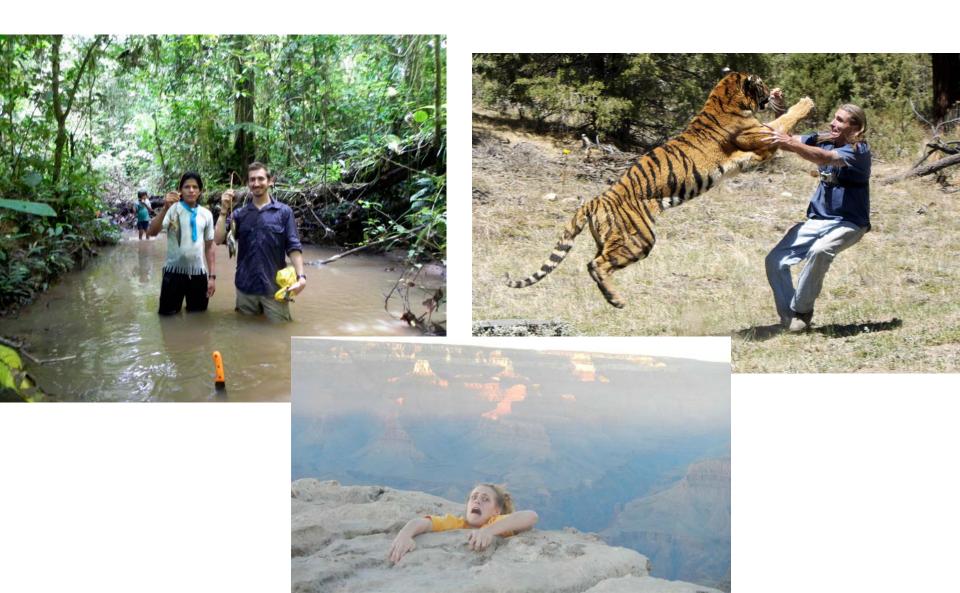
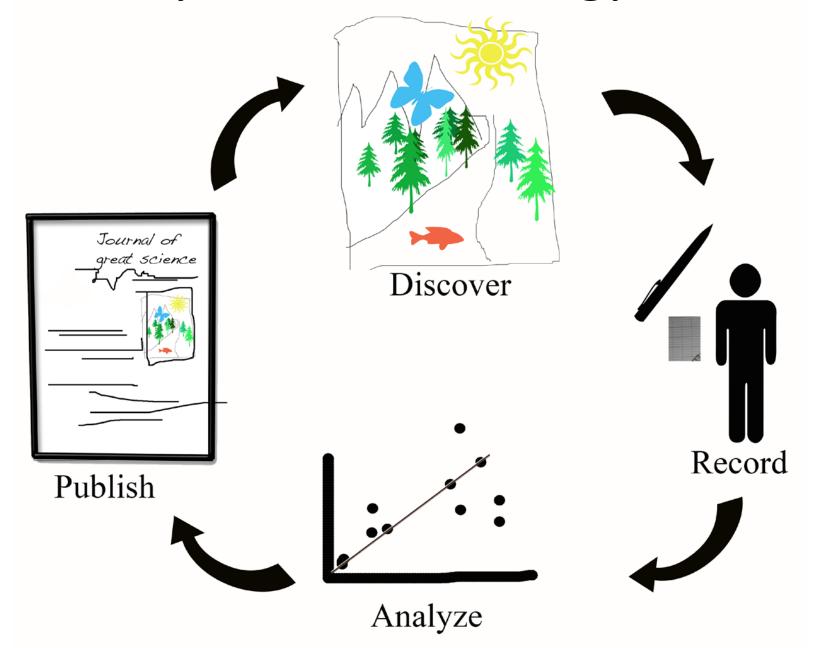
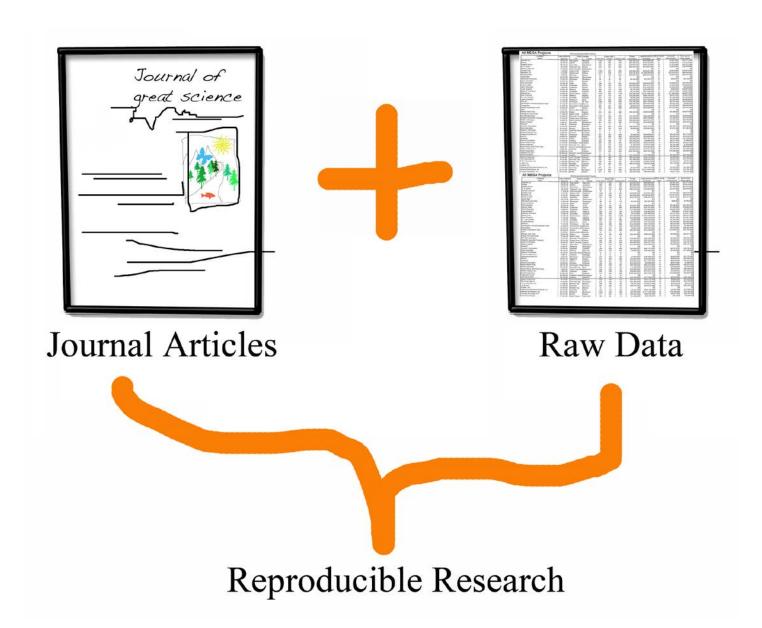
# why a field to db workshop?



