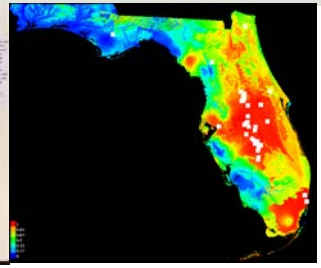
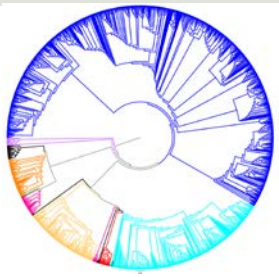
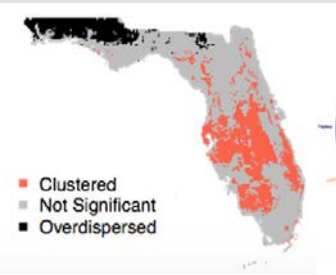
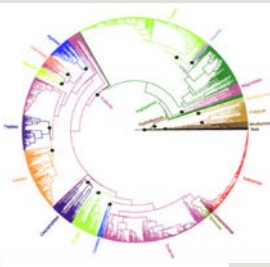









Linking Heterogeneous Resources for Biodiversity Research

Pamela S. Soltis

FLORIDA MUSEUM OF NATURAL HISTORY

iDigBio
Integrated Digitized Biocollections

UF UNIVERSITY of FLORIDA
The Foundation for The Gator Nation

Collections: The Library of Life

>1600 natural history collections in the US

1-2 billion specimens in the US

3-4 billion specimens worldwide

Systematics and Taxonomy

Linnea (twinflower)

Carl Linné, aka Carolus Linnaeus

Collections: The Library of Life

Genetics
Genomics
Chemistry
Species interactions
Phenology
Biogeography
More!

Collections: The Library of Life

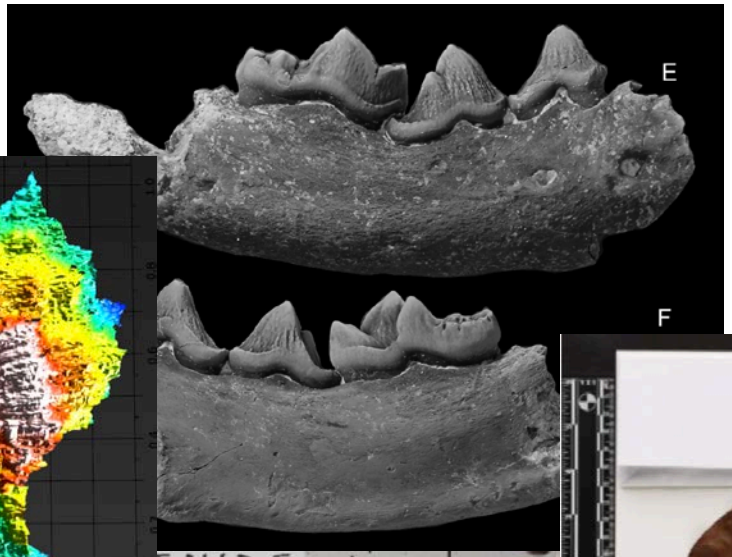
Most specimens locked away in cabinets, unavailable for general use.

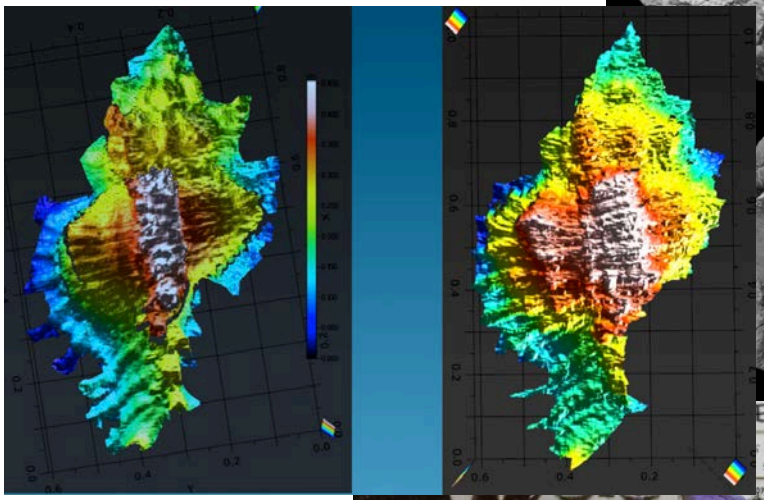
DIGITIZATION!!!!

Label Data

- Scientific name – including authority
- Date
- Collector
- Location – state, county, specific site, GPS coordinates
- Associated species
- Notes

Images





iDigBio
Integrated Digitized Biocollections
www.idigbio.org

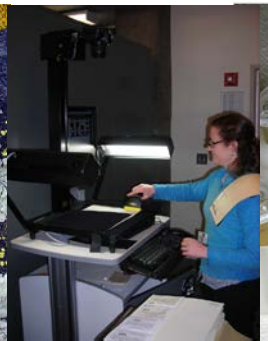
National Coordinating Center
For Digitization of Biodiversity Collections

Ingest, serve, integrate data:

