

BAYESIAN BIOGEOGRAPHIC ANALYSIS OF SQUIRREL EVOLUTION
SUGGESTS AN ASIAN ORIGIN

by

ANNA BANKS

A THESIS

Presented to the Environmental Studies Program of the University of Oregon

In partial fulfillment of the requirements

For the degree of

Bachelor of Science

University of Oregon

June, 2021

An Abstract of the Thesis of

Anna Banks for the degree of Bachelor of Sciences
In the Environmental Studies Program to be taken June 2021

Title: BAYESIAN BIOGEOGRAPHIC ANALYSIS OF SQUIRREL EVOLUTION
SUGGESTS AN ASIAN ORIGIN

Approved: _____

[note: Faculty Adviser signs here]

Dr. Edward Davis

Squirrels have an incredibly diverse lineage with a global distribution and have dispersed over millions of years from a common ancestor approximately 40-50 MYA. The assumption since the 1980's has been that squirrels originated in North America and evolved from North American tree squirrels about 34-39 MYA, but in a more recent study, the oldest giant flying squirrel fossil was found in Oregon suggesting a possible eastward migration from Asia. This study made me question the previously accepted hypothesis that squirrels originated in North America and led to an exploration of squirrel evolution and origins. We used an R package called BiogeoBEARS with Bayesian biogeographic analysis and a DEC+J model to get a better picture of where squirrels were located at different points in their evolution. To run this analysis, we input a phylogenetic tree file, representing 225 modern Sciuridae species, and a biogeography file with each of their locations sorted into eight regions. My analysis resulted in a few noteworthy findings regarding major biogeography changes and their timing: (1) we believe that squirrels originated in Asia, (2) North American and Southern African squirrels came from Asia, and (3) we found one long South American branch, genetically distant from all other South American Sciuridae species and stemming from near the oldest common squirrel ancestor, leading to just one species, *Sciurillus pusillus*, a modern species that has resided in South America for about 37 Ma.

TABLE OF CONTENTS

INTRODUCTION.....	p.4
Background information and current knowledge.....	p.4
Gaps in current knowledge.....	p.6
Research questions	p.7
METHODS.....	p.7
What is BiogeoBEARS?	p.7
Bayesian Analysis.....	p.8
Data Inputs	p.8
RESULTS.....	p.10
DISCUSSION.....	p.11
Intercontinental dispersal.....	p.11
Asian Origins.....	p.12
<i>Sciurillus pusillus</i>	p.13
CONCLUSIONS.....	p.16
REFERENCES.....	p.18
ACKNOWLEDGEMENTS	p.20
APPENDIX 1. R Script.....	p.21
APPENDIX 2. Biogeography data files.....	p.22
APPENDIX 3. Full analysis PDF.....	p.30

INTRODUCTION

Background Information

Squirrels have an incredibly diverse lineage and global distribution. Tree squirrels (*Sciurini*), ground squirrels (*Spermophilus*), flying squirrels (*Pteromyini* or *Petauristini*), chipmunks (*Tamiini*), and marmots (*Marmotini*) are all found in this group. The varying genera look quite different, have unique morphological features, and have a diverse range of physical characteristics. They are native to five continents, and somehow dispersed over millions of years from a common ancestor approximately 40-50 MYA. **This raises the question, where did squirrels come from?**

The assumption since the 1980's has been that squirrels originated in North America. A study that based their biogeographical analysis on fossil records has previously suggested that chipmunks and ground squirrels, marmots, and prairie dogs diverged from North American tree squirrels between the early Oligocene (34-39 MYA) and the Late Oligocene (approximately 25 MYA) (Ellis and Maxon, 1980). This understanding, although not entirely concrete, has allowed us to infer that the common ancestor of rodents in the Sciuridae was a North American tree squirrel about 34-39 MYA.

In 2016, new work found that the oldest giant flying squirrel fossil, *Miopetaurista neogrivensis*, was identified at the Cave Basin field site in Central Oregon. Fossil remains of this extinct giant flying squirrel were excavated dating to the mid-Miocene (11-16 million years ago) (Beidron, 2016). Prior to this study, the only other North American remains of this species were two instances identified from the Pliocene of Florida (about 2.5-5 million years ago) (Robertson, 1970; and Webb, 2008), but this is the first time a *Miopetaurista* specimen was found in Oregon.

The Oregon giant flying squirrel is also the oldest fossil evidence of *Miopetaurista neogrivensis* collected internationally. Comparatively, giant flying squirrel specimens collected in Europe and Asia date back to earlier in the Miocene (Daxner-Hock, 2004) such as a partial *Miopetaurista neogrivensis* skeleton recovered in Spain with an estimated age of 11.63 Ma (Casanovas-Vilar, 2018). The Cave Basin specimen from Oregon were collected from a depositional period estimated to be between 16.26 MA and 15.75 MA (Beidron, 2016), making this specimen significantly older than the European *Miopetaurista* collections.

Biedron (2016) made me question the previously accepted hypothesis that squirrels originated in North America. Because the North American fossils were found only in Oregon and Florida, we have to ask how they ended up in such different locations and were from such different time periods. Additionally, it is curious that no specimens have been identified between the two states. The environment between Oregon and Florida has not always been distinguished by grasslands, prairies, and deserts like we see today. During the early Miocene, the Northern hemisphere experienced high average temperatures peaking from about 17 to 15 MYA during a period called the Mid-Miocene Climatic Optimum (Zachos et. al., 2001). This period coincided with what some believe was the initial spread and diversification of flying squirrels in North America. Mid-latitudes of North America became characterized by a warm, humid climate and broadleaf and mixed forests that provided a sustainable habitat for flying squirrels (Casanovas-Vilar, 2018). In theory, if the land between Oregon and Florida was once wooded, flying squirrels would have been able to comfortably inhabit this space. Taphonomy studies have indicated that bone fragments and teeth on forest floors tend to persist longer and be better-preserved than animal remains in non-wooded areas because the shade and humidity may slow weathering processes and reduce fragmentation (Kerbis Peterhans, 1993; and Tappen, 1994). If

the environment between Oregon and Florida was previously inhabited by *Miopetaurista neogrivensis* during the Mid-Miocene Climatic Optimum, then there are good chances that their teeth are well preserved and have either been misidentified as other species or have yet to be discovered.

