Pilosa: Evolutionary Diversity Visualized through Locational Data

Kristen Lee

Introduction

The order Pilosa is a group of placental mammals. The two orders under are Vermilingua and Folivora- more commonly known as Anteaters and Sloths. Originally, I was focusing on just sloths; however to work with a larger data set, I moved up a taxonomic level.

Sloths have always intrigued me, so I was interested in learning more about trends in their evolution. One biological characteristic that plays a role in shaping the ecological patterns of many species is metabolism. Basal metabolic rate (BMR) is a primary factor shaping body size evolution (Toledo *et al.* 2017). Xenarthrans (the superoder to Pilosa) are known to have much lower body temperatures compared to other mammals their size. Accordingly, their metabolic rates are also much lower, between 40-60% lower than what is expected based on their mass (Gilmore 2000).

One way in which I could explore this evolutionary characteristic was through the distribution of the species. Poor thermoregulation and metabolism confine a species to live in warmer temperatures. Extinct ground sloths were probably able to tolerate colder temperatures due to their greater size and muscle mass (Gilmore 2001). But after evolving to much smaller sizes, their distribution altered as well. The current brown-throated three-toed sloth is one of the most common medium-sized mammals in the tropical forests of Central and South America (Gilmore 2000).

Ultimately, as the species diversified from larger creatures to smaller, they became confined to different temperatures and locations. Using the PaleoBiology Database, I wanted to manipulate data on Pilosa to see if their locational data would ultimately parallel and support this trend.

Description

I utilized the Paleobiology Database which houses fossil occurances along with a variety of parameters. Since, I was focusing on primarily the location of pilosa, I used locational data(latitude & longitude), age data (in MYA), and taxonomic data (order, species). I first manipulated the data to seperate/organize taxonomic and age information. Then after analyzing my narrowed down data, I combined it with locational data to visualize with matplotlib. And then I drew conclusions from analyzing my results.

Methods & Data Manipulation

First I ran `tr -d "\n\r" < pbdb_data_pilosa_full.csv` in the terminal to elimante empty rows in the csv file. Then I wrote this function to print a list of all the species in my dataset. I created a for loop that printed the species name only if it was listed as a "species" in a seperate column.

```python
in_file = open("pbdb_data_pilosa_full.csv", 'r', encoding= "ISO-8859-1")
read_file = in_file.readlines()[18:] #starting where the occurance data starts

list_species = [] #empty list to store species in

for line in read_file:
    species = line.split('","')[11] #delimiter ","
    name_of_species = line.split('","')[9]
    if "species" in species:
        list_species.append(name_of_species)
print("\n".join(list_species)) #check to see if species list printed (works)
```

Secondly, I wrote a function regarding age.

$$MeanAge = (MinimumAge + MaximumAge)/2 \tag{1}$$

Using this simple equation, I was able to calculate the mean age of the fossils with the following function:

```python
mean_age_list = []


for line in read_file:
    minage = line.split('","')[15] #'""' gets rid of quotes
    maxage = line.split('","')[16]
    mean_age = (float(minage)+float(maxage))/ 2

    mean_age_list.append(float(mean_age))
print(mean_age_list)
```

I then created a set of unique ages

```python
unique_means = set() #creating a set to get the unique mean age
for item in mean_age_list:
    unique_means.add(item)

unique_means_list = list(unique_means) #converting back to list
```

```python
print(sorted(unique_means_list))  #sort numerically
```

Next I created a dictionary that printed out the number of occurances per time period.

```python
#finding which mean age had the most species present
from collections import Counter

ageDict = dict(Counter(mean_age_list)) #counts the occurences of each item in means lis
print("{" + "\n".join("{}: {}".format(k, v) for k, v in ageDict.items()) + "}")

#0.06885: 127 most occurances late pleistocene
#1.29985: 110 second most pleistocene
```

I did not print out the data from all these functions to maintain organization, but the results of these functions in general gave me a better sense of the data I was working with.

