A Practical Computing Approach to Newly Invasive Seagrass

Cooper Kass

Introduction

The species, genuses, and families found in this analysis all belong to the order Alismatales. The order is made up of 14 different, monophyletic families (Petersen *et al.* 2015). They are all angiosperms, or flowering plants (Petersen *et al.* 2015). Although most seagrass species are not invasive, a few more recent species have spread quickly in non-native waters (Tussenbroek *et al.* 2016). The inspiration for this project came from one of my professors, Dr. Peggy Fong. She just recently completed data collection for a study of invasive turtle grass in the Caribbean. Apparently, sea turtles in the area will eat the grass if it is the only source of food available. However, if other food sources are available, they will eat anything else. The problem is that this invasive seagrass is growing remarkably fast, outcompeting other primary producers. Hopefully, anazlyzing the biodiversity of the order over the course of its history will shed light on this recent, potentially harmful divergence. A recent study offers a possible explanation: "Rapid evolution of dispersal ability makes biological invasions faster and more variable" (Ochocki & Miller 2017).

Figure 1: A turtle grass in the Caribbean.

Methods

The data used in this project was downloaded from the Paleobiology Database (PBDB). The data is comma-separated, and it contains 324 fossil occurrences. Each occurrence included information on existence times and periods, the specific taxa it represents, as well as the latitude and longitude at which the fossil was found. A segment of the first 10 lines of the data file produced from the download of the order "Alismatales" is shown below:

```
system("head one-header_alismatales_data.csv", intern = T)
```

```
##  [1] "\"occurrence_no\",\"record_type\",\"reid_no\",\"flags\",\"collection_no\",\"acc
##  [2] "\"135643\",\"occ\",\"\",\"\",\"11326\",\"Alismataceae\",\"family\",\"55790\",\"
##  [3] "\"135760\",\"occ\",\"\",\"\",\"11340\",\"Alismataceae\",\"family\",\"55790\",\"
##  [4] "\"135949\",\"occ\",\"\",\"\",\"11359\",\"Alismataceae\",\"family\",\"55790\",\"
##  [5] "\"135956\",\"occ\",\"\",\"\",\"11360\",\"Alismataceae\",\"family\",\"55790\",\"
##  [6] "\"136056\",\"occ\",\"\",\"\",\"11370\",\"Alismataceae\",\"family\",\"55790\",\"
##  [7] "\"136320\",\"occ\",\"\",\"\",\"11393\",\"Potamogeton\",\"genus\",\"83632\",\"Ea
##  [8] "\"136404\",\"occ\",\"\",\"\",\"11400\",\"Alismataceae\",\"family\",\"55790\",\"
##  [9] "\"136956\",\"occ\",\"\",\"\",\"11480\",\"Alismataceae\",\"family\",\"55790\",\"
## [10] "\"136964\",\"occ\",\"\",\"\",\"11481\",\"Alismataceae\",\"family\",\"55790\",\"
```

I created a couple of functions to parse through this data, making several lists that aided in data visualization. Here is the first function, creating a list of accepted names:

```python
def makeAcceptedNameList(filename):
    # this function opens the seagrass data file and extracts
    # the accepted names
    AcceptedNameList = []
    # open the file
    acceptedname_object = open(filename, "r")
    # read the file and save contents as a list
    acceptedname_list = acceptedname_object.readlines()
    for line in acceptedname_list:
        acceptedname = line.split(",")[5]
        # stick accepted name into AcceptedNameList
        AcceptedNameList.append(acceptedname)
    return AcceptedNameList
```

Here is the second function, creating a list of periods:

```python
def makePeriodsList(filename):
    # this function opens the sea grass data file and extracts
    # the early period name for each occurrence
    PeriodsList = []
    # open the file
    periods_object = open(filename,"r")
    # read the file and save contents as a list
```

```python
    periods_list = periods_object.readlines()
    for line in periods_list:
        period = line.split(",")[8]
        # stick periods into PeriodsList
        PeriodsList.append(period)
    return PeriodsList
```