# lifecycle of field biology data



# publishing today













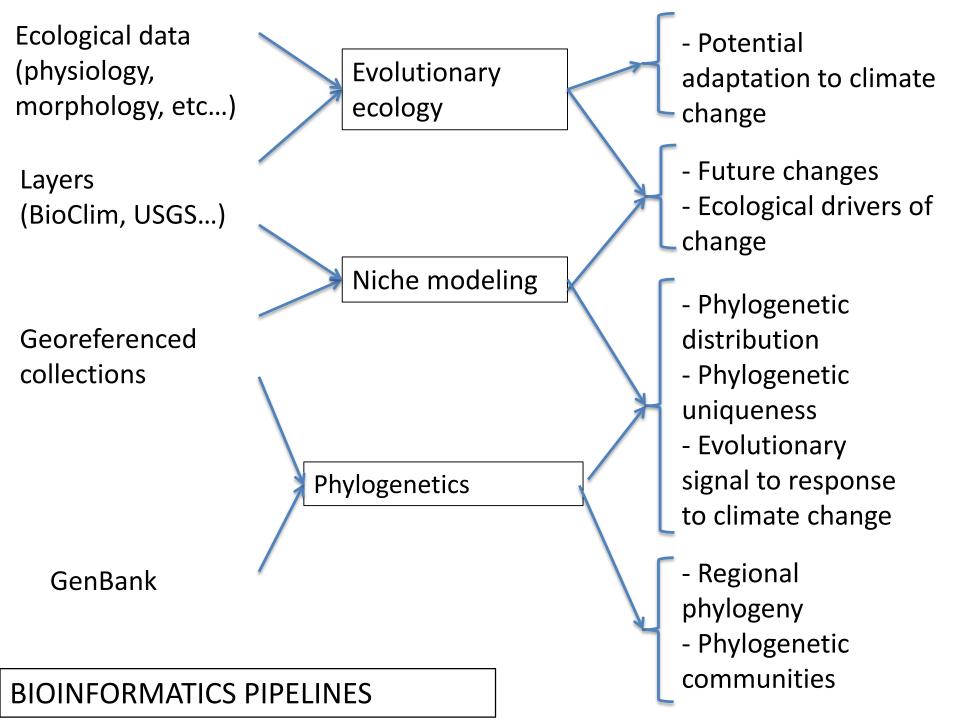
What do we do with all of the data?

# Using big data for big questions

- Distribution of biodiversity over landscape?
- Distribution of characteristics over landscape?
  - (trees, herbs, endemics, xeric-adapted, etc...)
- How will climate change impact this distribution?
- Will the impact be the same for all species/communities?
- Has it started already?

.... etc, etc.....