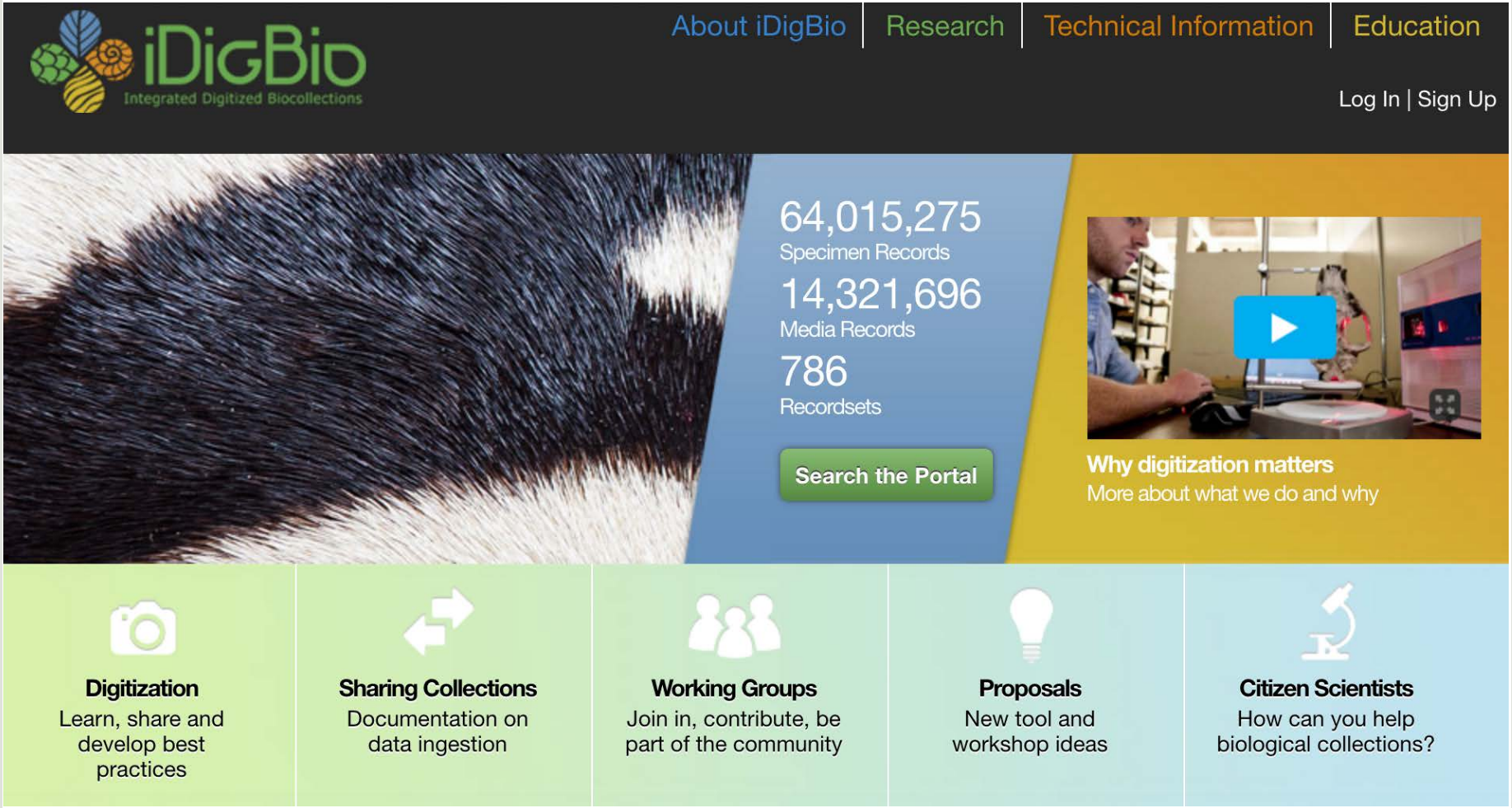
Localities

Dates

Images

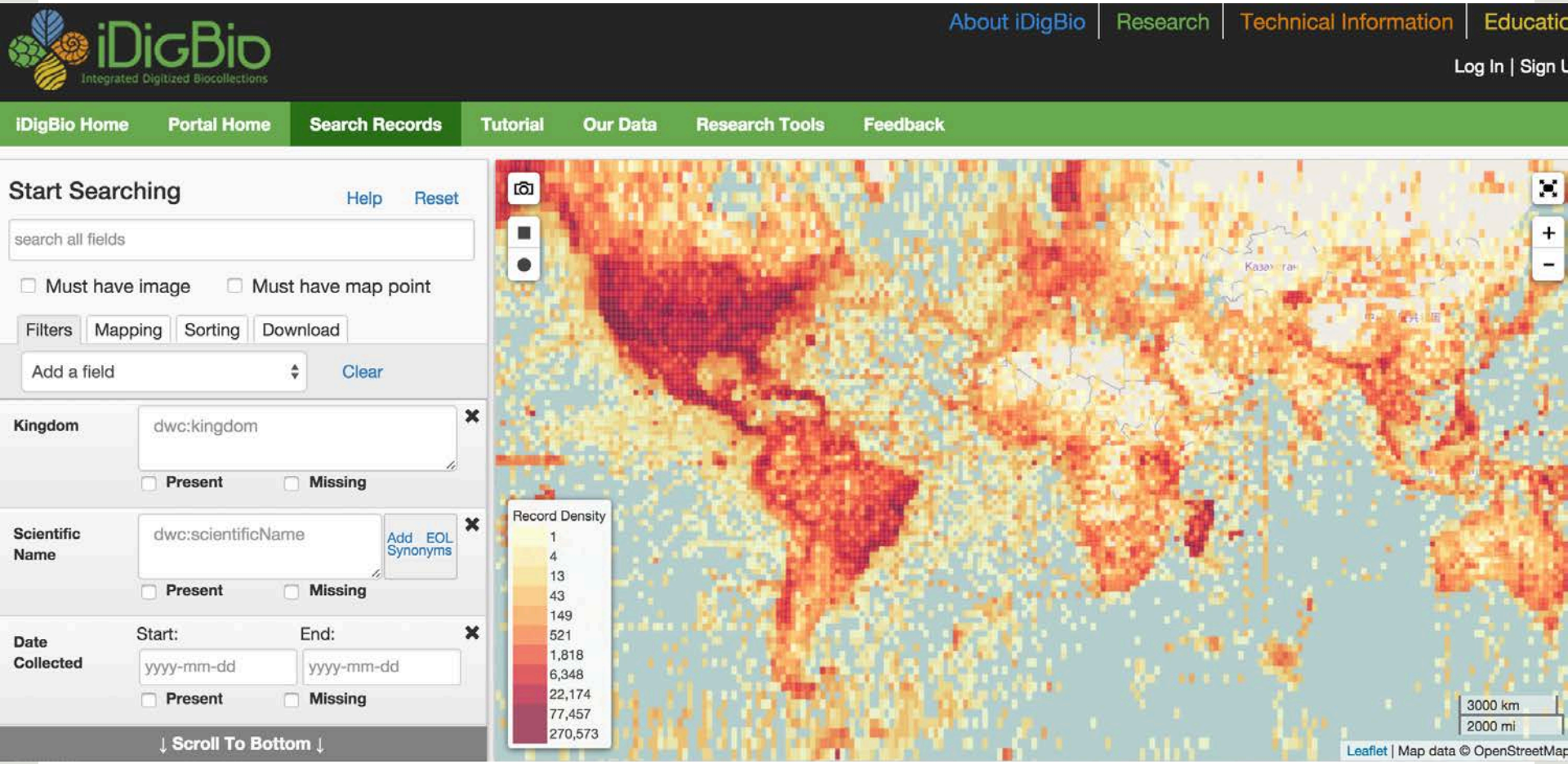


iDigBio Homepage

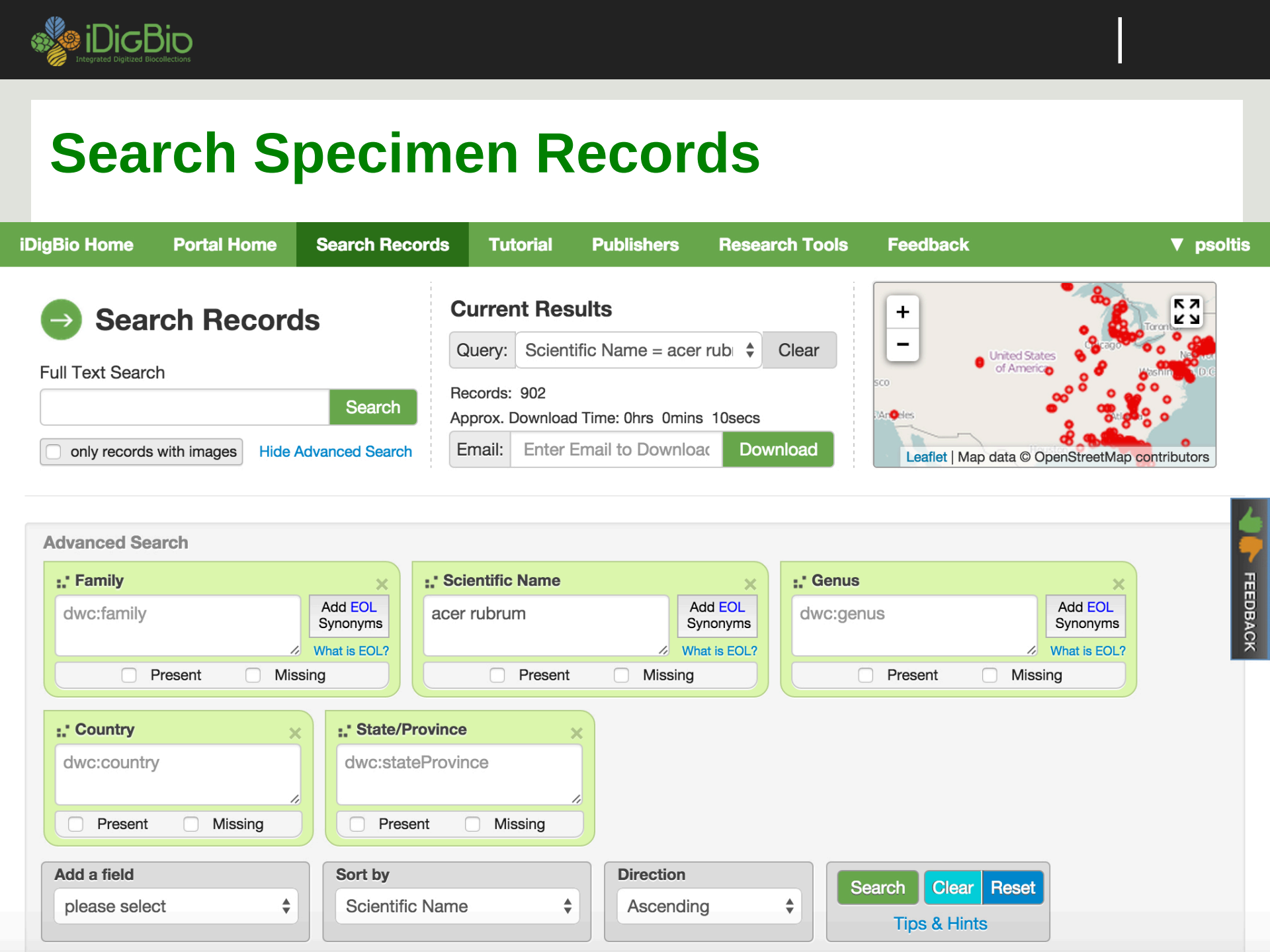


www.idigbio.org

Search Specimen Records



Search Specimen Records



Search Specimen Records

Media Record: *Acer rubrum*, Wats.

praae.idigbio.org/arch/nansh/ARCH01/ARCH01366_lg.jpg

Go To Specimen Record

Go To Recordset | View Raw Data

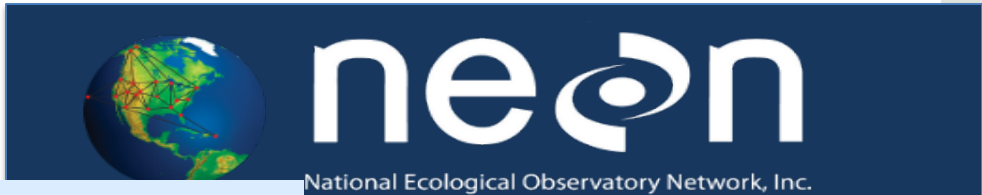
Media Metadata

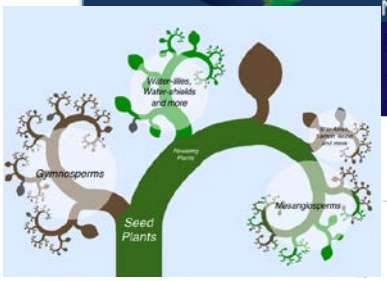
Associated Specimen Reference	http://nansh.org/portal/collections/individual/index.php?occid=4139508
Type of Resource	StillImage
Subtype	Photograph
Metadata Date	2014-02-19 00:58:28
Provider-managed ID	urn:uuid:fcedaa46-86c0-4150-8e4b-af6d430ca43a
Credit	Archbold Biological Station (ARCH-herbarium)
License Terms	CC BY-NC-SA (Attribution-NonCommercial-ShareAlike)
License URI	http://creativecommons.org/licenses/by-

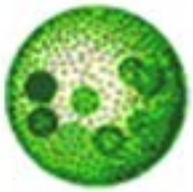
FEEDBACK

Linking Collections to…

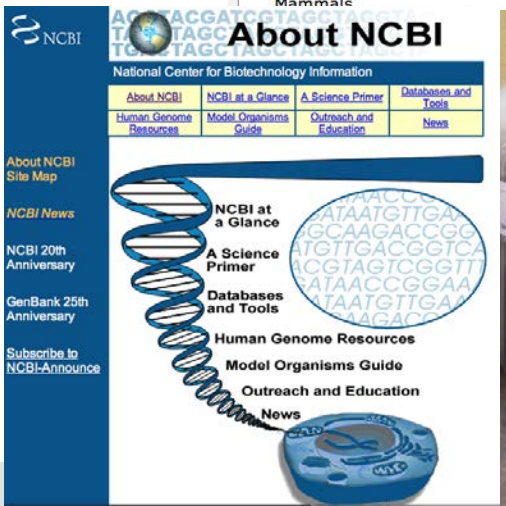
- Phylogeny
- Ecology
- Paleontology
- Living Collections
- Genomics
- Other Repositories







UTEX



Linking Heterogeneous Resources for Biodiversity Research

- Infrastructure
  - PhyloJIVE/PhyloLink
  - Connecting iDigBio, Open Tree of Life, Lifemapper
- Examples
  - Spatial patterns of phylogenetic diversity
    - Florida Plant Phylogeny Project
  - Community phylogenetics: case study
  - Plant traits and phylogeny

PhyloJIVE

Links biodiversity data to trees

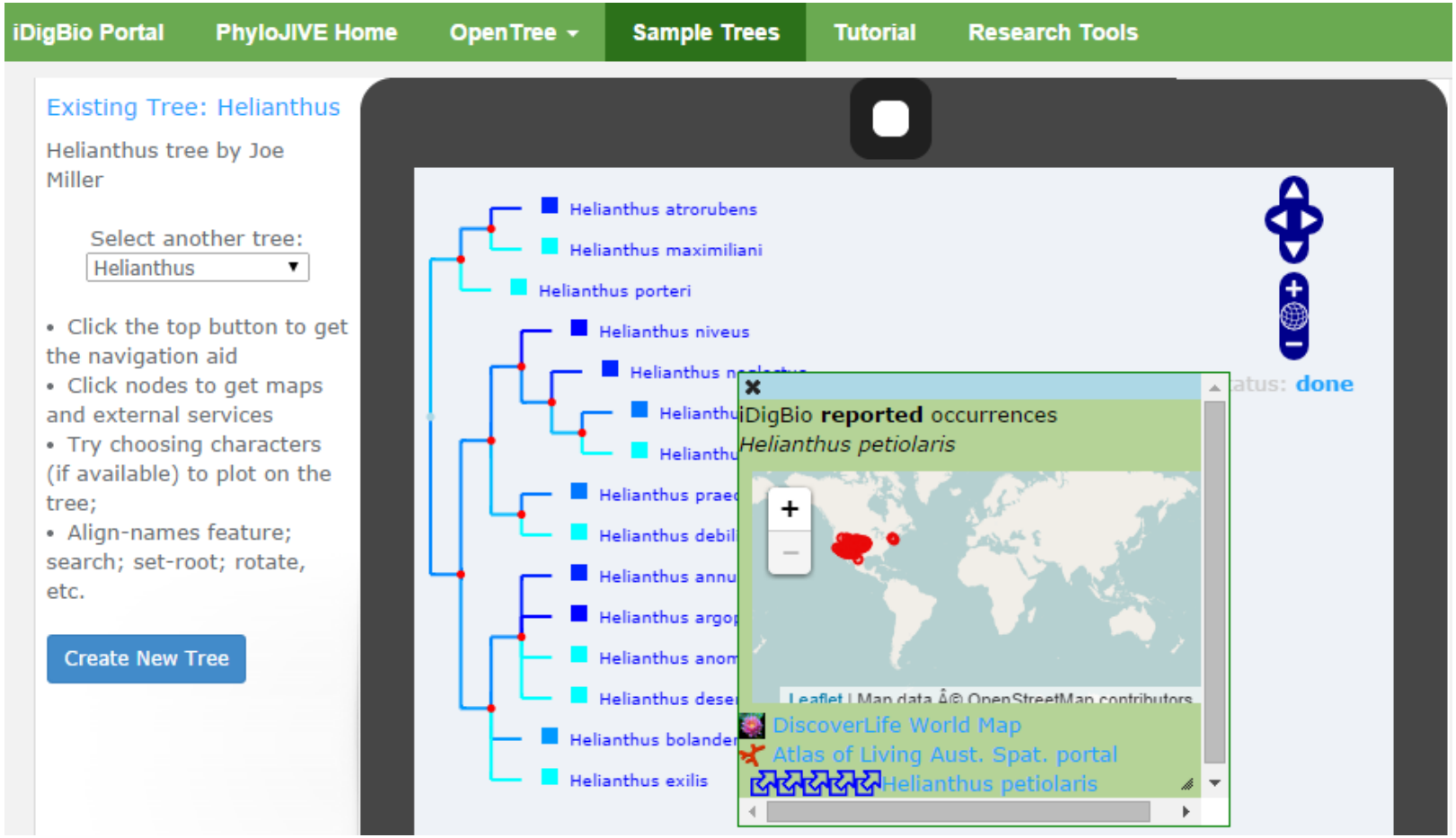
Joe Miller & Garry Jolley-Rogers

phylojive.ala.org.au/

A. buxifolia

Source: Australian Plant Image Index Image by: Macd

PhyloJIVE instance in iDigBio



- Developed by Garry Jolley-Rogers, Joe Miller, and Temi Varghese
- Integrates biodiversity data with phylogeny
- http://phylojive.acis.ufl.edu/

A. Matsunaga

Phylolink

Atlas Of Living Australia | ALA Apps | ALA Info | Search the Atlas | Search | User settings

You may see intermittent outages with Phylolink for the next few days, while Atlas of Living Australia moves to a new system.

Phylolink

Overview

Phylolink is a collection of tools through which biodiversity can be explored from a phylogenetic (or tree of life) perspective.