In terms of visualizing the data, I created dictionaries for both Accepted Names and for Periods where the values for each key were the number of times each key occurred in the data set. In Python, I used matplotlib to produce pie charts for each dictionary. In RStudio, I used ggplot2 to produce histograms for each dictionary. The work flow is shown below:

Dictionary and Pie Chart Code for Accepted Names–Python

```python
# importing matplotlib
import matplotlib.pyplot as plt
# setting font size
plt.rcParams['font.size'] = 100.0
# setting list variables
AcceptedNamesList = makeAcceptedNameList("no-header_alismatales_data.csv")
PeriodsList = makePeriodsList("no-header_alismatales_data.csv")

# creating empty accepted name dictionary
acceptedname_dict = {}

# filling accepted name dictionary
for name in AcceptedNamesList:
    acceptedname_count = AcceptedNamesList.count(name)
    acceptedname_dict[name] = acceptedname_count
unique_names = acceptedname_dict.keys()
unique_names_occurences = list(acceptedname_dict.values())

## making pie chart for order diversity
fig1, ax1 = plt.subplots()
ax1.pie(unique_names_occurences, labels=unique_names, autopct='%1.1f%%',
        shadow=True, startangle=0)
ax1.axis('equal')
fig = plt.gcf()
# make chart somewhat readable
fig.set_size_inches(200,200)
plt.show()
```

Dictionary and Pie Chart Code for Periods–Python

```python
# creating empty periods dictionary
periods_dict = {}
```

```python
# filling periods dictionary
for period in PeriodsList:
    periods_count = PeriodsList.count(period)
    periods_dict[period] = periods_count
unique_periods = periods_dict.keys()
unique_periods_occurences = list(periods_dict.values())

# making pie chart for periods diversity
fig2, ax1 = plt.subplots()
ax1.pie(unique_periods_occurences, labels=unique_periods, autopct='%1.1f%%',
        shadow=True, startangle=0)
ax1.axis('equal')
fig = plt.gcf()
fig.set_size_inches(200,200)
plt.show()
```

Faceted Histogram Code for Accepted Names–RStudio

```r
library(ggplot2)
ggplot(sea_grass_data, aes(x = accepted_name)) +
  geom_histogram(stat = "count", color = "blue") +
  facet_wrap(~accepted_name, nrow = 3)
```

Faceted Histogram Code for Periods–RStudio

```r
ggplot(sea_grass_data, aes(x = early_interval)) +
  geom_histogram(stat = "count", color = "red") +
  facet_wrap(~early_interval, nrow = 2)
```

In order to evaluate the geographical distribution of Alismatales, I created a world map with ggplot that plots the latitude and longitude coordinates of each fossil occurrence. Then, using the **cut** and **grep** commands in the shell, I created a csv file of the latitudes and longitudes of extant species. I then used this new data to create a world map that plots the latitude and longitude coordinates of fossil occurrence of stille existing taxa. The R code for each map is shown below:

Code for World Map of All Occurrences–RStudio

```r
mp <- NULL
# create a layer of borders
mapWorld <- borders("world", colour="gray50", fill="gray50")
mp <- ggplot() + mapWorld
# create map
mp <- mp +
  geom_point(aes(x=sea_grass_data$lng, y=sea_grass_data$lat),
             color="blue", size=1)
```

4

```
mp
```

Code for World Map of Occurrences of Extant Taxa–RStudio*

```
mp2 <- NULL
extant_alismatales <- read.csv("extant_alismatales.csv")
# see beginning of data file
head(extant_alismatales[1:3])
# create a layer of borders
mapWorld <- borders("world", colour="gray50", fill="gray50")
mp2 <- ggplot() + mapWorld
# create map
mp2 <- mp2 +
  geom_point(aes(x=extant_alismatales$lng, y=extant_alismatales$lat),
            color="blue", size=1)
mp2
```

Finally, I used the PyRate program to create speciation and extinction visuals. I ran the simulation 2,500,000 times.