To get an even better idea of the data that I had been manipulating (species and age), I used ggplot to build a bar graph that showed the occurances (x-axis) per species (y-axis), and the age (color). From my results, I could see what time period had the most fossil occurances.

I first created a new csv file with the species in one column and mean age in the next. Then I used ggplot to graph the file.

```python
import csv #creating a new csv file with only species and age for R

with open("mean_ages_for_r.csv", 'w') as f:
    writer = csv.writer(f, delimiter = ',')
    writer.writerows(list(zip(list_species, mean_age_list)))
```

```r
library(ggplot2)

setwd("/home/eeb177-student/Desktop/eeb-177/eeb-177-final-project/final-project/")

pilosa <- read.csv("mean_ages_for_r.csv", header = F, as.is = T)

names(pilosa) <- c("Species", "Age")
head(pilosa)
```

```
##                     Species      Age
## 1      Megalonyx curvidens  1.29985
## 2 Pliometanastes protistus 53.05000
## 3     Thinobadistes wetzeli  7.60000
## 4      Megalonyx curvidens  7.60000
## 5      Megalonyx curvidens  7.60000
## 6      Megalonyx curvidens  7.60000
```
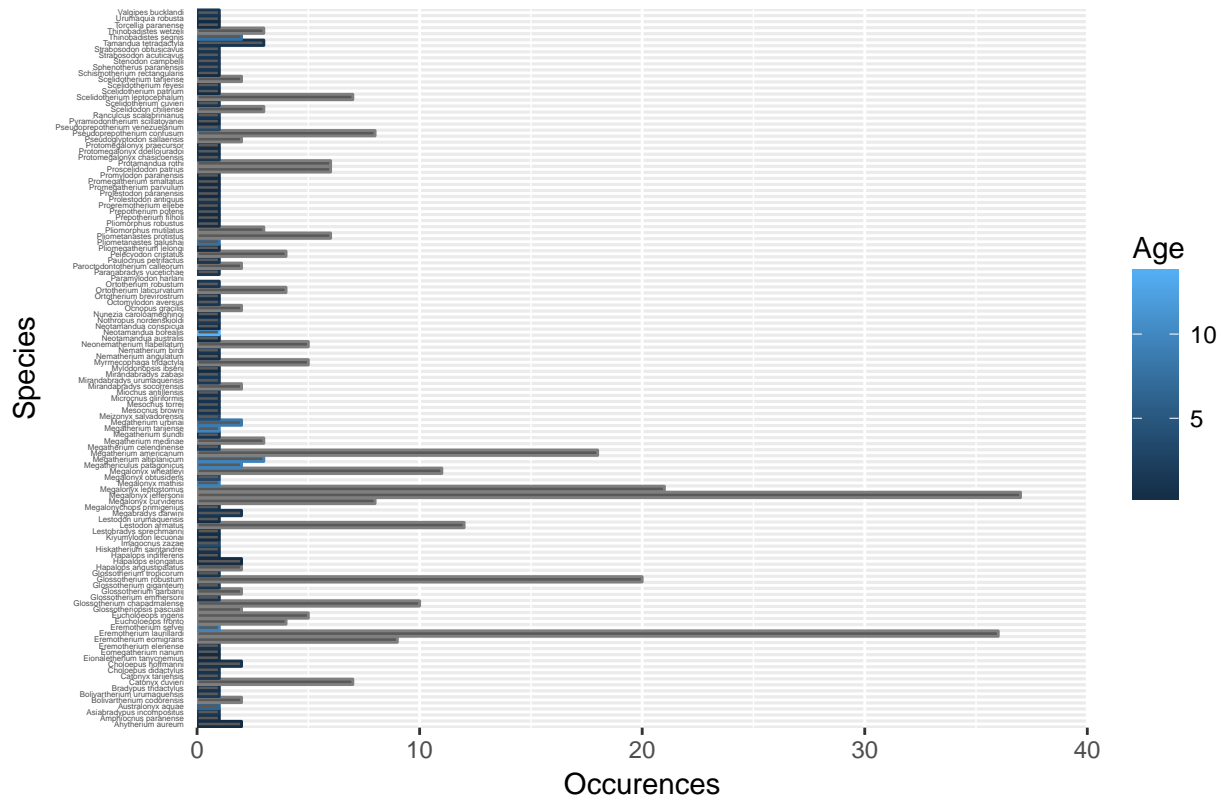
```
pilosa_occ <- ggplot(pilosa, aes(Species, color = Age))
pilosa_occ <- pilosa_occ + geom_bar()
pilosa_occ <- pilosa_occ + coord_flip() + theme(axis.text.y = element_text(size=3)) +the
pilosa_occ + scale_y_continuous(limits = c(0,40), expand = c(0, 0), breaks = c(0, 10, 20
```