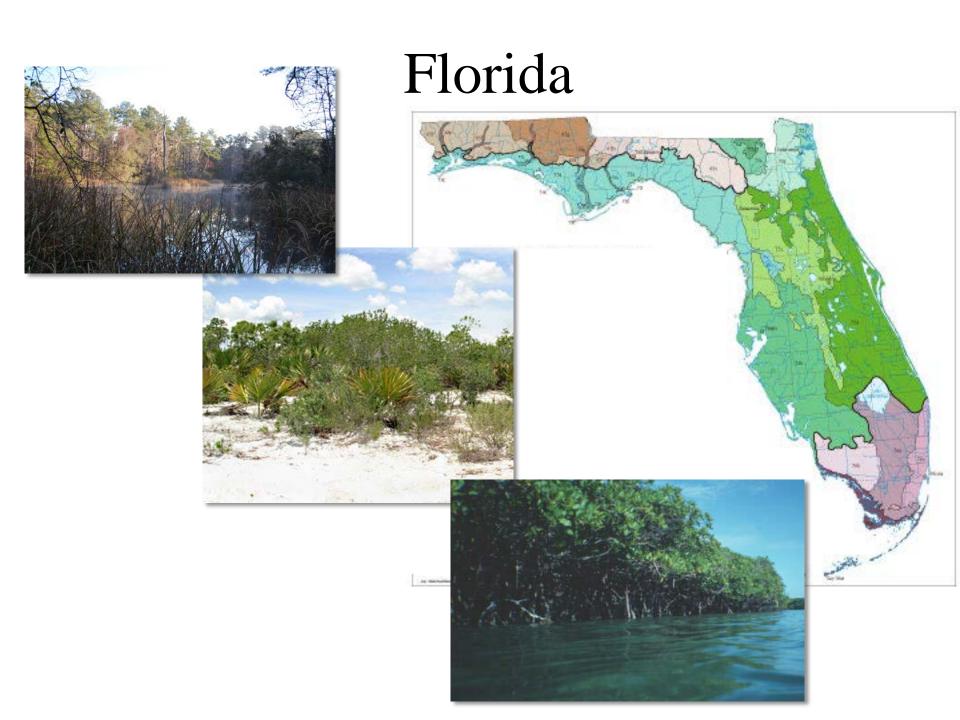
Combine with other sources of big data





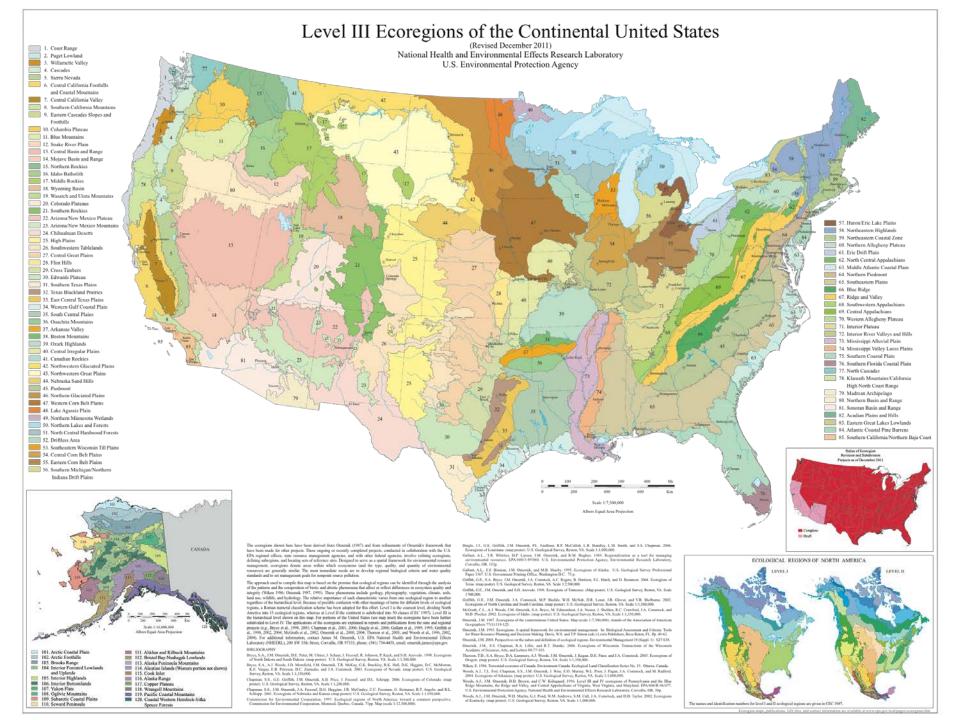
#### Research applications of museum data: The past, present and future of Florida plants

Julie Allen, Charlotte Germain-Aubrey, Douglas Soltis, Robert Guralnick, Jose Miguel Ponciano, Lucas Majure, Kurt Neubig, Pamela Soltis