One hypothesis is the East-from-Asia hypothesis that would account for how the *Miopetaurista* fossil ended up in Oregon. The genetic similarities between New World and Old World squirrels suggest that flying squirrels may have been able to cross the Bering Strait land bridge and could have migrated eastward across it from Asia to North America. The eastward migration hypothesis may also account for the age gap between the European and Floridan giant flying squirrel fossils. The European *Miopetaurista* specimens are much older than the Florida specimens which allow us to infer that giant flying squirrels went through an Eastward migration out of Europe. This evidence allows us to infer that flying squirrels did not originate in North America and possibly migrated east from Asia.

Gaps in Current Knowledge

Considering this background, there are still many gaps in knowledge that need to be addressed. Beidron's (2016) research alone raised many questions that I had not previously considered. For one, how did the same flying squirrel end up on opposite ends of North America? Did they disperse from one side of the continent to the other, or did they migrate from another continent? Is there evidence from the squirrel phylogeny and fossil record that these flying squirrels have had a wider distribution than we previously thought. If they dispersed across North America, further questions can be raised regarding the climate and the state of the environment of North America during the periods that these fossils were found (from the mid-Miocene to the Pliocene). We can hypothesize that *Miopetaurista* dispersed from Asia, from

Europe, or from both in two separate migrations to North America. The eastward migration hypothesis indicates that *Miopetaurista* dispersed eastward from Asia, while the westward migration hypothesis indicates that *Miopetaurista* likely dispersed westward from Europe.

Research Questions

On a much broader scale, these general questions can be applied to the larger Sciuridae phylogeny. It is still unclear where squirrels originated and how they spread around the world. We have very little knowledge about the biogeographical origins of these species, the timing of their dispersals and evolutions, and whether these changes in biogeography coincided with any large scale environmental or climatic changes.

In this body of work, I took an exploratory approach and conducted an inductive study to further explore these unanswered questions. I did not begin with any strong hypotheses but was curious to further explore these gaps in knowledge and develop a better understanding of squirrel evolution and biogeography.

METHODS

What is BiogeoBEARS?

I wanted to understand where different lineages of squirrels were located at different points in their evolution. To do this we used an R package called BiogeoBEARS, designed to analyze historical biogeography and assess species distributions (Matzke, 2013; <http://phylo.wikidot.com/biogeoBEARS>). It accounts for a variety of processes that could explain species' distributions. Prior to the creation of this software, there was no method to determine which particular model of evolution and which processes are most applicable to the biogeographic and phylogenetic data in biogeography analyses (Matzke, 2013). The program addresses this problem by incorporating several models, allowing for parameters to be turned on

or off. It is the first historical biogeography software to include the ability to include fossils as ancestors or tree tips and the option to factor in founder event speciation. The latter addition gives the program much higher accuracy in inferring species' ancestral states than other historical biogeography software. Matzke (2013, p. 242), designed the program to implement, "...many models in a common likelihood framework, so that standard statistical selection procedures can be applied to let the data choose the best model." By utilizing a common likelihood framework, the program essentially chooses the model best suited for the data. In doing this, the researcher's bias is accounted for because they do not have to make assumptions about what processes occurred in the species' evolution to produce the modern biogeographical distribution we can see today. Integrative models like this are preferable because they "integrate over uncertainty rather than assuming fixed parameters" (Ogilvie, 2018, p. 5).

Bayesian Analysis

The model I used is the DEC+J model. DEC stands for dispersal-extinction-cladogenesis. This model can factor for dispersal, extinction, vicariance, and sympatry. The "+J" parameter indicates that founder-event speciation is also factored into the model.

Data Inputs

To conduct this analysis, BiogeoBEARS required two main data inputs: a phylogenetic tree file and a biogeography file. The phylogenetic tree file gives the program our best estimation of squirrels' evolutionary history, and the biogeography file gives the program a summary of where each modern species is currently located. BiogeoBEARS combines these and analyzes the probability that each split in the phylogeny occurred in different regions.

I was given permission by Miriam Zelditch to use a phylogenetic tree created by her research group (Zelditch et al., 2015, 2017). This tree is a bifurcating tree, meaning that every

node divides into two new branches. This phylogeny was built using data from Sciuridae gene sequence analysis (Zeldich et al., 2015) and data from the known Sciuridae taxonomy (Koprowski et al. 2016), producing a phylogeny with 225 modern species. The branch lengths on the tree represent time in millions of years, and the timeline goes back about 45 million years to the middle Eocene (Fig. 1).

I used a script provided by Matzke and mainly kept the default parameters he set in place, making few changes to the code (See Appendix 1). Before reading in the phylogenetic tree file, I “pruned” *Sundasciurus tahan* and *Geosciurus inauris* from the phylogenetic tree because I could not accurately identify their biogeography, leaving the remaining 223 species.

I created the biogeography file myself, sourcing each modern taxa’s locations and assigning them regions. I found most of the biogeography data in the IUCN Red List of Endangered Species, with the exception of a few species that were more difficult to locate. As I mentioned above, species whose locations could not be determined from any available source were “pruned” or removed from the phylogenetic tree to allow the analysis to proceed.

I sorted the biogeography into eight regions: (1) North America, (2) Central America and the Caribbean, (3) South America, (4) Northern Africa, (5) Southern Africa, (6) Europe, (7) Middle East, and (8) Asia (See Appendix 2). There are no squirrels or related rodents native to Australia, Polynesia, or Antarctica, so these regions are excluded from the analysis (Mercer and Roth, 2003).

Initially, I hoped to run this analysis with 19 regions so the biogeography could be pinpointed more precisely (Also included in Appendix 2), but I had to reduce the number of regions to 8. The way the BiogeoBEARS works, the number of parameters inflates dramatically with the number of areas. If there are too many regions, the number of parameters that must be

tested increases dramatically, and the program cannot process the analysis (Matzke pers. comm., 2021). For example, if a species were input with two possible regions (A, B), there would be 4 resulting possible combinations of regions: Null, A, B, AB. An analysis of 19 regions would have to run billions of region combinations and would practically never finish to provide us with an analysis, so we limited the number of regions so that each species inhabits up to three regions maximum.

