no
Holochilus chacarius	MSB235493	Estancia Bolivar	Tarija	Bolivia	no
Holochilus chacarius	CTA1539	Base de Pesquisa do Pantanal, CENAP/IBAMA, 110 km SSW Poconé.	Mato Grosso	Brazil	no
Holochilus chacarius	MVZ198015	Base de Pesquisas do Pantanal, CENAP/IBAMA, 110 Km SSW Poconé.	Mato Grosso	Brazil	no
Holochilus chacarius	MDL13	Comodoro	Mato Grosso	Brazil	no
Holochilus chacarius	LBCE5361	Fazenda Alegria, Corumbá	Mato Grosso do Sul	Brazil	no
Holochilus chacarius	LBCE5409	Fazenda Alegria, Corumbá	Mato Grosso do Sul	Brazil	no
Holochilus chacarius	LBCE4432	Fazenda Rio Negro, Aquidauana	Mato Grosso do Sul	Brazil	no
Holochilus chacarius	LBCE4433	Fazenda Rio Negro, Aquidauana	Mato Grosso do Sul	Brazil	no
Holochilus chacarius	MZ35144	Parque Nacional do Pantanal	Mato Grosso do Sul	Brazil	no
Holochilus chacarius	MZ35145	Parque Nacional do Pantanal	Mato Grosso do Sul	Brazil	no
Holochilus chacarius	UMMZ166381	17 km N (by air) of Bahia Negra, W bank Rio Negro, Estancia Immaculada Concepcion,	Alto Paraguay	Paraguay	no
Holochilus chacarius	UMMZ166255	4 km N (by air) of Bahia Negra, Estancia Dona Julia, W bank of Rio Paraguay	Alto Paraguay	Paraguay	yes
Holochilus chacarius	UMMZ166256	6 km SE (by air) of Bahia Negra, W bank of Rio Paraguay along Riacho Ramos	Alto Paraguay	Paraguay	yes
Holochilus chacarius	TTU108452	Laguna Placenta	Alto Paraguay	Paraguay	no
Holochilus chacarius	TTU108467	Estancia San José, 5 km E of house	Ñeembucú	Paraguay	no
Holochilus chacarius	TTU108438	Estancia Yacare	Ñeembucú	Paraguay	no
Holochilus chacarius	TTU108439	Estancia Yacare	Ñeembucú	Paraguay	no
Holochilus chacarius	TTU108458	Estancia Yacare	Ñeembucú	Paraguay	no
Holochilus chacarius	TTU108460	Estancia Yacare	Ñeembucú	Paraguay	no
Holochilus chacarius	TTU108462	Estancia Yacare	Ñeembucú	Paraguay	no
Holochilus chacarius	UMMZ174829	Estancia Yacare, 1.68 km NW of Puesto San Fernando	Ñeembucú	Paraguay	yes

Holochilus chacarius	UMMZ165996	24 km NW (by air) of Villa Hayes, Estancia La Golondrina	Presidente Hayes	Paraguay	no
Holochilus chacarius	UMMZ166707	24 km NW (by air) of Villa Hayes, Estancia La Golondrina	Presidente Hayes	Paraguay	no
Holochilus chacarius	TTU108463	Estancia Loma Pora	Presidente Hayes	Paraguay	no
Holochilus chacarius	UMMZ166389	West bank of Rio Paraguay, 4 km NW of Puerto Fonciere,	Presidente Hayes	Paraguay	yes
Holochilus chacarius	UMMZ166214	Edge of Rio Manduvira, 53 km (by air) NW Asuncion	San Pedro	Paraguay	yes
Holochilus chacarius	UMMZ166250	Island in middle of Rio Paraguay, 10 km (by air) Nw of Rosario	San Pedro	Paraguay	yes
Holochilus chacarius	UMMZ166391	Island in middle of Rio Paraguay, 10 km (by air) NW of Rosario	San Pedro	Paraguay	no
Holochilus chacarius	UMMZ166375	Island in Rio Paraguay, 10 km (by air) NW of Rosario	San Pedro	Paraguay	yes
Holochilus chacarius	UMMZ166409	Island in Rio Paraguay, 10 km (by air) NW Rosario	San Pedro	Paraguay	no
Holochilus sp alagoas	MUFAL0023	ESEC Murici	Alagoas	Brazil	yes
Holochilus sp amazonia	MSB236670	Independencia, 3720m	Pando	Bolivia	no
Holochilus sp amazonia	MSB210547	San Miguel Rincon, 0300m	Santa Cruz	Bolivia	no
Holochilus sp amazonia	MVZ193736	Altamira, right bank Rio Juruá	Amazonas	Brazil	no
Holochilus sp amazonia	MVZ190357	near Miranda. Left bank Rio Juruá	Amazonas	Brazil	no
Holochilus sp amazonia	MVZ193734	Penedo, right bank Rio Juruá	Amazonas	Brazil	no
Holochilus sp amazonia	MJ742	UHE Jirau, Porto Velho	Rondônia	Brazil	no
Holochilus sp amazonia	BRFMP10668	UHE Jirau, Porto Velho	Rondônia	Brazil	no
Holochilus sp amazonia	MC164	UHE Santo Antônio, Porto Velho	Rondônia	Brazil	no
Holochilus sp amazonia	SA88	UHE Santo Antônio, Porto Velho	Rondônia	Brazil	no
Holochilus sp amazonia	T4581	Kaw		French Guiana	no
Holochilus sp amazonia	T4595	Kaw		French Guiana	no
Holochilus sp amazonia	T6822	Macouria: RN1 pK-35		French Guiana	no
Holochilus sp amazonia	T6876	Sinnamary: Piste de l Anse		French Guiana	no
Holochilus sp amazonia	TTU75634	Iquitos	Loreto	Peru	no
Holochilus sp amazonia	TTU75641	No locality		Peru	no
Holochilus sp amazonia	TTU76303	No locality		Peru	no
Holochilus sp amazonia	TTU98745	Maynas	Loreto	Peru	no
Holochilus sp bolivia	MSB208404	Mamore, San Ramon	El Beni	Bolivia	no
Holochilus sp bolivia	MSB211485	Rio Matos, 6 KM E of Estacion Biologica de Beni	El Beni	Bolivia	no

Holochilus sp bolivia	MSB211440	Rio Tijamuchi, 240m	El Beni	Bolivia	no
Holochilus sp bolivia	MSB55298	3.5 KM W Estacion Pailon	Santa Cruz	Bolivia	yes
Holochilus sp bolivia	MSB236995	Ayacucho, 250m	Santa Cruz	Bolivia	no
Holochilus sp bolivia	MSB55296	Estancia Cachuela Esperanza	Santa Cruz	Bolivia	no
Holochilus sp bolivia	MSB99051	Estancia San Marcos 6 km w OS ascension	Santa Cruz	Bolivia	no
Holochilus sp bolivia	MSB55988	San Rafael de Amboro, 400m	Santa Cruz	Bolivia	yes
Holochilus sp central	MCNU4117	Eldorado dos Carajás	Para	Brazil	no
Holochilus sp central	M24	Ribeiro Gonçalves	Piauí	Brazil	no
Holochilus sp central	PCH4155	PCH Guará, Guará	São Paulo	Brazil	no
Holochilus sp central	PNA153	Parque Nacional do Araguaia	Tocantins	Brazil	no
Holochilus sp central	RGR514	Fazenda Lago Verde, Lagoa da Confusão	Tocantins	Brazil	yes
Holochilus sp central	RGR526	Fazenda Lago Verde, Lagoa da Confusão	Tocantins	Brazil	yes
Holochilus sp central	RGR528	Fazenda Lago Verde, Lagoa da Confusão	Tocantins	Brazil	no
Holochilus sp central	RGR135	Margem esquerda do Rio Javaés, Parque Estadual do Cantão, Pium	Tocantins	Brazil	no
Holochilus sp rj	LBCE1222	Silva Jardim	Rio de Janeiro	Brazil	yes
Holochilus venezuelae	AMNH257336	Arismendi, Finca Vuelta Larga, 9.7 km by road southeast of Guaraunos	Sucre	Venezuela	no
Holochilus vulpinus	MSB204443	Pergamino, FARM INT, GRID X	Buenos Aires	Argentina	no
Holochilus vulpinus	UMMZ166479	6 km S (by road) of Puerto Ibicuy	Entre Ríos	Argentina	no
Holochilus vulpinus	UMMZ166524	Las Cuevas, 35 Km (by air) SSE of Diamante in Flood Plain of Rio Parana	Entre Ríos	Argentina	no
Holochilus vulpinus	UFSC5167	UN-SIX Petrobrás	Paraná	Brasil	no
Holochilus vulpinus	MCNU1946	Arroio Grande	Rio Grande do Sul	Brazil	no
Holochilus vulpinus	MCNU3424	Barra do Ribeiro	Rio Grande do Sul	Brazil	no
Holochilus vulpinus	MCNU3425	Barra do Ribeiro	Rio Grande do Sul	Brazil	no
Holochilus vulpinus	MCNU3426	Barra do Ribeiro	Rio Grande do Sul	Brazil	no
Holochilus vulpinus	MCNU1943	Dom Pedrito	Rio Grande do Sul	Brazil	no
Holochilus vulpinus	MCNU2746	Ilha da Pintada, POA	Rio Grande do Sul	Brazil	no
Holochilus vulpinus	MCNU3428	Palmares do Sul	Rio Grande do Sul	Brazil	no
Holochilus vulpinus	MCNU3427	Rio Grande	Rio Grande do Sul	Brazil	no
Holochilus vulpinus	MCNU2342	São Gabriel	Rio Grande do Sul	Brazil	no

Holochilus vulpinus	MCNU2595	Viamão (Águas Claras)	Rio Grande do Sul	Brazil	no
Holochilus vulpinus	UFSC5074	Anita Garibaldi	Santa Catarina	Brasil	no
Holochilus vulpinus	UMMZ174820	610 m S Hotel Centu Cue	Misiones	Paraguay	no
Holochilus vulpinus	UMMZ175059	Margin of Rio Tebicuary, 1 km Upstream (E) Hotel Centu Cue	Misiones	Paraguay	no
Holochilus vulpinus	UMMZ174824	Margin of Rio Tebicuary, 1.2 km Upstream (E) Hotel Centu Cue (Opposite Margin)	Paraguarí	Paraguay	yes
Oligoryzomys nigripes	EEB714	Estação Ecológica do Bananal	São Paulo	Brazil	no
Nectomys squamipes	LGA3261	Reserva Biológica Córrego do veado, Pinheiros	Espírito Santo	Brazil	no
Lundomys molitor	MCNU2302	Dom Pedrito	Rio Grande do Sul	Brazil	no
Necromys lasiurus	MJ092	UHE JIRAU, right bank of Rio Madeira, Porto Velho	Rondônia	Brazil	no
Nectomys squamipes	TTU108465	No locality	Canindeyu	Paraguay	no
Nectomys squamipes	TTU108466	No locality	Canindeyu	Paraguay	no
Pseudoryzomys simplex	PNA022	Parque Nacional do Araguaia	Tocantins	Brazil	no

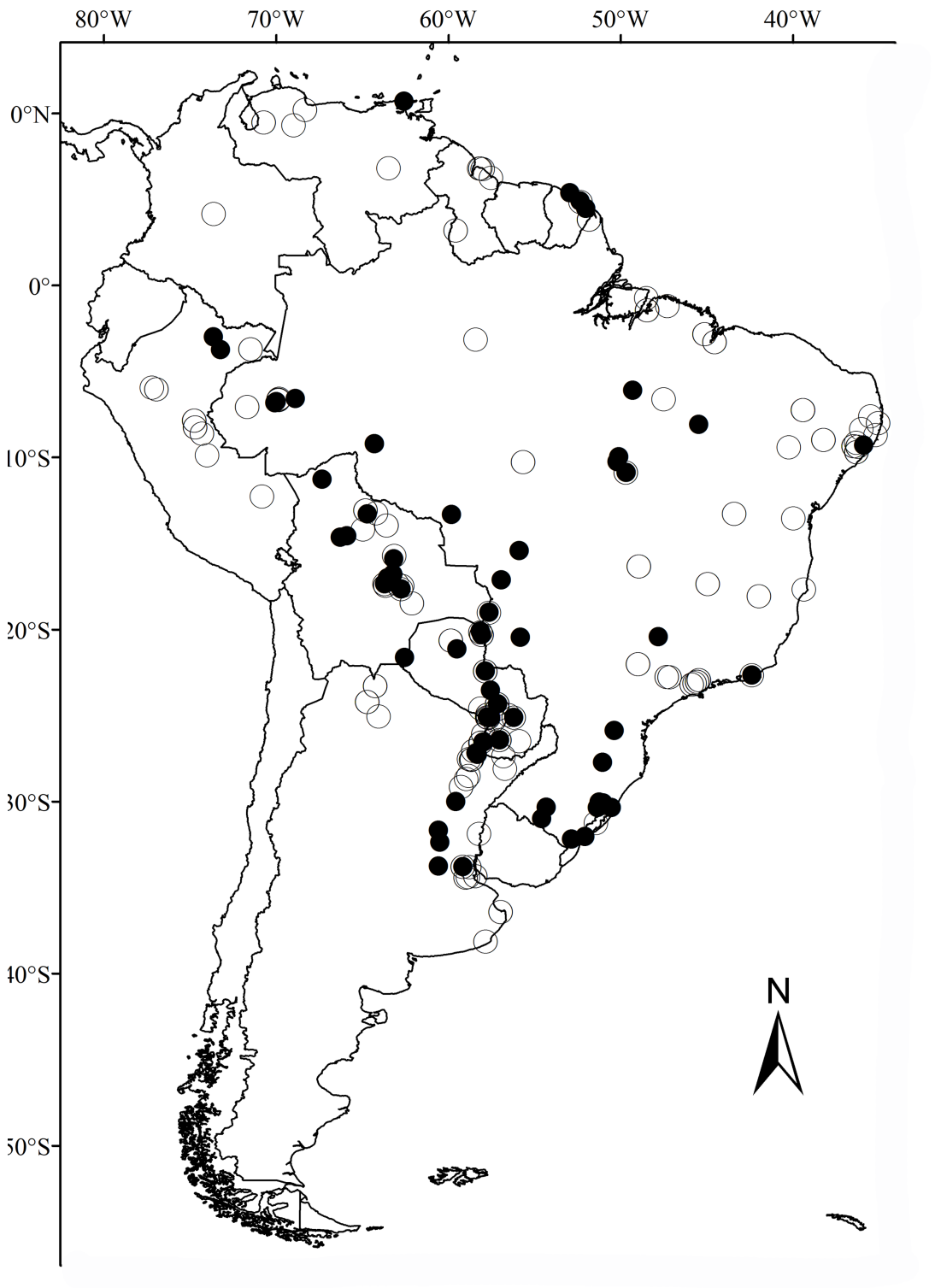


Figure I.1. Distributional records of samples analyzed in this study. Empty circles represent samples morphologically analyzed and full circles shows the samples with next generation sequencing data produced.

APPENDIX J. Morphometric Variables

Descriptions of the morphometric variables used in the present study, as well as a figure showing the variables drawn in a rodent skull are provided in this appendix, Appendix J.

Length of the upper molar series (LM): measured from the anterior surface of the first upper molar to the posterior surface of the third upper molar, at the crown of the molars;

Breadth of M1 (BM1): greatest breadth of the first upper molar measure of the base crown, the height of the protocone;

Length of incisive foramen (LIF): the greatest length measured from the anterior edge to posterior edge of incisive foramen;

Breadth of incisive foramen (BIF): the greatest internal breadth, measured on the lateral margins of the incisive foramen;

Breadth of the incisor tips (BIT): breadth measured between the side faces of the two upper incisors attached;

Breadth of palate (PB): measured in the external lateral portion of the maxillary, between the second and third molar;

Length of nasal (LN): measured from the anteriormost end of the nasal to the naso-frontal suture;

Breath of nasal (BN): breadth measured from the widemost part of the lateral of the nasal;

Least interorbital breadth (LIB): shortest distance through the frontals in the orbital fossa;

Breadth of braincase (BB): greater width of the braincase, measured posterior to the squamosal root of the zygomatic arch;

Breadth of zygomatic plate (BZP): the shortest distance between the anterior and posterior margin of the inferior zygomatic root or zygomatic plate;

Depth of incisor (DI): corresponds to the depth of the incisor measured from the anterior to the posterior face, near the base of the incisor;

Breadth of the occipital condyles (BOC): greater breadth through lateral sides of the occipital condyles;

Length of palatal bridge (LPB): measured from the posterior margin of incisive foramen to the anterior margin of mesopterygoid fossa;

Length of interparietal (LI): greatest length (anteroposterior) of interparietal bone;

Breadth of interparietal (BI): greatest breadth of interparietal bone;

Lambdoidal breadth (LB): measure between lambdoidal ridges;

Condylo-zygomatic length (CZL): shortest distance between the posteriormost point of occipital condyle and the posteriormost point of the upper edge of the zygomatic notch.