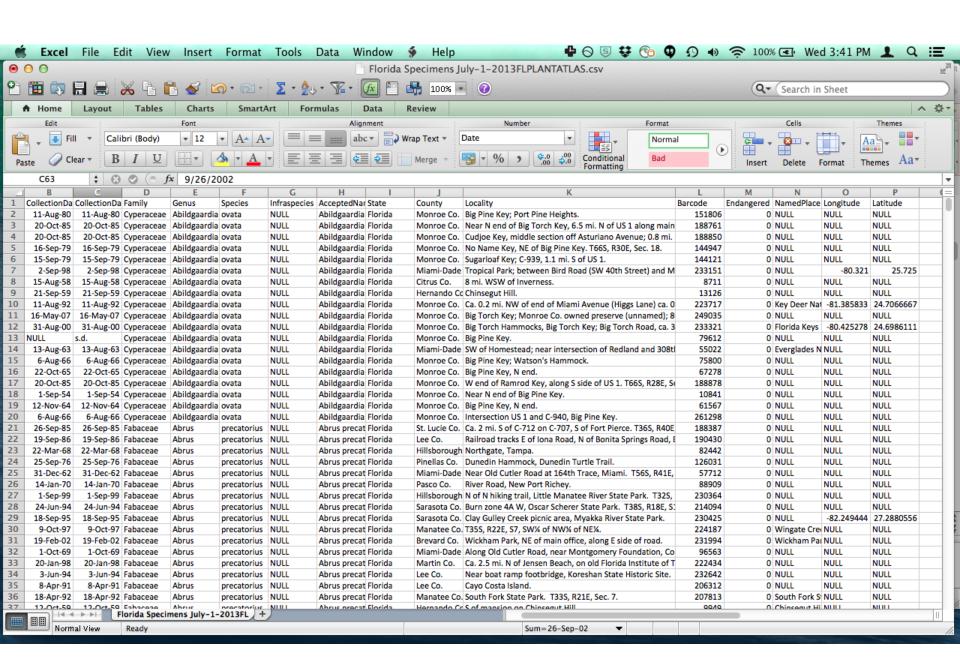
# Road Map

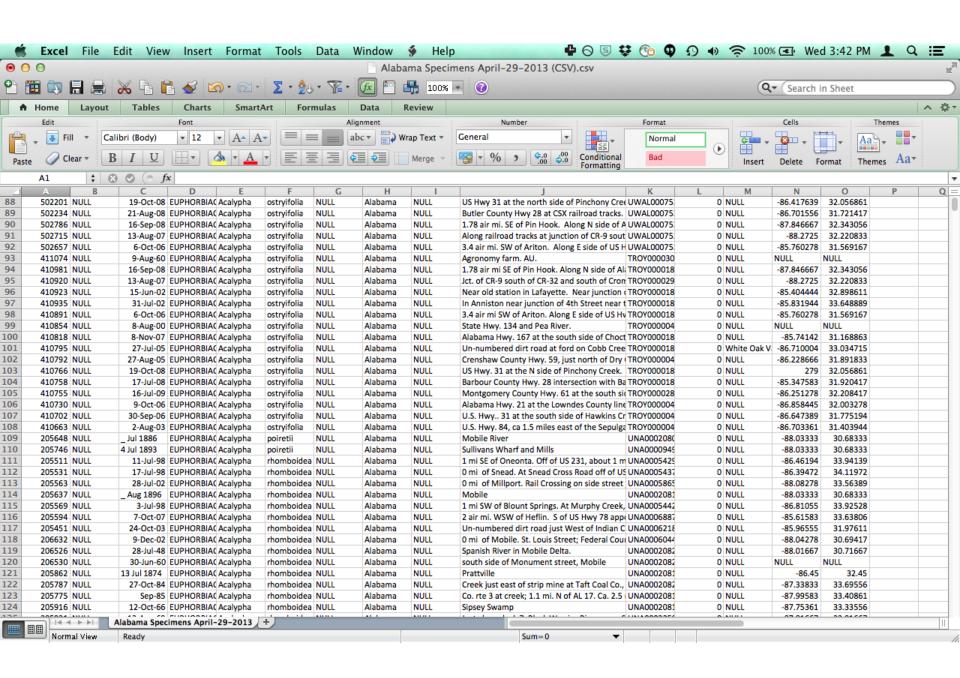
- Collect museum data
- Ecological niche models for each species
  - understand diversity of Florida
    - Past, Present, Future
- Phylogenetic Tree of these same species
  - explore the phylogenetic diversity of Florida