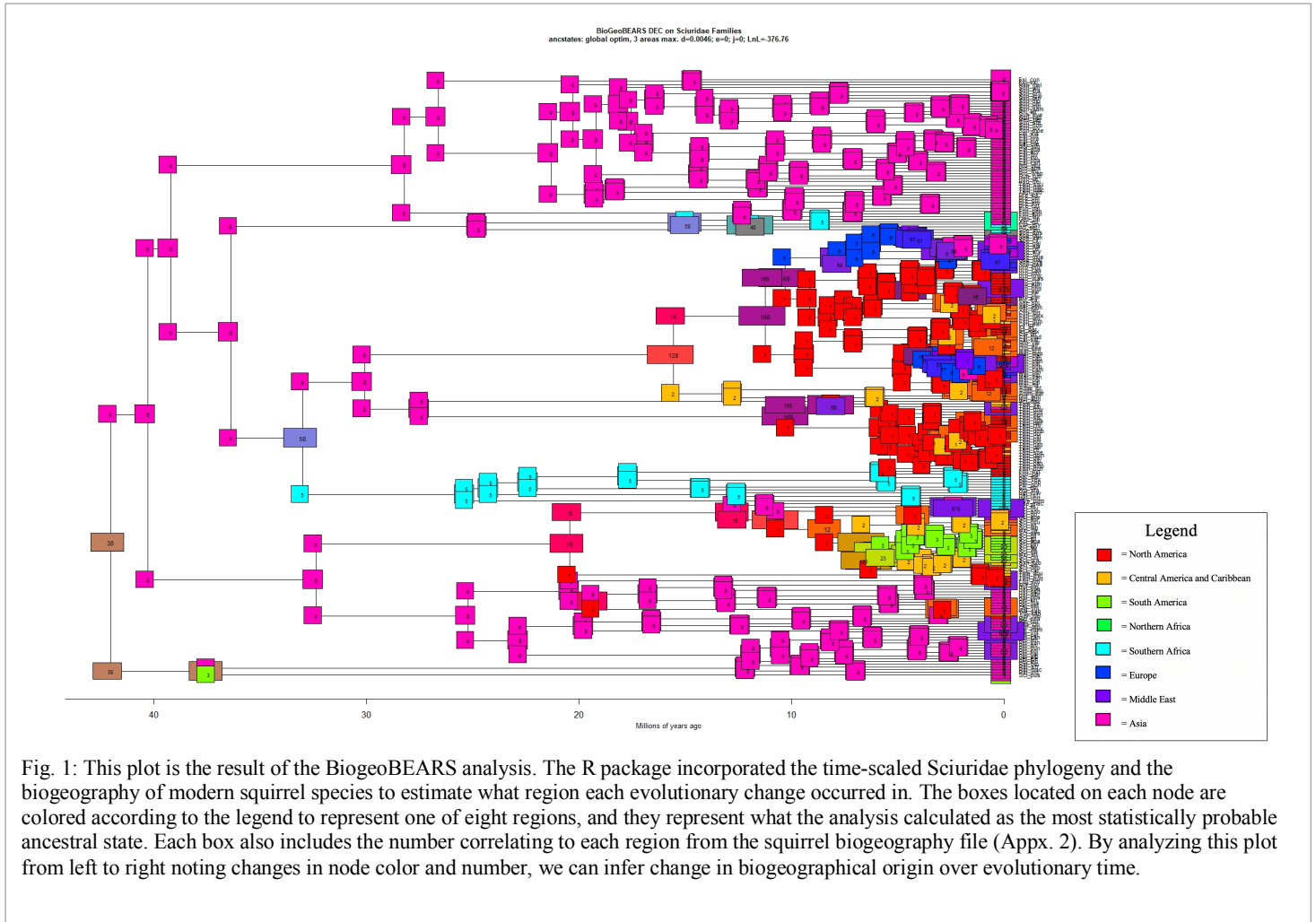
Because I had to limit the number of regions in the biogeographic analysis to eight, there were limited options for ways to divide up Earth's geography into regions. I had to make regions large enough that no species would be found in more than three regions, so the eight regions are really dictated by the distribution of the few squirrel species that are found across multiple continental regions.

Eventually, I hope that a similar analysis of the Sciuridae phylogeny can be applied to smaller regions to obtain a more precise biogeographic analysis. This would entail running separate analyses for each continent divided into smaller regions to get a more accurate inference of ancestral origin.

RESULTS

Figure 1 presents the primary results of my analysis, showing the phylogenetic tree input with colors to indicate the most probable region for each node. The x-axis represents time in millions of years, so the oldest nodes are on the left, and the modern taxa are on the right. When the colors change over evolutionary time (from left to right), this means that the species evolved and moved to a different region.

As you can see in the more comprehensive analysis attached in appendix 3, each node has a pie chart, with proportions of colors representing the probability that the split or evolutionary event at that node occurred in different regions. If the pie chart is mainly one color, there is high probability that that node formed in one region.



DISCUSSION

Intercontinental Dispersal

I found several noteworthy results regarding major biogeographic changes and their timing. First, my analysis suggested that North American Scuridae ancestors came from Asia. In

Fig. 1, you will notice that North American species are represented by the color red and the number one. The main groups of these red nodes first appear in two separate clusters, about 20 MYA and 15 MYA. We can infer that these two separate dispersals to North America came from Asia because their ancestors to the left are represented by pink nodes and the number eight, meaning they were Asian species.

Sciuridae species from southern Africa also appeared to have evolved from Asian squirrels about 34 MYA. This is visible in my analysis where pink nodes (Asian squirrels) transition to a large clade of light blue or teal nodes, representing squirrels with southern African roots. Since these southern African squirrels appear to have evolved from species represented by pink nodes, we can infer that their ancestral state was somewhere in Asia.

Most importantly, my analysis strongly suggested that squirrels came from Asia, not North America as previous research had suggested. In the colored phylogenetic tree, the majority of species directly branching from the oldest common ancestor are represented by pink nodes, indicating Asian origins. This supports the “east-from-Asia” hypothesis discussed earlier.

Asian Origins

The oldest common ancestor of squirrels and related rodents was not conclusively assigned a biogeographical location by my analysis, because its first branch, curiously forms one species in South America and numerous species in Asia. The overwhelming majority of the species that branched from the common ancestor can be traced to Asia, but only one species can be traced to South America. These findings suggest that squirrels may have originated in Asia because the overwhelming majority of direct descendants from the common ancestor are Asian squirrel species.