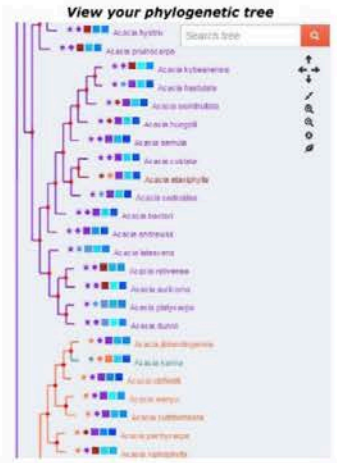
At the core of these tools is the ability to easily intersect a phylogenetic tree with species occurrence records, environmental data, and species character information.

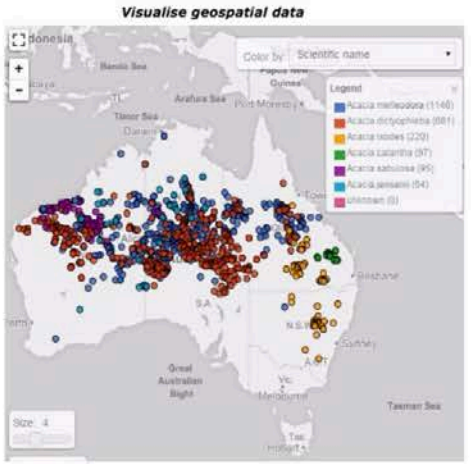
The result is powerful ways of combining data to generate flexible and customisable visualisations, profiles and metrics for biodiversity.

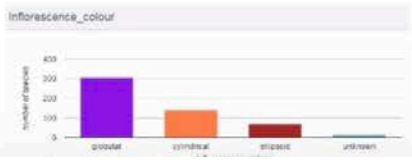
View an example demonstration here. Or, view screencast on how to view phylolink here.

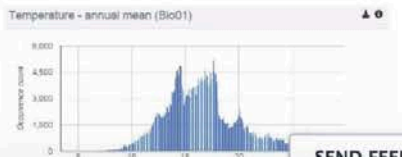
Start Phylolink

Explore Phylogenetic Diversity



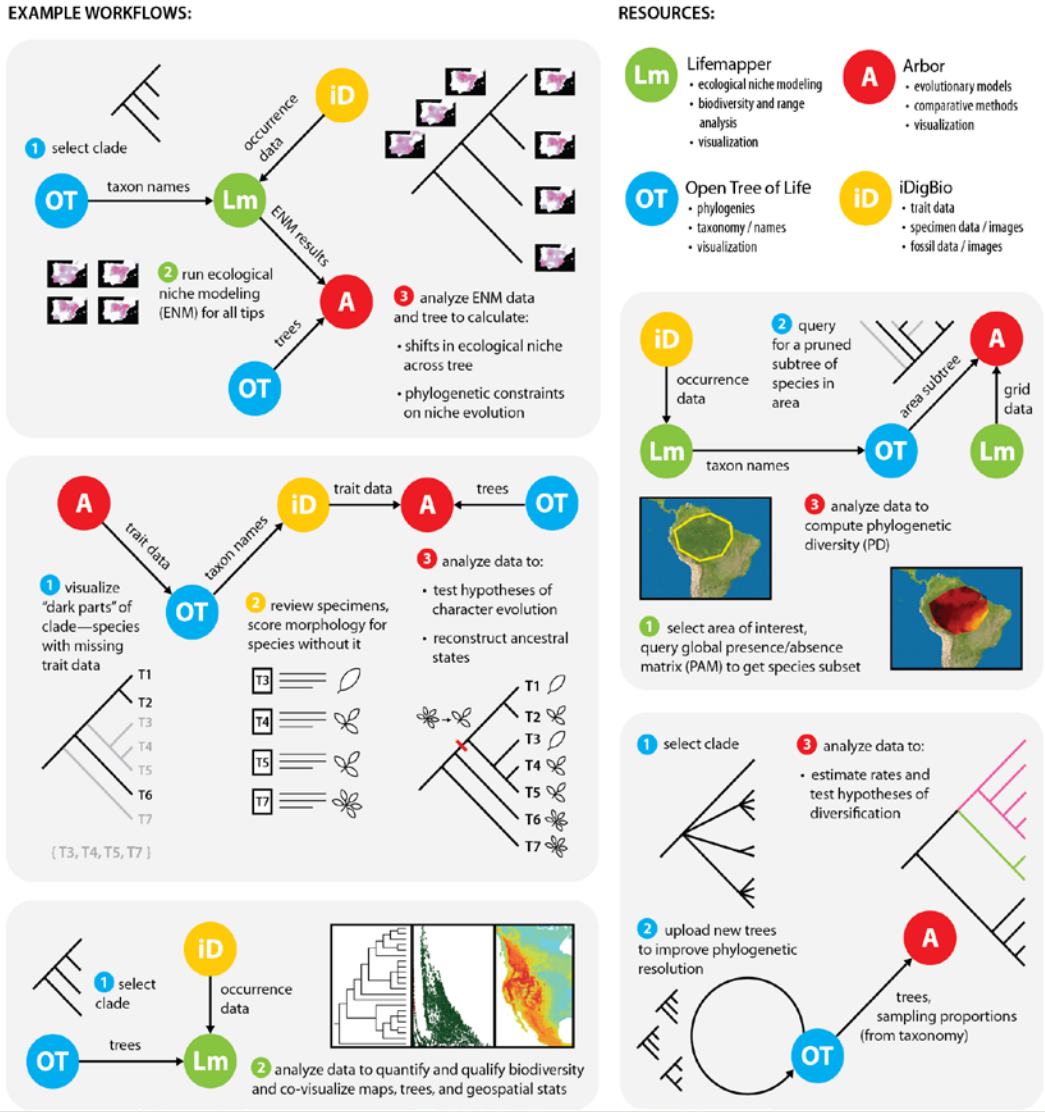






Connecting Trees, Specimens, Tools

Connecting Trees, Specimens, Tools



Connecting Trees, Specimens, Tools

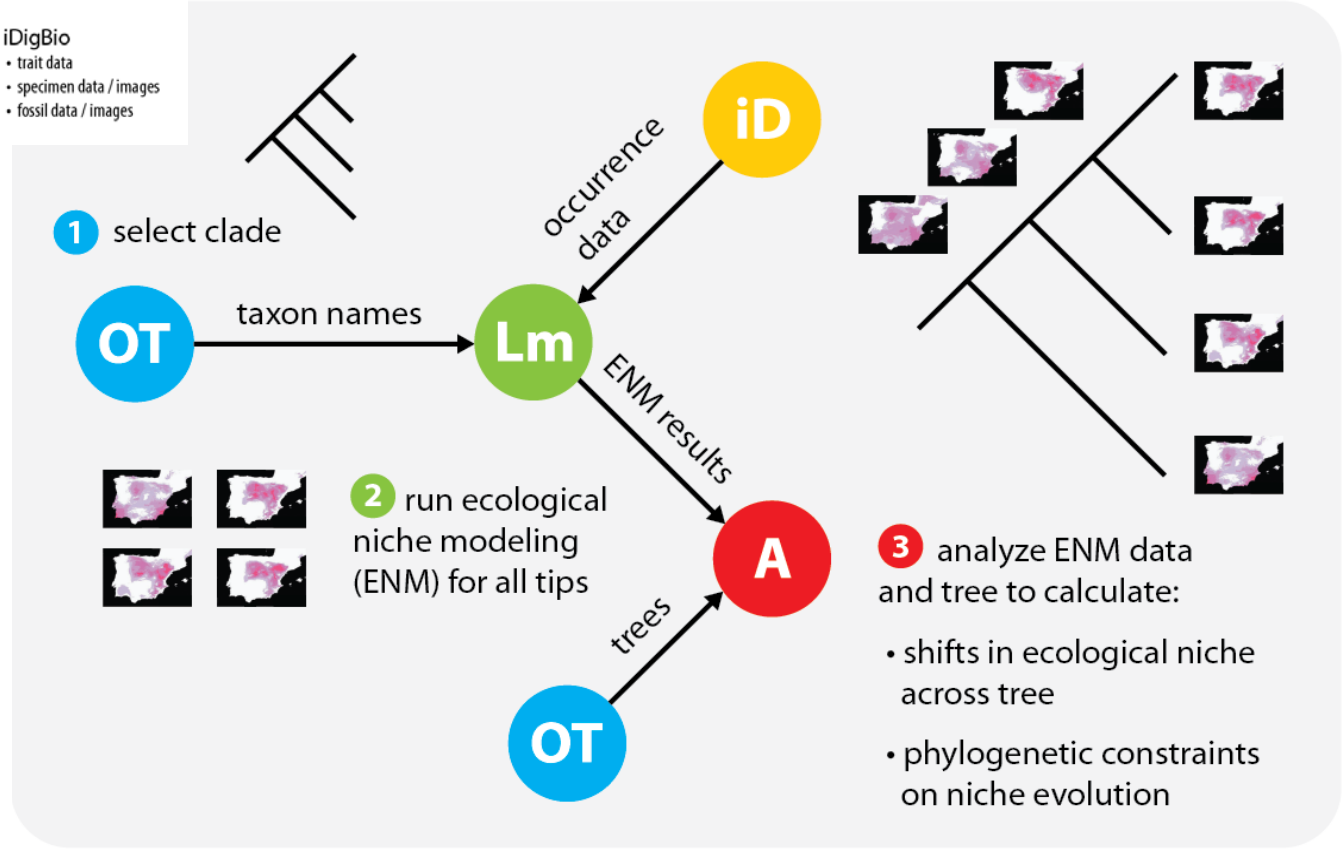
RESOURCES:

Lm Lifemapper
- ecological niche modeling
- biodiversity and range analysis
- visualization

A Arbor
- evolutionary models
- comparative methods
- visualization

OT Open Tree of Life
- phylogenies
- taxonomy / names
- visualization

iD iDigBio
- trait data
- specimen data / images
- fossil data / images



Connecting Trees, Specimens, Tools

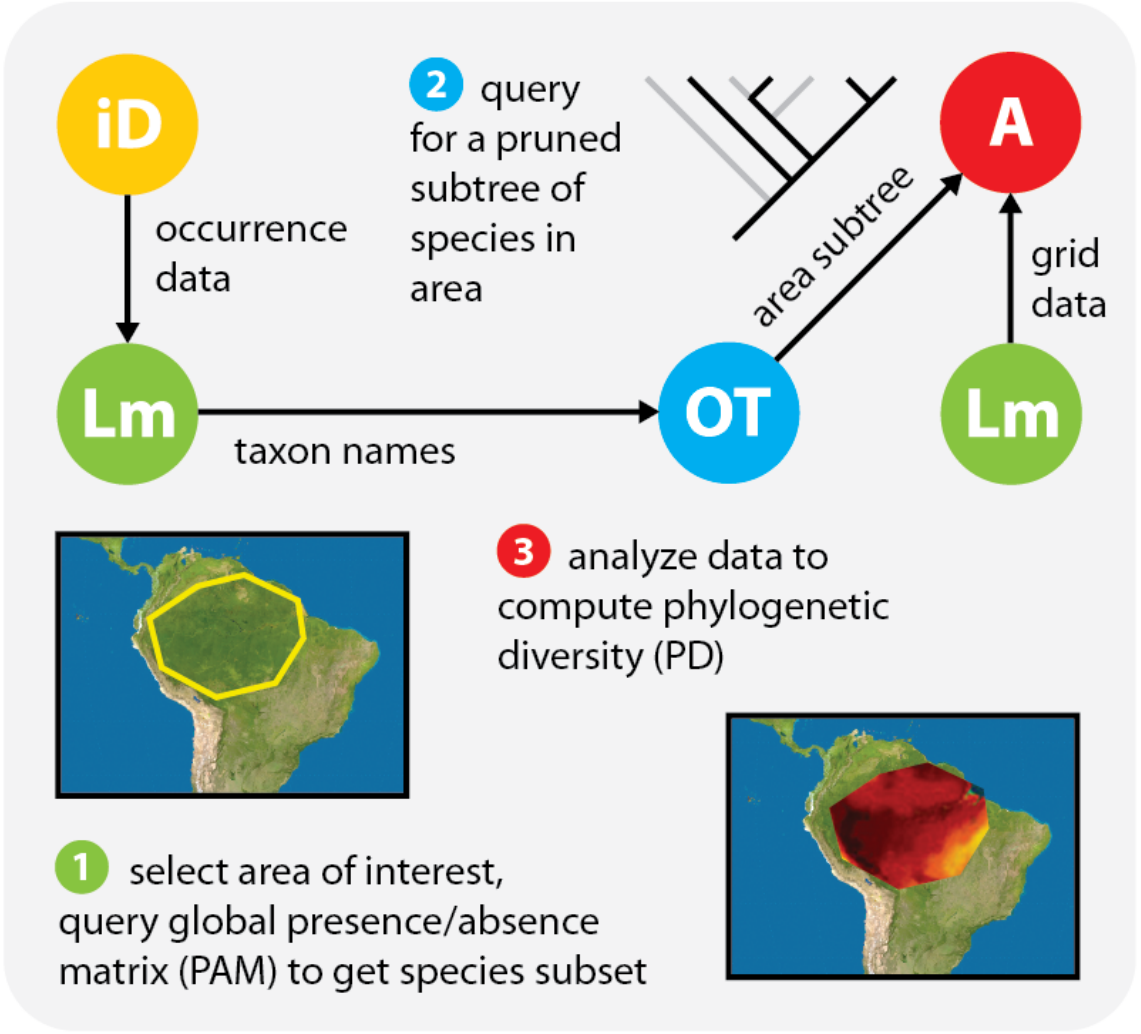
RESOURCES:

Lm Lifemapper
- ecological niche modeling
- biodiversity and range analysis
- visualization

A Arbor
- evolutionary models
- comparative methods
- visualization

OT Open Tree of Life
- phylogenies
- taxonomy / names
- visualization

iD iDigBio
- trait data
- specimen data / images
- fossil data / images



Connecting Trees, Specimens, Tools

ABI Innovation: BiotaPhy Project
Connecting resources
to enable large-scale biodiversity analyses

D. Soltis, P. Soltis, J. Fortes, A. Matsunaga,
J. Beach, J. Soberon, S. Smith

RESOURCES:

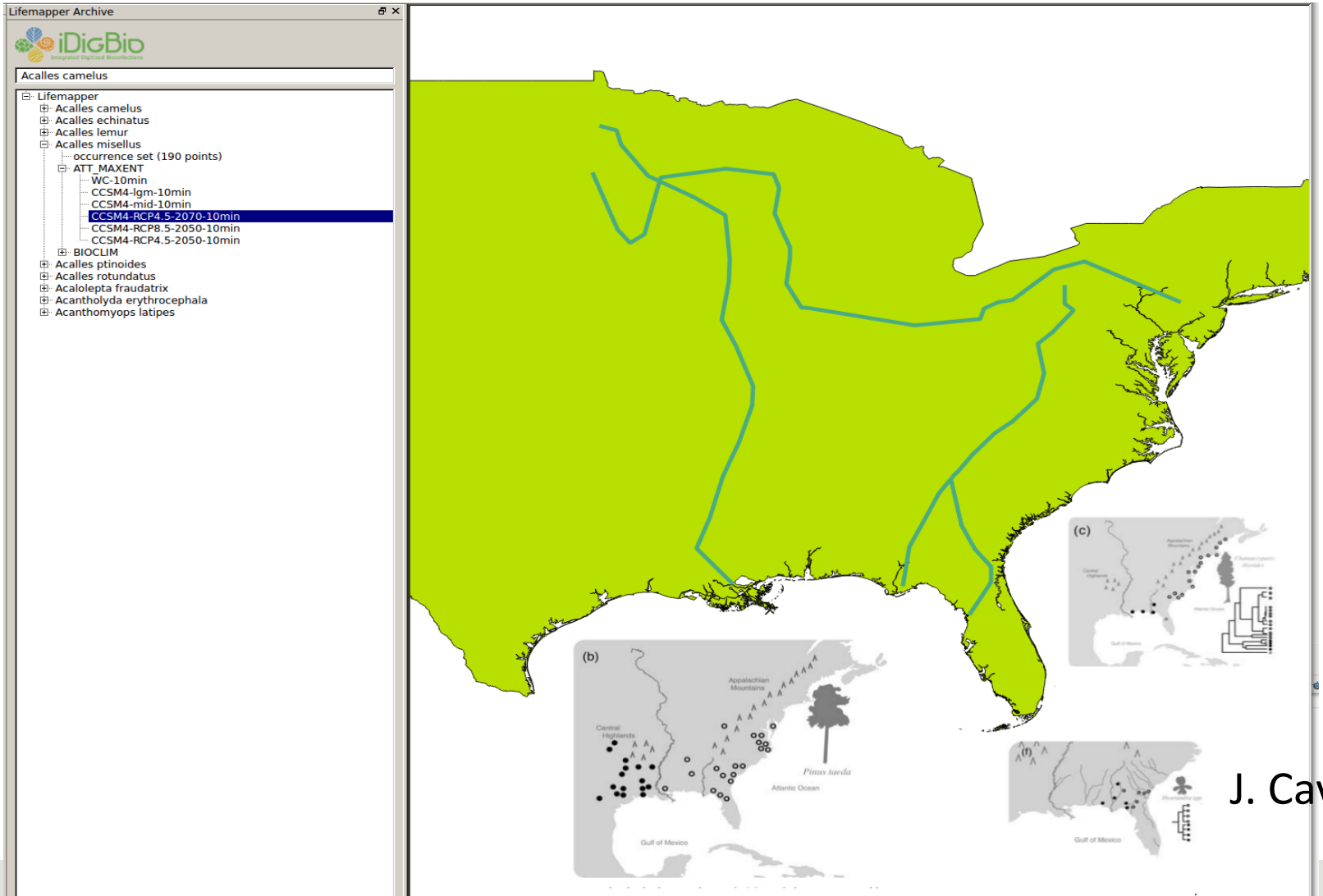
Lm Lifemapper
- ecological niche modeling
- biodiversity and range analysis
- visualization

A Arbor
- evolutionary models
- comparative methods
- visualization

OT Open Tree of Life
- phylogenies
- taxonomy / names
- visualization

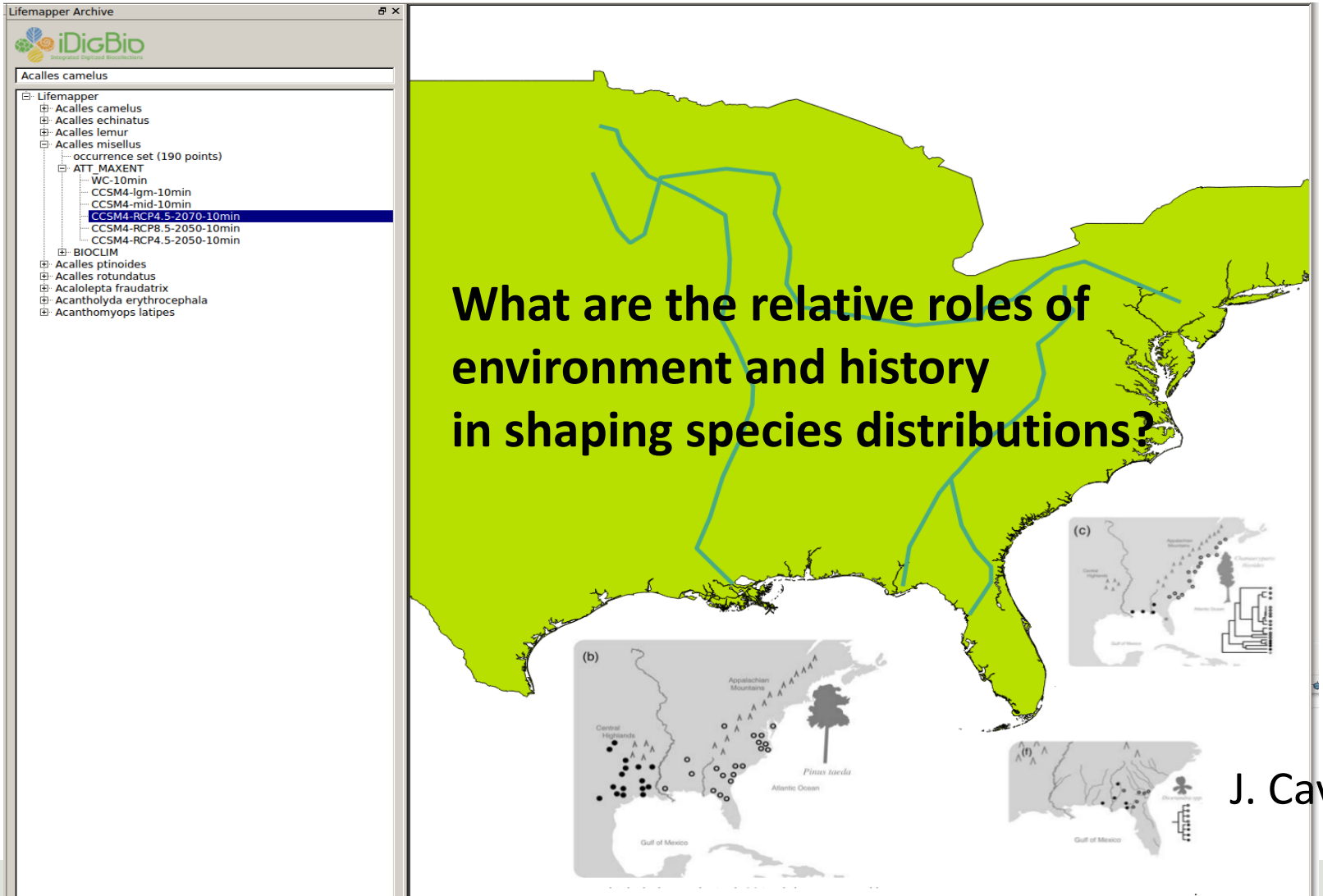
iD iDigBio
- trait data
- specimen data / images
- fossil data / images

Mapping physiographic discontinuities to breaks in the phylogeny



J. Cavner

Mapping physiographic discontinuities to breaks in the phylogeny



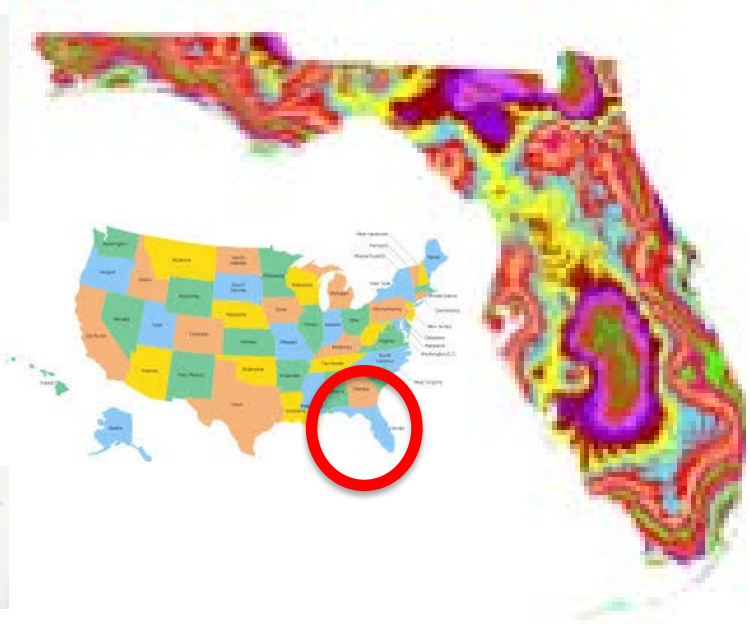
J. Cavner

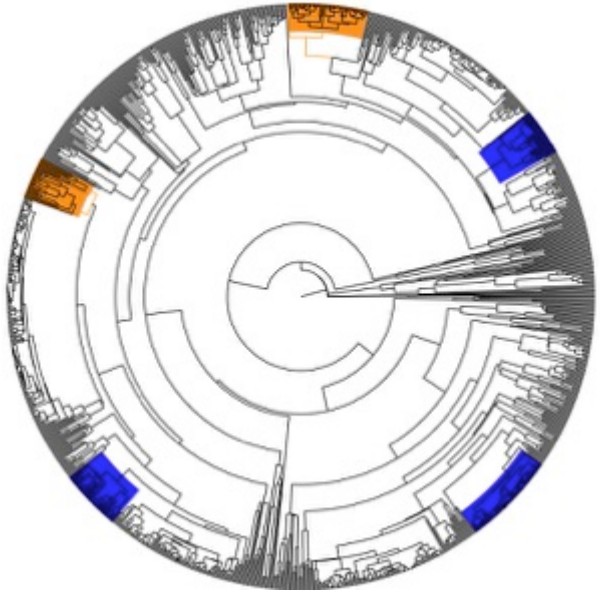
Linking Heterogeneous Resources for Biodiversity Research