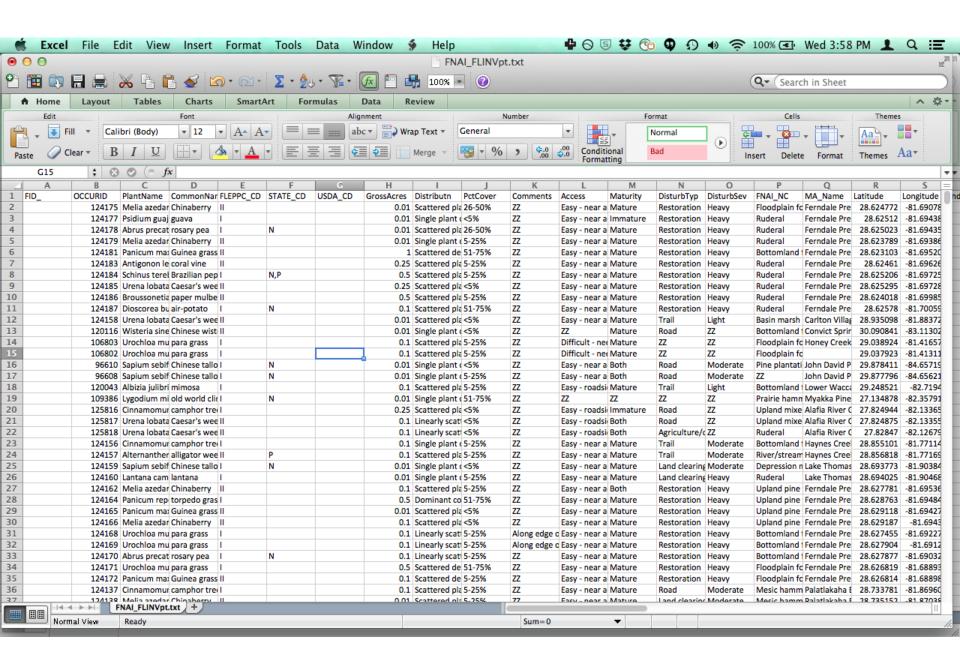


# Data collecting

- Florida Plant Atlas
- Florida Native Area Inventory
- Global Information Facility
- Florida State University Herbarium
- Louisana State University Herbarium
- University of North Carolina Herbarium
- Alabama Plant Atlas
- Mississippi State University Herbarium
- Florida Museum of Natural History Herbarium
  - >500,000 georeferenced points







# Formatting challenges

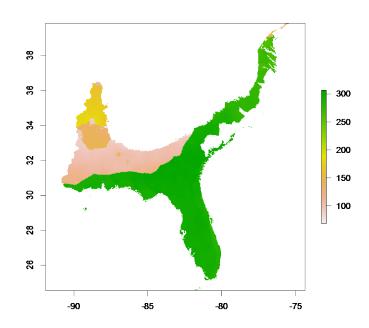
- Reconciling datasets within each institution
- Reconciling datasets between institutions
- Georeferencing specimens
- Transforming Lat/Long
- UNCERTAINTY info missing !!!!
- Dates format (had to make assumptions)
- Taxonomic Name Reconciliation
- Pool datasets together large files

# **Data collecting**

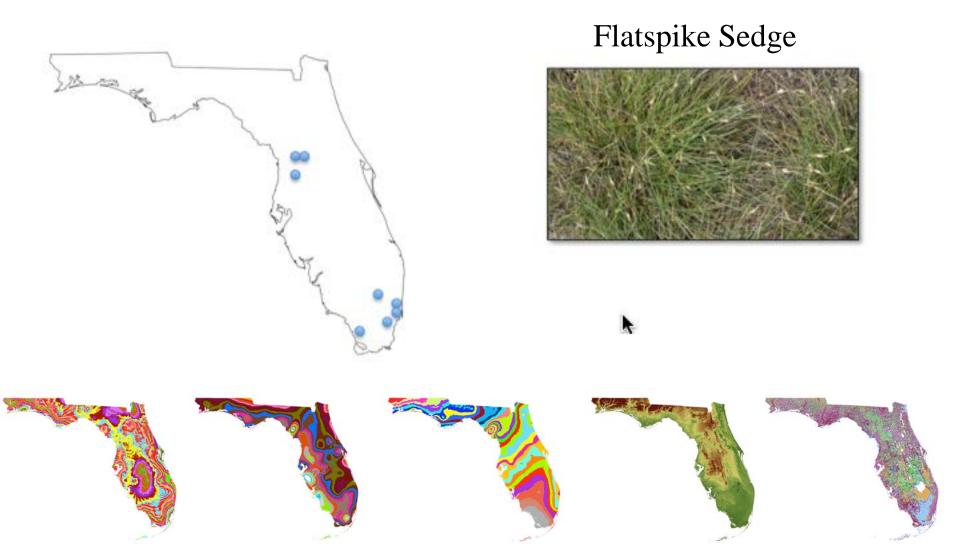
- Florida Plant Atlas
- Florida Native Area Inventory
- Global Information Facility
- Florida State University Horbarium
- Louisana State University Herbarium
- University of North Carolina Herbarium
- Alabama PlamAtias
- Mississiphate University Herbarium
- Florida Juseum of Natural History Herbarium
   500,000 georeferenced points