Significance of the Neotropical Pygmy Squirrel (*Sciurillus pusillus*)

The earliest common ancestor's biogeography on the phylogenetic tree is represented mainly by the color brown- which was not a color assigned to a region. As it branches off one node away, its biogeography estimate splits into two regions represented by pink (Asia) and green (South America). We believe that the brown color represents a **dispersal event**. Brown is a mix of the pink and green regions, and the nearest descendants diverge and separate to each of these regions at the nearest node, so we hypothesize that an early dispersal event found this lineage moving to South America. The early South American lineage exists today as only one long branch leading to one species: **the Neotropical pygmy squirrel (*Sciurillus pusillus*)**. While the portion of the node branching out to Asian squirrel species evolved into numerous subspecies and new species, the South American branch has no other living subspecies besides this taxon. The one South American species, *S. pusillus*, appears to be the earliest living squirrel in South America, and is genetically distant from all other South American rodent species (Jessen et al. 2013). *S. pusillus* is also the only living species in the genus *Sciurillus* and subfamily *Sciurillinae*. Based upon our results, this species or its extinct relatives are very closely related to the common ancestor and may have lived in South America for approximately 37 million years. The questions about this lineage will only be addressed through new paleontological research because of its lack of extant diversity.

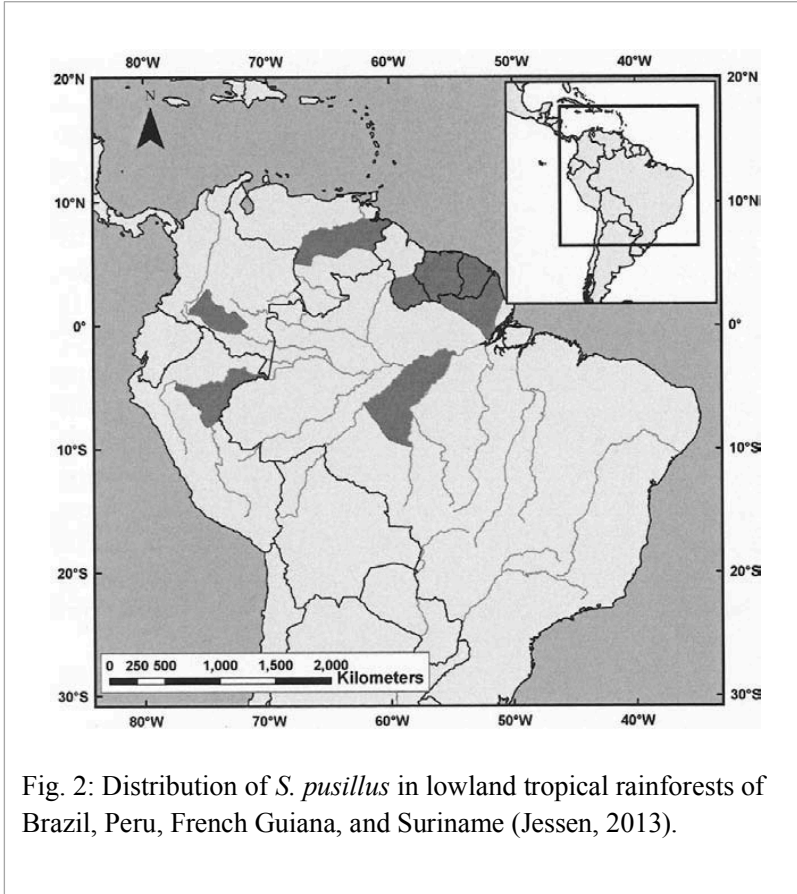


Fig. 2: Distribution of *S. pusillus* in lowland tropical rainforests of Brazil, Peru, French Guiana, and Suriname (Jessen, 2013).

The Neotropical Pygmy Squirrel is distributed in South American lowland tropical forests, and can be found in Brazil, French Guiana, Peru, and Suriname, shown in Fig. 2 (Jessen, 2013). This distribution may explain why *S. pusillus* has theoretically existed in South America for so long without dispersal, because these lowland tropical rainforests have unique features that make the land difficult to alter like table-like mountains called tepuis and the Andes mountains.

This is a noteworthy finding for several reasons:

(1) It confirms previous findings that the neotropical pygmy squirrel, the only remaining species in the subfamily *Sciurillinae*, is unusual in its biogeographic distribution and lineage structure. I can to confirm the understanding that *S. pusillus* makes up a monotypic lineage that diverged early in history and has no documented genetic polymorphism (Mercer and Roth 2003), meaning that the species is fairly genetically isolated. In the same way that mammals evolved to stop laying eggs but the egg-laying platypus is still present, this particular squirrel seems to have stagnated while its relatives steadily evolved over millions of years. The neotropical pygmy squirrel is a living relic from the past that has seemingly stayed constant while its relatives went extinct or evolved to look and function differently.

(2) *S. pusillus* may be the closest living relative to the common ancestor of all squirrels. Because *S. pusillus* is closely related to the common ancestor, doesn't have any living subspecies, and still exists in the wild, it may be the most genetically-similar species to squirrels' common ancestor. Intriguingly, there are practically no fossils of *S. pusillus* (Simpson, 1980) even though they do not appear to have changed geographical location in about 37 million years. Although the fossil record of *S. pusillus* is classified as "data deficient" (Jessen et al. 2013), it resides in several relatively undisturbed regions of South America, and fossil recovery may be possible.

(3) This finding further suggests that *S. pusillus* may have existed in South America for much longer than we previously believed. In South America, there has been no evidence of any squirrels before North and South America became connected through the Panamanian land bridge, about 3.1 million years ago (Marshall and Sempere, 1993 and Coates and Obando, 1996). My work indicates that *S. pusillus* and/or any extinct relatives of *S. pusillus* have existed in South America for approximately 37 million years, which is about 34 million years older than any South American squirrels identified to date. Currently, there is not enough physical evidence or published research available to confirm this, but my historical biogeography analysis suggest that squirrels have lived in South America for much longer than 3.1 million years. Utilizing this analysis and adding in additional fossil records and further archeological evidence all have great potential to pave the way for further understanding of the origin of Sciuridae.