Warning: Removed 1 rows containing missing values (geom_bar).

Visualization

Finally, I was able to visualize my data. First, I created lists for the latitudes and longitudes of each species:

```
latitude = [] #empty list for lat
longitude = [] #empty list for long

for line in read_file:
    lats = line.split('","')[30]
    lons = line.split('","')[29]
    latitude.append(float(lats)) #convert to float
    longitude.append(float(lons))
print(latitude + longitude) # checking to see it works
```

In lab we worked with matplotlib for a bit, so with the help of google and this tutorial http://introtopython.org/visualization_earthquakes.html, I was able to use a package called Basemap. Basemap is a library for plotting 2D data on maps.

The first code and map is just a general plotting of all the occurances.

```python
from pylab import rcParams

%matplotlib inline
rcParams['figure.figsize'] = (8,6)

# Make the plot larger.
plt.figure(figsize=(16,12))


from mpl_toolkits.basemap import Basemap
import matplotlib.pyplot as plt
import numpy as np

eq_map = Basemap(projection='robin', resolution = 'l', area_thresh = 1000.0,
                 lat_0=0, lon_0=-130)  #parameters for map
eq_map.drawcoastlines()
eq_map.drawcountries()
eq_map.fillcontinents(color = 'gray')
eq_map.drawmapboundary()
eq_map.drawmeridians(np.arange(0, 360, 30))
eq_map.drawparallels(np.arange(-90, 90, 30))

x,y = eq_map(longitude, latitude)
eq_map.plot(x, y, 'ro', markersize=6)


plt.figtext(.5,.9,'Stratigraphic Location: Pilosa', fontsize=40, ha='center')
```
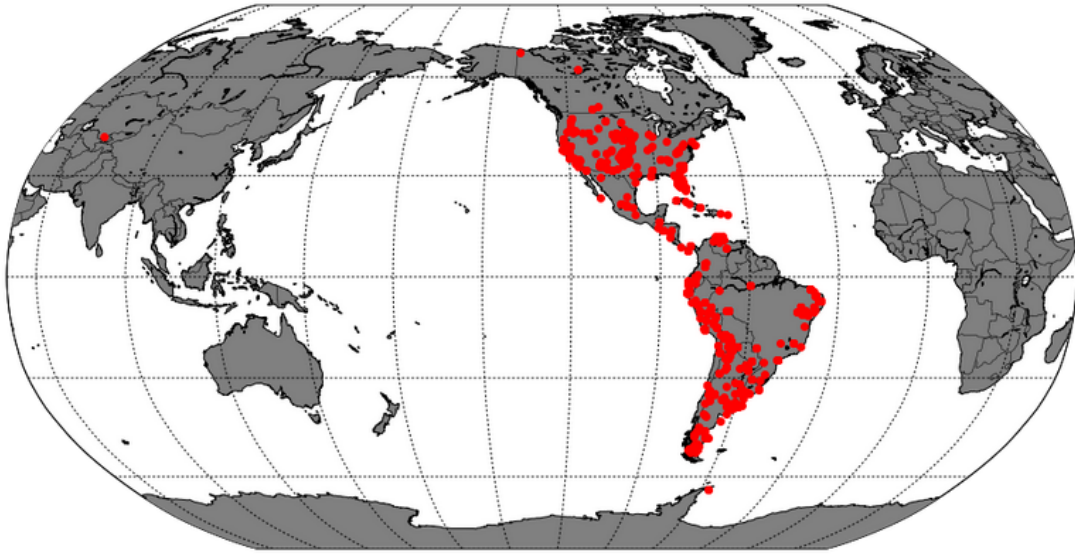
Stratigraphic Location: Pilosa