# Data cleaning

- Wunderlin list of 4,094 species of Florida plants
  - Check list against Tropicos accepted names
- All non-Florida species removed
- Duplicates removed
- 3 EPA ecoregions
   391,937 points
   343,266 dated points
- 30+ points per species
  - 372,241 pts and 1,548 species



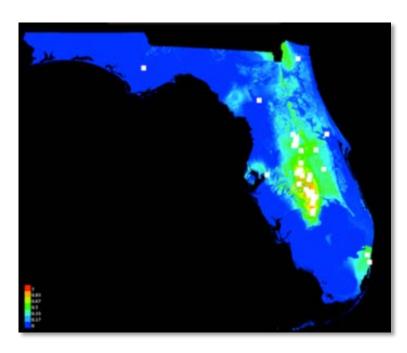
# Ecological Niche Modeling



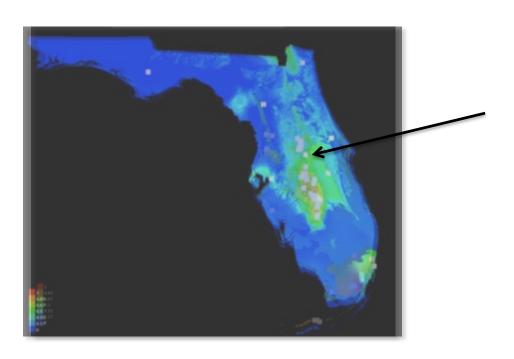
#### Flatspike Sedge



#### Scrub Palm

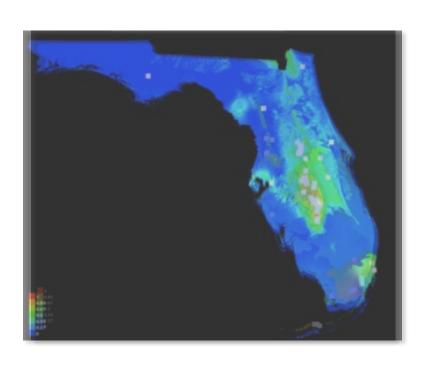


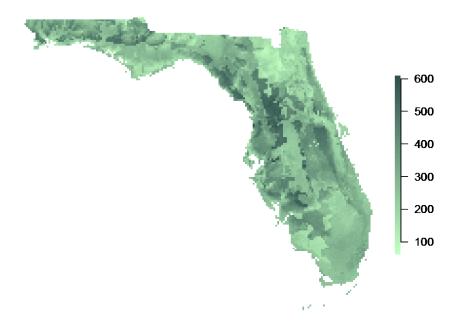




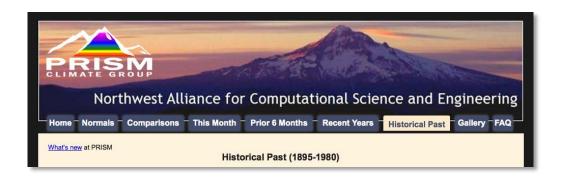
How many species are predicted to reside in this point?

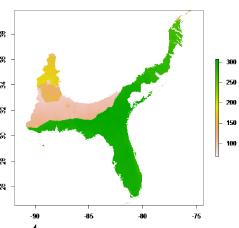
# All Plant Diversity





## Museum Specimens and Climate Data



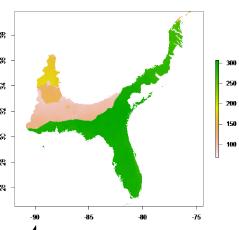


 Extraction R – package – dismo to create bioclim layers from monthly PRISM data

• Associate each species record with the climate data for the correct year.