- Infrastructure
  - PhyloJIVE/PhyloLink
  - Connecting iDigBio, Open Tree of Life, Lifemapper
- Examples
  - Spatial patterns of phylogenetic diversity
    - Florida Plant Phylogeny Project
  - Community phylogenetics: case study
  - Plant traits and phylogeny

Florida Plant Diversity in a Changing Climate

Integrating herbarium specimen data, ENM, climate change models, and phylogeny



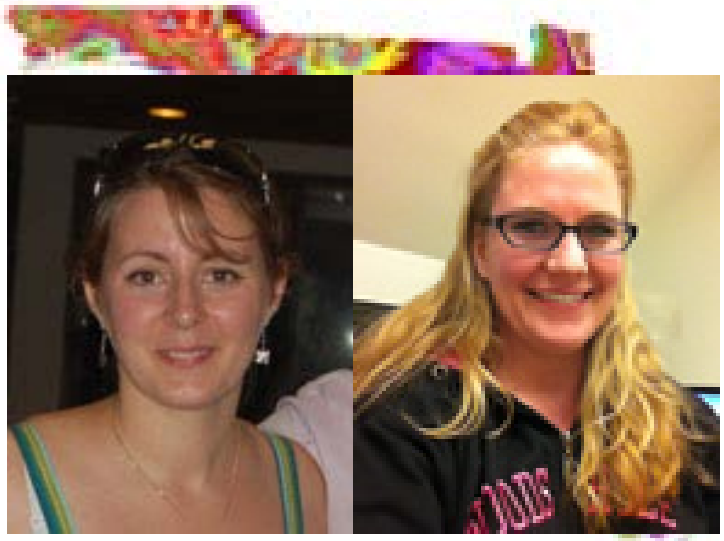


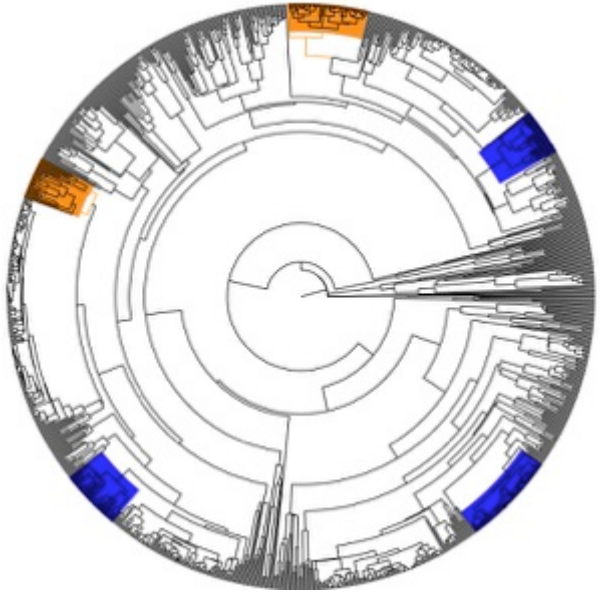
Charlotte Germain-Aubrey, Julie Allen

K. Neubig, L. Majure, R. Abbott, M. Whitten, J. M. Ponciano, B. Mishler, S. Laffan, T. Lamy, R. Guralnick, D. Soltis

Florida Plant Diversity in a Changing Climate

Integrating herbarium specimen data, ENM, climate change models, and phylogeny





Charlotte Germain-Aubrey, Julie Allen

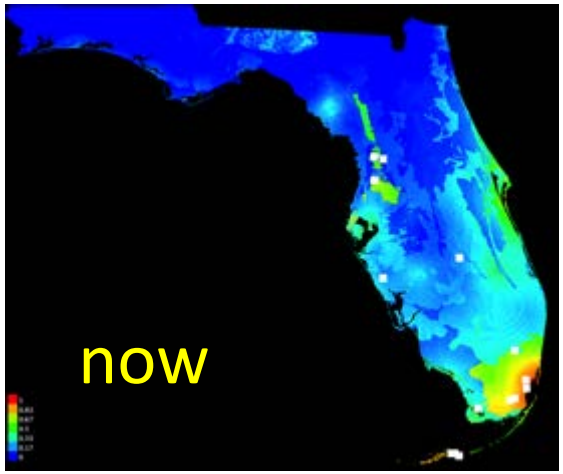
K. Neubig, L. Majure, R. Abbott, M. Whitten, J. M. Ponciano, B. Mishler, S. Laffan, T. Lamy, **R. Guralnick, D. Soltis**

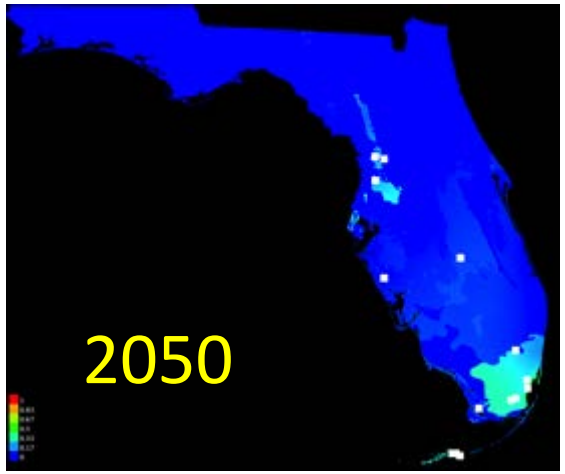
Modeling the Distribution of Species

- Location information and environmental data
- Software to model the range of each species
- For Florida plants:
  - ~1600 plant species (of 4200 species)
  - >511,000 georeferenced points (GPS)
  - Environmental features: temperature, precipitation, soil, etc.
- Project onto future climate conditions

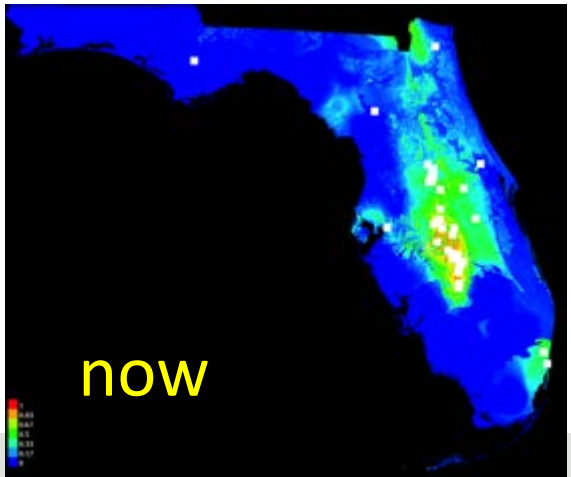
Responses to Climate Change: past, present, future

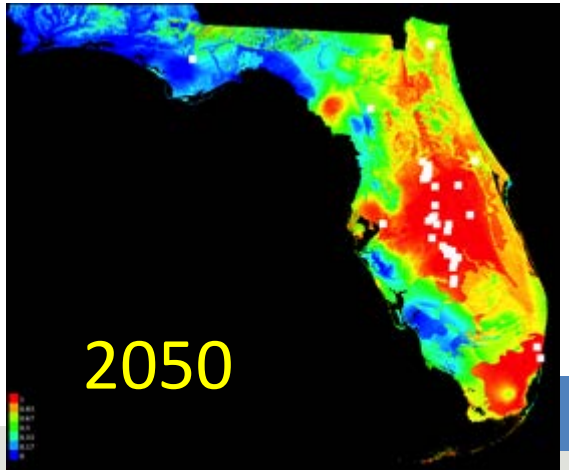
Abildgaardia ovata (flatspike sedge)





Prunus geniculata (scrub plum)