Results

This section includes the actual visualizations of the data. For clarity's sake, only the Python Pie Charts of Accepted Names and Periods, the RStudio World Distribution Maps, and the PyRate Speciation/Extinction Visuals are included.
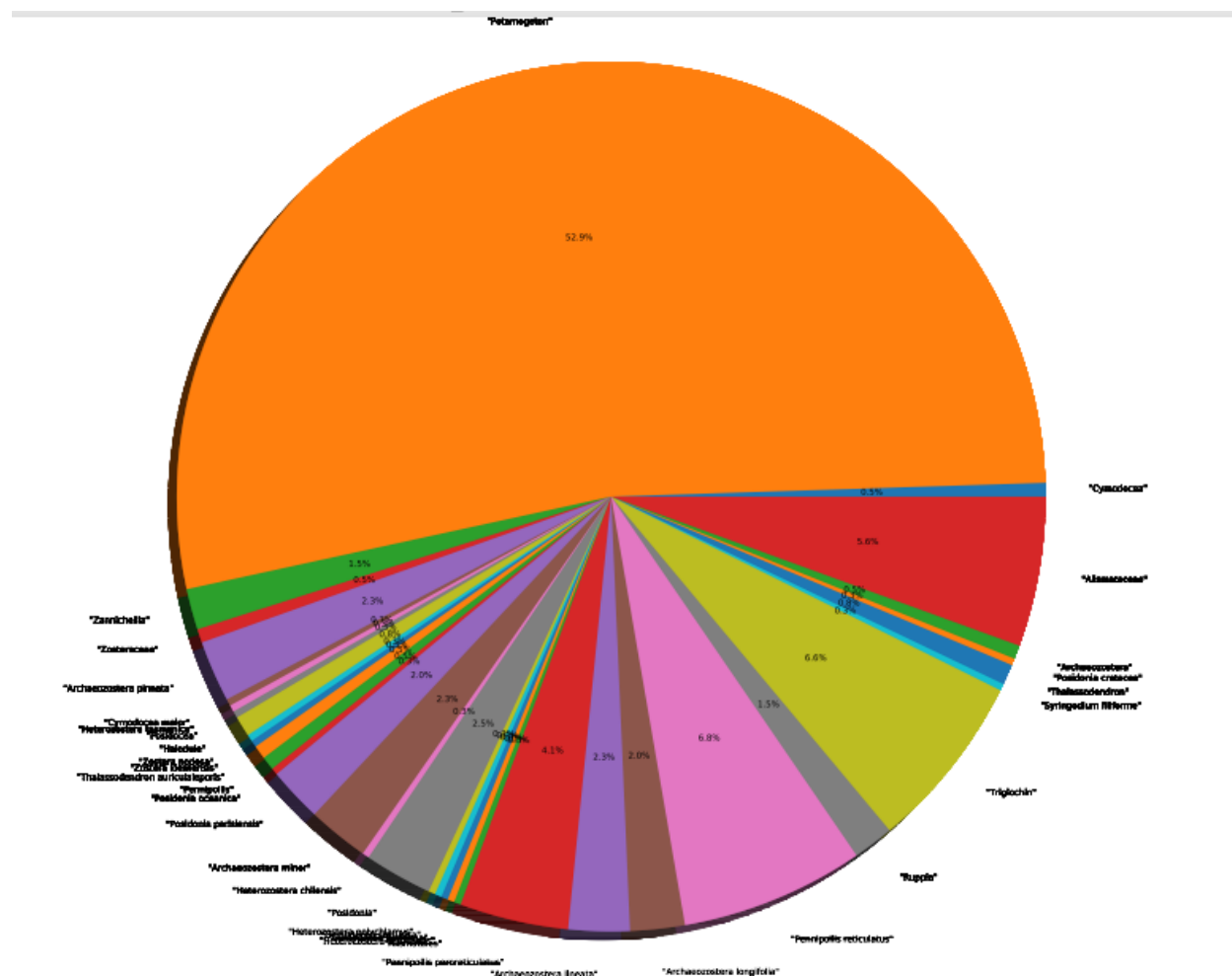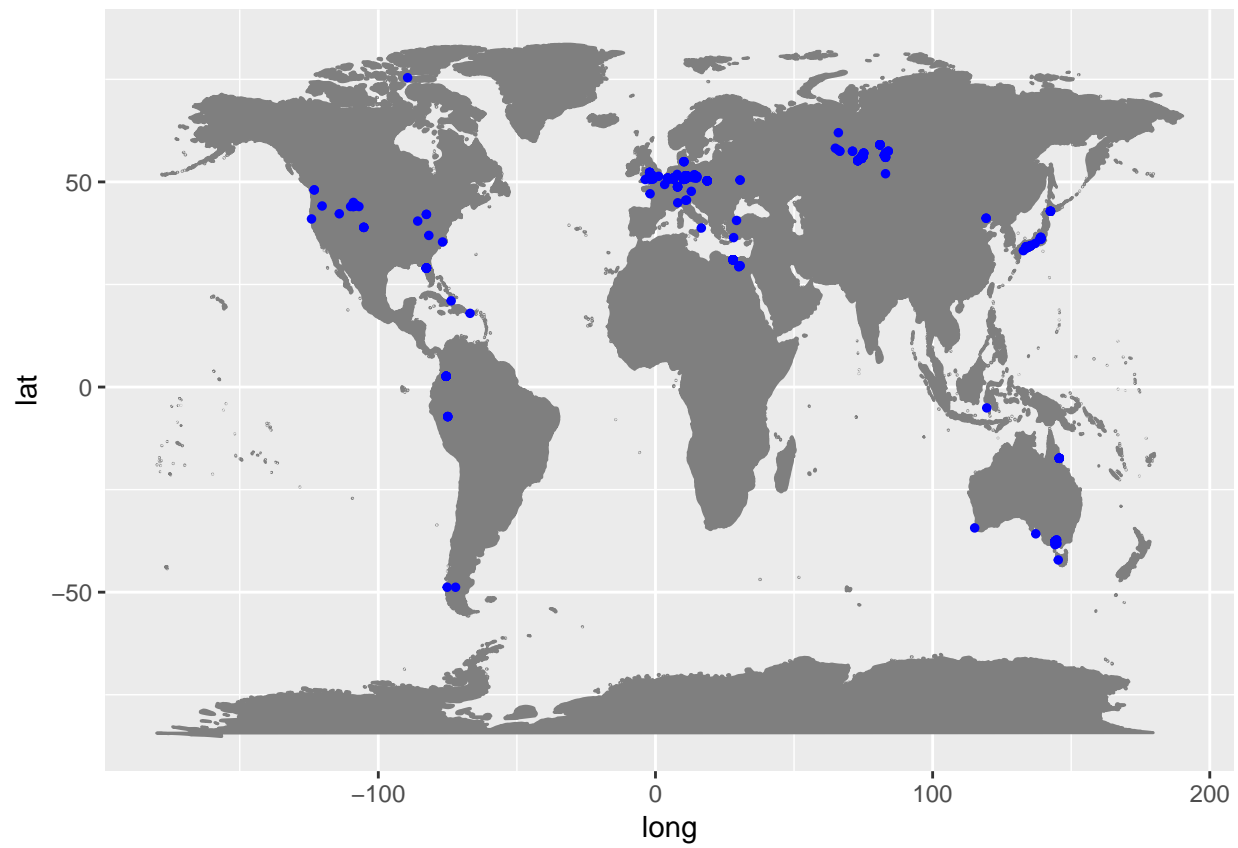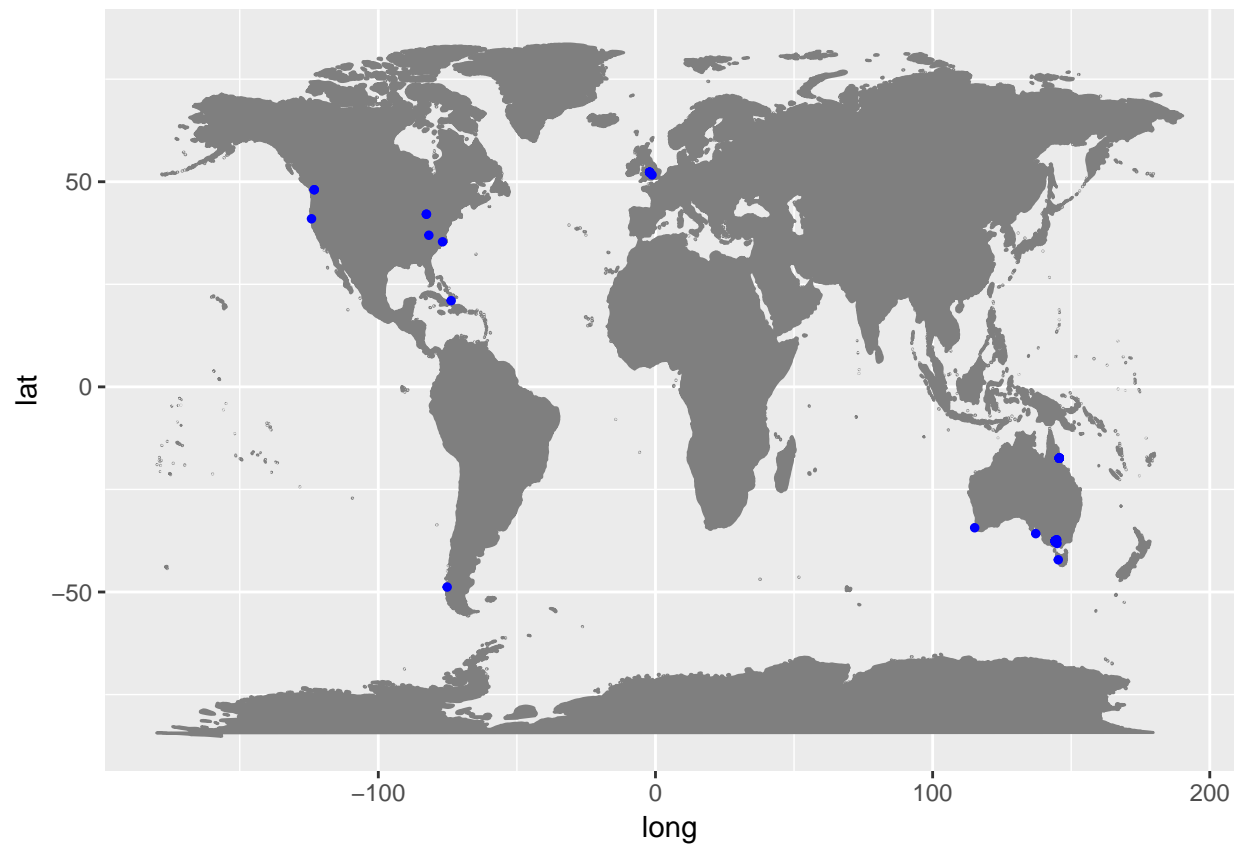