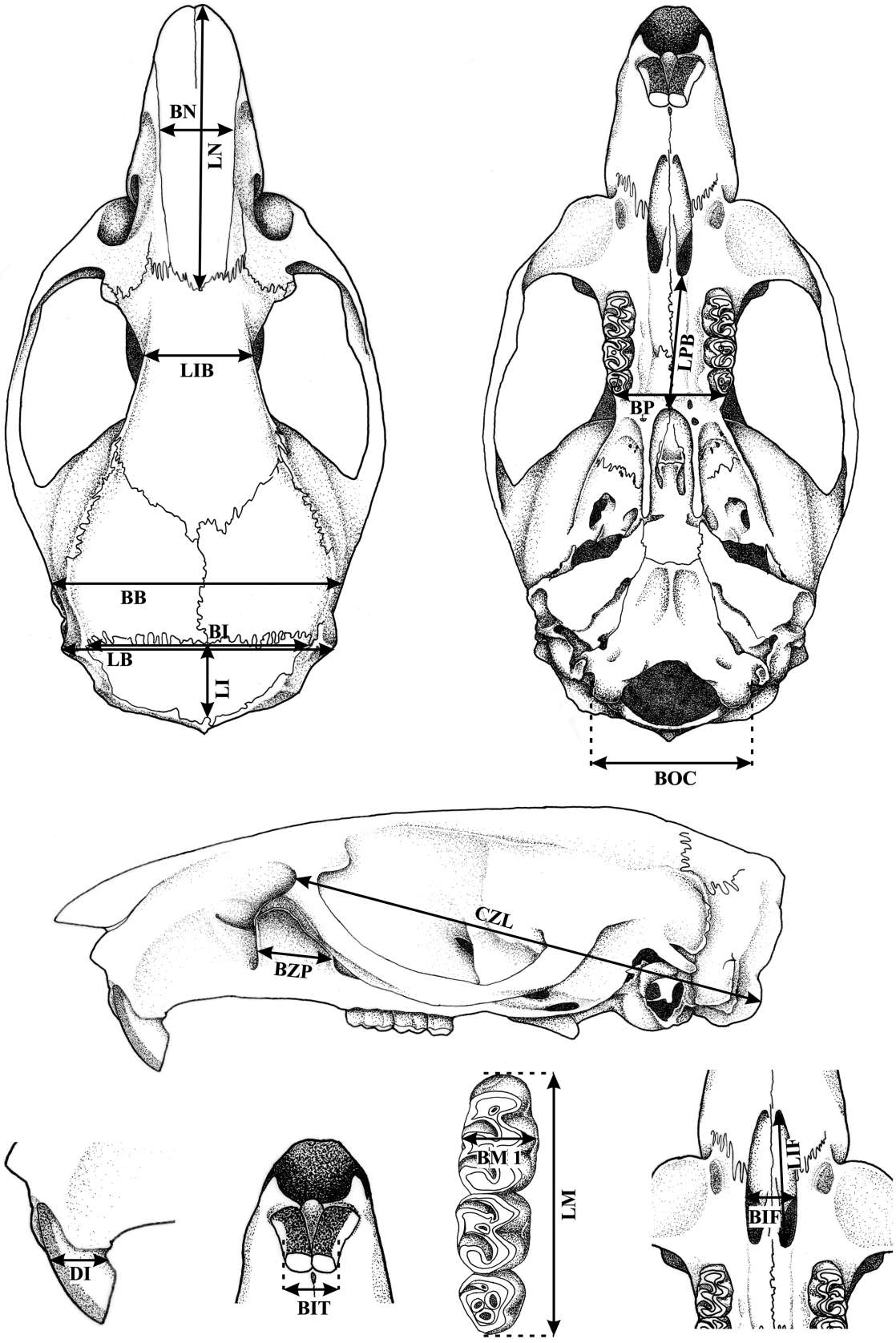


Figure J.1. Representation of the cranial-dental variables extract from the *Holochilus* specimens (illustration of a *Euryoryzomys russatus* skull made by Gustavo S. Libardi).

APPENDIX K. Summaries of genomic data

Processing information and pyRAD summary statistics for species sequenced on the Illumina platform (Table K.1), as well as the number and position of segregating sites (before and after edition; Figure K.1 and K.2), number of taxa per locus, interspecific pairwise divergence (Figure K.2), and the proportion of loci shared among individuals (Figure K. 3), are provided in this appendix, Appendix K.

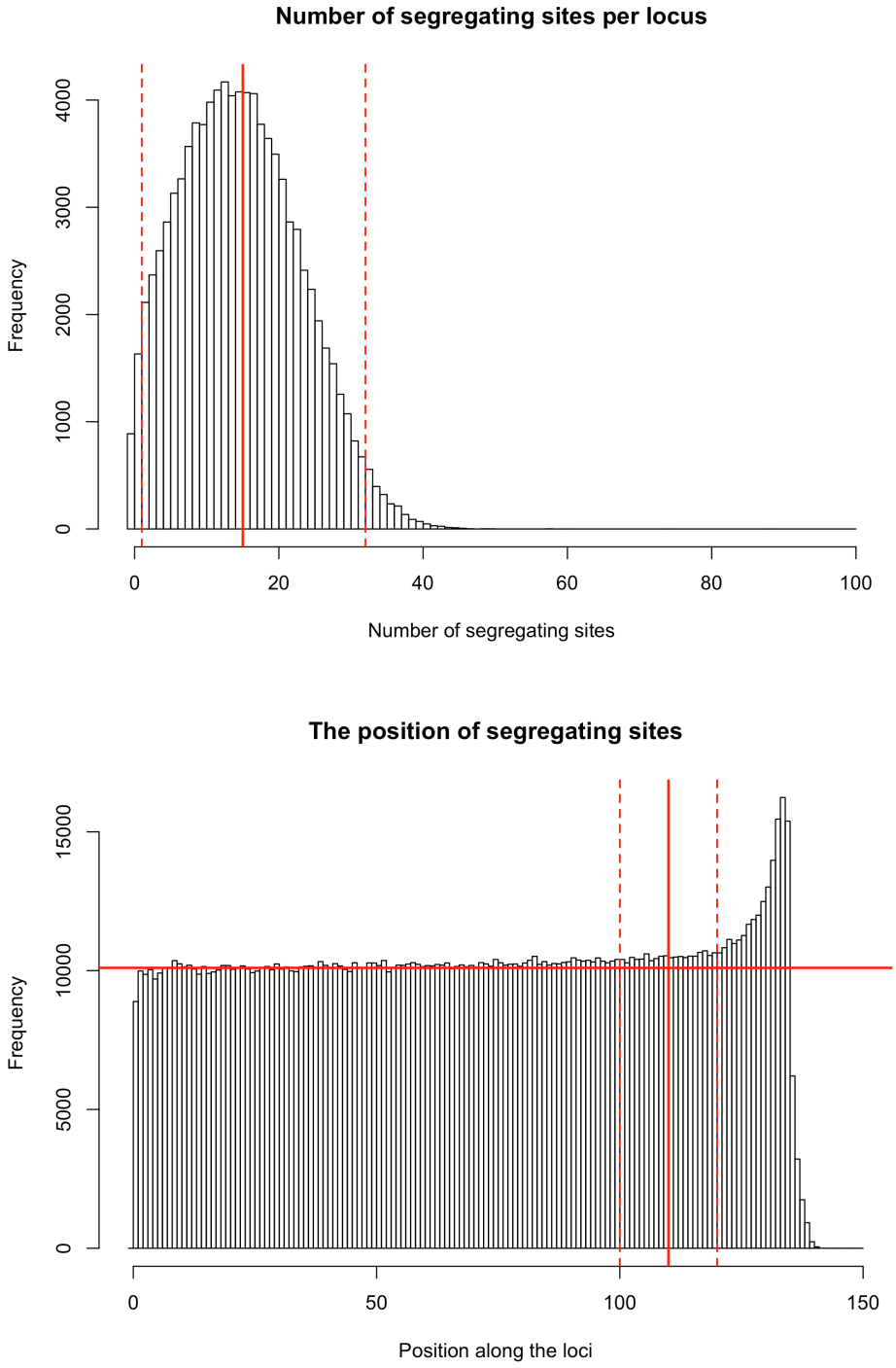
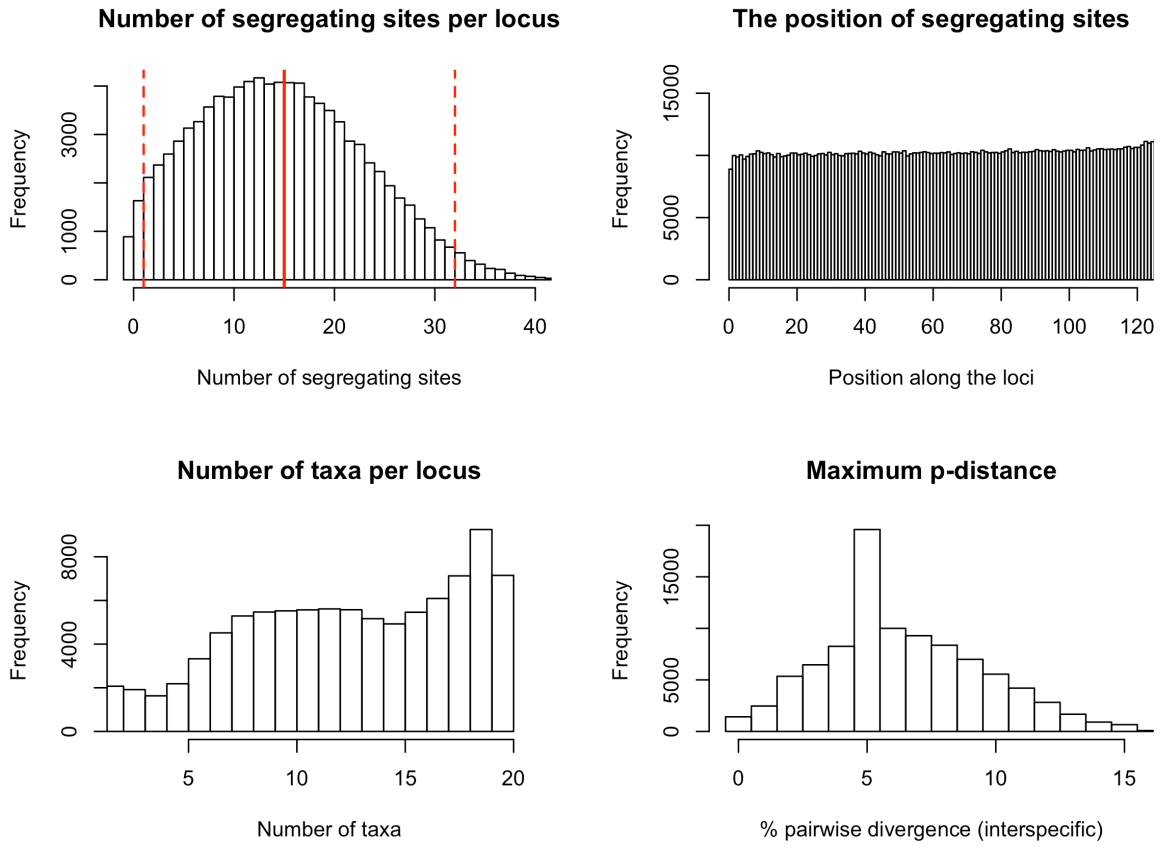


Figure K.1. Results from the original pyRAD output file. A systematic increase in the frequency of variable site (SNP) can be found close to the end among the aligned sequences (around position 110 [a red solid line]).



Figures K.2. Summaries of molecular data in the final dataset (after edited). Sites after position 110 in all loci are excluded, and loci that have sequence data from less than 4 different taxa and those that exhibit more than 15% maximum pairwise sequence divergence (based on uncorrected p-distance) between taxa are discarded in the final dataset.

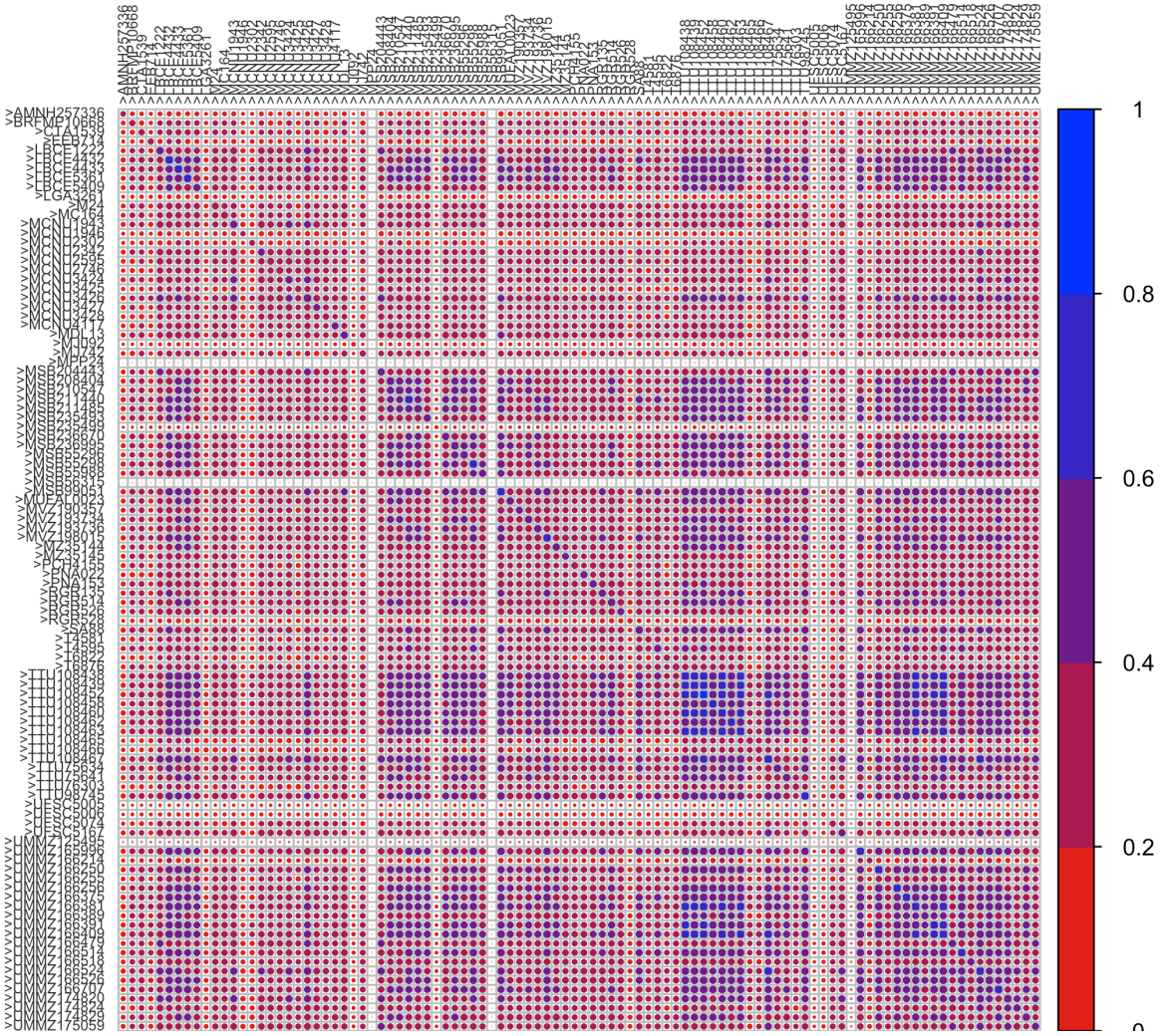


Figure K.3. The number of shared loci between sequenced samples. A larger and darker circle between pairs of samples indicates a higher number of shared loci that can be found in the final molecular dataset.

Table K.1. Processing information and pyRAD summary statistics for specimens sequenced. Raw reads refers to the total reads produced during Illumina sequencing; post-processing reads are those that remained after filtering for quality, adaptor contamination, ambiguous barcodes, and mitochondrial sequences; total clusters are the number of homologous sequences clusters created with the post-processing reads; mean depth is the mean depth of coverage of a cluster. Heterozygosity (H) and error-rate (E) were estimated across clusters, and consensus sequences were created for each cluster. Variable and invariable DNA sites were summed across all loci (Total sites), and the percentage of polymorphic sites (% poly) is presented. Consensus sequences were clustered across specimens, and loci that passed filtering parameters were included in the final data matrix (Final loci).

Specimen	Raw reads	Post-processing reads	% retained	Total clusters	Mean depth of clusters	H	E	Total sites	% poly	Final loci
AMNH257336	994049	832123	83.71	279539	2.892	0.00404096	0.00117328	4780262	0.0013855	20578
BRFMP10668	1091492	1004838	92.06	256792	3.797	0.00855754	0.0007382	6376490	0.005933	28525
CTA1539	1052044	940994	89.44	197454	4.692	0.0081235	0.0007333	6434046	0.0057415	29698
EEB714	6816815	5684637	83.39	558439	9.617	0.00954388	0.00092773	18091089	0.0059166	27601
LBCE1222	3473491	3059655	88.09	303672	9.851	0.00272448	0.00037465	10169334	0.0010824	44643
LBCE4432	4677508	4284873	91.61	437847	9.578	0.00806451	0.0003415	13239715	0.0054389	58879
LBCE4433	6640210	6062176	91.29	448718	13.155	0.0079268	0.00039559	13532530	0.0052874	60517
LBCE5361	5359667	4852534	90.54	449616	10.507	0.0080967	0.00032485	12660711	0.0051902	56917
LBCE5409	2808052	2534167	90.25	287996	8.639	0.00765758	0.00045	9615708	0.0051676	43480
LGA3261	994217	920671	92.60	193306	4.687	0.00353182	0.00081783	6564461	0.0017302	16710
M24	1715967	1540926	89.80	267891	5.657	0.0027633	0.00060203	8677697	0.0011528	37631
MC164	1833434	1669071	91.04	258498	6.346	0.00582822	0.00053741	8551708	0.0037429	37821
MCNU1943	2430868	2255127	92.77	312496	7.092	0.0050657	0.0004399	10675057	0.0032484	45989
MCNU1946	616199	561339	91.10	134124	4.111	0.00490172	0.00075916	4166497	0.002979	19031
MCNU2302	2561802	2335296	91.16	291246	7.841	0.00351978	0.00067052	9869420	0.0016637	24251
MCNU2342	1644103	1522980	92.63	251048	5.987	0.00493386	0.0005812	9041553	0.0031638	39780
MCNU2595	1697441	1583986	93.32	260316	5.953	0.00512251	0.00053447	8338354	0.0032492	37143
MCNU2746	1512138	1384511	91.56	239904	5.647	0.00503895	0.00061623	8296495	0.0032083	36902
MCNU3424	1967356	1815353	92.27	304779	5.849	0.00521635	0.00057592	9952367	0.003335	44119
MCNU3425	1120034	1026857	91.68	209496	4.795	0.00514547	0.00062659	6820718	0.0031607	30058
MCNU3426	4783383	3909789	81.74	481642	7.908	0.00561712	0.00053981	13912567	0.0031955	55484