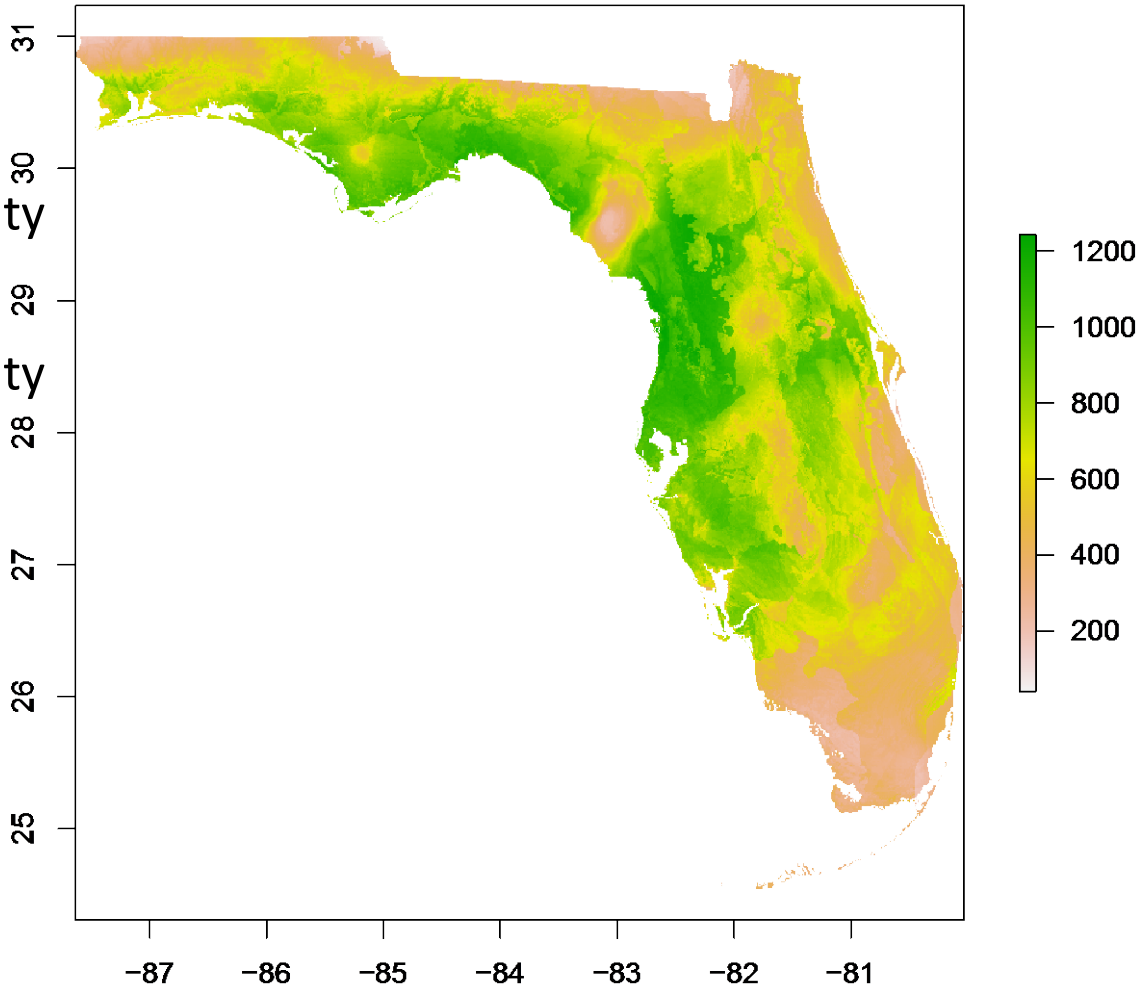




Florida Plant Diversity: Now

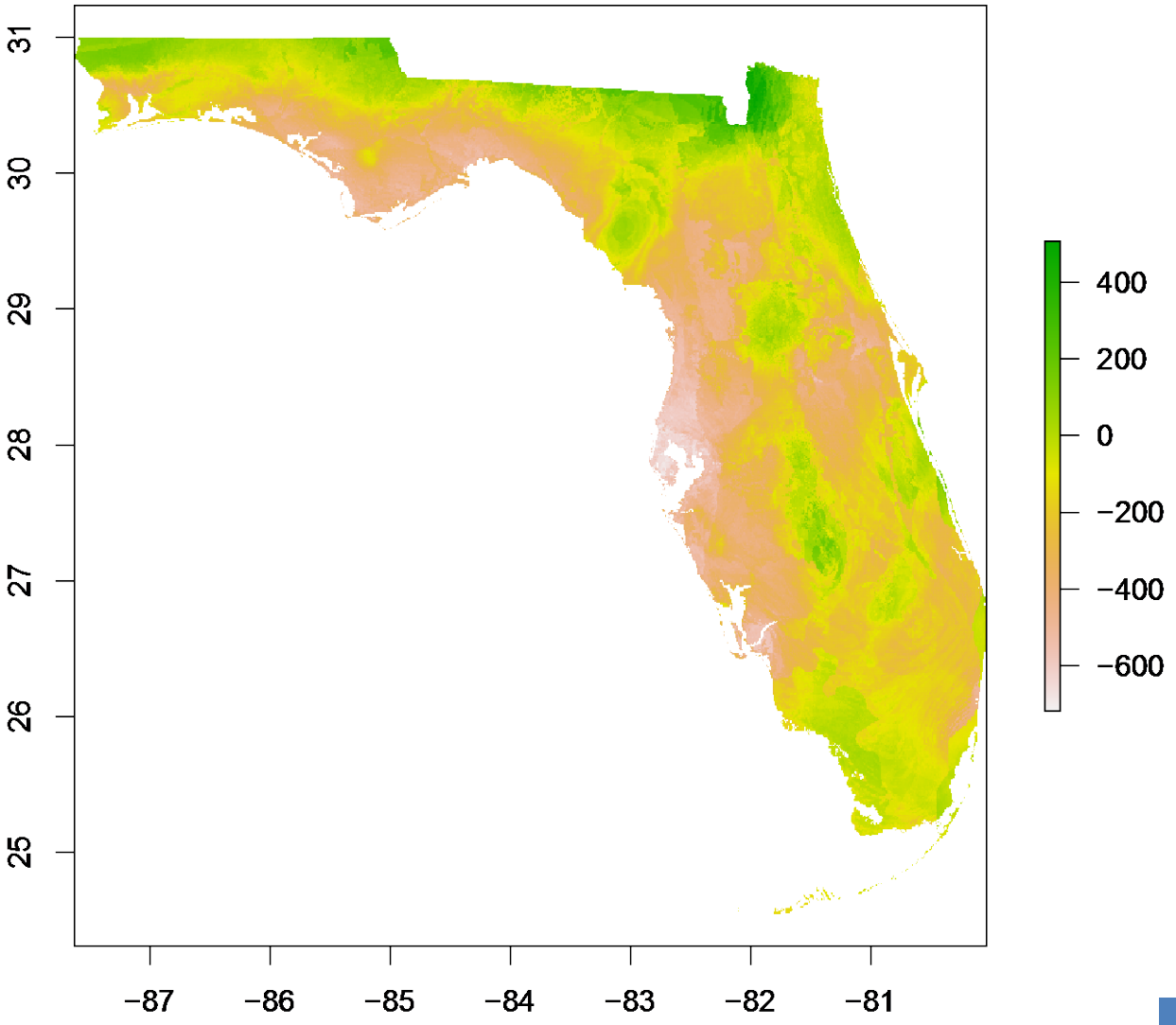
High species diversity

Low species diversity



Between Now and 2050…

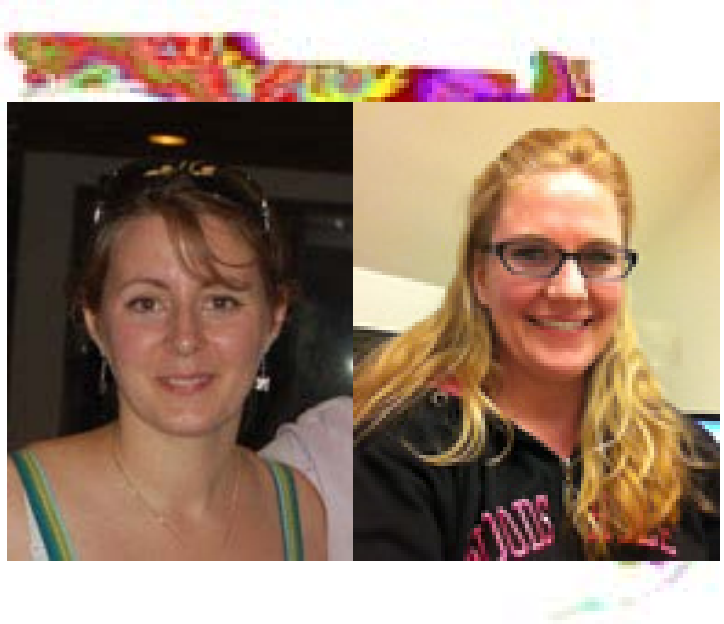
- Panhandle species moving NORTH!
- Peninsula species moving SOUTH!

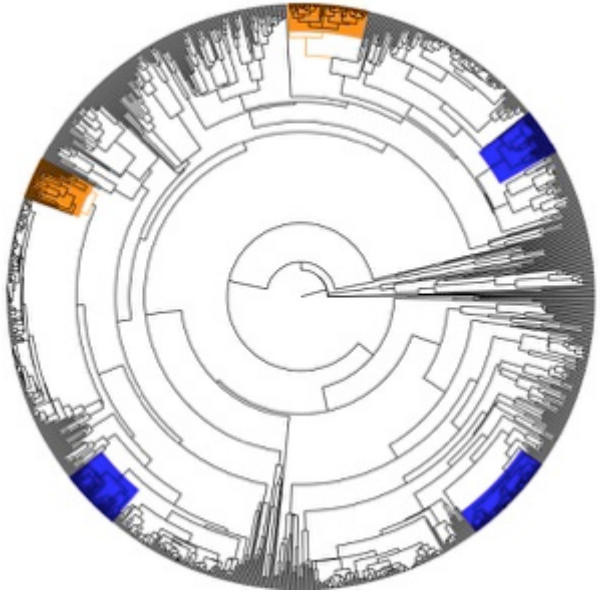


spp 2050 - #spp now

Florida Plant Diversity in a Changing Climate

Integrating herbarium specimen data, ENM, climate change models, and phylogeny



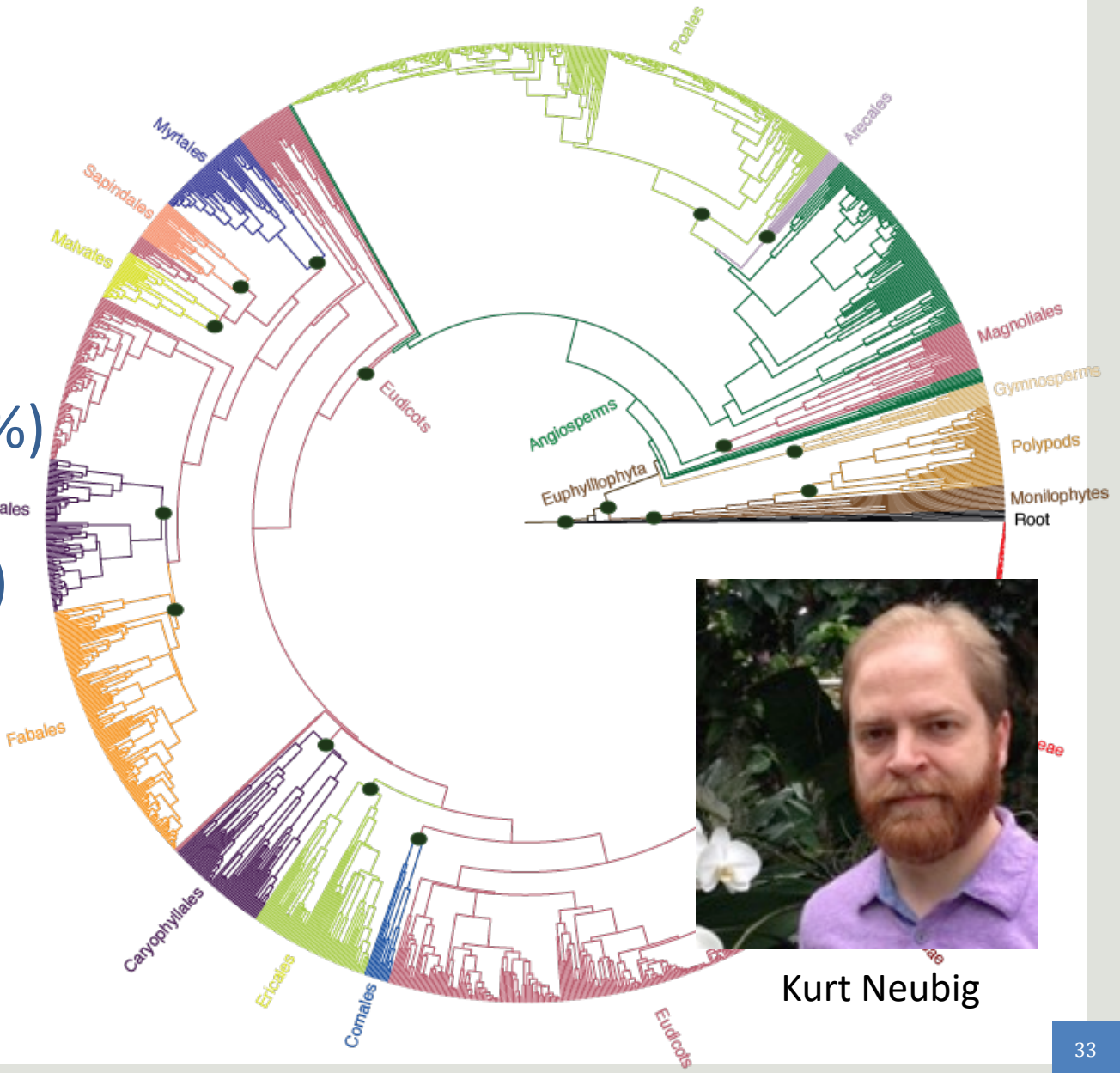


Charlotte Germain-Aubrey, Julie Allen

K. Neubig, L. Majure, R. Abbott, M. Whitten, J. M. Ponciano, B. Mishler, S. Laffan, T. Lamy, R. Guralnick, D. Soltis

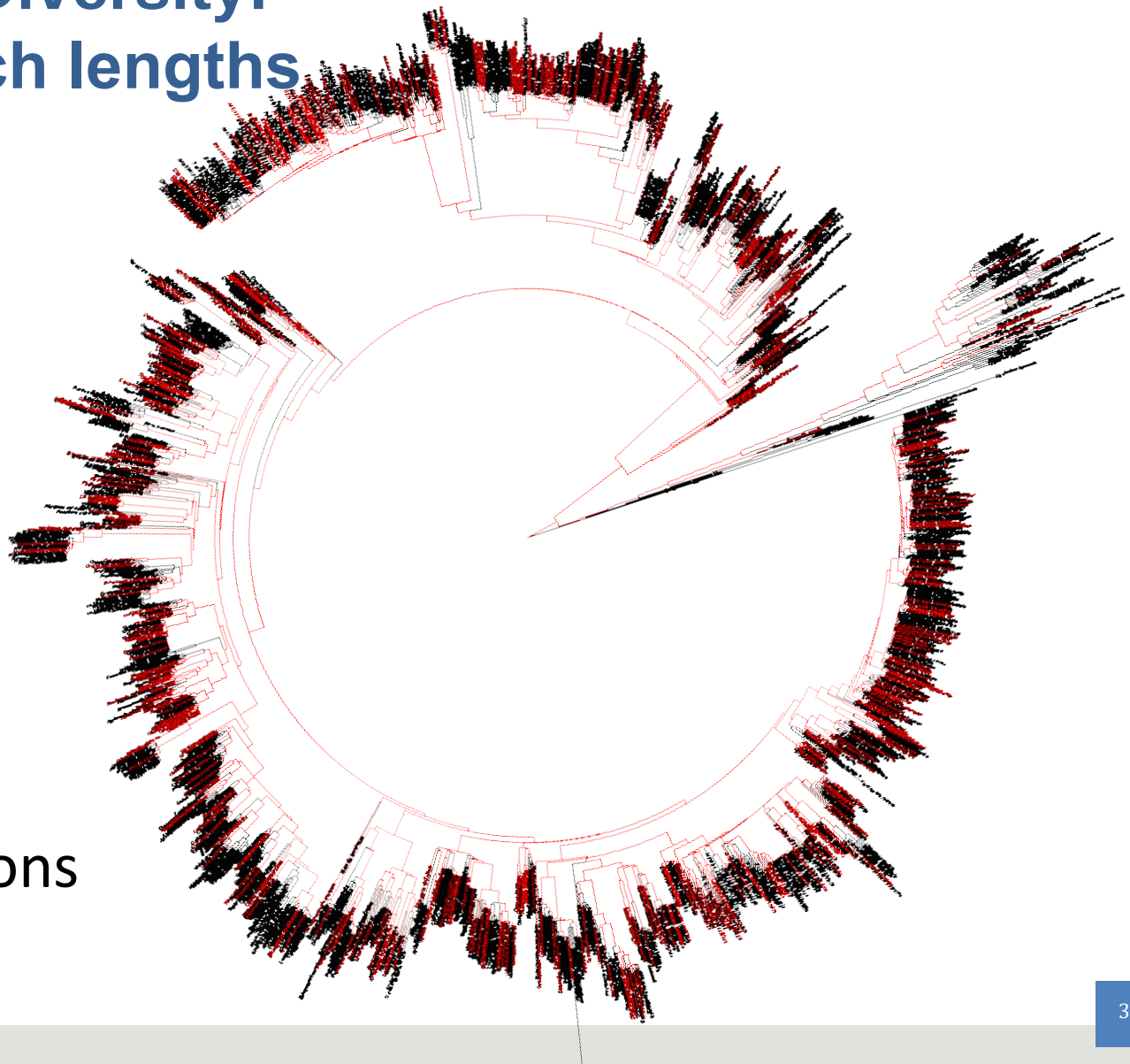
Florida Plant Phylogeny

1,548 species (38%)
685 genera (44%)
185 families (78%)
rbcL, *matK*
GenBank & new
RAxML
Dated with r8s