Figure 1: Map of Pilosa Occurances

Finally, I combined all the parameters to create a final map. Using the information I was able to extract from the age data, I divided the occurences into three categories: beginning of cenozoic to oligocene, miocene, and pliocene to the present. I assigned each of these ranges to a different marker color and plotted the data.

```python
def get_marker_color(mean_age_list):
    if mean_age_list < 5.33:   #pliocene to present
        return ('go')
    elif mean_age_list < 23.03:   #miocene
        return ('yo')
    else:
        return ('ro')   #oligocene to the beginning of Cenozoic (all of this is in Cenoz


eq_map = Basemap(projection='robin', resolution = 'l', area_thresh = 1000.0,
                lat_0=0, lon_0=-130)
eq_map.drawcoastlines()
eq_map.drawcountries()
eq_map.fillcontinents(color = 'gray')
eq_map.drawmapboundary()
eq_map.drawmeridians(np.arange(0, 360, 30))
eq_map.drawparallels(np.arange(-90, 90, 30))
```

```
for lon, lat, age in zip(longitude, latitude, mean_age_list):

    x,y = eq_map(lon, lat)
    marker_string = get_marker_color(age)
    eq_map.plot(x, y, marker_string, markersize=8)

plt.figtext(.5,.9,'Stratigraphic Location: Pilosa 2', fontsize=30, ha='center')

plt.show()
```