Figure 2: Python Pie Chart of Accepted Names

Figure 3: Python Pie Chart of Periods

RStudio ggplot World Map Seagrass Distributions

All occurrences

Occurrences of Extant Species
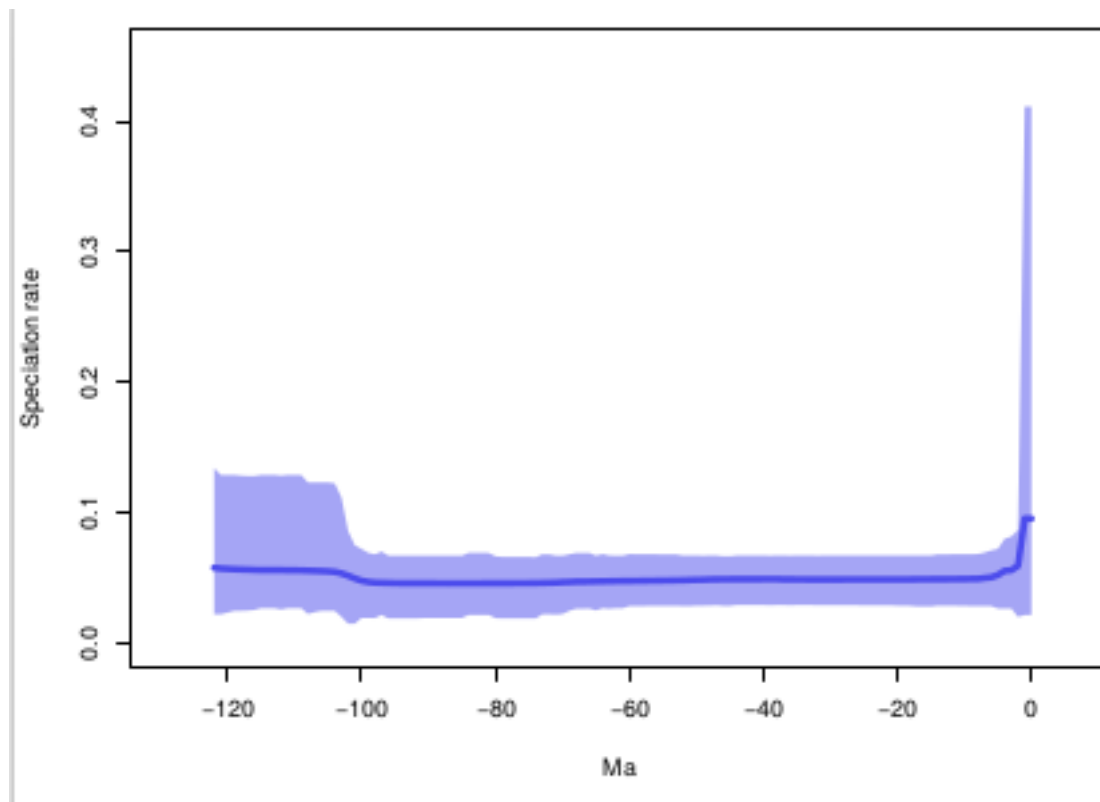
PyRate Speciation/Extinction Visualizations
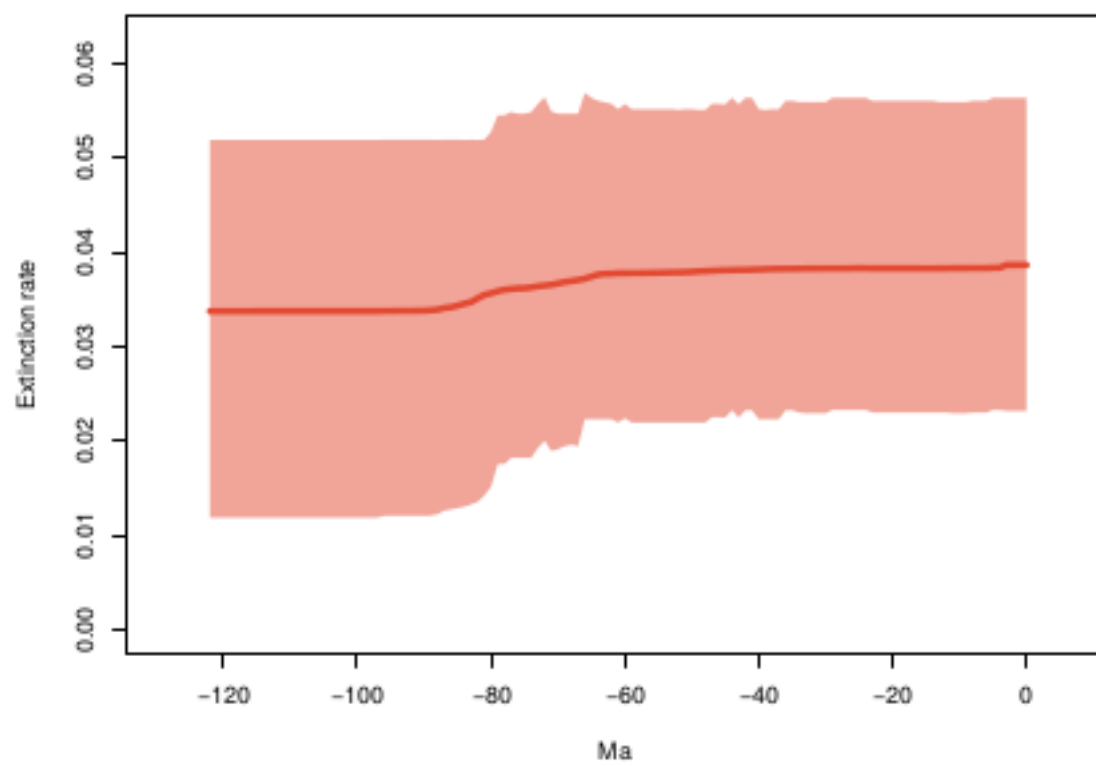