## Museum Specimens and Climate Data





 Extraction R – package – dismo to create bioclim layers from monthly PRISM data

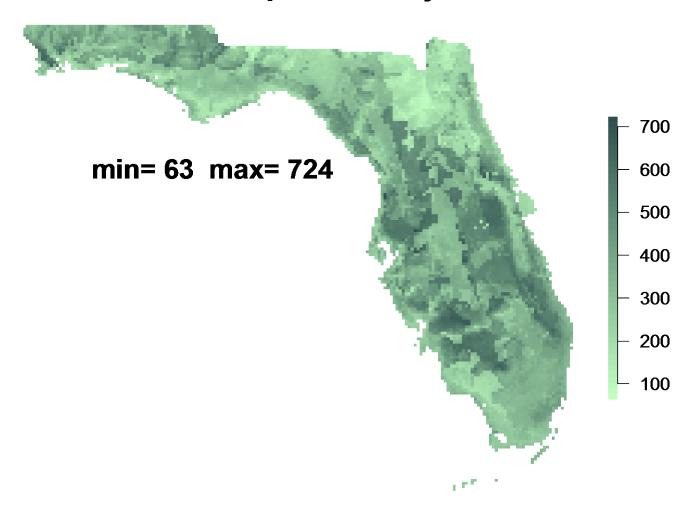
• Assoc clima R package in progress clima r. vith the

#### Climate Data

- Bioclim correlation 8 layers < 0.85
- Altitude
- Geology

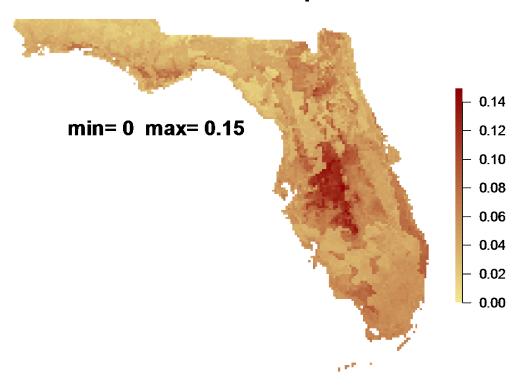
Ran models for all 1,548 species. Combined maps to create a heat map of all species and then cropped it to Florida.

#### **Present Alpha Diversity**



# Endemism hotspots

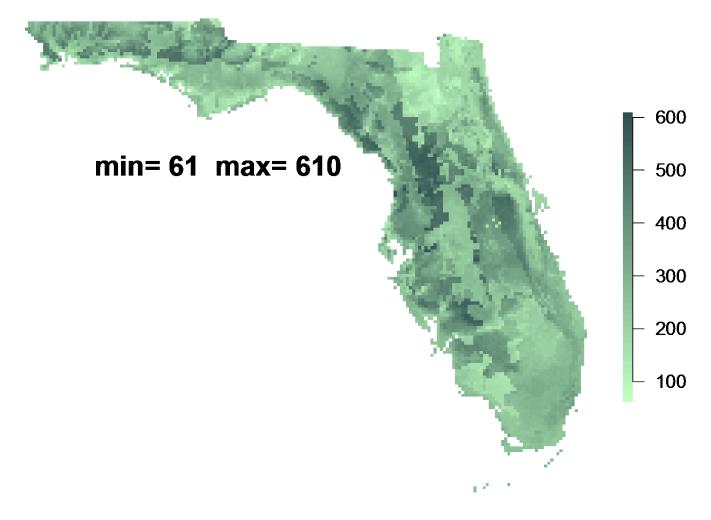
#### **Present Endemic Hotspot**



#### Florida's Past

• Take models of each species and project it to the climate layers created from 1902 – 1912.

### Florida's Past



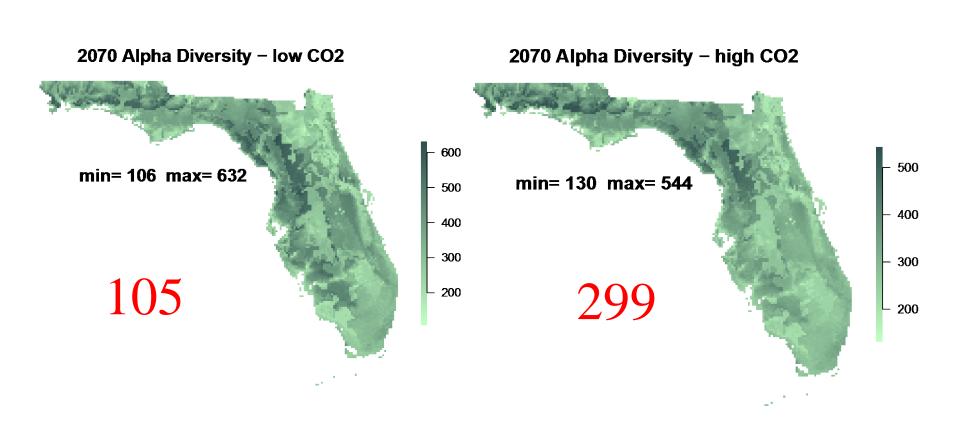
1902 - 1912

#### Future Models

- Consensus models from 3 hand-picked future scenarios:
  - CCSM4 -Community Climate System Model
  - MPI Max Plank Institute
  - MIROC5 Model for the Interdisciplinary Research on Climate
- Both highest and lowest estimates of CO<sub>2</sub>
- 2050 and 2070