MCNU3427	1418830	1290503	90.96	239982	5.28	0.0051238	0.00055881	8370134	0.0032534	37874
MCNU3428	1173679	1084000	92.36	197613	5.356	0.00483483	0.00046887	6992537	0.0031555	31276
MCNU4117	1865975	1672841	89.65	262619	6.223	0.00401478	0.00070897	8875798	0.0021174	38651
MDL13	2583798	2302660	89.12	320887	7.049	0.00789406	0.00063534	9768412	0.0052699	43160
MJ092	7750275	6503737	83.92	608946	10.384	0.00593569	0.0004858	19202884	0.0032786	12575
MJ742	965143	884735	91.67	198989	4.344	0.00798081	0.00064882	6402715	0.0056654	28657
MSB204443	4289467	3878039	90.41	319088	11.916	0.00479387	0.00037414	10316416	0.002877	44802
MSB208404	5071247	4668783	92.06	396640	11.485	0.00898536	0.00043422	12412419	0.0063018	53950
MSB210547	4778990	4407564	92.23	400684	10.724	0.0089112	0.0004151	12280995	0.0062342	54022
MSB211440	5066613	4666747	92.11	370866	12.274	0.00891779	0.00037821	13438870	0.0062792	57765
MSB211485	5920725	5461809	92.25	417822	12.787	0.00936205	0.00036365	12110156	0.0060694	51675
MSB235493	3518878	3108082	88.33	268841	11.378	0.00609572	0.00036858	9822532	0.0040955	44607
MSB236670	4477978	3934033	87.85	359242	10.746	0.00777103	0.00040118	10681931	0.0049122	46303
MSB236995	5796474	5309196	91.59	355402	14.56	0.00869267	0.00042733	12895710	0.0059307	56571
MSB55296	4096218	3707255	90.50	343564	10.486	0.00872128	0.00042761	11424220	0.0062217	50513
MSB55298	4838917	4410764	91.15	383911	11.235	0.00882397	0.00031201	13348121	0.0061402	57792
MSB55988	3498030	3185412	91.06	316387	9.862	0.00845901	0.00042823	10147083	0.0056417	43910
MSB99051	5861231	5426141	92.58	450280	11.82	0.00910491	0.00034699	14157765	0.0062882	59505
MUFAL0023	3006330	2719433	90.46	317931	8.348	0.00404567	0.00056541	10584281	0.0020688	46676
MVZ190357	2166281	1914865	88.39	242013	7.768	0.00612044	0.00043375	8991962	0.0039519	39739
MVZ193734	2821388	2539254	90.00	302322	8.207	0.00570446	0.00048811	10747596	0.0036566	48641
MVZ193736	3659228	3352979	91.63	325687	10.09	0.00671293	0.00039332	10332302	0.004261	45635
MVZ198015	4174428	3832891	91.82	369943	10.153	0.00790413	0.0004277	12748722	0.0054626	57013
MZ35144	2597221	2383589	91.77	310107	7.487	0.00800026	0.00064144	9942081	0.0050023	46232
MZ35145	1660478	1498336	90.24	244580	6.003	0.00775297	0.00056389	8406425	0.0053613	38339
PCH4155	1329587	1224091	92.07	236193	5.112	0.00227675	0.00054178	8252381	0.000863	36328
PNA022	3862319	3562406	92.23	380001	9.217	0.00617284	0.00045554	12804671	0.004095	42021
PNA153	3209918	2901487	90.39	331895	8.584	0.00471781	0.00052232	10356040	0.0025561	44118
RGR135	1969952	1824119	92.60	306980	5.798	0.0044853	0.0004993	10528093	0.0026219	45319

RGR514	3724992	3382802	90.81	369017	8.934	0.00469066	0.0004621	11560310	0.0026584	50131
RGR526	2342720	2123906	90.66	273370	7.631	0.0042455	0.00048222	9440827	0.0024715	39677
RGR528	782112	652388	83.41	217402	2.943	0.00484565	0.00087982	4081206	0.0025198	17811
SA88	2880878	2650884	92.02	362780	7.153	0.00645011	0.00047128	11363978	0.0039807	50016
T4581	1071707	971505	90.65	223489	4.255	0.00601189	0.00074267	6709867	0.0037352	30058
T4595	1647724	1525651	92.59	267239	5.577	0.00538207	0.00061872	9842763	0.0034972	44587
T6822	989403	915942	92.58	198607	4.507	0.00598168	0.0007509	6441913	0.0037649	28819
T6876	1661887	1515012	91.16	258638	5.737	0.00560413	0.00048185	8089775	0.0034467	35620
TTU108438	12264098	10034025	81.82	715510	13.622	0.00877031	0.0003883	23997022	0.0053455	74315
TTU108439	9972896	8151918	81.74	603860	13.116	0.00895082	0.00038744	19395161	0.0052062	71361
TTU108452	14143581	11714163	82.82	674063	16.96	0.00874415	0.00032841	25151427	0.0055628	75659
TTU108458	6538040	5886481	90.03	414623	13.823	0.00762385	0.00037661	13236293	0.0050633	59399
TTU108460	10623103	8718667	82.07	533535	15.733	0.00805912	0.00036114	19057764	0.0051069	71777
TTU108462	5493553	5020967	91.40	405391	12.172	0.00746322	0.00036668	13499649	0.0050288	59779
TTU108463	11655216	9437412	80.97	653863	14.03	0.00909206	0.00044686	22878972	0.0052943	71495
TTU108465	6426668	5890311	91.65	449235	12.8	0.00366217	0.00041836	12966516	0.001177	31274
TTU108466	5539184	5117664	92.39	405863	12.434	0.0032184	0.00037355	13059564	0.0011648	31872
TTU108467	12452926	10190111	81.83	666971	14.842	0.00612702	0.00036778	23661724	0.0030709	69264
TTU75634	2984869	2733834	91.59	291783	9.208	0.00708256	0.00041627	11139308	0.0047836	48542
TTU75641	3553870	3220345	90.62	262110	12.059	0.00705978	0.00041694	10958789	0.0047005	47764
TTU76303	1224043	1117736	91.32	239555	4.568	0.00732739	0.0007539	6988528	0.004876	31138
TTU98745	6028376	5524302	91.64	448031	12.039	0.00750499	0.00032425	14429389	0.0047985	59825
UFSC5074	1297680	1169980	84.71	205001	5.578	0.00450674	0.00050279	7212859	0.002837	29973
UFSC5167	1599213	1464515	90.16	265955	5.35	0.00442445	0.00058423	9193226	0.0025953	41475
UMMZ165996	5225241	4803766	91.58	394287	11.924	0.00757522	0.00049276	13591576	0.0050858	60205
UMMZ166214	1092158	991589	91.93	115173	8.463	0.00710554	0.00030013	4946259	0.0052824	21986
UMMZ166250	6314019	5784805	90.79	491575	11.519	0.00799237	0.00044473	14225107	0.0050093	61464
UMMZ166255	1401343	1257790	91.62	163053	7.585	0.00745929	0.00041906	7644939	0.0053786	34784
UMMZ166256	6859588	6283181	89.76	485387	12.608	0.00811808	0.00038624	13917968	0.0053833	61706

UMMZ166375	5260017	4871297	91.60	415603	11.511	0.00753124	0.00029222	14192493	0.0051745	60244
UMMZ166381	8200123	6854510	92.61	580418	11.396	0.00852344	0.00036314	19155871	0.0055698	69802
UMMZ166389	3125312	2881107	83.59	292485	9.669	0.00754402	0.00034154	9761289	0.0050965	43864
UMMZ166391	6919255	6283612	92.19	533300	11.474	0.00803974	0.00042534	14889041	0.0052095	63070
UMMZ166409	13257835	10874258	90.81	637302	16.592	0.00865426	0.00034377	23730594	0.005449	74803
UMMZ166479	4220721	3880578	82.02	348011	10.931	0.00512994	0.00038148	10565953	0.0029071	45488
UMMZ166514	5865636	5364273	91.94	472939	11.101	0.00697985	0.00035007	14370801	0.0044418	61086
UMMZ166518	2354474	2108148	91.45	258105	8.04	0.00707085	0.00042646	8971849	0.0048806	40535
UMMZ166524	11645819	10683912	89.54	813991	12.821	0.00828003	0.00044593	20843030	0.0029347	62579
UMMZ166526	4272073	3945791	91.74	344844	11.201	0.00728241	0.00034047	12340378	0.0049574	56215
UMMZ166707	5245729	4772443	92.36	428239	10.862	0.00709488	0.00039275	12838398	0.0046749	58386
UMMZ174820	3198698	2943133	90.98	338212	8.517	0.00490356	0.00044527	12339197	0.0030545	53284
UMMZ174824	3251982	2943906	92.01	324992	8.82	0.0061569	0.00104503	9696922	0.0032235	42759
UMMZ174829	2697934	2454352	90.53	299467	8.007	0.00746187	0.00045031	10544888	0.0051543	48528
UMMZ175059	2803615	2541994	90.97	299800	8.294	0.00505144	0.00047278	9805398	0.0029753	43078

APPENDIX L. Discriminant Morphometric Analysis

Description of clusters used in the Discriminant Analysis (Table L.1), as well as the steps followed to diagnose distinguishable morphometric clusters (Figs. L.1, L.2, L.3), are provided in this appendix, Appendix L.

Table L.1. Information about the collected localities used in the Discriminant Function Analysis (DFA). Cluster 1 refers to the groups used in the first round of DFA; cluster 2 refers to the groups used in the second round of DFA; and final clusters refers to the main morphometric clusters found according to the previous rounds of analyzes and used in the species delimitation section.

Final Cluster	Cluster2	Cluster1	Locality	Depto/State	Country	N Ind
Bolivia	Bolivia	El Beni	Itenez, Lago Victoria	El Beni	Bolivia	2
			Magdalena, La Granja	El Beni	Bolivia	1
			Rio Tijamuchi	El Beni	Bolivia	1
			San Joaquin	El Beni	Bolivia	5
		Santa Cruz	3.5km W Estacion El Pailon	Santa Cruz	Bolivia	1
			6km W (by road) of Ascension	Santa Cruz	Bolivia	1
			8km SE Tita	Santa Cruz	Bolivia	2
			Buenavista	Santa Cruz	Bolivia	5
			San Rafael de Amboro	Santa Cruz	Bolivia	1
			Uruma	Santa Cruz	Bolivia	1
			Warnes	Santa Cruz	Bolivia	1
Central	Belem	Belem	Belem	Para	Brazil	7
			Belem, Utinga	Para	Brazil	1
			Belem, Utinga, Bacia Agua Preta	Para	Brazil	4
			Ilha do Marajó Soure	Para	Brazil	1
			Peixe Boi	Para	Brazil	1
		Maranhão	São Bento	Maranhão	Brazil	7
			São Jeronimo	Maranhão	Brazil	4
	Central	Lagoa Confusão	Faz Lago Verde, Lagoa da Confusão	Tocantins	Brazil	2
		Mato Grosso	Ilha de Taimã	Mato Grosso	Brazil	3
		Tocantins	Usina Hidreletrica Estreito, Palmeiras do	Tocantins	Brazil	1

			Tocantins			
		Goias	Anapolis	Goias	Brazil	4
	São Paulo	São Paulo	São Paulo	São Paulo	Brazil	5
Minas Gerais		Rio de Janeiro	Silva Jardim	Rio de Janeiro	Brazil	1
	Minas Gerais	Bom Jesus da Lapa	Bom Jesus da Lapa	Bahia	Brazil	2
		Caravelas	Caravelas	Bahia	Brazil	1
		Jaguaquara	Jaguaquara	Bahia	Brazil	1
		Lagoa Santa	Lagoa Santa	Minas Gerais	Brazil	2
		Pirapora	Pirapora	Minas Gerais	Brazil	4
	Northeast	Juazeiro	Fazenda do Horto, Juazeiro	Bahia	Brazil	1
Amazon	Amazon	Amazonas	Eirunepe	Amazonas	Brazil	1
			Igarape Grande	Amazonas	Brazil	1
			Ipixuna, Purus River	Amazonas	Brazil	2
			João Pessoa, Juruá River	Amazonas	Brazil	1
		Moyomamba	Moyobamba, Mayo River	San Martin	Peru	3
			Yurac Yacu	San Martin	Peru	1
		Peru	Cumeria	Loreto	Peru	1
			San Jeronimo	Loreto	Peru	8
			Yarinacocha	Loreto	Peru	4
			Manu	Madre de Dios	Peru	1
			Masisea, Ucayali River	Ucayali	Peru	2
	Peruate	Rio Peruate	Boca Rio Peruate	Loreto	Peru	3
Guyana	Guyana	Guyana	Blairmont	Berbice	Guyana	1
			Nonpareil Plantation	Demerara	Guyana	1
			Versailles	Demerara	Guyana	1
			Hyde Pack	Demerara	Guyana	1
		Kanuku	Kanuku Uts	Upper Takutu-Upper Essequibo	Guyana	2

		Itacoatiara	Itacoatiara	Amazonas	Brazil	1
		Oiapoque	Vila Velha do Caciporó	Amapa	Brazil	12
			Ile de Cayenne		French Guiana	1
Northeast	Northeast	Alagoas	Quebrangulo	Alagoas	Brazil	29
			Viçosa	Alagoas	Brazil	110
			Anadia	Alagoas	Brazil	9
		Crato	Crato	Ceará	Brazil	13
		Sao Lourenco	Sao Lourenco	Pernambuco	Brazil	2
			Serra dos Cavalos, Caruaru	Pernambuco	Brazil	1
Paraguay	Paraguay	Ayolas	Laguna Galarza	Corrientes	Argentina	1
			5 km ENE Ayolas	Misiones	Paraguay	1
		Bahia Negra	Passo de Lontra	Mato Grosso do Sul	Brazil	1
			4 km N (by air) of Bahia Negra, Estancia Dona Julia, W bank of Rio Paraguai	Alto Paraguai	Paraguay	1
			6 km SE (by air) of Bahia Negra, W bank of Rio Paraguai along Riacho Ramos	Alto Paraguai	Paraguay	6
			50 km WNW Fortin Madrejon	Chaco	Paraguay	7
		Bermejo	Moreno Cue, Estancia Las Rosas	Chaco	Argentina	1
			Rio de Oro	Chaco	Argentina	1
		Estancia Yacare	Estancia Yacare, 1.68 Km NW of Puesto San Fernando	Neembucu	Paraguay	1
			Estancia Yacare, 3.59 Km NNE of Puesto San Fernando	Neembucu	Paraguay	1
		Estero Valenzuela	Estero Valenzuela	Corrientes	Argentina	8
			Cañada del Pirayu, Ruta Nacional 12	Corrientes	Argentina	1
			San Cayetano, CAPRIM	Corrientes	Argentina	1
		Golondrina	15.5 km NNW Chaco-I	Presidente	Paraguay	10

				Hayes		
			24 km NW (by air) of Villa Hayes, Estancia La Golondrina	Presidente Hayes	Paraguay	23
			24 km NW Villa Hayes	Presidente Hayes	Paraguay	7
		Itapua	Estancia San Isidro, 5.56 Km NW of houses	Itapua	Paraguay	1
		Luque	17 km E (by road) Luque	Central	Paraguay	4
			Compania Minas, Cue, Emboscad	Cordillera	Paraguay	1
		Manduvira	24 km NNW Carayao, Estancia San Ignacio	Caaguazu	Paraguay	1
			Edge of Rio Manduvira, 53 km (by air) NW Asuncion	San Pedro	Paraguay	2
		Oran	Villa Carolina, Rio Lavallen	Jujuy	Argentina	10
			P.N. Rio Pilcamayo	Formosa	Argentina	1
			Tabaca	Salta	Argentina	8
		Puerto Fonciere	west bank of Rio Paraguai, 4 km NW of Puerto Fonciere	Presidente Hayes	Paraguay	9
		Riacho Pilaga	Kilometro 182, Riacho Pilaga, 10 Mi NW	Formosa	Argentina	2
		Rosario	east bank of Rio Paraguai, 22 km (by air) SSW Rosario	San Pedro	Paraguay	4
			Island in middle of Rio Paraguai, 10 km (by air) NW of Rosario	San Pedro	Paraguay	10
			Island in Rio Paraguai, 10 km (by air) NW of Rosario	San Pedro	Paraguay	8
		Saladas	Bella Vista	Corrientes	Argentina	1
			Goya	Corrientes	Argentina	2
			Loma Alta	Corrientes	Argentina	1
		Santa Teresa	Estancia Santa Teresa, 3.88 Km S of Puesto Anastacio	Neembucu	Paraguay	1

South	South	Ajo	Ajó	Buenos Aires	Argentina	3
		Buenos Aires	Buenos Aires	Buenos Aires	Argentina	1
			Buenos Aires, 35km N.O. de Ayacucho	Buenos Aires	Argentina	1
			Delta I.N.T.A	Buenos Aires	Argentina	2
			Isla Ella, Delta del Parana	Buenos Aires	Argentina	2
			Zelaya	Buenos Aires	Argentina	1
			6 km S (by road) of Puerto Ibicuy	Entre Rios	Argentina	3
		Cento Cue	Centu Cue	Misiones	Paraguay	1
			Margin of Rio Tebicuary, 1.2 km upstream (E) Hotel Centu Cue (opposite margin)	Paraguari	Paraguay	2
		Lagoa dos Patos	Lagoa dos Patos	Rio Grande do Sul	Brazil	1
		Otamendi	Otamendi, I.N.T.A-Delta	Buenos Aires	Argentina	2
		Parque Nacional El Palmar	Parque Nacional El Palmar	Entre Rios	Argentina	1
Southeast	Southeast	Paulinia	Córrego Argentini	São Paulo	Brazil	1
			Faz. Saltinho	São Paulo	Brazil	1
			Faz. Sto Angelo, Represa de Americana	São Paulo	Brazil	2
		Taubate	Faz Antonio Taino, Bairro barranco, Taubaté	São Paulo	Brazil	1
			Faz Honda, Canal Principal, lado esquerdo, Distrito Eugênio Mello, São José dos Campo	São Paulo	Brazil	1
			Faz Honda, São José dos CampoS	São Paulo	Brazil	1
			Faz. Antonio Tavares, Taubaté	São Paulo	Brazil	1
			Faz. Boa Esperança, Pindamonhangaba	São Paulo	Brazil	1

			Faz. Instituto Agronômico, Pindamonhangaba	São Paulo	Brazil	2
			Faz. José Nane, Caçapava	São Paulo	Brazil	2
			Faz. Kangutti, Bairro do Remédio, Taubaté	São Paulo	Brazil	1
			Gleba Paulo Japonês, bairro barranco, Taubaté	São Paulo	Brazil	10
			Linha de captura, Gleba Paulo Japonês, Bairro barranco, Taubaté	São Paulo	Brazil	1
			Pindamonhangaba	São Paulo	Brazil	2
Venezuela	Venezuela	Bolivar	Hato San Jose, 20 Km W La Paragua	Bolivar	Venezuela	1
		Meta	Villavicencio, Finca Buque	Meta	Venezuela	1
		Trujillo	Nulita, 29 km SSW Santo Domingo, Selvas de San Camilo	Apure	Venezuela	1
			Montalban, 2 km SE Montalban, Potrerito	Carabobo	Venezuela	1
			El Cruce, La Esperanza Farm	Portuguesa	Venezuela	1
			Valera, 30 km NW Valera, Nr. El Dividive	Trujillo	Venezuela	4
Mendoza	Mendoza	Mendoza	El Nihuil, San Rafael	Mendoza	Argentina	1