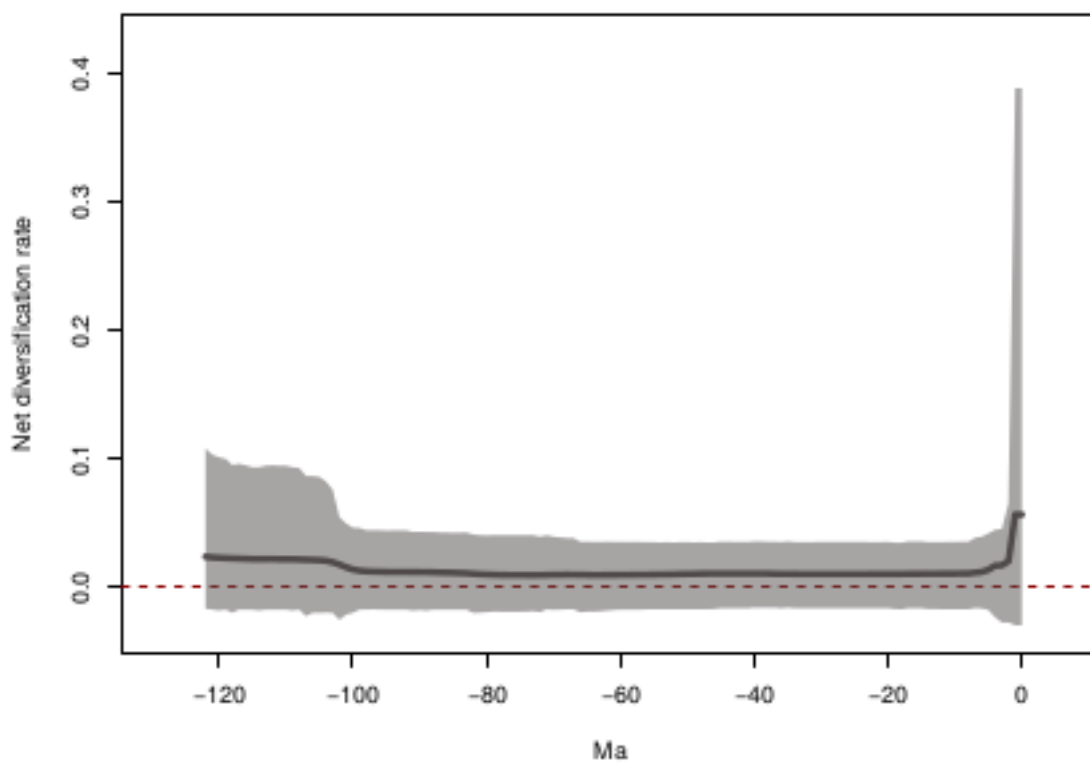Figure 4: Speciation graph

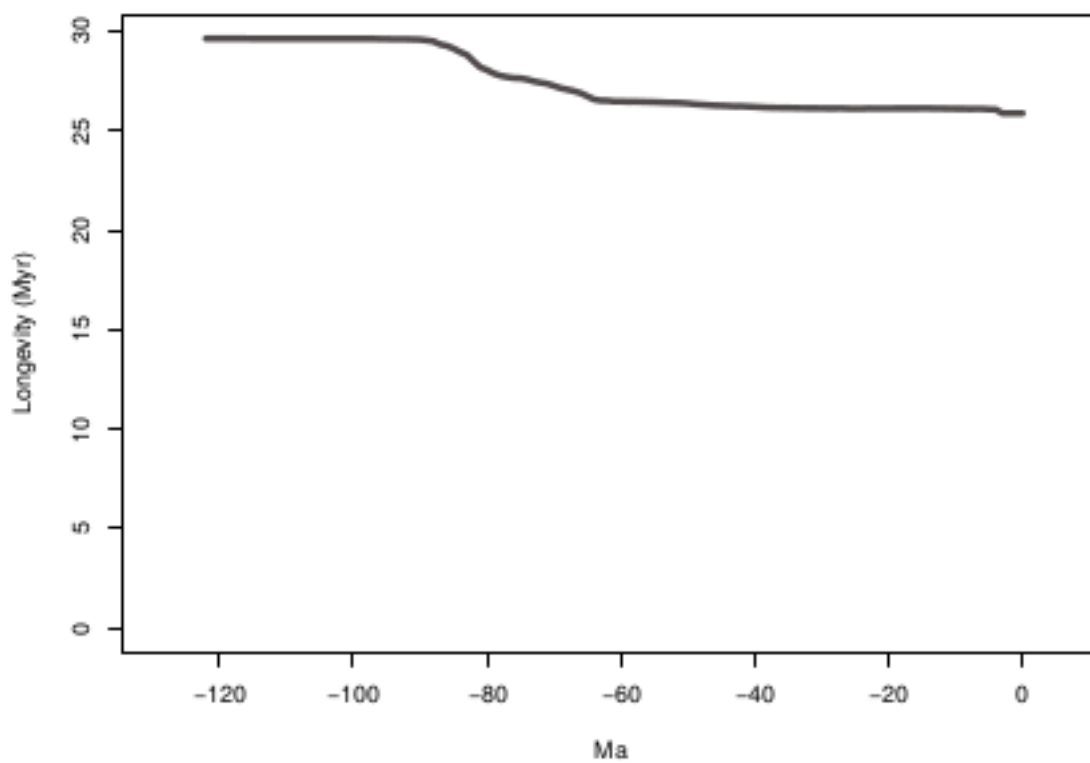Figure 5: Extinction graph

Figure 6: Net diversification graph

Figure 7: Longevity graph

Discussion and Conclusion

According to the PyRate visuals, the extinction rate of this order is kept relatively constant over the last 120 million years. The same goes for the speciation rate, save for the last approximately 2 million years where the rate spikes up just a bit. These two rates come together into the net diversification rate, which is shown as staying constant just above 0 for most of the last 120 million years. It spikes up just a little bit at the same point as the speciation rate. This information implies that there may have been a recent increase in the biodiversity of this order.

According to the ggplot World Map Distributions, compared side by side, the Alismatales order used to be more widely spread across the European, Asian, and North American continents. However, it appears that the habitation of the Australian continent is fairly new, as only occurrences of extant taxa are plotted in Australia.