Based upon these findings, future directions could include further archaeological and paleontological research and fossil identification of *S. pusillus* and other South American squirrels. It is very likely that Miocene-era or older fossil specimens of *S. pusillus* and its relatives exist but are largely undiscovered. In South America, efforts to excavate and preserve

fossils are typically hindered by acidic soils because acidic water or sediments can selectively dissolve calcium in shells and bones (Clayburn et al., 2004). Portions of the habitat of *S. pusillus* do have topography with higher pH such as the Andes mountain range and regions with volcanic deposits, so in these areas, it may be possible to locate and identify specimens ranging back approximately 35 million years. If found and further studied, fossil *S. pusillus* remains may closely resemble modern individuals because of its possibly stagnant evolutionary history. *S. pusillus* remains may also give us insight to the physiological and morphological characteristics of the common ancestor of squirrels because of their close relation, allowing for further hypotheses to be tested. Further archeological and paleontological work is needed in South America to confirm and explore the implications suggested by my analysis.

CONCLUSIONS

In this body of work, I wanted to explore the origins, evolution, and distribution of Sciuridae. To do this, I conducted Bayesian analysis of squirrel evolution based on current phylogenetic and biogeographic data through an R package called BiogeoBEARS. Running this analysis provided a comprehensive result calculating the most probable ancestral state of each species in the Sciuridae phylogeny, organized by region.

Further analysis allowed us to conclude that North American and southern African species both originated in Asia, with two separate dispersals to North America about 15 MYA and 20 MYA and one main dispersal from Asia to southern Africa about 34 MYA.

The results also revealed one very long branch stemming from the oldest common ancestor leading to just one extant South American squirrel, *Sciurillus pusillus*. This species has been around and not dispersed from South America for about 37 Ma, is still living, and has no other living relatives. The oldest evidence of squirrels in South America dates back about 3.1 Ma

to after the origination of the Panamanian land bridge, but this finding suggests that *S. pusillus* and other Sciuridae ancestors may have lived in South America for about 34 Ma longer than we previously believed. Further analysis of *S. pusillus* may provide insight to the characteristics of the oldest common ancestor, why this one species has remained practically unchanged for ~37 Ma, and may reveal new fossil specimens.

Most importantly, this analysis led us to conclude that squirrels likely originated in Asia, not in North America like previous research has suggested. Squirrels could provide another tool for studying the biogeographic connections between Asia and the other continents through the tectonic and climate changes of the Miocene.

REFERENCES

- Biedron, E.M. 2016. *The Sciuridae (Rodentia: Mammalia) of Cave Basin (Oregon), a new middle Miocene microfossil locality*. Sr. Thesis. University of Oregon.
- Casanovas-Vilar, I., et al. 2018. Oldest skeleton of a fossil flying squirrel casts new light on the phylogeny of the group. *Elife* 7: e39270.
- Clayburn, J.K., Smith, D.L., and Hayward, J.L. 2004. Taphonomic effects of pH and temperature on extant avian dinosaur eggshell." *Palaios* 19.2 (2004): 170-177.
- Coates, A.G. and Obando, J.A. 1996. The geologic evolution of the Central American isthmus, p. 21-56. In J.B.C. Jackson, A.F. Budd & A.G. Coates (eds.). *Evolution and environment in tropical America*. Chicago University, Chicago, USA.
- Daxner-Höck, G. 2004. Flying squirrels (Pteromyinae, Mammalia) from the upper Miocene of Austria. *Annalen des Naturhistorischen Museums in Wien. Serie A für Mineralogie und Petrographie, Geologie und Paläontologie, Anthropologie und Prähistorie* pp: 387-423.
- Ellis, L.S., and Maxson, L.R. 1980. Albumin evolution within new world squirrels (Sciuridae). *American Midland Naturalist* (1980): 57-62.
- Jessen, R.R., Gwinn, R.N., and Koprowski, J.L. 2013. *Sciurillus pusillus* (Rodentia: Sciuridae). *Mammalian Species* 45.903: 75-79.
- Kerbis Peterhans, J. C. et al. 1993. A contribution to tropical rain forest taphonomy: retrieval and documentation of chimpanzee remains from Kibale Forest, Uganda. *Journal of Human Evolution* 25(6): 485-514.
- Koprowski, J., Goldstein, E., Bennett, K., and Pereira, C. 2016. Family Sciuridae. Pp. 648–987 in D. E. Wilson, T. E. Lacher Jr, and R. A. Mittermeier, eds. *Handbook of the Mammals of the World Volume 6. Lagomorphs and Rodents*.

- Matzke, Nicholas. 2013. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography*. 5(4): 242-248.
- Marshall, L.G., and Sempere, T. 1993. Evolution of the Neotropical Cenozoic Land Mammal Fauna in its geochronologic, stratigraphic, and tectonic context. In: P. Goldblatt (Ed.), *Biological Relationships between Africa and South America*, Yale University Press, New Haven (1993), pp. 329-392.
- Mercer, J.M., and Roth, V.L. 2003. The effects of Cenozoic global change on squirrel phylogeny. *Science* 299.5612: 1568-1572.
- Ogilvie, Huw A., et al. 2018. Inferring species trees using integrative models of species evolution. *bioRxiv*: 242875.
- Robertson, J. S. Jr. 1970. Blancan mammals from Haile XVA, Alachua County, Florida. Ph.D. Dissertation, University of Florida, Gainesville, Florida, 172 pp.
- Simpson, G. G. 1980. *Splendid isolation: the curious history of South American mammals*. Yale University Press, New Haven, Connecticut.
- Tappen, M. 1994. Bone weathering in the tropical rain forest. *Journal of Archeological Science* 21(5): 667-673.
- Webb, S. D. et al. 2008 Terrestrial mammals of the Palmetto Fauna (early Pliocene, latest Hemphillian) from the Central Florida Phosphate District. *Geology and Vertebrate Paleontology of Western and Southern North America: Contributions in Honor of David P. Whistler*. Natural History Museum of Los Angeles County Science Series 41: 293-312.
- Zachos, J., et al. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* 292.5517: 686-693.

Zelditch, M.L., et al. 2017. Rare ecomorphological convergence on a complex adaptive landscape: body size and diet mediate evolution of jaw shape in squirrels (Sciuridae).