Kurt Neubig

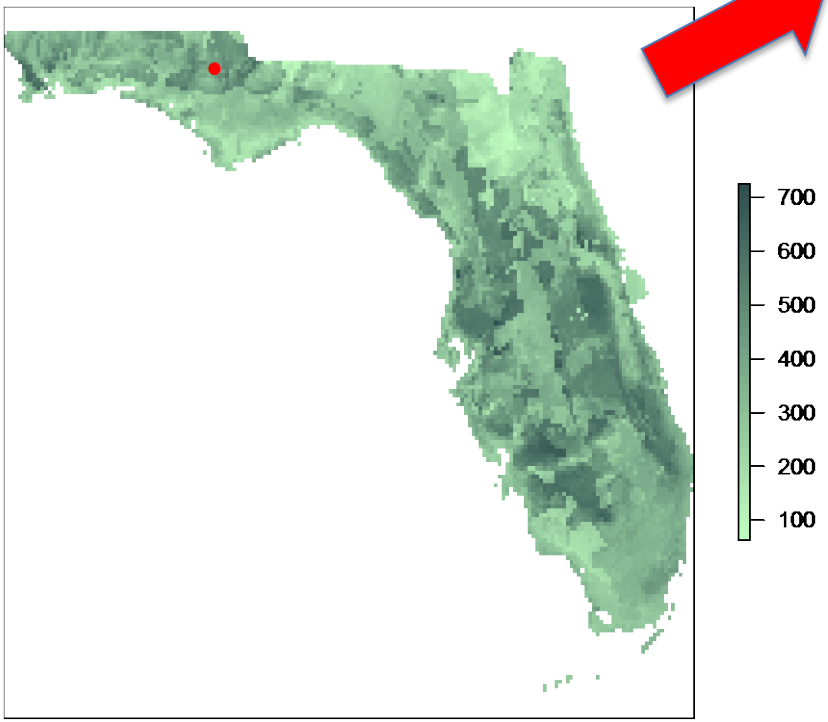
Phylogenetic Diversity:
≈ sum of branch lengths

