Mapping with Style

Welcome to my repo for Computational Biology focusing on my independent project!

Biological Question/Introduction

Understanding species boundaries is still a hot topic in ecological and evolutionary biology. This opens your eyes to understanding the species niches and trait adaptation. Most importantly, this helps establish what reproductive isolation mechanisms are at play when two closely related taxa are in close proximity to one another. Hybrid zones are the final key in the equation to understanding reproductive isolation between two taxa. If one had the locations of parental species and hybrids then one maybe able to establish the species boundaries and hybrid zone, the first steps in understanding RI. My driving question for this project was to see if I could visually display species boundaries and hybrid zones.
I worked with a dataset that consist of locations, subspecies/hybrid, GPS locations, latitudes, longitudes, and various other information that is of little important with respect to this project.

Methods

The dataset is from a larger grant Dr. Rebecca Safran has been working on that has not yielded any published papers (plenty that are in the works). The grant was meant to look at whole genome sequences of hybrids and parental species and address multiple biological concepts/questions, for example, speciation, hybridization, evolutionary mechanisms, species boundaries, and more. Dr. Safran's graduate and postdocs student went to remote locations to collect the samples that are represented in the dataset.

Originally I deemed this project a challenge because I knew I would have to subset the data and work with the filtered subset. Later after speaking with Dr. Flaxman, I learned of an easier method to use that would allow me to use ggplots own built in grouping system to my advantage.

As far as I know, the original collector and Dr.Safran did not us this data to plot the locations on terrain maps or any maps of that nature. Seriously, I did not see any other maps or graphical plots that did what I did using ggmaps.

The goal of the project was just showcase the hybrid zones and parental species boundaries using mapping software and I believe I successfully achieved that goal. My code shows multiple global locations where the bird samples were taken and also shows a region around China, Russia,

and Mongolia where the a great depiction of parental species boundaries and hybrid zones. These maps including the hybrid map isn't suppose to show defined lines but rather an idea of where this different organism are found and their sympatric lifestyle.

Although my data set proved to be easier to use and manipulate than I expected originally, one challenge that I faced with filtering the hybrids from the parental organisms. For an advanced coded, this task would seem simplistic but due to the novice stage I consider my self in, the challenge ruled over several days. I basically did not know how the filter function truly worked. After alternating between multiple coding scenarios, I was able to figure our the function and complete my for loop (which I am also proud of).

Results and conclusions

I am unable to add the images to this markdown file but they are located in the *Assignment 9 image file* located in the same directory. Again, my project wasn't to replicate what someone else did. It was to master the art of visual representation of biological species boundaries using a dataset that was otherwise used for Identifying DNA samples. To see the work that was done and see the boundaries you must look at the maps in the described file. Lastly, please note that the dataset used is not open to the public at the moment, I apologize for the inconvenience.

Sam I will email the data to you.

thank you