Evolution 71.3: 633-649.

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

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Edward Davis, for his mentorship, guidance, and help in the completion of this work, and Nick Matzke for his assistance and support working with BiogeoBEARS. Thank you to the College of Arts and Sciences and the Environmental Science department for the opportunity to continue and share this work.

APPENDIX 1. R Script

```
#Load ape and BiogeoBEARS by checking the boxes under "packages" or running codes below
library(BioGeoBEARS)
library(ape)
#Set working directory:session>set working directory>to source file location

#set up tree file
trfn = "squirrel_phylogeny.tree"
try_result = read.tree(trfn)

#set up geography file
geogfn = "8regionbiogeo.txt"
tipranges = getranges_from_LagrangePHYLIP(lgdata_fn=geogfn)
#geogcsv = "squirrel_biogeography.csv"
#squirrel_biogeography <- read.csv("squirrel_biogeography.txt")

#prune tree-----
#Read in species file with species names and turning them into row names
#geog<-read.csv(geogfn,row.names="Species")
#Finding species in tree file that aren't in the geog file
#spdrop<-as.vector(setdiff(try_result$tip.label,row.names(geog)))
#New tree with dropped tips
#prunedtrs<-drop.tip(try_result,spdrop)
#saving new tree to working directory
#write.tree(prunedtrs, file="prunedtrs.tre")

#load files
phylip = geogfn
tree_file = "prunedtrs.tre"
try_result = read.tree(tree_file)

#Run DEC model
#define maximum range size
max_range_size = 3
#Initialize a default model (DEC)
BioGeoBEARS_run_object = define_BioGeoBEARS_run(tree_file)
#Give BiogeoBEARS the location of the phylogeny Newick file
BioGeoBEARS_run_object$trfn = tree_file
#Give BiogeoBEARS the location of the geography text file
BioGeoBEARS_run_object$geogfn = phylip

#Input the maximum range size
BioGeoBEARS_run_object$max_range_size = max_range_size
#identify minimum branch length to treat tip as a direct ancestor (no speciation event)
BioGeoBEARS_run_object$min_branchlength = 0.000001
#Set null range for DEC model
BioGeoBEARS_run_object$include_null_range = TRUE

#Speed options and multi-core processing
BioGeoBEARS_run_object$on_NaN_error = -1e50
BioGeoBEARS_run_object$speedup = TRUE
BioGeoBEARS_run_object$use_optimx = "GenSA"
BioGeoBEARS_run_object$num_cores_to_use = 1
```

```

#No sparse matrix exponentiation (<600 regions/ranges)
BioGeoBEARS_run_object$force_sparse = FALSE

#Load dispersal multiplier matrix etc from the text files into the model object
BioGeoBEARS_run_object = readfiles_BioGeoBEARS_run(BioGeoBEARS_run_object)

#Default settings to get ancestral states
BioGeoBEARS_run_object$return_condlikes_table = TRUE
BioGeoBEARS_run_object$calc_TTL_loglike_from_condlikes_table = TRUE
BioGeoBEARS_run_object$calc_ancprobs = TRUE
#list of settings
BioGeoBEARS_run_object
#contains model object
BioGeoBEARS_run_object$BioGeoBEARS_model_object
#This table contains the parameters of the model
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table

#Run this to check inputs-----
check_BioGeoBEARS_run(BioGeoBEARS_run_object)

#Runs analysis
resDEC = bears_optim_run(BioGeoBEARS_run_object)
#Saves results
save(resDEC, file="resDEC.txt")

#Title analysis
analysis_titledtxt = "BioGeoBEARS DEC on Sciuridae Families"
#Setup
results_object = resDEC
scriptdir = np(system.file("extdata/a_scripts", package="BioGeoBEARS"))

#States
res2 = plot_BioGeoBEARS_results(results_object, analysis_titledtxt, addl_params=list("j"), plotwhat="text",
label.offset=0.45, tipcex=0.7, statecex=0.7, splitcex=0.6, titlecex=0.8, plotsplits=TRUE, cornercoords_loc=scriptdir,
include_null_range=TRUE, tr=try_result, tipranges=tipranges)
#Pie chart
plot_BioGeoBEARS_results(results_object, analysis_titledtxt, addl_params=list("j"), plotwhat="pie",
label.offset=0.45, tipcex=0.25, statecex=0.1, splitcex=0.1, titlecex=0.8, plotsplits=TRUE,
cornercoords_loc=scriptdir, include_null_range=TRUE, tr=try_result, tipranges=tipranges)

```

APPENDIX 2. Biogeography data files

8 Regions (.txt file)

```

223 8 (1 2 3 4 5 6 7 8)
Aer_mel 00000001
Aer_tep 00000001
Amm_har 11000000
Amm_int 11000000
Amm_leu 11000000
Atl_get 00010001
Bel_pea 00000001
Cal_ada 00000001
Cal_can 00000001
Cal_ery 00000001
Cal_fin 00000001

```

Cal_ino 00000001
Cal_nig 00000001
Cal_not 00000001
Cal_ore 00000001
Cal_pha 00000001
Cal_pre 00000001
Cal_lat 10000000
Cal_mad 01000000
Cal_sat 10000000
Cyn_gun 10000000
Cyn_leu 10000000
Cyn_lud 11000000
Cyn_mex 01000000
Cyn_par 10000000
Dre_gul 00000001
Dre_lok 00000001
Dre_per 00000001
Dre_pyr 00000001
Dre_ruf 00000001
Eog_fim 0000101
Epi_ebi 00001000
Eup_cin 0000101
Eux_ery 00001000
Exi_con 00000001
Exi_exi 00000001
Fun_lay 00000001
Fun_pal 00000001
Fun_pen 00000001
Fun_sub 00000001
Fun_car 00001000
Fun_pyr 00001000
Geo_pri 00001000
Gla_sab 10000000
Gla_vol 11000000
Gly_sim 00000001
Hel_ruf 00001000
Hel_ruw 00001000
Hel_und 00001000
Hyl_alb 00000001
Hyl_nig 00000001
Hyl_pha 00000001
Hyl_sag 00000001
Hyl_spa 00000001
Hyo_hei 00000001
Hyo_ile 00000001
Ict_mex 11000000
Ict_par 11000000
Ict_tri 10000000
Iom_hor 00000001
Lar_ins 00000001
Mar_bai 00000111
Mar_bob 00000110
Mar_bro 10000000
Mar_cal 10000000
Mar_cam 00000100
Mar_cau 00000011
Mar_fla 10000000
Mar_him 00000001
Mar_mar 00000100
Mar_men 00000010
Mar_mon 10000000
Mar_oly 10000000
Mar_sib 00000101
Mar_van 10000000
Men_ber 00000001
Mic_alf 01100000
Mic_fla 00100000
Myo_pum 00001000
Nan_mel 00000001
Not_ado 01000000