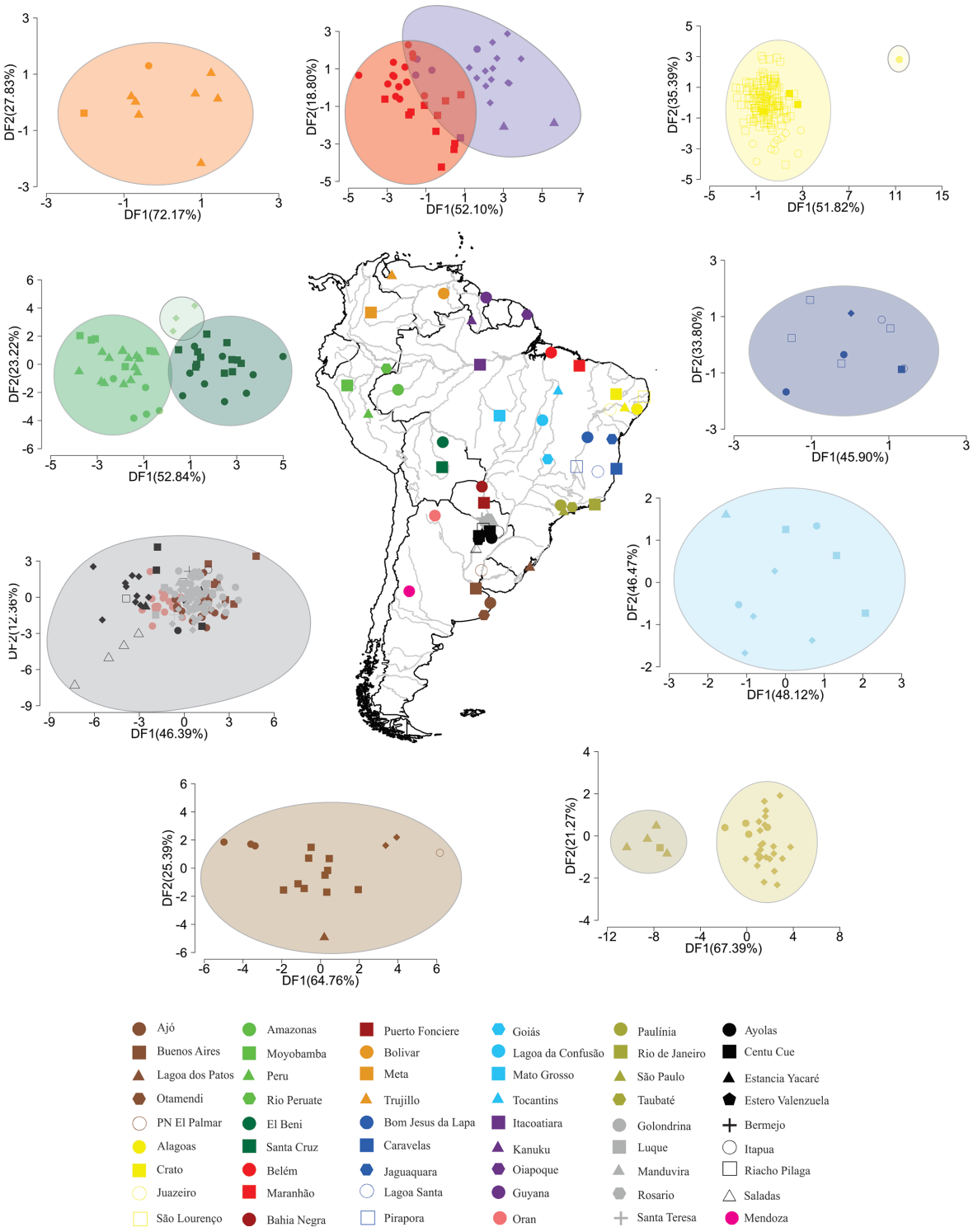


Figure L.1. Scatterplots with the scores of First and Second Discriminant Function (DF1 and DF2) with 18 morphometric variables accros the frist set of samples clusters (third column of Table L.1).

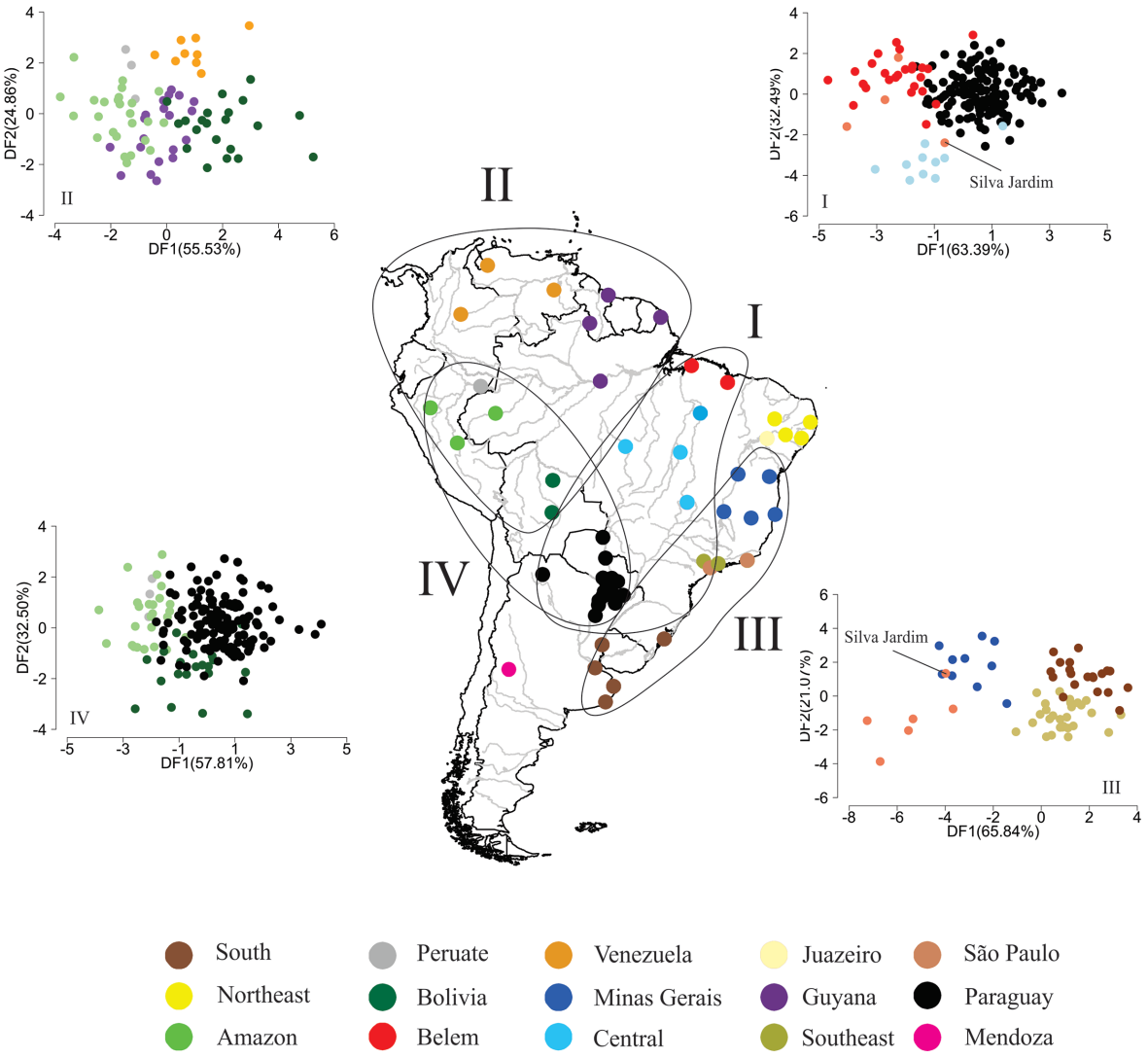


Figure L.2. Scatterplots with the scores of First and Second Discriminant Function (DF1 and DF2) with 18 morphometric variables accros the second set of samples clusters (second column of Table L.1).

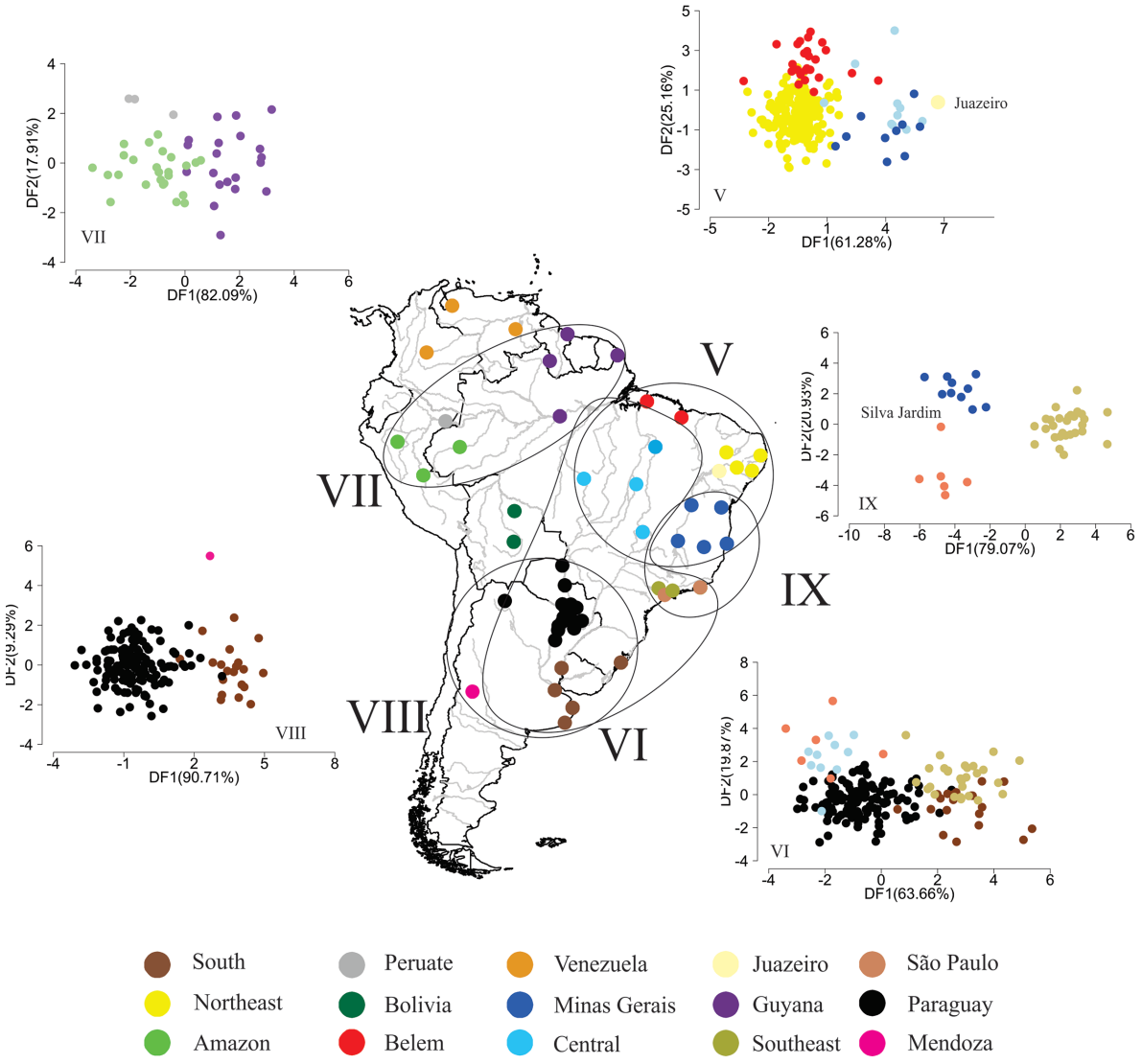


Figure L.3. Scatterplots with the scores of First and Second Discriminant Function (DF1 and DF2) accros the second set of samples clusters (second column of Table L.1). DFs numbered as V, VI, VII were performed with 18 morphometric variables, and DF numbered as VIII was performed with 16 morphometric variables (see Methods section).

5. THE CLADE D OF TRIBE ORYZOMYINI AND THE PHYLOGENETIC POSITION OF GENUS *HOLOCHILUS* BRANDT, 1835 (RODENTIA, SIGMODONTINAE, ORYZOMYINI)

ABSTRACT

Identifying mechanisms that promote lineage diversification is essential to understand biogeographic patterns. Here we investigate the biogeographic processes related to the extant genera of the Oryzomyini clade D and access the position of genus *Holochilus* inside this clade. This clade includes the most comprehensive generic diversity of the tribe, with lineages occupying distinct phytophysiognomies. The current distributional pattern and the presence of tetralophodonts genera in this clade suggest that there is a close relationship between these genera and the evolution of South Amercia open areas. Using a genomic dataset we construct a time calibrated tree, and estimate the range evolution across phylogeny obtained through statistical model testing among different biogeographic models. The origin of clade D occurred approximately 8.8 Ma, and most of diversification events occurred during Pliocene. Geographic range estimates suggest that the ancestors of this clade were open area inhabitants of the Cis-Andean part, and the retraction and expantion of the open enviroments, together with the formation of Isthmus of Panama, favored the dispersion and diversification on this clade. Regarding to the position of genus *Holochilus*, it appears as a sister group of *Pseudoryzomys*, but the tetralophodonts genera did not show a monophyletic relationship, as proposed by other studies.

Keywords*:* RADSeq; Ancestral range estimate; Neotropics; Climatic changes

5.1. Introduction

Identifying mechanisms that promote lineage diversification is essential to understand speciation processes and biogeographic patterns across different spatial and temporal scales (Werneck et al. 2012). Vegetation changes associated with climate fluctuations and geologic events (tectonics, structural arches, river genesis) are often cited as keys factors in structuring diversity of organisms and drivers of evolutionary change (Taberlet et al. 1998; Soltis et al. 2006; Carnaval et al. 2009; Woodburne, 2010; Morgan et al. 2011). The history and evolution of the open vegetation biomes in South America has taken place under different perspectives and scales for mammals, both on time and space (for a recent review, see Carmignotto et al. 2012). Previews hypotheses proposed that savannas would be an Eocene innovation in South America (Janis, 2003), with posterior expansions of these environments after the end of Eocene to the Oligocene, bringing specialized grazing mammals lineages to this continent (Vivo, 2008). Subsequently, expansion and retraction of these areas modeled the current diversity of the Sour American mammalian fauna (Vivo, 2008).

Hershkovitz (1962, 1966, 1972) offered the first efforts to understand the patterns of differentiation among open-vegetation South America mammals, especially rodents of the subfamily Sigmodontinae. According to Hershkovitz (op. cit.), the tribe Oryzomyini was the basal sigmodontine lineage and contained only forest-dwelling and pentalophodont taxa, being the base of an evolutionary

transition to non-forest habitats. However, nowadays the definition of Oryzomiyni tribe includes not only forest pentalophodont taxa, but also several lineages of pastoral and tetralophodont genera (Weksler, 2015).

Another issue constantly addressed in the biogeography of Oryzomyini is the geographical origin of this lineage, being three vicariance scenarios frequently evaluated, the separation of Cis, Trans and Andean lineages by the Andean uplift (van der Hammen, 1974; Weksler, 2006; Prado and Percequillo, 2013, among others). Modern studies shows that this pattern does not show correspondence with the oryzomyine phylogeny, instead each major lineage presents all of the three patterns, proposing that the vicariance framework include cases of over-Andean dispersal (Weksler, 2006).

Phylogenetic studies on the tribe Oryzomyini expanded considerably over recent years (Weksler, 2015). Although its composition, which was originally defined and diagnosed by Voss and Carleton (1993), is often corroborated by more comprehensive phylogenetic studies (Steppan, 1995; Weksler, 2003, 2006; D'Elía et al. 2007; Turvey et al. 2010; Percequillo et al. 2011; Pine et al. 2012), the relationships among the four major Oryzomyini clades, A, B, C and D (see Weksler, 2006) and among the genera within these clades are still not clear (Weksler, 2006; Percequillo et al. 2011; Pine et al. 2012). Additionally, the relationships among taxa of species group were not adequately addressed and are uncertain as well as the time and diversification mechanisms are inconclusive for most genera and species (Steppan, 1995; Weksler, 2006; Parada et al. 2015).

The clade D of the tribe Oryzomyini is a typical example of the issues presented previously. This clade has the most comprehensive generic diversity of the entire tribe, comprising 15 genera and 41 species (Weksler, 2015), that represents about 50% and 30% of the tribe diversity, respectively. Their representatives are distributed throughout South America, inhabiting both forested environments and open areas (Prado and Percequillo, 2013). The clade D is also the clade with largest number of monotypic genera in the tribe, as there are eight genera with only one species. Is interesting to note that most of these monotypic genera are forest habitants, except by *Lundomys* and *Pseudoryzomys*. Within the more diversified taxa of clade D, most genera are open areas dwellers (Weksler, 2015). This fact can indicate that the transition of a forest environment to a more open habitat in South America landscape marked the diversification history of clade D. Another evidence that sustain this hypothesis is the presence of tetralophodont (taxa that have molars with four ridges or lophs and reduced mesoloph/id) genera in clade D; this allegedly derived character is observed only in the genera *Holochilus, Lundomys, Pseudoryzomys* (members of clade D), *Zygodontomys* (member of clade A; Voss and Carleton, 1993), and *Microakodontomys* (genus with position still not determined within the tribe), although some species of genera *Aegialomys*, *Cerradomys* and *Eremoryzomys* (all three members of clade D) are polymorphic for the presence of the mesoloph and mesolophid. In the most recent phylogeny for clade D (Machado et al. 2013) the tetralophodont genera of clade D appears as paraphyletic, but these oryzomyines have usually been recovered as a monophyletic group in

previous studies (Carleton and Olson, 1999; Weksler, 2003, 2006; Percequillo et al. 2011; Pine et al. 2012). Actually, it is the sister group relationship of *Holochilus* that changes depending on the gene and method used in phylogenetic reconstruction (Machado et al. 2013).

Therefore, there are several key issues yet to be studied regarding clade D like its monophyly, and the internal relationship among genera, as well as hypotehsis on the origins of open/forest dwellers. Here we apply an coalescent-based genomic method to reconstruct a clade D phylogeny, including the most expressive specific sample published until the moment, to investigate temporal and spatial aspects of the diversification of this clade, as well to access the phylogenetic position of genus *Holochilus.* We also performed phylogeny-guided model selection procedures to evaluate the relative importance of historical processes in the distribution patterns of clade D lineages.

5.2. Material and methods

5.2.1. Samples and Collections

Genomic DNA data was generated for 35 individuals of Oryzomyini clade D species. Tissues were requested from museums (see Appendix M for full details of holdings) and all the living genera of this clade are represented, except by *Amphinectomys* and *Sigmodontomys* for which we were not able to obtain adequate genomic data. In addition to those specimens from clade D, we also generated genomic data for 4 outgroup genera, namely *Scolomys, Zygodontomys, Oligoryzomys* (oryzomyines rodent genera that belong to other clades) and *Necromys* (a member of tribe Akodontini). See Appendix M for a complete list of the samples used in the genomic analysis, and the complete examined material with information about the collected localities.

5.2.2. DNA extraction, amplification, and sequencing

DNA was extracted from liver, muscle or skin of each individual using the Qiagen DNeasy Blood and Tissue Kit. Reduced representation libraries, were built using the Double Digest RADseq method (Peterson et al. 2012). DNA was double digested with the EcoR1 and MseI restriction enzymes, unique barcodes (10 bp) and Illumina adapter sequences were ligated to the digested fragments, and the fragments were amplified by PCR. Barcoded product from each individual were pooled to select fragment with size between 350 and 450 bp and the libraries were sequenced in the Illumina *HiSeq2000*, according to manufacturer's instructions to generate 150 base pairs, single-end reads.