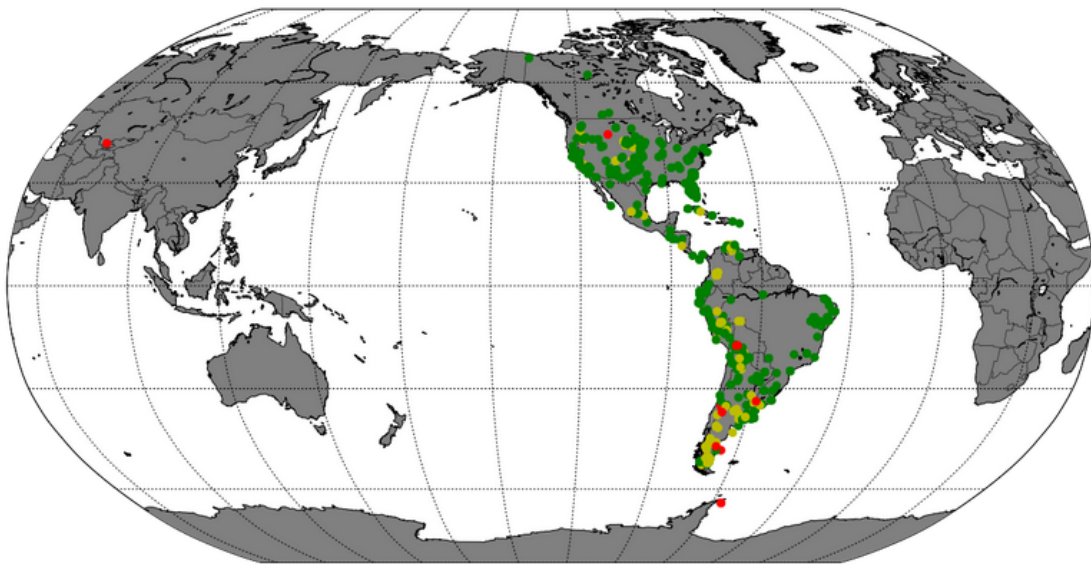Stratigraphic Location: Pilosa 2

Figure 2: Map of Pilosa Occurances based on Age

Conclusion

As mentioned earlier, sloths are poor regulators of body temperature because they have little ability to increase their metabolism due to their small muscle masses. After plotting my data, I looked for any trends that would indicate this shift from cooler temperatures to more tropic temperatures. To help do so, I compared it to a world map that depicts the tropics around the equator(see figure 3 below). In general, there seems to be a trend of occurances moving north. The red and yellow markers remain primarily in the south while the green markers are spread in both regions, but heavily in North America. I was hoping to see a more obvious movement towards the equator in the Basemap, for I do not think it can lead

to any conclusive data. Rather it could lead to further inevestigation with different maps, perhaps differentiating every individual age rather than a range of ages.
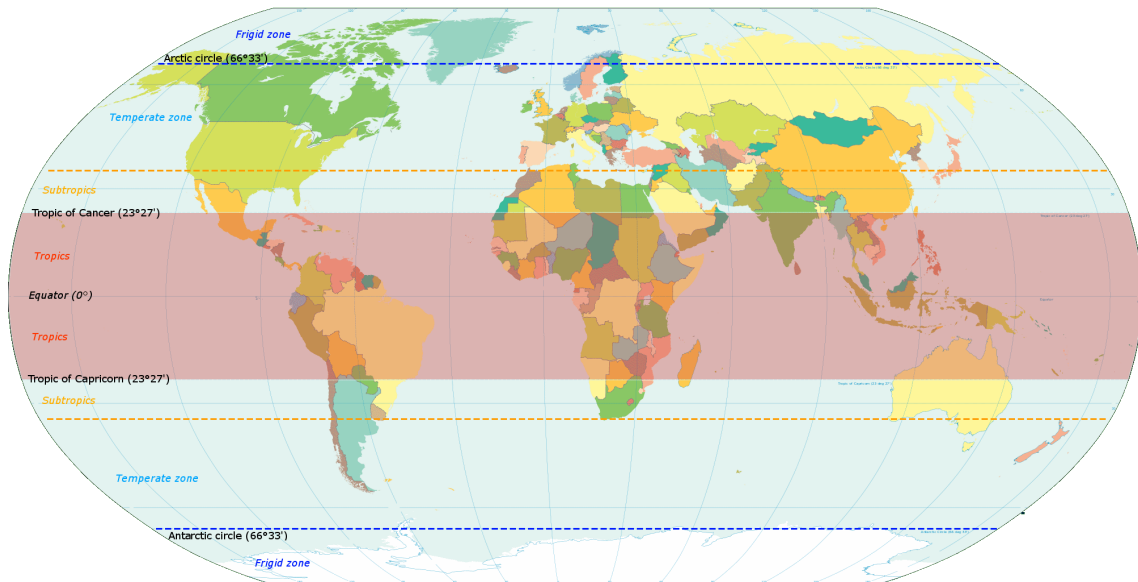

Figure 3: Current Map of Tropics

Here is the link to my github: {https://github.com/Kristenlee22/eeb-177-final-project} More detail of my project like my python notebooks with full code and output can be found here!

References

1.Gilmore, C.A.D., D.P. AND Da Costa. (2001). Sloth biology: an update on their physiological ecology, behavior and role as vectors of arthropods and arboviruses. *Brazilian Journal of Medical and Biological Research*, 34, 9–25.

2.Gilmore, C.A.D., D.P. AND Da-Costa. (2000). An update on the physiology of two- and three-toed sloths. *Brazilian Journal of Medical and Biological Research*, 33, 129–146.

3.Toledo, N., Bargo, M., Vizcaíno, S., De Iuliis, G. & Pujos, F. (2017). Evolution of body size in anteaters and sloths (xenarthra, pilosa): Phylogeny, metabolism, diet and substrate preferences. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh*, 106, 289–301.