Not_ann 01000000
Oto_atr 01000000
Oto_bee 11000000
Oto_var 11000000
Par_ale 00001000
Par_cep 00001000
Par_och 00001000
Par_vex 00001000
Pet_kin 00000001
Pet_alv 00000001
Pet_alb 00000001
Pet_can 00000001
Pet_ele 00000001
Pet_hai 00000001
Pet_len 00000001
Pet_leu 00000001
Pet_pet 00000101
Pet_phi 00000101
Pet_xan 00000001
Pet_yun 00000001
Pet_fus 00000001
Pet_set 00000001
Pol_fra 10000000
Pro_abs 00000001
Pro_leu 00000001
Pro_mur 00000001
Pro_web 00000001
Pro_sta 00001000
Pte_mom 00000001
Pte_vol 00000111
Pte_pul 00000001
Rat_aff 00000001
Rat_bic 00000001
Rat_ind 00000001
Rat_mac 00000001
Rhe_mac 00000001
Rhi_lat 00000001
Rub_rub 00000001
Sci_pus 00100000
Sci_dav 00000001
Sci_abe 11000000
Sci_aes 00100000
Sci_all 01000000
Sci_ano 00000011
Sci_aur 01000000
Sci_car 10000000
Sci_col 01000000
Sci_dep 01000000
Sci_fla 00100000
Sci_gil 00100000
Sci_gra 01100000
Sci_gri 11000000
Sci_igt 00100000
Sci_igv 01100000
Sci_lis 00000001
Sci_nig 11000000
Sci_ocu 01000000
Sci_pyr 00100000
Sci_spa 00100000
Sci_str 00100000
Sci_var 01000000
Sci_vul 00000111
Spe_lep 00000011
Spe_ala 00000001
Spe_cit 00000110
Spe_dau 00000101
Spe_ery 00000111
Spe_ful 00000110
Spe_maj 00000110
Spe_mus 00000100

Spe_pal 0000001
Spe_pyg 00000110
Spe_rel 00000011
Spe_sus 00000100
Spe_tau 00000010
Spe_xan 00000011
Sun_bro 00000001
Sun_eve 00000001
Sun_fra 00000001
Sun_hip 00000001
Sun_hoo 00000001
Sun_jen 00000001
Sun_juv 00000001
Sun_low 00000001
Sun_min 00000001
Sun_moe 00000001
Sun_phi 00000001
Sun_rab 00000001
Sun_sam 00000001
Sun_ste 00000001
Sun_ten 00000001
Syn_bro 01000000
Tam_alp 10000000
Tam_amo 10000000
Tam_bul 01000000
Tam_can 10000000
Tam_cin 10000000
Tam_dor 11000000
Tam_dur 01000000
Tam_mer 11000000
Tam_min 10000000
Tam_obs 11000000
Tam_och 10000000
Tam_pal 10000000
Tam_pan 10000000
Tam_qum 10000000
Tam_quv 10000000
Tam_rfi 10000000
Tam_rfu 10000000
Tam_sen 10000000
Tam_sib 00000101
Tam_sis 10000000
Tam_son 10000000
Tam_spe 10000000
Tam_str 10000000
Tam_tow 10000000
Tam_umb 10000000
Tam_dou 10000000
Tam_fre 10000000
Tam_hud 10000000
Tam_mar 00000001
Tam_mac 00000001
Tam_rod 00000001
Tam_swi 00000001
Tro_xan 00000001
Uro_arm 10000000
Uro_bel 10000000
Uro_bru 10000000
Uro_can 10000000
Uro_col 10000000
Uro_ele 10000000
Uro_mol 10000000
Uro_par 10000100
Uro_ric 10000000
Uro_tow 10000000
Uro_und 00000111
Uro_was 10000000
Xer_moh 10000000
Xer_per 01000000
Xer_spi 11000000

Xer_ter 11000000
Xer_rut 00011000

19 Regions (.txt file)

223 19 (1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19)

Aer_mel 000000000000001000
Aer_tep 000000000000000100
Amm_har 110000000000000000
Amm_int 110000000000000000
Amm_leu 110000000000000000
Atl_get 000000001000000000
Bel_pea 000000000000001100
Cal_ada 000000000000000100
Cal_can 0000000000000001100
Cal_ery 000000000000001100
Cal_fin 000000000000000100
Cal_ino 000000000000001100
Cal_nig 000000000000000100
Cal_not 000000000000000100
Cal_ore 000000000000000100
Cal_pha 000000000000001100
Cal_pre 000000000000000100
Cal_lat 100000000000000000
Cal_mad 010000000000000000
Cal_sat 100000000000000000
Cyn_gun 100000000000000000
Cyn_leu 100000000000000000
Cyn_lud 110000000000000000
Cyn_mex 010000000000000000
Cyn_par 100000000000000000
Dre_gul 000000000000001100
Dre_lok 000000000000010100
Dre_per 000000000000001100
Dre_pyr 000000000000001100
Dre_ruf 000000000000001100
Eog_fim 0000001000000010000
Epi_ebi 0000000001100000000
Eup_cin 0000001000000010000
Eux_ery 0000000001010000000
Exi_con 000000000000000100
Exi_exi 000000000000000100
Fun_lay 0000000000000010000
Fun_pal 0000000000000010000
Fun_pen 0000000000000011000
Fun_sub 0000000000000010000
Fun_car 0000000000110000000
Fun_pyr 0000000001110000000
Geo_pri 0000000000101000000
Gla_sab 1000000000000000000
Gla_vol 1100000000000000000
Gly_sim 000000000000000100
Hel_ruf 0000000001110000000
Hel_ruw 0000000001100000000
Hel_und 0000000000010000000
Hyl_alb 000000000000001100
Hyl_nig 000000000000000100
Hyl_pha 000000000000001100
Hyl_sag 000000000000000100
Hyl_spa 000000000000000100
Hyo_hei 000000000000000100
Hyo_ile 000000000000000100
Ict_mex 1100000000000000000
Ict_par 1100000000000000000
Ict_tri 1000000000000000000
Iom_hor 000000000000000100
Lar_ins 000000000000000100
Mar_bai 0000001000000101000