5.2.3. Processing Illumina Data

Data was processed using *pyRAD* v3.0.66 (http://pyrad.googlecode.com) pipeline. First, sequences were identified allowing for one base mismatch in their sample-specific barcode. The restriction site and barcode were trimmed from each sequence, and bases with a FASTQ quality score below 20 were replaced with N. Sequences having more than >5 Ns were discarded. Then, for each sample, sequences were clustered using a threshold of 90%. Error-rate and heterozygosity were estimated from the base counts in each site across all clusters, and the averages were used to establish consensus sequences. Clusters with less coverage than a minimum depth of 5 were excluded in order to ensure accurate base calls. Consensus sequences from all samples were clustered by sequence similarity, with their input order randomized, using the same similarity threshold as the within-sample clustering (90%). The resulting clusters were then aligned with Muscle in the pyRAD pipeline. Any locus appearing heterozygous at the same site across more than 5 samples were discarded. The remaining clusters are treated as RAD loci, i.e., multiple alignments of putatively orthologous sequences, which were assembled into phylogenetic data matrices.

Customized R scripts (see Huang, 2016 for further details) were applied to evaluate, visualize and filter even more the results. Sequences were chopped to keep the first 110bp in order to exclude sites with high variation towards the end of the alignments. All invariable loci and loci that contained samples from <4 species were also excluded.

5.2.4. Tree estimation

Another customized R script was used to convert the edited loci file to a file that contained aligned unlinked SNP data from each locus. We randomly choose one parsimony-informative SNP from each locus to produce a concatenated data set (see Huang, 2016). The output was then manually edited, saved in *nexus* format and imported into PAUP* (ver. 4.0d147; Swofford, 2002).

We include a total of 126,260 loci (each SNP was considered as a independent locus) and 40 individuals for the species tree estimation using the coalescent-based program SVDQuartets (implemented in PAUP*). We evaluate all possible quartets, selecting tree using the QFM quartet assembly, and we also performed bootstrapping with 1000 replicates to calculate branches support.

5.2.5. Divergence time estimation

Divergence times were estimated using standard models of evolution implemented in BEAST v. 2.4.5 (Drummond et al., 2012). Simultaneous analysis of the complete dataset (126,260 loci) was not computationally feasible, so we made a subset from our total dataset, using only the loci

present in at least 35 individuals (1,943 loci) in a concatenated supermatrix. We force the program to use the same tree topology provided by SVDQuartets. We used a lognormal relaxed clock-model with a HKY model of substitution. The Markov chain Monte Carlo with 30^7 million generations each, with sampling every 60,000 generations, and we discarded the first 25 % as burnin. Convergence statistics were examined using Tracer v1.6.0 (Rambaut et al. 2014). A time tree was obtained with TreeAnnotator v2.4.5 (Drummond et al. 2012) using the maximum clade credibility (MCC) tree from all post-burnin trees and without posterior limit for each node.

We used three calibration points. The first was based on the origin of the Galapagos archipelago about 4 Ma (Grehan, 2001). We set the divergence between *Nesoryzomys* (endemic to Galapagos) and *Aegialomys* (sister taxa), by placing a mean of 3 ± 1 Ma in a normal distribution. For the second calibration point we used a uniform distribution and it was based on the fossil records of *Nectomys squamipes* (Pardiñas et al. 2002) dated to Middle Pleistocene (1.2 - 0.4 Ma) at the Ensenadan age. The last calibration point was based on the fossil records of *Holochilus vulpinus* (Pardiñas et al. 2002) dated to the Middle Pleistocene (0.78 Ma), with uniform distribution.

5.2.6. Ancestral range estimation

Ancestral ranges inference was obtained using package BioGeoBEARS (Matzke, 2013), available in R software environment (R Development Core Team, 2014). We compared different biogeographic models to investigate the cladogenetic pattern of clade D members: DIVA-like (the likelihood version of the parsimony-based method implemented in Dispersal-Vicariance Analysis; Ronquist, 1997), DEC (Dispersal-Extinction-Cladogenesis; Ree and Smith, 2008), and BayArea-like (Landis et al. 2013). These models consider dispersal, vicariance and extinction events in different perspectives, allowing the exploration of distinct biogeographic scenarios (Pavan and Marroig, 2017). We also considered an additional cladogenetic event, the founder-event speciation (+J), which permit that the new species jumps to a range outside of the ancestral range. By the end, six probabilistic models were tested: DIVA-like, DIVA-like+J, DEC, DEC+J, BayArea-like and, BayArea-like+J. The models were compared by the Akaike Information Criteria (AIC; Burnham and Anderson, 2002), using the Akaike weights (AICwt; Wagenmakers and Farrell, 2004).

The ultrametric tree employed in the analyses was the one generated with Beast, and we coded the presence or absent of each species based in two different set of taxa occurrence. In the first area matrix the species were coded by their presence in three areas: CIS Andean, Andean and TRANS Andean. CIS Andean comprehends the vast area of South America on the oriental side of Andes Cordillera, including most typical biomes of the continent, as Atlantic Forest, Amazon Forest, Caatinga, Chaco, Cerrado, Chaco, Pampa, Llanos, Gran Sabana, Patagonian Estepe, and a few others. Andean area corresponds to the Cordillera, above 1,000 m of elevation, and comprehends different

Montane Forests, Páramo and Puna habitats. TRANS Andean corresponds to the part of South America Continent that is localized in the occidental side of the Andes Cordillera, and includes Chilean and Peruvian Arid Deserts, Lomas, and Chocó Forests. In the second area matrix, species were coded by their the presence in five main habitats: OH (Humid open area), OD (Dry open areas), MF (Montane Forest), LF (Lowland Forest), DF (Dry Forest). Humid open areas comprehend periodically or permanent flooded grasslands. Dry open areas correspond to the dry grasslands and shrublands without a pronounced flooded period. Montane Forest here is defined as the super-humid forests occurring greater than 1,000m above sea level. Lowland Forest comprehends the evergreen broadleaf forests present throughout most of South America. Dry Forest is defined by the Deciduous and semi-deciduous tropical forest. We identified this areas based on distributional data available for these species (Prado and Percequillo, 2013; D'Elía and Pardiñas, 2015). Outgroups were removed from the phylogeny used for estimating ancestral ranges, since their long branch lengths in phylogeny would negatively bias the results.

5.3. RESULTS

5.3.1. Sequencing

A total of 131,593,244 million sequence reads were generated, of which 118,276,519 million passed stringent quality filters. We retained an average of 2,956,912 million reads per sample. Clustering within samples produced an average of 47,592 loci per individual, with coverage greater than five, yielding a mean coverage depth of 7.13 (Appendix N, Table N.1). Sequence clustering, alignment across samples and editing using customized R scripts provided a final data set containing 126,260 loci variable loci.

5.3.2. Phylogenetic Analyses and Diversification times

The RAD data set provided unprecedented resolution of species relationships within members of clade D, nearly all nodes in the tree had bootstrap values of 95 or greater (Fig. 27). Our results recovered with the SVDQuartets indicate that the first cladogenetic event separated the clade composed by *Drymoreomys albimaculatus* and *Eremoryzomys polius* from the other members of the clade D. The second split separated the species of genus *Cerradomys*. Then a clade composed by the genera *Lundomys*, *Microakodontomys* and *Sooretamys* diverged from the others, followed by the split between the clade formed by species from genus *Holochilus* and a clade composed by the genera *Aegialomys, Nesoryzomys, Melanomys, Tanyuromys, Oryzomys* and *Nectomys*. This last large clade diverged in two main smaller clades, one clustering species of genus *Nectomys* and the other grouping genera *Aegialomys, Nesoryzomys* and *Oryzomys* and the genera *Tanyuromys* and *Melanomys*.

The divergence analysis included the concatenated dataset and estimated the rising of clade D in the late Miocene, ca. 8.8 Ma (95% HPD = 11.5–6.5 Myr). Diversification events giving rise to the main clades are not evenly distributed along the phylogeny. Origin and diversification of most extant genera occurred during Pliocene, about 5 to 2 Myr. The most recent genus is *Melanomys*, dating ca. 0.9 Ma (95% HPD = 1.5–0.5 Myr), followed by *Aegialomys* (ca. 1.2, 95% HPD = 1.8–0.7 Myr).

5.3.3. Ancestral geographic ranges

Log-likelihood values of the six biogeographic models implemented in BioGeoBEARS (Table 4) show that BayArea-like+J is the model best fitted the data in both ancestral geographic range analysis, indicating that the geographic range evolution made by dispersal-vicariant method is the most probable. BayArea-like has two free parameters (d and e) specifying the rate of "dispersal" (range expansion) and "extinction" (range contraction), and at cladogenesis events it assumes that no range evolution occurs (Matzke, 2013). The assignment of the free parameter J in the model, specifying weights for jump-dispersal events in the cladogenesis matrix, produced significant improvement of the biogeographic model (Table 4).

Figure 28 shows the most probable states for the ancestral range estimated with BayArea-like+J between the five vegetation areas (OH, OD, MF, LF, DF). According to this figure the ancestral (node A) of the members of clade D inhabited a range of vegetation formation including open humid and dry areas. The first split separated taxa from dry forest (the ancestral of *D. albimaculatus* and *E. polius*; node B) and open areas (both humid and dry; node C). The next cladogenesis split taxa from open and dries forested areas (the ancestor of genus *Cerradomys*; node D), and taxa from lowland forest (node E). The ancestors of the clade composed by *Sooretamys, Microakodontomys* and *Lundomys* (node F) were lowland forest inhabitants that posteriorly dispersed to more open humid areas (node H). The next split (node G) divided into two main clades, one that the ancestor is an open dweller occupant (node J), and the other with the ancestors occupying the lowland forests of South America (node I). This ancestor gave origin to two lineages, one lowland forest dwellers (node K), and other open humid dwellers (node M). The later ancestral originated other two main lineages, one still an open humid occupant (node R), and the other an open dry inhabitant (node O). The next cladogenetic step originated other two main lineages, one with the ancestors distributed throughout the montane forest (clade N), and the other inhabiting the open dry areas of South America (clade O).

Results of range estimation of early nodes showed high probabilities, demonstrating lower uncertainty of biogeographic inferences, only the ancestral area of the clades composed by *S. angouya, L. molitor* and *M. transitorius* (MFLF), *Cerradomys* (DFODOH), and *Pseudoryzomys* and *Holochilus* (OH) presented the highest probability lower than 80% (Table 5). Other two most likely alternative states of ancestral ranges are listed in Table 5.

Figure 29 shows the most probable states for the ancestral range estimated with BayArea-like+J among the three distributional patterns (CIS-ANDES-TRANS). According to this figure the ancestral (node A) of the members of clade D inhabited the Cis-Andes part of South America. Most of the diversification history of clade D was in the Cis-Andean part of the continent, and the invasion of the Trans-Andean part of South America happened two times; first by the ancestors of the genera *Eremoryzomys* and *Drymoreomys* (node B), and then by the radiation composed by the genera *Oryzomys, Tanyuromys, Melanomys, Nesoryzomys,* and *Aegialomys* (node M), that dispersed to the Andean and Trans-Andean part of South America. Again, the results of range estimation of early nodes showed high probabilities, only the ancestral area of the clade composed by *Melanomys* and *Tanyuromys* (ANDES) presented the highest probability lower than 80% (Table 6). Other two most alternative states of ancestral ranges are listed in Table 6.

5.4. DISCUSSION

5.4.1. Phylogeny

The phylogeny of clade D recovered here presents some differences in the topology from the phylogenies previously available (Weksler, 2006; Turvey et al. 2010; Percequillo et al. 2011; Pine et al. 2012; Machado et al. 2013). *Eremoryzomys* and *Drymoreomys* are still sister genera and appear as the basal clade in the phylogeny. However, other topological relationships are quite distinct. First, *Cerradomys*, a genus widespread through Central South America, is the sister lineage of an inclusive clade including all other genera of the clade D, and not the sister group of *Sooretamys* as recovered in all other studies (op. cit.). On the other hand, *Sooretamys* is the sister taxa of a clade composed by *Lundomys* and *Microakodontomys*, a previously uncovered phyletic relationship. The reasons for that may be twofold: i) this is the first time that a study includes samples and evaluates the position of the genus *Microakodontomys* within the tribe Oryzomyini, ii) our comprehensive sampling both on taxonomic coverage and database allowed us to recover a more statistically supported tree. It is also important to note that the position of *Lundomys* is highly variable among all the proposed phylogenetic hypotheses presented, being typically recovered near *Holochilus* and *Pseudoryzomys*, other tetralophodont genera. This clade comprehends elements that are associated with eastern South America, one from the Atlantic Forest from Argentina, Brazil and Paraguay, one from the core Brazilian Cerrado and one from the Pampas and Atlantic Forest/Pampas transition on Brazil and Uruguay.

Another important difference is the position of the genus *Oryzomys*, that in other hypotheses sometimes appears: *i*) as a sister group of a clade composed by *Holochilus*, *Lundomys*, and *Pseudoryzomys* (Weksler, 2006); *ii*) as a sister genus of a larger clade composed by Andean genera *Melanomys*, *Sigmodontomys* and *Tanyuromys*, Trans-andean genera *Aegialomys* and *Nesoryzomys* and

Cis-Andean genus *Nectomys* (Pine et al. 2012); and *iii*) as the sister group of a clade formed by *Sooretamys* and *Cerradomys* (Machado et al. 2013). The relationship recovered here is the same as recovered by Percequillo et al (2011), *Oryzomys* as a sister group of a clade composed by Andean genera *Melanomys* and *Tanyuromys* and Trans-Andean genera *Aegialomys* and *Nesoryzomys*. The North American component of the tribe, *Oryzomys*, is sister to genera from the northern portion of the continent, and all these northern taxa belong to a clade sister to genus *Nectomys*, an irradiation of water rats widespread in South America.

Finally, genus *Holochilus* appears as a sister group of *Pseudoryzomys,* and this clade shows a sister relationship with a larger clade composed by Trans-Andean genera *Oryzomys*, *Aegialomys* and *Nesoryzomys,* Andean genera *Melanomys* and *Tanyuromys* and a Cis-Andean genus, *Nectomys.* The relationship of *Holochilus* was most of the time linked to the other tetralophodont genera of clade D (*Pseudoryzomys* and *Lundomys*), although the relationship among them is variable. Studies shows that using morphological data only *Holochilus* has a sister relationship with *Lundomys*, and then *Pseudoryzomys* (Weksler, 2006). Analyses based on the combination of IRBP and morphology (Weksler, 2006; Turvey et al. 2010), or IRBP, Cytb and morphology (Percequillo et al. 2011) also presented *Lundomys* as sister group to *Holochilus,* and this clade sister of *Pseudoryzomys*. Although, when IRBP data only (Weksler, 2006) or IRBP, Cytb and 12S (Pine et al. 2012) are used, *Pseudoryzomys* appears as sister to *Holochilus,* and this clade sister of *Lundomys*. Machado et al. (2013) was the only study that proposed the sister relationship of *Holochilus* and *Pseudoryzomys* and the non-sister relationship of this clade with *Lundomys.*

5.4.2. Tempo and mode of evolution

In general the diversifications dates presented in this study are older than the dates proposed by others studies (e.g. Parada et al. 2013; Machado et al. 2013; Leite et al. 2014), although the confidence intervals are overlapped. For example, the crown age of the clade D here is set as 8.8 Ma in the Late Miocene, although others studies estimated this date between 3.31- 5.5 Ma (Parada et al. 2013; Machado et al. 2013; Leite et al. 2014). The crown age of *Holochilus* date from 2.12 Ma. Parada et al. (2013) suggested a date for the split between *Holochilus* and *Pseudoryzomys* of *ca*. 3.5 Ma, Machado et al. (2013) suggested *ca.* 1.9 Ma, Leite et al., (2014) suggested something between 2-3 Ma. The origin of the clade composed by Andean and Trans Andean genera dates from 5.02 Ma, which is also older than the dates provided by Parada et al. (*ca* 4.8 Ma; 2013), Machado et al. (*ca* 2-2.58 Ma; 2013), and Leite et al. (*ca* 3-4 Ma; 2014).

This discrepancy can be due to several distinct factors that can influence divergence-time estimation, as uncertainty in the phylogenetic tree, complexity of the model for the molecular clock, among others (Smith et al. 2010; Dornburg et al. 2012; Parham et al. 2012; Heath and Moore 2014;

Beaulieu et al. 2015; Kumar and Hedges 2016). Besides, there is also some discussion in the literature about the utility of the genomic datasets to divergence-time estimation, however Smith et al. (2017) conducted a "gene shopping" approach and found that every empirical dataset examined by them included gene with clock-like behavior, which allows the use of this kind of dataset to confidently estimate ages.

5.4.3. Impacts of climate change in the diversification of clade D

Even thought the ancestors of clade D started the diversification during the Late Miocene, the period with higher number of cladogenetic events in this clade is recent, occurring during the Quaternary. This is potentially related with the climatic fluctuations during this period, that was one of the most radicals compared to any other period (Vivo, 2008).

According to our analyses of ancestral geographic range, the ancestral of clade D was a taxon that inhabited the open humid and dry areas in the Cis-Andean part of South America about 8.8 Ma ago. In that period, the continent was experimenting significant palaeogeographic and tectonic changes, which can have lead to the arising of new dispersal routes and geographical barriers (Pavan and Marroig, 2017). During the Neogene, most of the cladogenetic events happened in the Cis-Andean part of the continent with inhabitants occurring in open and forested areas. The geographical composition of South America during the Middle-Late Miocene until Early Pliocene (*ca*. 17–3 Ma) changed considerably due to the successive marine transgressions (Ortiz-Jaureguizar and Cladera, 2006). The formation of several "under water surfaces", as the *Paranean* and *Amazonian Sea*, with the posterior replacement of these widespread flooded areas by likewise spread plains (Marshall et al. 1983; Pascual et al. 1996), possibly facilitates the diversification process of clade D, and the occupation and expansion to large humid areas by its ancestors during this period.

Apart from the Cis-Andean cladogenetic events, the ancestral taxa of the clade composed by *Oryzomys*, *Tanyuromys, Melanomys, Nesoryzomys* and *Aegialomys*, and the ancestor of the clade composed by *Eremoryzomys* and *Drymoreomys* left the Cis-Andean region and dispersed to the Andes and Trans-Andean part of the continent. There are others evidences in the literature of this east-west pattern of relationship among mammals separated by the Andes (e.g. Cortes-Ortiz et al. 2003; Weksler, 2006; Solari et al. 2009; Gutiérrez et al. 2010), suggesting that dispersal events throughout this regions is not uncommon, and can be related with numerous geologic and tectonic, as well as paleoclimatic events.