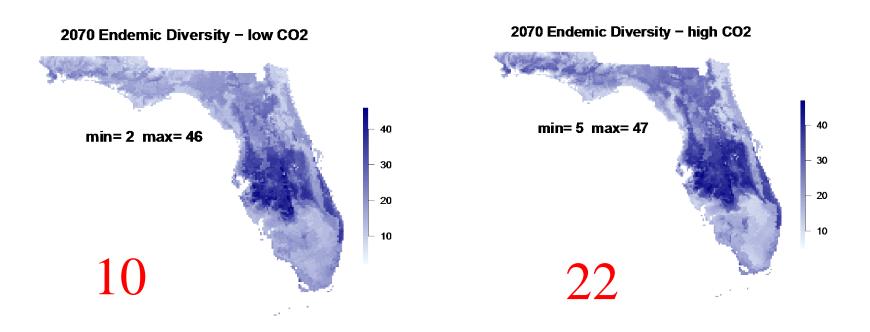
12 models for all 1,548 species averaged for a consensus of these models

# Future Projections



Species Extinct from Florida

#### **Future Endemics**

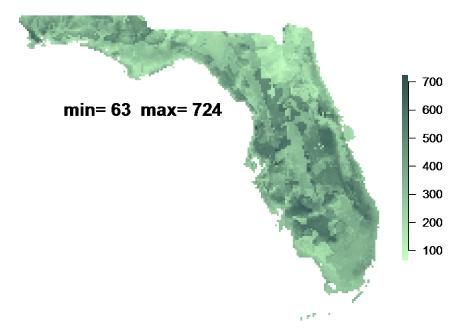


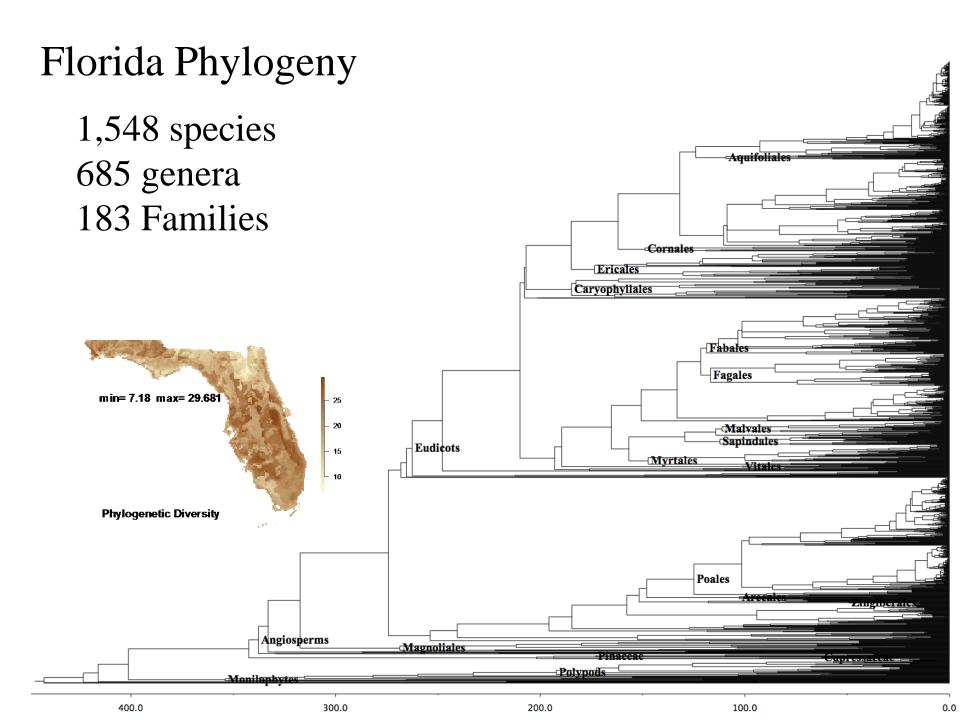
#### **Endemic Species Extinct**

# Next Steps

- Species clusters
- Movement Analysis
- Sea level rise change







# 'basic' niche modeling

- R package Biomod2
- Use Bioclim climatic layers
- Use altitude and geology
- 9 species
- Run 3 algorithms: general linear model, random forest, maximum entropy