Mar_bob 0000001100000100000
Mar_bro 1000000000000000000
Mar_cal 1000000000000000000
Mar_cam 0000001000000000000
Mar_cau 000000000000111000
Mar_fla 1000000000000000000
Mar_him 000000000000011000
Mar_mar 0000011000000000000
Mar_men 00000000000100000
Mar_mon 1000000000000000000
Mar_oly 1000000000000000000
Mar_sib 000000100000001000
Mar_van 1000000000000000000
Men_ber 000000000000000100
Mic_alf 0101000000000000000
Mic_fla 0001000000000000000
Myo_pum 000000000100000000
Nan_mel 000000000000000100
Not_ado 0100000000000000000
Not_ann 0100000000000000000
Oto_atr 0100000000000000000
Oto_bee 1100000000000000000
Oto_var 1100000000000000000
Par_ale 000000000110000000
Par_cep 000000000111000000
Par_och 000000000010000000
Par_vex 000000000010000000
Pet_kin 000000000000000100
Pet_alv 000000000000010000
Pet_alb 000000000000001100
Pet_can 000000000000011100
Pet_ele 000000000000011100
Pet_hai 00000000000001000
Pet_len 000000000000001000
Pet_leu 000000000000001000
Pet_pet 0000001000000011100
Pet_phi 0000001000000011100
Pet_xan 000000000000001000
Pet_yun 000000000000001000
Pet_fus 000000000000010000
Pet_set 000000000000000100
Pol_fra 1000000000000000000
Pro_abs 000000000000000100
Pro_leu 000000000000000100
Pro_mur 000000000000000100
Pro_web 000000000000000100
Pro_sta 0000000001110000000
Pte_mom 0000000000000001000
Pte_vol 0000101000000101000
Pte_pul 000000000000000100
Rat_aff 000000000000000100
Rat_bic 000000000000011100
Rat_ind 000000000000010000
Rat_mac 000000000000010000
Rhe_mac 000000000000000100
Rhi_lat 000000000000000100
Rub_rub 000000000000000100
Sci_pus 0001000000000000000
Sci_dav 000000000000001000
Sci_abe 1100000000000000000
Sci_aes 0001000000000000000
Sci_all 0100000000000000000
Sci_ano 0000000100000010000
Sci_aur 0100000000000000000
Sci_car 1000000000000000000
Sci_col 0100000000000000000
Sci_dep 0100000000000000000
Sci_fla 0001000000000000000
Sci_gil 0001000000000000000
Sci_gra 0101000000000000000

Sci_gri 110000000000000000
Sci_igt 000100000000000000
Sci_igt 001100000000000000
Sci_lis 0000000000000001000
Sci_nig 110000000000000000
Sci_ocu 010000000000000000
Sci_pyr 000100000000000000
Sci_spa 000100000000000000
Sci_str 000100000000000000
Sci_var 010000000000000000
Sci_vul 0000111000000101000
Spe_lep 0000000000000110000
Spe_ala 0000000000000001000
Spe_cit 0000011100000000000
Spe_dau 0000001000000001000
Spe_ery 0000001000000101000
Spe_ful 0000001100000100000
Spe_maj 0000001000000100000
Spe_mus 0000001000000000000
Spe_pal 0000000000000001000
Spe_pyg 0000001000000100000
Spe_rel 0000000000000101000
Spe_sus 0000001000000000000
Spe_tau 0000000100000000000
Spe_xan 0000000100000010000
Sun_bro 0000000000000000100
Sun_eve 00000000000000001100
Sun_fra 0000000000000000100
Sun_hip 0000000000000000100
Sun_hoo 0000000000000000100
Sun_jen 0000000000000000100
Sun_juv 0000000000000000100
Sun_low 0000000000000000100
Sun_min 0000000000000000100
Sun_moe 0000000000000000100
Sun_phi 0000000000000000100
Sun_rab 0000000000000000100
Sun_sam 0000000000000000100
Sun_ste 0000000000000000100
Sun_ten 0000000000000000100
Syn_bro 0100000000000000000
Tam_alp 1000000000000000000
Tam_amo 1000000000000000000
Tam_bul 0100000000000000000
Tam_can 1000000000000000000
Tam_cin 1000000000000000000
Tam_dor 1100000000000000000
Tam_dur 0100000000000000000
Tam_mer 1100000000000000000
Tam_min 1000000000000000000
Tam_obs 1100000000000000000
Tam_och 1000000000000000000
Tam_pal 1000000000000000000
Tam_pan 1000000000000000000
Tam_qum 1000000000000000000
Tam_quv 1000000000000000000
Tam_rfi 1000000000000000000
Tam_rfu 1000000000000000000
Tam_sen 1000000000000000000
Tam_sib 0000111000000001000
Tam_sis 1000000000000000000
Tam_son 1000000000000000000
Tam_spe 1000000000000000000
Tam_str 1000000000000000000
Tam_tow 1000000000000000000
Tam_umb 1000000000000000000
Tam_dou 1000000000000000000
Tam_fre 1000000000000000000
Tam_hud 1000000000000000000
Tam_mar 0000000000000001100

Tam_mac 000000000000011100
Tam_rod 000000000000000100
Tam_swi 000000000000011100
Tro_xan 00000000000001000
Uro_arm 10000000000000000
Uro_bel 10000000000000000
Uro_bru 10000000000000000
Uro_can 10000000000000000
Uro_col 10000000000000000
Uro_ele 10000000000000000
Uro_mol 10000000000000000
Uro_par 10000100000000000
Uro_ric 10000000000000000
Uro_tow 10000000000000000
Uro_und 00000100000101000
Uro_was 10000000000000000
Xer_moh 10000000000000000
Xer_per 01000000000000000
Xer_spi 11000000000000000
Xer_ter 11000000000000000
Xer_rut 00000001001000000