According to Machado et al. (2013), the migration route took by these taxa was by the extreme northern South America, during the time when the latest Andes uplift events occurred, and the Andes could have played a role as a post dispersal barrier. Of all the genera that inhabits the Trans-Andean region, *Eremoryzomys, Aegialomys* and *Nesoryzomys* are endemic to the northern portion of

South America continent, and *Oryzomys, Melanomys* and *Tanyuromys*, also northern genera, expanded their distribution throughout Central America and even North America (as the case of *Oryzomys*). Our results shows that the ancestors of *Melanomys* and *Tanyuromys* were distributed by the Andes region about 3.2 Ma, right before the complete formation of the Isthmus of Panama (*ca*. 2.8 Ma; O'Dea et al. 2016), and the ancestors of genus *Oryzomys* inhabited the Trans-Andean part of the Neotropics about 2.07 Ma, right after the complete formation of the Isthmus of Panama. This fact suggest the Isthmus, with its gradually emergence of land and seaways constrictions, could had played an important role in the diversification pattern of clade D taxa, opening a new dispersal route for these rodents. The ancestor of *Eremoryzomys* and *Drymoreomys* inhabited the dry-forested areas in the Cis-Andean part of the continent about 5.1 Ma, possibly the climatic oscillations presented during the Neogene (described above) was the responsible to open dispersal routes to these ancestors cross the Andes and reach the Trans-Andean portion.

Even with the all events occurred during the Neogene, was during the Pleistocene that the clade experienced its larger diversification activity, with the emergence of most of the current genera. This period was marked by severe climatic oscillations, due to successive glacial–interglacial cycles, instead of geological events as anterior periods (Ortiz-Jaureguizar and Cladera, 2006). At this time, the cyclical expansion and shrink of arid (savannas, steppes) and humid (tropical and subtropical forests) biomes gave opportunities for species occupy new habitats, while at the same time favored great extinctions as a consequence of the reduction of suitable areas (Marshall and Cifelli, 1990). Our results show that most of the lineages were also Cis-Andean inhabitants during Pleistocene-Holocene. The phytophysiognomies occupied by the ancestors that lived during this time were also more diverse then during the Neogene. Our analyzes show that the ancestors were distributed thought the most part of vegetation types, as Open humid and dry areas, and Lowland and Montane forests.

The genera are also current distributed throughout different types of vegetation. The most generalist genus is *Cerradomys*, that is current find in Open and Forested areas in the Cis-Andean part of South America. But all the other genera are more specialists to each type of environment, as *Holochilus*, *Microakodontomys*, *Lundomys* and *Oryzomys* that occurs mostly in the open humid areas of the continent; *Aegialomys* is an open dry area inhabitant; *Drymoreomys*, *Eremoryzomys, Sooretamys, Nectomys, Tanyuromys*, and *Melanomys* are forest specialist genera.

Although we did not measure the diversification rate inside the clade, there is a strong association between open areas dwellers and number of species, which is another important evidence of how the climatic changes and the expansion of open areas played an important role in the diversification of this clade. The clade presents sixteen genera, which eight are monotypic (Weksler, 2015; Uturunco and Pacheco, 2016), and inside the monotypic genera only *Lundomys, Microakodontomys* and *Pseudoryzomys* are open areas inhabitants.

In summary, the diversification pattern inside the clade D of the tribe Oryzomyini suggest a complex biogeography history highly related with the expansions and retractions of the open areas in

South America mainly during the Pleistocene period, and the latest Andes uplift events occurred. This scenario is in agreement with past studies (Machado et al. 2013), and the large volume of molecular data presented here provides reliable and detailed knowledge of the evolutionary relationship among genera, with strong phylogentic signal. Our also study highlights the potential use of RADseq markers to estimate fine-scale divergence times presenting an interesting model of speciation for this rodents.

Acknowledgements

This work was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP; Proc: 2014/22444-0, 2012/24099-3, 2009/16009-1). We are also grateful to the curators and researchers that generously granted us tissues samples for the next generation sequencing. We also would like to thanks Andrea Thomaz, Luciana Resende, Anna Papadopoulou and Jen-Pan Huang for all the help with libraries preparation and bioinformatics processing of genomic data, Jeronymo Dalapicolla and Edson F. de Abreu-Junior for the help with divergence time estimation, and Ana Carolina Pavan for the help with biogeographic analyzes.

REFERENCES

Beaulieu, J. M., O'Meara, B. C., Crane, P., Donoghue, M. J., 2015. Heterogeneous rates of molecular evolution and diversication could explain the triassic age estimate for angiosperms. Syst. Bio. 64, 869–878.

Carleton, M. D. Olson, S.L., 1999. Amerigo Vespucci and the rat of Fernando de Noronha: a new genus and species of Rodentia (Muridae: Sigmodontinae) from a volcanic island of Brazil's continental shelf. Am. Mus.Nov. 3256, 1–59.

Carmignotto, A.P., Vivo, M. de, Langguth, A., 2012. Mammals of the Cerrado and Caatinga: distribution patterns of the tropical open biomes of Central America, in: Patterson, B.D., Costa, L. (Eds), Bones, clones, and biomes. University of Chicago Press, pp. 307–350.

Carnaval, A.C., Hickerson, M. J., Haddad, C.F.B., Rodrigues, M.T., Moritz, C. 2009. Stability Predicts Genetic Diversity in the Brazilian Atlantic Forest Hotspot. Science. 323, 785–789.

Cortés-Ortiz, L., Bermingham, E., Rico, C., Rodríguez-Luna, E., Sampaio, I., Ruiz-García, M., 2003. Molecular systematics and biogeography of the Neotropical monkey genus, Alouatta. Mol Phylogenet Evol. 26, 64–81.

D'Elía, G., Pardiñas, U.U.F.J., Teta, P., Patton, J.L.J., 2007. Definition and diagnosis of a new tribe of sigmodontine rodents (Cricetidae: Sigmodontinae), and a revised classification of the subfamily. Gayana. 71, 187–194.

D'Elía, G., Pardiñas, U.F.J., 2015. Subfamily Sigmodontinae Wagner, 1843, in: Patton J.L, Pardiñas U.F.J., D'Elía G. (Eds.), Mammals of South America, Vol. 2, Rodents. University of Chicago Press, Chicago, pp. 63–73.

Dornburg, A., Brandley, M. C., McGowen, M. R., Near, T. J., 2012. Relaxed clocks and inferences of heterogeneous patterns of nucleotide substitution and divergence time estimates across whales and dolphins (mammalia: Cetacea). Mol. Biol. Evol. 29, 721–736.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973.

Grehan, J., 2001. Biogeography and evolution of the Galapagos: integration of the biological and geological evidence. Biol J Linn Soc. 74, 267–287.

Gutiérrez, E.E., Jansa, S.A., Voss, R. S., 2010. Molecular systematics of mouse opossums (Didelphidae: Marmosa): assessing species limits using mitochondrial DNA Sequences, with comments on phylogenetic relationships and biogeography. Am. Mus.Nov. 3692, 1–22.

van der Hammen, T. 1974. The Pleistocene Changes of Vegetation and Climate in Tropical South America. J. Biogeogr. 1, 3–26.

Heath, T. A., Moore, B.R., 2014. Bayesian inference of species divergence times, in: Chen, M.H., Kuo, L., Lewis, P. O. (Eds.), Bayesian Phylogenetics: Methods Algorithms, and Applications. CRC Press, Boca Raton, pp 277–318.

Hershkovitz, P., 1962. Evolution of Neotropical cricetine rodents (Muridae) with special reference to the phyllotine group. Fieldiana Zoology. 46, 1–524.

Hershkovitz, P., 1966. South American swamp and fossorial rats of the scapteromyine group (Cricetinae, Muridae), with comments on the glans penis in murid taxonomy. Z. Saugetierkd. 31, 81–149.

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

Huang, J-P., 2016. Parapatric genetic introgression and phenotypic assimilation: testing conditions for introgression between Hercules beetles (Dynastes, Dynastinae). Mol Ecol. 25, 5513–5526.

Janis, C., 2003. Tectonics, Climate Change, and the Evolution of Mammalian Ecosystems, in: Rothschild, L., Lister, A. (Eds.), Evolution on the Planet Earth: Impact of the Physical Environment. London Academic Press, pp. 319–338.

Kumar, S., Hedges, S. B., 1998. A molecular timescale for vertebrate evolution. Nature. 392, 917–920.

Landis, M.J., Matzke, N.J., Moore, B.R., Huelsenbeck, J.P., 2013. Bayesian analysis of biogeography when the number of areas is large. Syst. Biol. 62, 789–804.

Leite, R.N., Kolokotronis, S.O., Almeida, F.C., Werneck, F.P., Rogers, D.S., Weksler, M., 2014. In the wake of invasion: tracing the historical biogeography of the South American cricetid radiation (Rodentia, Sigmodontinae). PLoS One. 9, e100687.

Machado, L.F., Leite, Y.L.R., Christoff, A.U., Giugliano, L.G., 2013. Phylogeny and biogeography of tetralophodont rodents of the tribe Oryzomyini (Cricetidae: Sigmodontinae). Zool Scripta. 43, 119–130.

Marshall, L.G., Cifelli, R.L., 1990. Analysis of changing diversity patters in Cenozoic land mammal age faunas, South America. Palaeovertebrata. 19, 169–210.

Marshall, L.G., Hoffstetter, R., Pascual, R., 1983. Mammals and stratigraphy: geochronology of the continental mammal-bearing Tertiary of South America. Palaeovertebrata, 1–93.

Matzke, N., 2013. BioGeoBEARS: Biogeography with Bayesian (and Likelihood) Evolutionary Analysis in R Scripts. University of California, Berkeley, Berkeley, CA.

Morgan, K., O'Loughlin, S.M., Chen, B., Linton, Y.M., Thongwat, D., Somboon, P., Fong, M.Y., Butlin, R., Ver- ity, R., Prakash, A., Htun, P.T., Hlaing, T., Nambanya, S., Socheat, D., Dinh, T.H., Walton, C., 2011. Comparative phylogeography reveals a shared impact of Pleistocene environmental change in shaping genetic diversity within nine Anopheles mosquito species across the Indo-Burma biodiversity hotspot. Mol. Ecol. 20, 4533–4549.

Ortiz-Jaureguizar, E., Cladera, G.A., 2006. Paleoenvironmental evolution of southern South America during the Cenozoic. J Arid Environ. 66, 498–532.

Parada, A., Pardiñas, U. F. J., Salazar-Bravo, J., D'Elía, G., 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.

Parada, A., D'Elía, G., Palma, R.E., 2015. The influence of ecological and geographical context in the radiation of Neotropical sigmodontine rodents. BMC Evol. Biol. 15, 1–17.

Pardiñas, U.F.J., D'Elía, G., Ortiz, P.E. 2002. Sigmodontinos fósiles (Rodentia, Muroidea, Sigmodontinae) de América del Sur: estado actual de su conocimiento y prospective. Mastozool. Neotrop. 9, 209–252.

Parham, J. F., Donoghue, P. C. J., Bell, C. J., Calway, T. D., Head, J. J., Holroyd, P. A., Inoue, J. G., Irmis, R. B., Joyce, W. G., Ksepka, D. T., Patane, J. S. L., Smith, N. D., Tarver, J. E., van Tuinen, M., Yang, Z., Angielczyk, K. D., Greenwood, J. M., Hipsley, C. A., Jacobs, L., Makovicky, P. J., Müller, J., Smith, K. T., Theodor, J. M., Warnock, R. C. M., Benton, M. J., 2012. Best practices for justifying fossil calibrations. Syst. Bio. 61, 346–359.

Pascual, R., Ortiz-Jaureguizar, E., Prado, J.L., 1996. Land mammals: paradigm of Cenozoic South American geobiotic evolution, in: Arratia, G. (Eds.), Contribution of Southern South America to Vertebrate Paleontology. Müncher Geowissenschaftliche Abhandlungen. pp. 265–319.

Pavan, A.C., Marroig, G., 2017. Timing and patterns of diversification in the Neotropical bat genus Pteronotus (Mormoopidae). Mol Phylogenet Evol. 108, 61–69.

Percequillo, A.R.; Weksler, M.; Costa, L.P. 2011. 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. 161, 357–390.

Peterson, B. K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PloS One. 7, e37135.

Pine, R.H., Timm, R.M., Weksler, M., 2012. A newly recognized clade of trans-Andean Oryzomyini (Rodentia: Cricetidae), with description of a new genus. J. Mammal. 93, 851–870.

Prado, J.R., Percequillo, A.R. 2013. On the geographic distribution of the genera of Tribe Oryzomyini on South America, with some comments on the patterns of diversity and Reig's Areas of Original Differentiation. Arq. Zool. 44, 1–120.

R Development Core Team. 2014. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available at: http:// www.R-project.org.

Rambaut, A., Suchard, M.A., Xie, D., Drummond, A.J., 2014. Tracer v1. 6.

Ree, R.H., Smith, S.A., 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst. Biol. 57, 4–14.

Ronquist, F., 1997. Dispersal-vicariance analysis: a new approach to them quantification of historical biogeography. Syst. Biol. 46, 195–203.

Smith, S. A., Beaulieu, J. M., Donoghue, M. J., 2010. An uncorrelated relaxed-clock analysis suggests an earlier origin for owering plants. Proc. Natl Acad. Sci. USA. 107, 5897–5902.

Smith, S., Brown, J.W., Walker, J.F., 2017. So many genes, so little time: comments on divergence-time estimation in the genomic era. bioRxiv. Not Published.

Solari, S., Hoofer, S.R., Larsen, P.A., Brown, A.D., Bull, R.J., Guerrero, J.A., Ortega, J., Carrera, J.P., Bradley, R.D., Baker, R.J., 2009. Operational criteria for genetically defined species: an analysis of the diversification of the small fruit- eating bats, Dermanura (Phyllostomidae: Stenodermatinae). Acta chiropterol. 11, 279–288.

Soltis, D.E., Morris, A.B., McLachlan, J.S., Manos, P.S., Soltis, P.S., 2006. Comparative phylogeography of unglaciated eastern North America. Mol. Ecol. 15, 4261–4293.

Steppan, S.J. 1995. Revision of tribe Phyllotini (Rodentia: Sigmodontinae), with a phylogenetic hypothesis for the Sigmodontinae. Fieldiana Zoology. 80, 1–112.

Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Ver. 4. Sinauer, Sunderland, Massachusetts.

Taberlet, P., Fumagalli, L., Wust-Saucy, A.-G., Cosson, J.-F., 1998. Comparative phylogeography and postglacial colonization routes in Europe. Mol. Ecol. 7, 453–464.

Turvey, S.T., Weksler, M., Morris, E.L., Nokkert, M., 2010. Taxonomy, phylogeny, and diversity of the extinct Lesser Antillean rice rats (Sigmodontinae: Oryzomyini), with description of a new genus and species. Zool. J Linn Soc. 160, 748–772.

de Vivo, M. 2008. Mamíferos e mudanças climáticas, in: Buckeridge, M.S. (Eds.), Biologia e Mudanças Climáticas no Brasil. Rima Editora, pp. 207-23.

Voss, R.S., Carleton, M.D. 1993. A new genus for Hesperomys molitor Winge and Holochilus magnus Hershkovitz (Mammalia, Muridae) with an analysis of its phylogenetic relationships. Am. Mus.Nov. 3085, 1–39.

Wagenmakers, E.-J., Farrell, S., 2004. AIC model selection using Akaike weights. Psychon. Bull. Rev. 11, 192–196.

Weksler, M., 2003. Phylogeny of Neotropical oryzomyine rodents (Muridae: Sigmodontinae) based on the nuclear IRBP exon. Mol Phylogenet Evol. 28, 331–349.

Weksler, M., 2006. Phylogenetic relationships of oryzomyine rodents (Muroidea: Sigmodontinae): separate and combined analyses of morphological and molecular data. Bull. Am. Mus. Nat. Hist. 296, 1–149.

Weksler, M., 2015. Tribe Oryzomyini Vorontsov, 1959, in: Patton J.L, Pardiñas U.F.J., D'Elía G. (Eds.), Mammals of South America, Vol. 2, Rodents. University of Chicago Press, Chicago, pp. 63–73.

Werneck, F.P., Gamble, T., Colli, G.R., Rodrigues, M.T., Sites, J.W., 2012. Deep diversification and long-term persistence in the South American 'Dry Diagonal': integrating continent-wide phylogeography and distribution modeling of geckos. Evolution. 66, 3014–3034.

Woodburne, M.O., 2010. The Great American Biotic Interchange: Dispersals, Tectonics, Climate, Sea Level and Holding Pens. J Mamm Evol. 17, 245–264.

Figures

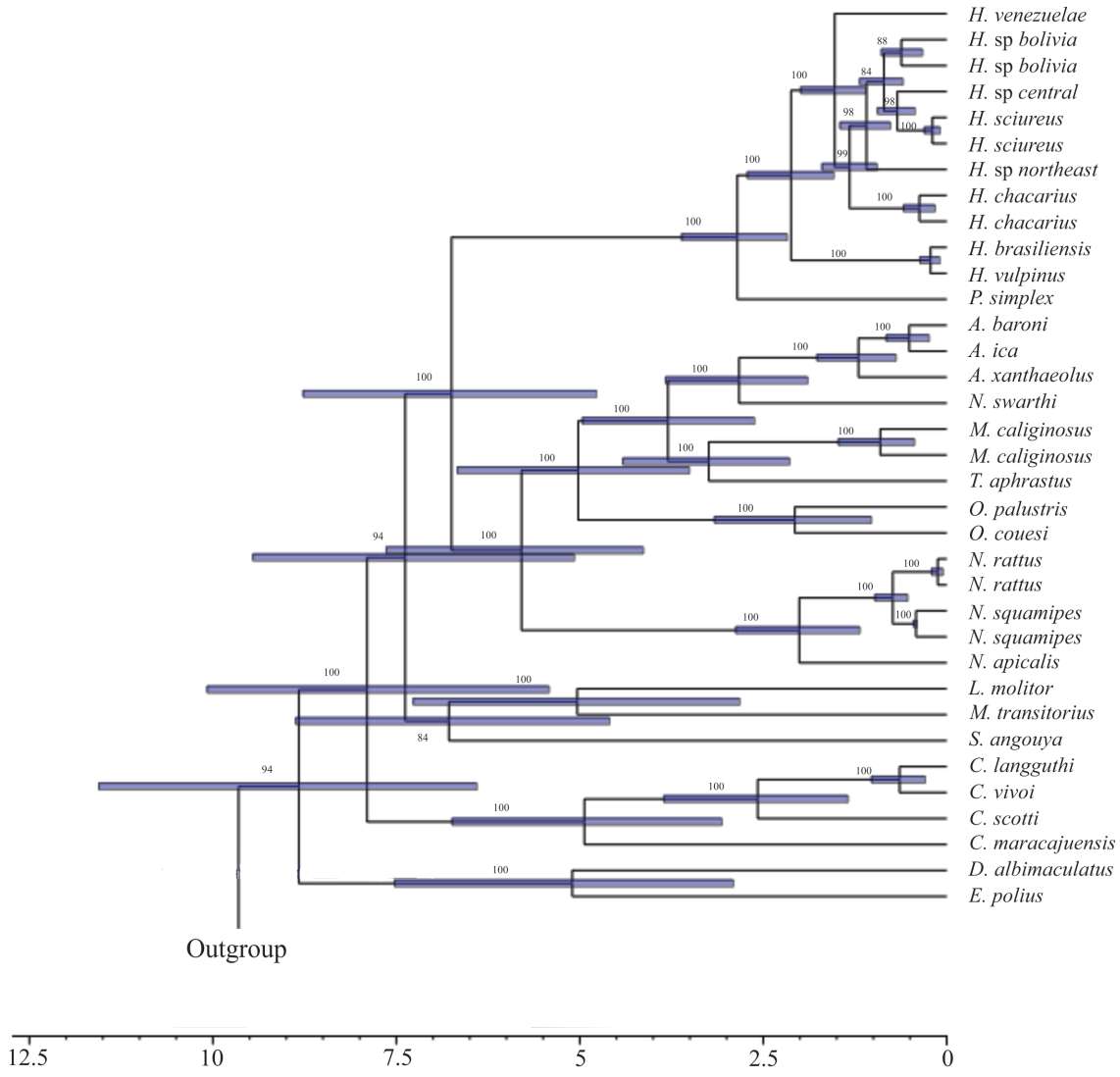


Figure 27. Time calibrated phylogeny in the Oryzomyini clade D. Bars correspond to the 95% High Posterior Density (HPD) time. Valeus on the branches correspond to the bootstrap values of SVDQuartets analysis.