One of the extant species, Heterozostera nigricaulis, is said to be "widely distributed in temperate Australia" (Kuo 2005). Therefore, it is possible that this rapid dispersal across Australia may be implicative of rapid dispersals in other parts of the world, such as the Caribbean. Seeing as the order has suddenly become more diverse and more rapidly distributed, it is reasonable that it is becoming more invasive.

Github Repository Links

https://github.com/cooperkass

https://github.com/cooperkass/final_project_eeb177

References

1.Kuo, J. (2005). A revision of the genus heterozostera (zosteraceae). *Aquatic Botany*, 81, 97–140.

2.Ochocki, B.M. & Miller, T.E.X. (2017). Rapid evolution of dispersal ability makes biological invasions faster and more variable. *Nature Communications*, 8, 14315.

3.Petersen, G., Seberg, O., Cuenca, A., Stevenson, D.W., Thadeo, M. & Davis, J.I. *et al.* (2015). Phylogeny of the alismatales (monocotyledons) and the relationship ofAcorus(Acorales?). *Cladistics*, 32, 141–159.

4.Tussenbroek, B. van, Katwijk, M. van, Bouma, T., Heide, T. van der, Govers, L. & Leuven, R. (2016). Non-native seagrass halophila stipulacea forms dense mats under eutrophic conditions in the caribbean. *Journal of Sea Research*, 115, 1–5.