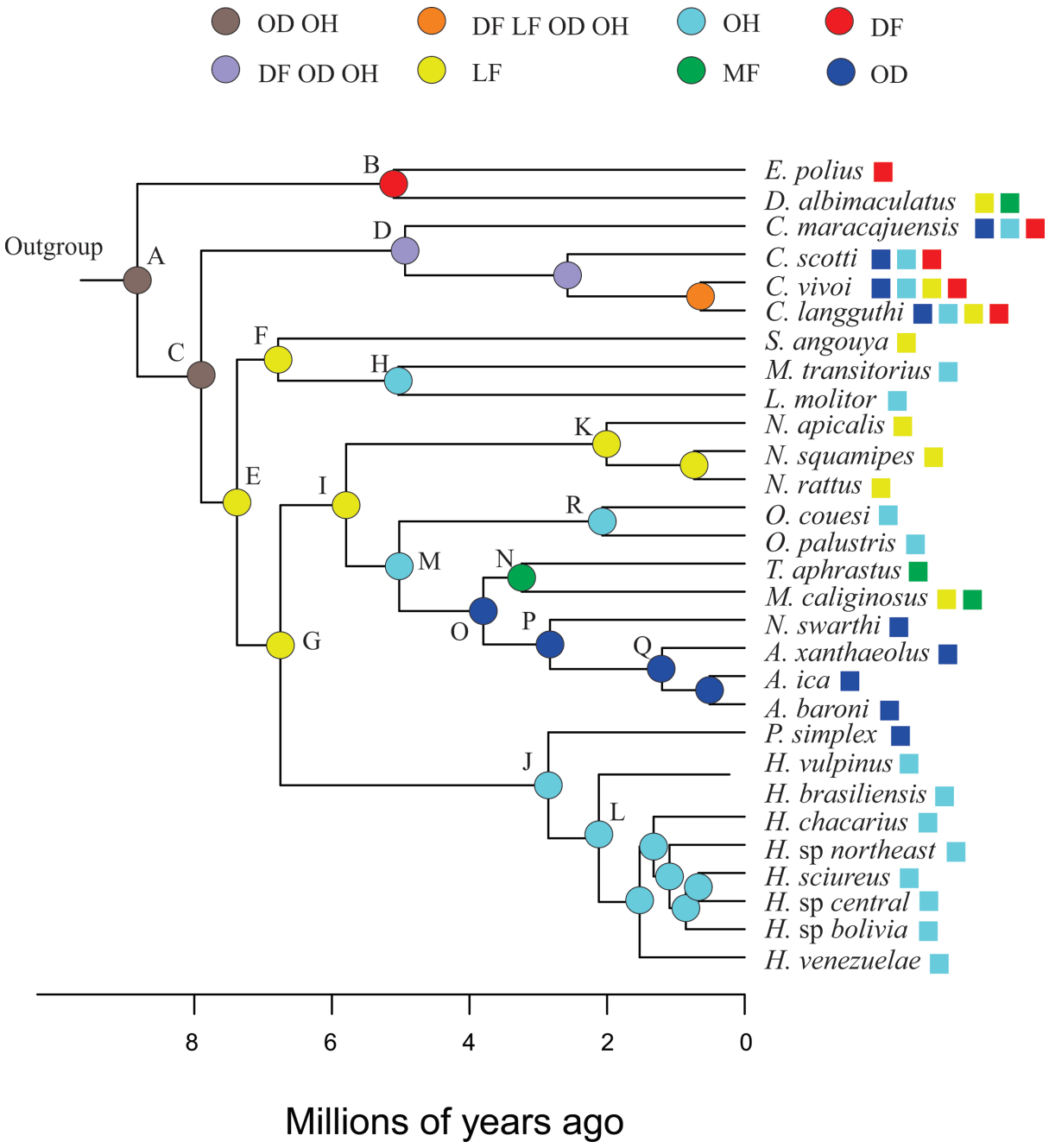


Figure 28. Geographic range evolution in the Oryzomyini clade D. Geographic range estimates with the highest marginal probabilities for ancestral nodes according to BayArea-like+J model. OD= Open Dry; OH= Open Humid; DF= Dry Forest; LF= Lowland Forest; MF= Montane Forest.

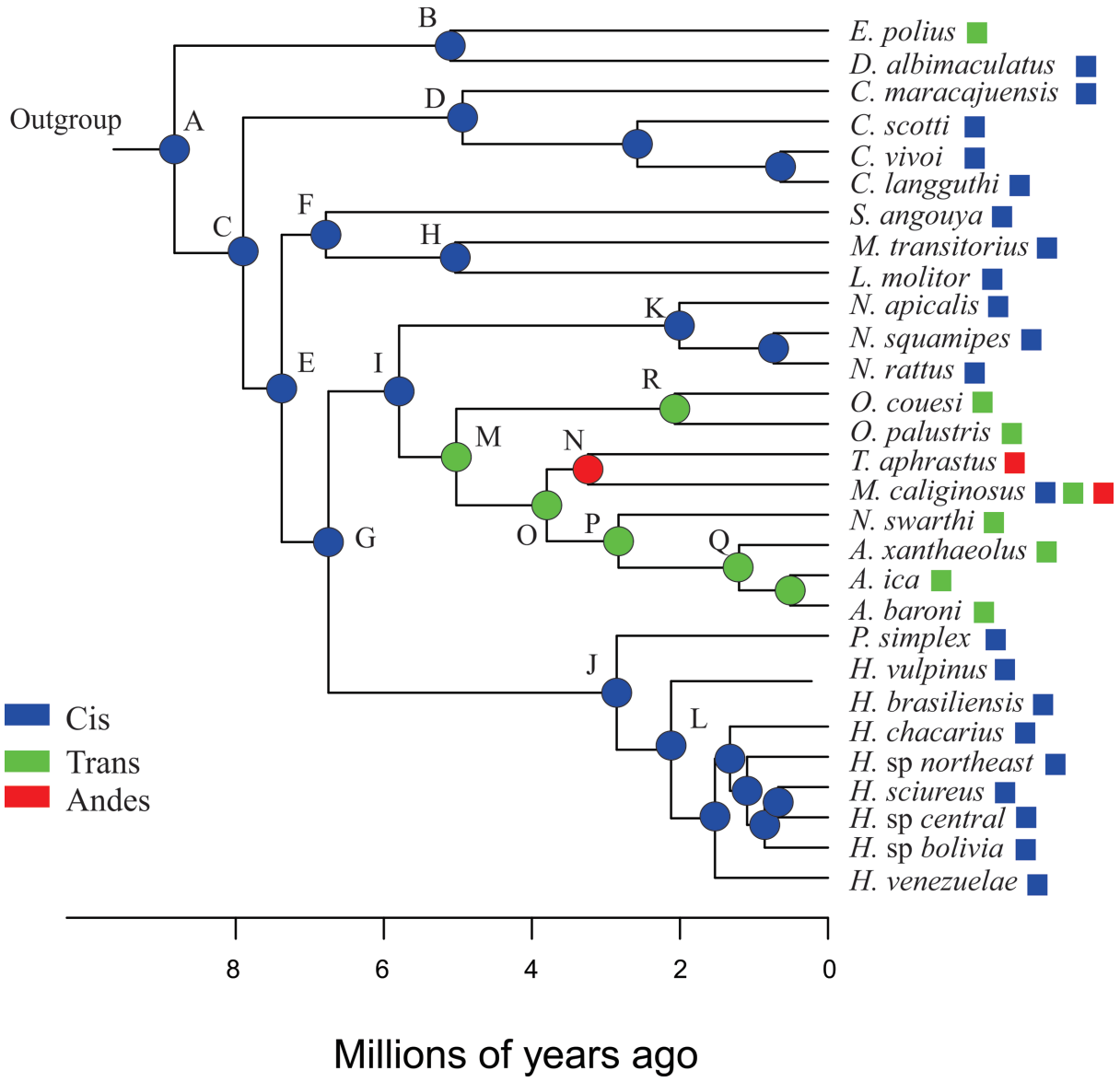


Figure 29. Geographic range evolution in the Oryzomyini clade D. Geographic range estimates with the highest marginal probabilities for ancestral nodes according to BayArea-like+J model.

Tables

Table 4. Models performance and Log-likelihood values of the six biogeographic models tested in the present study for Analysis 1 (five vegetation areas) and Analysis 2 (3 distributional areas). The Akaike Information Criterion (AIC) and Akaike weight (AICwt) are presented, the last being estimated just between the two better AIC values.

Analysis	Model	N of Parameter	LnL	AIC	AICwt
1	DEC	2	-87.269	178.53	0.645
	DIVA	2	-91.311	186.62	0.730
	BAYAREA	2	-72.495	148.99	0.142
	DEC+J	3	-86.869	179.73	0.354
	DIVA+J	3	-91.311	188.62	0.269
	BAYAREA+J	3	-69.700	145.40	**0.857**
2	DEC	2	-33.815	71.63	0.723
	DIVA	2	-35.823	75.64	0.71
	BAYAREA	2	-36.847	77.69	0.003
	DEC+J	3	-33.778	73.55	0.276
	DIVA+J	3	-35.739	77.47	0.28
	BAYAREA+J	3	-30.063	66.12	**0.99**

Table 5. Alternative ancestral ranges presenting the highest marginal probabilities in the BAYAREAj model for each node within clade D phylogeny. Dry+Ere = *D. albimaculatus* and *E. polius*; *Cerradomys* = *C. maracajuensis, C. scotti, C. vivoi*, and *C. langguthi*; Soo+Lun+Mic = *S. angouya, L. molitor* and *M. transitorius*; Lun+Mic = *L. molitor* and *M. transitorius; Nectomys = N. squamipes, N. rattus* and *N. apicalis*; *Oryzomys* = *O. palustris* and *O. couesi*; Mela+Tany = *M. caliginosus* and *T. aphrastus*; Aegi+Neso = *A. xanthaeolus, A. baroni, A. ica* and *N. swarthy*; Pseu+Holo = *P. simplex, H. vulpinus*, *H.* sp rj, *H. chacarius*, *H.* sp. alagoas, *H. sciureus*, *H.* sp. central, *H.* sp bolivia, *H. venezuelae*. *Holochilus = H. vulpinus*, *H.* sp rj, *H. chacarius*, *H.* sp. alagoas, *H. sciureus*, *H.* sp. central, *H.* sp bolivia, *H. venezuelae.* Area combinations: OH = Humid open area; OD = Dry open areas; MF = Montane Forest; LF= Lowland Forest; DF= Dry Forest.

Molecular dating (myr)	Phylogeny node	Highest probability	2nd highest	3rd highest
5.10	Dry+Ere	90,66% LFMF	3,08% DFOH	2,32% LFMFOH
4.93	*Cerradomys*	69,89% DFODOH	9% DFOH	8,60% DFOD
6.78	Soo+Lun+Mic	51,88% LFMF	29,04% OH	16,62% LF
5.03	Lun+Mic	98,62% OH	0,3% LF	0,2% ODOH
2	*Nectomys*	98,16% LF	1,71% MFLF	-
2.07	*Oryzomys*	99,70% OH	0,06% LFOH	-
3.24	Mela+Tany	97,83% LFMF	0,9% MF	0,9% LF
2.82	Aegi+Neso	99,79% OD	-	-
2.85	Pseu+Holo	51,59% OH	45,29% OD	0,95% LFOD
2.12	H*olochilus*	99,85% OH	-	-

Table 6. Alternative ancestral ranges presenting the highest marginal probabilities in the BAYAREAj model for each node within clade D phylogeny. Dry+Ere = *D. albimaculatus* and *E. polius*; *Cerradomys* = *C. maracajuensis, C. scotti, C. vivoi*, and *C. langguthi*; Soo+Lun+Mic = *S. angouya, L. molitor* and *M. transitorius*; Lun+Mic = *L. molitor* and *M. transitorius; Nectomys* = *N. squamipes, N. rattus* and *N. apicalis*; *Oryzomys* = *O. palustris* and *O. couesi*; Mela+Tany = *M. caliginosus* and *T. aphrastus*; Aegi+Neso = *A. xanthaeolus, A. baroni, A. ica* and *N. swarthy*; Pseu+Holo = *P. simplex, H. vulpinus*, *H.* sp rj, *H. chacarius*, *H.* sp. alagoas, *H. sciureus*, *H.* sp. central, *H.* sp bolivia, *H. venezuelae*. *Holochilus* = *H. vulpinus*, *H.* sp rj, *H. chacarius*, *H.* sp. alagoas, *H. sciureus*, *H.* sp. central, *H.* sp bolivia, *H. venezuelae.* Area combinations: CIS = Cis-Andes; TRANS= Trans-Andes, ANDES= Andes.

Molecular dating (myr)	Phylogeny node	Highest probability	2nd highest	3rd highest
5.10	Dry+Ere	99,94% CIS	-	-
4.93	*Cerradomys*	100% CIS	-	-
6.78	Soo+Lun+Mic	100% CIS	-	-
5.03	Lun+Mic	100% CIS	-	-
2	*Nectomys*	100% CIS	-	-
2.07	*Oryzomys*	100% TRANS	-	-
3.24	Mela+Tany	69,37% ANDES	28,80% TRANS	1,52% CISTRANS
2.82	Aegi+Neso	100% TRANS	-	-
2.85	Pseu+Holo	100% CIS	-	-
2.12	H*olochilus*	100% CIS	-	-

APPENDIX M. Summaries of samples collected

Details about vouchers (locality and Catalog Number) of sequenced individuals are provided in this appendix, Appendix M.

List of specimens used in the molecular analyses: Aegialomys baroni: **PERU:** Lambayeque: Las Juntas, in Quebrada La Pachinga, *ca*. 14 km N, 25 km E Olmos: M944. *Aegialomys ica*: **PERU:** Arequipa: 8.5 mi NNE Bella Union: MVZ145540. *Aegialomys xanthaeolus*: **ECUADOR:** El Oro: Zaruma: TK135790. *Cerradomys langguthi*, **BRAZIL:** Ceará: Sítio Páscoa, 5 km NW Crato: MVZ197804. *Cerradomys maracajuensis*: **BRAZIL:** Mato Grosso do Sul: Eldorado: MBR26. *Cerradomys scotti*: **BRAZIL:** Tocantins: Santa Rita do Tocantins: LTXTO21R. *Cerradomys vivoi*: **BRAZIL:** Bahia: Fazenda Bolandeira, 10 km S Una: MVZ197607. *Drymoreomys albimaculatus*: **BRAZIL:** São Paulo: Estação Ecológica do Bananal: EEB674. *Eremoryzomys polius*: **PERU:** Amazonas: Chachapoyas, Balsas, 19 km by rd E: FMNH129243. *Holochilus brasilienesis*: **BRAZIL:** Rio de Janeiro: Imbaú, Silva Jardim: LBCE1222. *Holochilus chacarius*: **ARGENTINA:** Corrientes: 0.5 km N of Itati, Island in Rio Parana: UMMZ166514. **PARAGUAY:** Presidente Hayes: 24 km NW (by air) of Villa Hayes, Estancia la Golondrina: UMMZ165996. *Holochilus sciureus*: **BRAZIL:** Amazonas: Altamira, right bank Rio Juruá: MVZ193736. **FRENCH GUIANA**: Régina: Kaw: T4595. *Holochilus* sp northeast alagoas: **BRAZIL**: Alagoas: ESEC Murici, Murici: MUFAL0023. *Holochilus sp bolivia*: **BOLIVIA:** Santa Cruz: 6 KM by road W of Ascencion, 240m: MSB211433. 3.5 KM W Estacion Pailon: MSB55298. *Holochilus sp central*: **BRAZIL**: Piauí: Ribeiro Gonçalves: M24. *Holochilus venezuelae*: **VENEZUELA**: Sucre: Arismendi, Finca Vuelta Larga, 9.7 km by road southeast of Guaraunos: AMNH257336. *Holochilus vulpinus*: **BRAZIL:** Rio Grande do Sul: Rio Grande: MCNU3427. *Lundomys molitor*: **BRAZIL**: Rio Grande do Sul: Dom Pedrito: MCNU2302. *Melanomys caliginosus*: **ECUADOR:** Esmeraldas: Comuna San Francisco de Bogota: TK135640. **NICARAGUA**: Atlantico Norte: Siuna, Rosa Grande: TK121431. *Microakodontomys transitorius*: **BRAZIL:** Distrito Federal: Brasilia, Centro de Instrução e Adestramento de Brasilia (CIAB): APC878. *Necromys lasiurus*: **BRAZIL**: Rondonia: UHE JIRAU, right bank of Rio Madeira, Porto Velho: MJ092. *Nectomys apicalis*: **PERU**: Amazonas: Huampami, Rio Cenepa: MVZ153539. *Nectomys rattus*: **BRAZIL**: Para: Vitória do Xingu, Paratizão, left bank of lower Rio Xingu: CN119. Piauí: Estação Ecológica Uruçuí-Una: MZUSP 30331. *Nectomys squamipes*: **BRAZIL:** Bahia: Jussari, RPPN Serra do Teimoso: MZUSP 29788. Rio Grande do Sul: Nonoai: MCNU1528. *Nesoryzomys swarthy*: **ECUADOR**: Galapagos Island: Isla Santiago, La Bomba: FMNH179527. *Oligoryzomys nigripes*: **BRAZIL**: São Paulo: Estação Ecológica do Bananal: EEB714. *Oryzomys couesi*: **MEXICO:** Veracruz: LSUMZ7796. *Oryzomys palustres*: **UNITED STATES:** Louisiana: LSUMZ135. *Pseudoryzomys simplex*: **BRAZIL:** Tocantins: Parque Nacional do Araguaia: PNA022. *Scolomys ucayalensis*: **BRAZIL:** Amazonas: Vila de Santa Fé, Japurá, right bank of Rio Japurá: JAP196, JAP214. *Sooretamys angouya*: **BRAZIL:** Rio Grande do Sul: Nova Roma do Sul: MCNU1230. *Tanyuromys aphrastus*: **COSTA RICA:** Alajuela: Monteverde Cloud Forest Reserve, Sendero a Penas Blancas: KU161003. *Zygodontomys brevicauda*: **VENEZUELA**: Portuguesa: Guanarito, La Arenosa: TK53539.

APPENDIX N. Summaries of genomic data

Processing information and pyRAD summary statistics for species sequenced on the Illumina platform (Table N.1), as well as the proportion of loci shared among individuals (Figure N.1), are provided in this appendix, Appendix N.

Table N.1. Processing information and pyRAD summary statistics for species sequenced. Raw reads refers to the total reads produced during Illumina sequencing; post-processing reads are those that remained after filtering for quality, adaptor contamination, ambiguous barcodes, and mitochondrial sequences; total clusters are the number of homologous sequences clusters created with the post-processing reads; mean depth is the mean depth of coverage of a cluster. Heterozygosity (H) and error-rate (E) were estimated across clusters, and consensus sequences were created for each cluster. Variable and invariable DNA sites were summed across all loci (Total sites), and the percentage of polymorphic sites (Percent poly) is presented. Consensus sequences were clustered across species, and loci that passed filtering parameters were included in the final data matrix (Final loci).

Species	Raw reads	Post-processing reads	Percent retained	Total clusters	Mean depth of clusters	H	E	Total sites	Percent poly	Final loci
A. baroni	1056151	962795	91.16	217144	4.326	0.00749575	0.00147416	8343003	0.0042	34356
A. ica	2024766	1847848	91.26	335746	5.347	0.00622221	0.00123986	12323377	0.0032	50021
A. xanthaeolus	5480230	4901827	89.44	466272	9.756	0.00881117	0.0012841	16413192	0.0043	58642
C. langguthi	2461487	2249993	91.40	377406	5.817	0.00423247	0.00074717	12044014	0.0021	40664
C. maracajuensis	2964223	2706753	91.31	376757	7.087	0.00360083	0.00070832	13902738	0.0017	47513
C. scotti	444615	407523	91.65	156402	2.54	0.00526906	0.00233417	3835073	0.0019	13495
C. vivoi	3402134	3062540	90.01	349286	8.562	0.00450354	0.00066758	13308698	0.0024	44727
D. albimaculatus	2253274	2046828	90.83	363576	5.44	0.00669448	0.00142775	12404229	0.0032	34566
E. polius	3009261	2710802	90.08	328325	8.125	0.00383148	0.00074047	12118305	0.0014	33409
H. chacarius	5451086	4848398	88.94	421952	11.247	0.00767362	0.00051686	16924009	0.0047	72109
H. chacarius	6128086	5415474	88.37	505325	10.488	0.00743361	0.00036344	18899462	0.0041	75069
H. sciureus	833947	729325	87.45	200092	3.569	0.00582442	0.00075233	7488156	0.0031	37262
H. sciureus	887557	723202	81.48	293397	2.401	0.00535621	0.0006368	12147789	0.0031	58363
H. venezuelae	1410252	1238723	87.83	253970	4.786	0.00401055	0.00133896	6837831	0.0011	31649
H. vulpinus	1728533	1550859	89.72	287036	5.279	0.00503872	0.00057909	10266131	0.0028	46777
H. sp. alagoas	2919063	2556888	87.59	324029	7.71	0.00405603	0.00058514	12834992	0.0019	57696
H. sp. bolivia	5458159	4924676	90.22	412733	11.613	0.00864072	0.00054971	16212782	0.0054	67399
H. sp. bolivia	5028161	4427286	88.04	412034	10.508	0.00899452	0.00038257	16635498	0.0056	71802
H. sp central	926967	823421	88.82	199964	4.024	0.00767926	0.00066052	8025910	0.0047	39229
H. sp. rj	3058424	2674000	87.43	321673	8.101	0.0061386	0.00106525	12023777	0.033	53316
L. molitor	2141130	1938262	90.52	309262	6.136	0.00350795	0.00082344	11429189	0.0016	40548

M. caliginosus	4983442	4414702	88.58	489126	8.689	0.00556874	0.00087499	17762328	0.0023	62337
M. caliginosus	4917026	4457300	90.65	446827	9.634	0.00804638	0.00076583	16028576	0.0043	60015
M. transitorius	2931808	2679672	91.39	455862	5.765	0.00385851	0.00081334	19182489	0.0019	47852
N. lasiurus	2714311	2442261	89.97	356652	6.667	0.00531613	0.00083623	13096264	0.0030	14796
N. apicalis	2695459	2403370	89.16	353392	6.714	0.00333949	0.00075696	12888888	0.0013	54487
N. rattus	3266416	2975530	91.09	378942	7.678	0.00404027	0.00077441	13371710	0.0017	56676
N. rattus	8054098	7280558	90.39	738479	9.657	0.00667785	0.0007782	27062971	0.0015	78981
N. squamipes	711766	650264	91.35	207625	3.078	0.00323682	0.00133408	6698475	0.0010	32017
N. squamipes	2037951	1866271	91.57	310495	5.844	0.00374515	0.00096844	11540542	0.0016	50690
N. swarthy	7063533	6400505	90.61	538238	11.76	0.00312532	0.00051789	21803244	0.0008	72036
O. nigripes	1176663	1075257	91.38	246439	4.161	0.00882576	0.00133109	9433548	0.0052867	26158
O. couesi	4985791	4482464	89.90	459653	9.431	0.00922612	0.00096239	16210448	0.005107	54179
O. palustris	6830047	6125236	89.68	571632	10.396	0.00625117	0.00076043	20312548	0.0025843	61035
P. simplex	2776057	2539069	91.46	367816	6.749	0.00624694	0.00069401	13197134	0.0037944	48975
S. ucayalensis	2791977	2557111	91.58	370456	6.302	0.00622504	0.00183422	12900081	0.0027884	20443
S. ucayalensis	1662608	1515048	91.12	284594	4.906	0.00561127	0.00184162	10381882	0.0026926	17117
S. angouya	1644863	1500654	91.23	305242	4.796	0.00543032	0.00110968	10632923	0.0031146	35276
T. aphratus	4342841	3921898	90.30	412947	9.21	0.00423726	0.00076377	14719853	0.0014463	55384
Z. brevicauda	6939081	6241926	89.95	535740	11.11	0.00534115	0.0005564	19906514	0.0026027	46633

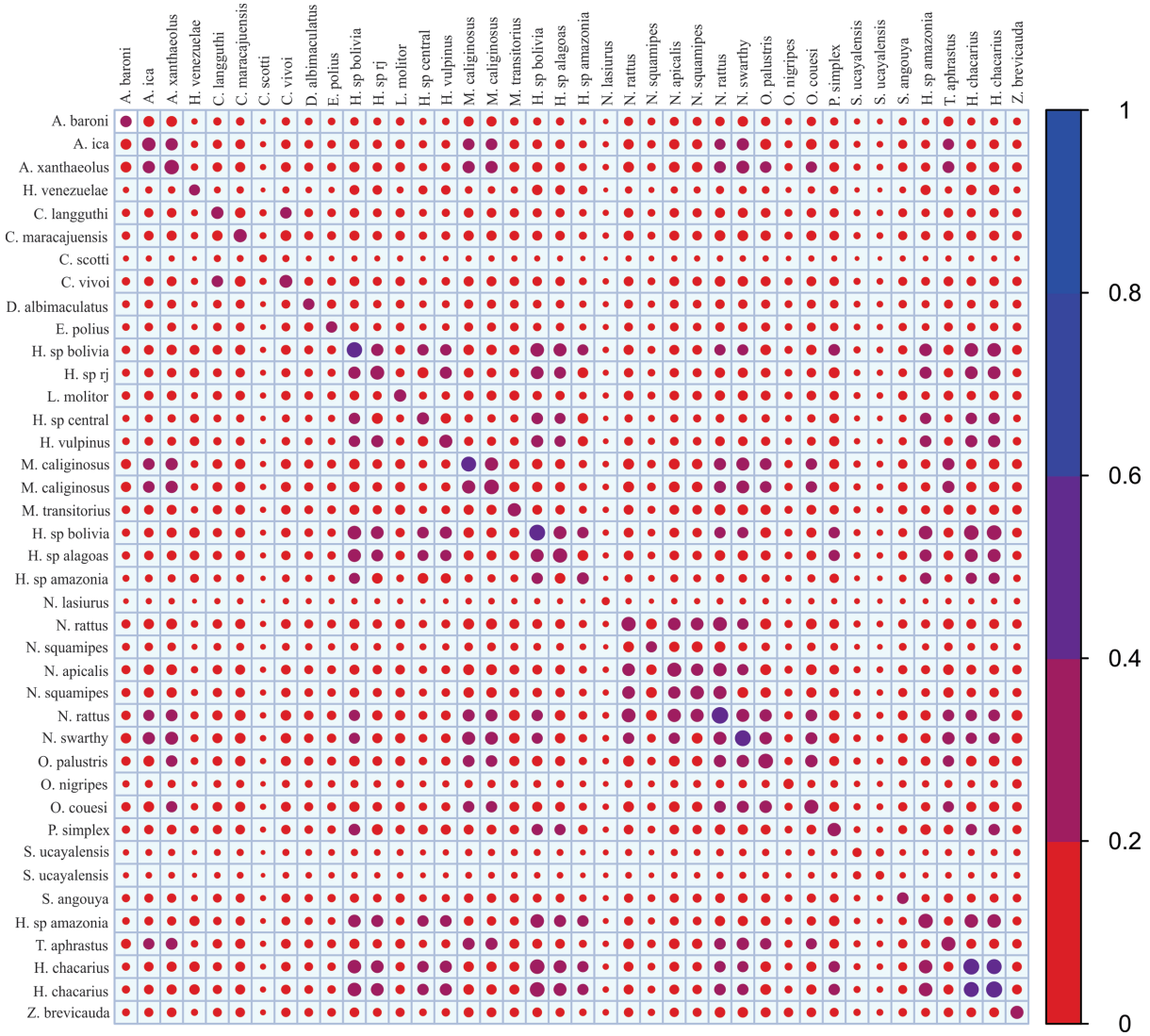


Figure N.1. Loci shared among individuals, where the areas and colors of circles show the proportion of shared loci (from 0 to 1) between individuals (off-diagonal cells) or successfully amplified within an individual (diagonal cells).

6. SYNTHESIS AND RECOMMENDATIONS

This thesis attempted to apply an interdisciplinary approach integrating micro and macroevolution, genomic and morphometric/morphological variation, systematics, quantitative genetics, and biogeography to investigate the evolutionary processes in genus *Holochilus*.

Based on the most comprehensive sampling already gathered, I applied powerful methods to study these evolutionary processes, such as: qualitative, univariate, and multivariate methods to compare morphometric/morphological variation; multivariate methods to compare genomic variation; ecological niche modeling to identify suitable areas along time; coalescent species delimitation methods integrating dataset of genomic and morphometric data; ancestral geographic ranges estimation, among others.

I focused on two major approaches to the problems aforementioned, evaluating intraspecific and interspecific variation on the genus, as well as the position of this genus inside the major Oryzomyini clade D.

6.1. Intra and Interspecific Variation Patterns

Intraspecific Variation was accessed by morphological/morphometric (Chapter 1) and genomic approaches (Chapter 2). The patterns were then compared among populations/species and correlations with extrinsic and intrinsic factors that could explain the patterns were made.

The non-geographic variation analysis were focused in two different species plus a captive population, and showed that the percentage of sexual variation is low in most of the variables and in all populations, and it is inside of the range of individual variation, and therefore not represents real sexual dimorphism, corroborating others studies with different Oryzomyini rodents. I also show that significant differences between age classes are more conspicuous in the younger age classes (1 and 2) in all populations. Variables responsible for the most part of the variation are all related with skull size, nasal, rostrum, braincase and incisors, regions directly related with the bite apparatus.

Regarding to the ontogenetic trajectories, I observed different patterns of ontogenetic development among populations, differing both in magnitude and direction of the variation. The direction of the trajectories from the captive and wild populations (Par_Wild and Par_Captive) is slightly different, as well as the magnitude. The ontogenetic trajectories of Alagoas_Wild and Par_Wild populations showed similar direction but strong different magnitudes.

The genomic intraspecific analysis was focused in three species, and the values of population genetic summary statistics were broadly overlapping among them. The correlation between pairwise F_{ST}-values and the Euclidean geographic distances among populations was not significant in either species. The association between geography and genomic variation proved to be significant, although

there are differences in the strength of the association across species. Regarding to the geographic distribution of genomic variation within each species, several patterns are reinforced.

Results from the non-geographic variation analysis permitted select individuals with lower degree of variation, controlling this variation prior to the geographic morphometric analysis used in the species delimitation process. Data from the population's genomic analyzes provided important information about how this variation is structured along the geography, and gave me more confidence to assign individuals for the species delimitation model.

6.2. Species Delimitation

A detailed analysis of independent quantitative and genomic variation recovered congruent groups of populations, and these groups were used as the hypothesis in the species delimitation analysis. Species A in the iBPP analyzes comprehended individuals from *Minas Gerais* morphometric cluster and specimens from clade A of the species tree; species B included individuals from *South* morphometric cluster and individuals from clade B of the species tree; species C included samples from *Paraguay* morphometric cluster and individuals from clade C of the species tree; species D contained specimens from *Amazon* and *Guyana* morphometric clusters and specimens from clade D of the species tree; species E comprised the individuals from *Bolivia* morphometric cluster and specimens from clade E of the species tree; clade F comprised samples from the *Northeast* morphometric cluster and individual from clade F of the species tree; clade G comprehended samples from *Venezuela* morphometric cluster and individual from clade G of the species tree; and clade H encompassed individuals from *Central* morphometric cluster, and samples from clade H of the species tree.

Our results arouse from an integrative approach to objectively identify eight evolutionary independent lineages by analyzing different data types under a common statistical framework (phenotypic data and molecular data). Our results not only recovered more highly supported independent lineages, but also presented resolution among tips, things that were not achieved by previous studies.

This increase in the diversity (from six to eight species) and reshape of known distribution of collection localities for these species of South American marsh rats, brings important implications for the conservation status in this group, allowing future species descriptions and the formulation of conservation strategies.

6.3. Diversification Drivers

The most important factors structuring genomic variation within the genus *Holochilus,* considered by this study, are the environmental and climatic differences across space and time. I identified that the distribution of species of *Holochilus* recovered here is congruent with the distribution of the major terrestrial South America biomes/ecoregions, being the allopatric speciation with niche conservatism one of the possible explanation to lead the diversification in this group.

Regarding to the ontogenetic development, although the bioclimatic associations tests did not show that current differences among the climatic conditions is enough to explain the ontogenetic variance observed, there is a conspicuous variation in the degree of correlation between morphometric and environmental variables among populations and in each age classes.

Additionally, the genomic tests showed that the environmental characteristics of the biomes the three species inhabit clearly differ, not only geographically and environmentally (based on past climatic conditions). There is also significant association between the environmental space and the genetic variation that is not related with geography. The ENMs suggested difference in the size of stables habitats among species, where biomes with larger areas of stability also presented more genomic structure, suggesting that historical dimension impacted population isolation/connectivity.

Regarding to clade D, results presented here show that the climatic oscillations and environmental changes (expansions and retractions of open areas) are important sources of variation and diversification drivers. Although here I cited others factors as geologic and tectonics process (Andes uplift events and formation of Panama Isthmus) that were also important to explain the current biogeographic pattern in this clade.

6.4. Next Generation Sequencing

This study highlights how next-generation sequencing can provide information for delimit species and resolving relationships among taxa over a variety of temporal depths (within *Holochilus* and within clade D), including cases where the conflict between subsets of Sanger sequencing data is common.

For example, previous studies considered the samples from Santa Cruz (Bolivia) and Entre Rios (Argentina) as part of the species called by them as *H. sciureus*. In their concept *H. sciureus* comprehends all the specimens localized in Amazon biome, including individuals from Amazonas state in Brazil, Suriname, Peru, Bolivia and one individual from Entre Rios (Argentina). Here I show that samples from that region are actually split in three different species. Individuals from Entre Rios (Argentina) belong to species B, individuals from the northern part of Amazon biome belong to species D, and specimens from Bolivia belong to species E.

Regarding to divergence time estimation, in general the diversifications dates presented in this study are older than the dates proposed by others studies, but the credibility intervals are always overlapped. Thus even with some discussion in the literature about the utility of the genomic datasets to divergence-time estimation, I consider that the concatenated dataset used was able to confidently recover accurately dates.

6.5. Futures Perspectives

Future studies on the genus *Holochilus* might employ quantitative genetic approaches to estimate the relative contributions of genetic and environmental factors in the variation observed in the current study, as well as explain differences related with the ontogenetic trajectory between population, by comparing data from captive and wild populations generating by phenotypic P (collected in the wild without known genealogy) and a genotypic G (constructed from specimens grown in colony) matrices.

Additionally, further studies including methodologies that integrate distributional, demographic and genomic (e.g. iDDC) data, and the tempo of evolution, will be able to provide a framework to test the hypotheses raised in this thesis and other possibilities that connect difference in population dynamics over space, time and species-specific traits.

Formal taxonomic descriptions of the species recognized in this thesis, with the correct assignment of names taking in consideration the occurrence of type specimens available and providing a completely diagnose for each species are pending. Hopefully, with the evidence recovered in this thesis, I will soon be able to clarify the taxonomic and nomenclatural problems attributed to genus *Holochilus*.

Moreover, other studies on morphological variation, as binary, meristic, and geometric morphometric variables within *Holochilus* remain to be investigated. Future analysis with ancestral trait reconstruction, will clarify the origin and evolution of the mesoloph/id characters, and elucidate the relationship of this trait and the tetralophodonts genera of clade D.