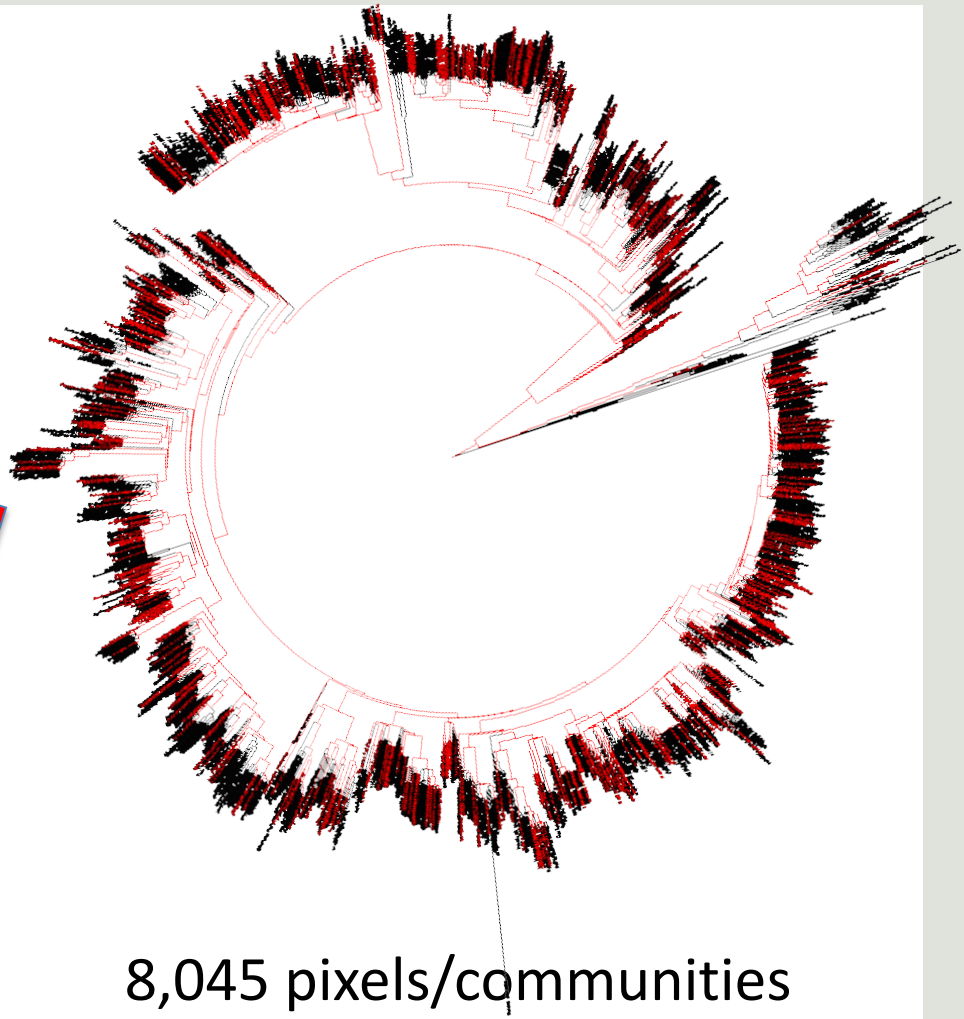


- Total diversity
- Compare regions

Phylogenetic Diversity: ≈ sum of branch lengths

Species list at each pixel
Generated from ENMs

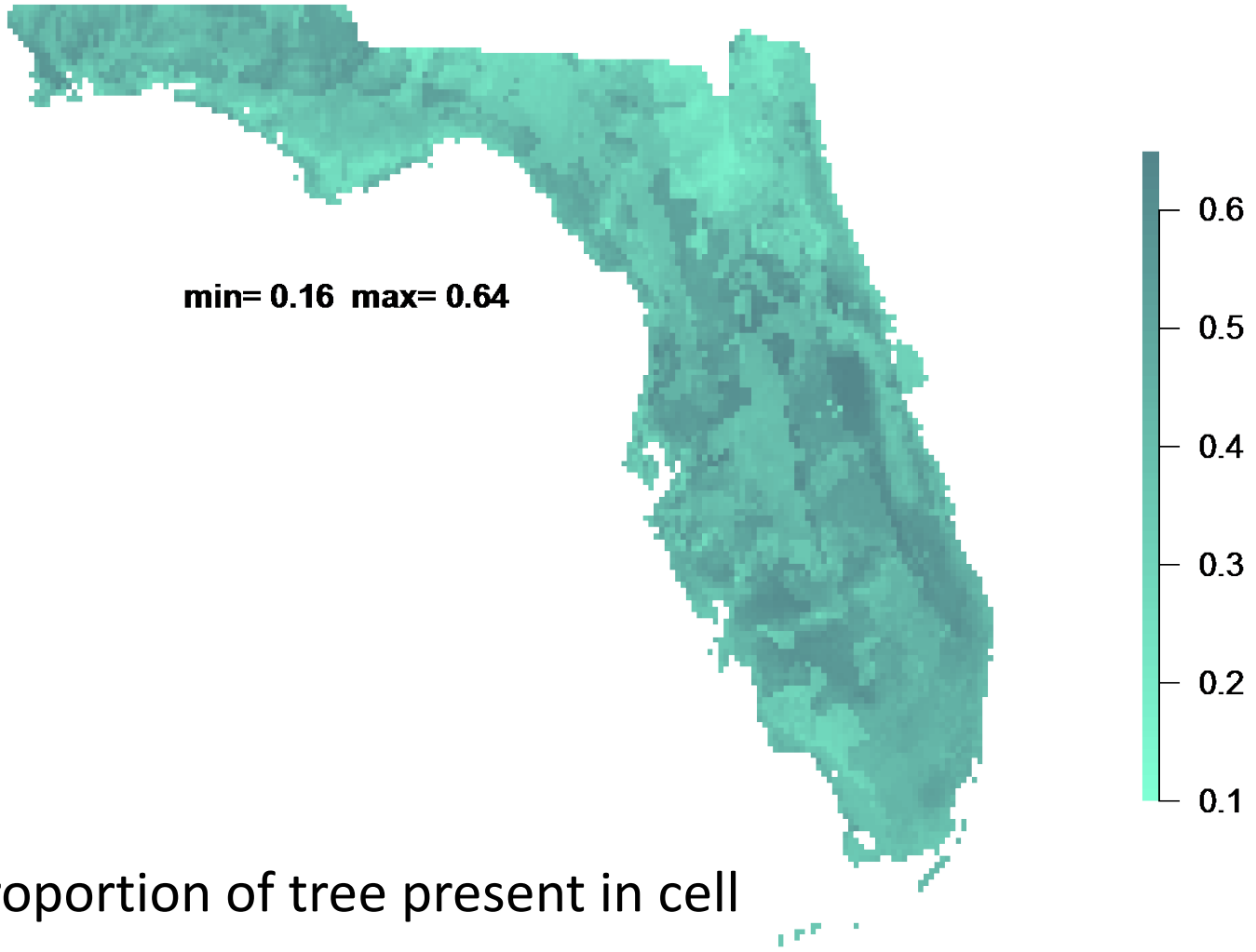




8,045 pixels/communities

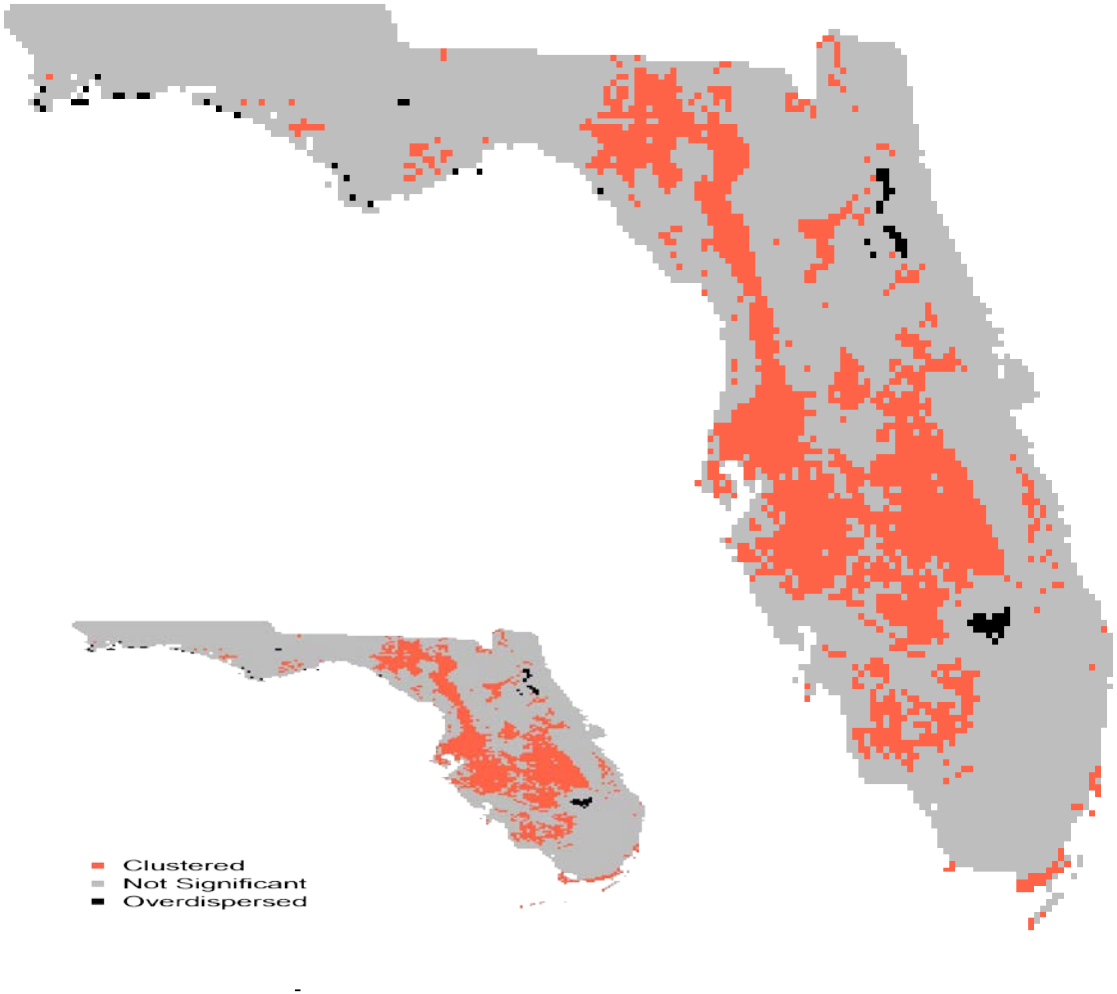
16 km^2 per pixel

Phylogenetic Diversity - angiosperms



PD = proportion of tree present in cell

Phylogenetic Diversity - angiosperms: clustering vs. overdispersion



Linking Heterogeneous Resources for Biodiversity Research

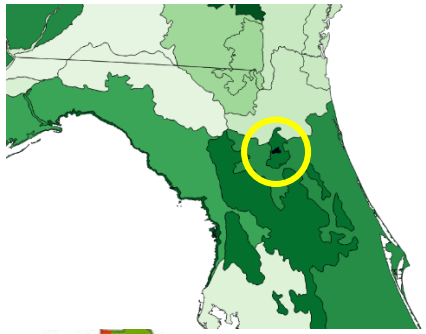
- Infrastructure
  - PhyloJIVE/PhyloLink
  - Connecting iDigBio, Open Tree of Life, Lifemapper
- Examples
  - Spatial patterns of phylogenetic diversity
    - Florida Plant Phylogeny Project
  - Community phylogenetics: case study
  - Plant traits and phylogeny

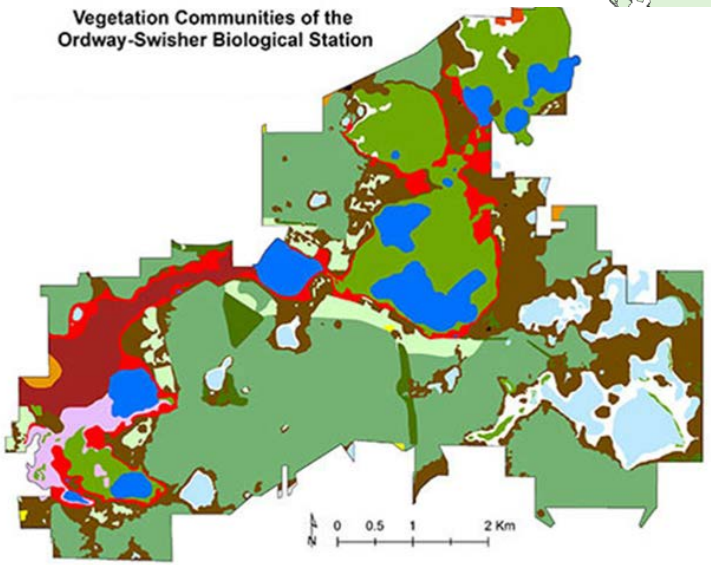
The Effects of Taxonomic and Spatial Scale on Measures of Phylogenetic Diversity Using a Test Case in Florida

Johanna Jantzen

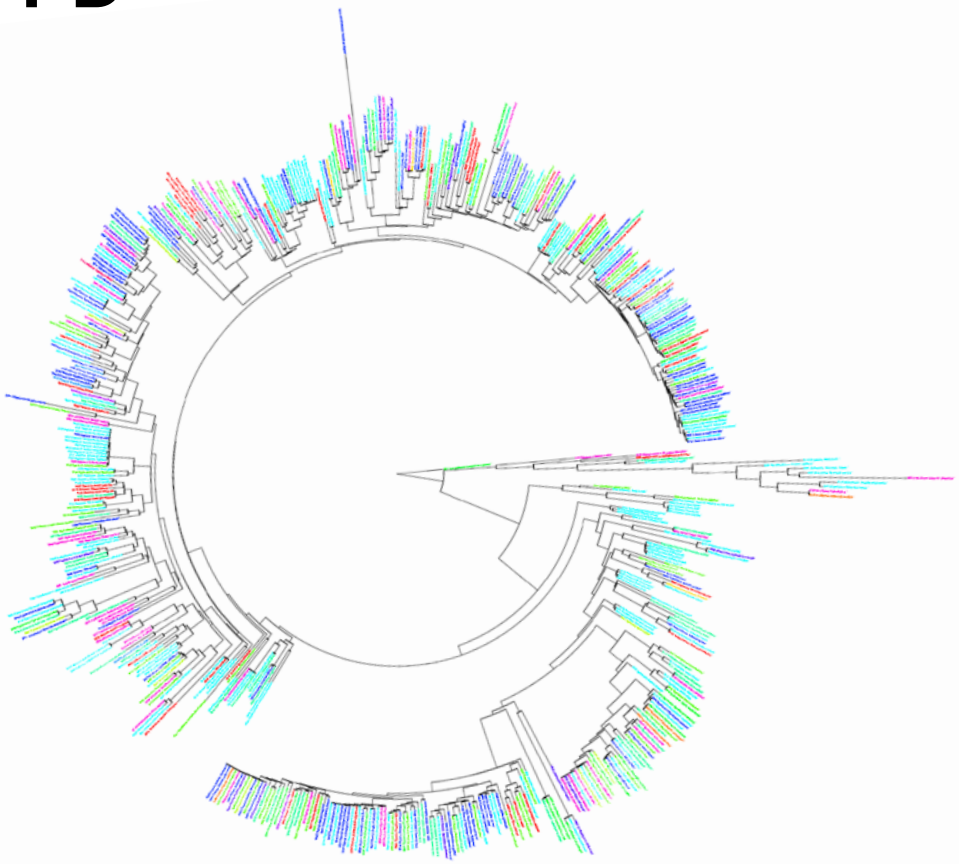
Community phylogenetics:

Ordway-Swisher Biological Station (OSBS)





Phylogeny Reconstruction and Calculation of PD



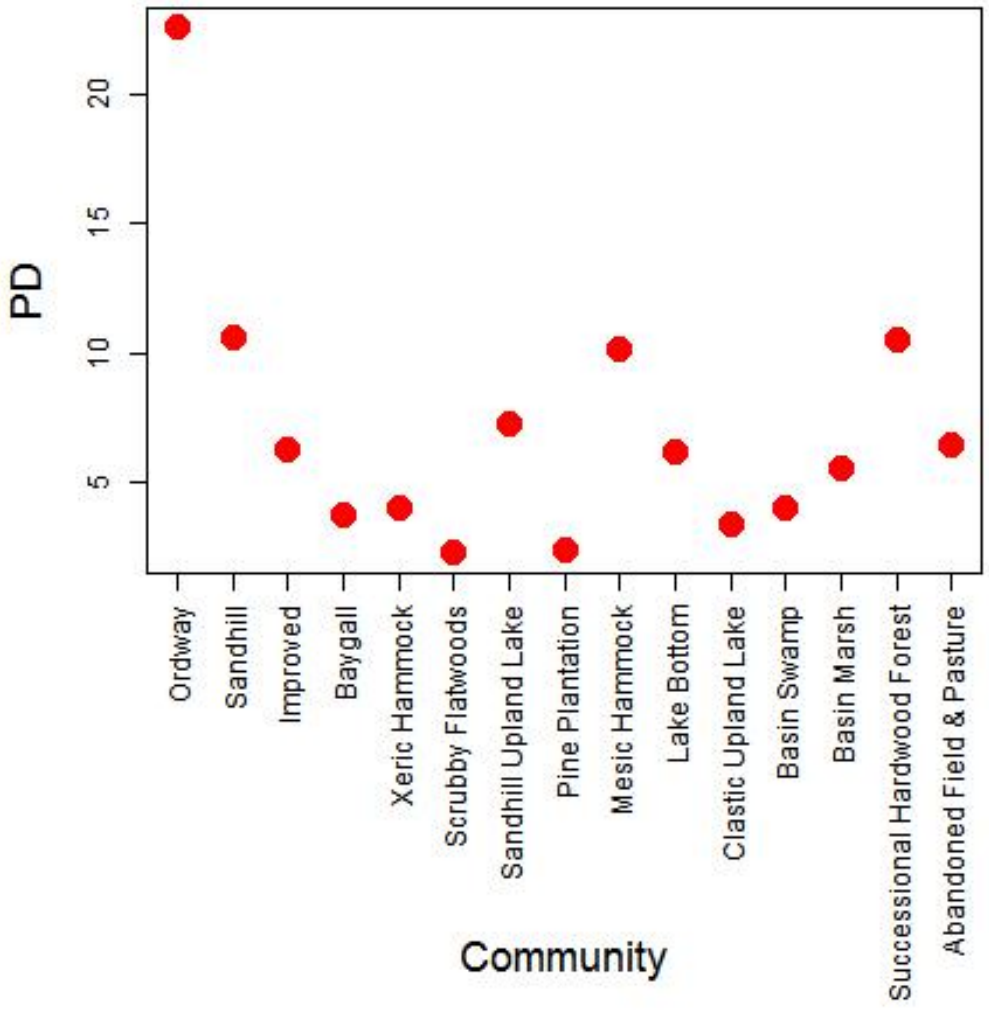
572 taxa

matK and *rbcL*

ML phylogeny reconstruction (RAxML)

PD calculations for 14 communities (Biodiverse)

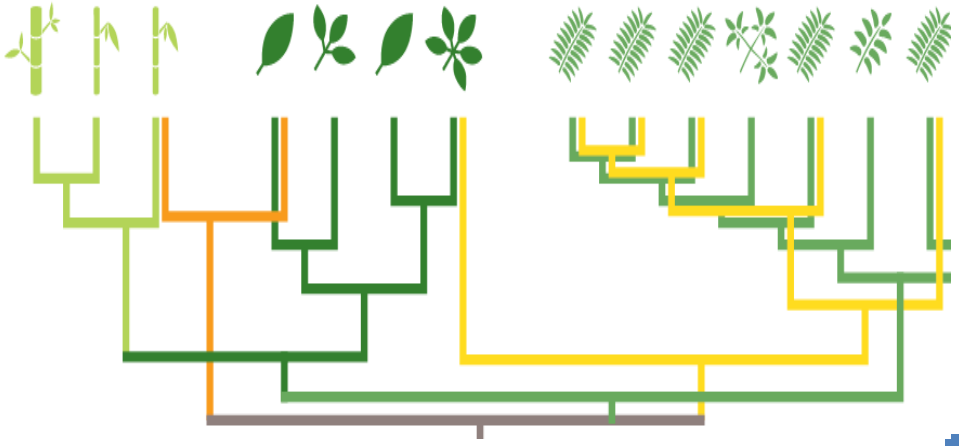
Phylogenetic diversity for 14 communities at OSBS



Linking Heterogeneous Resources for Biodiversity Research

- Infrastructure
  - PhyloJIVE/PhyloLink
  - Connecting iDigBio, Open Tree of Life, Lifemapper
- Examples
  - Spatial patterns of phylogenetic diversity
    - Florida Plant Phylogeny Project
  - Community phylogenetics: case study
  - Plant traits and phylogeny

Images, Functional Traits, Phylogeny

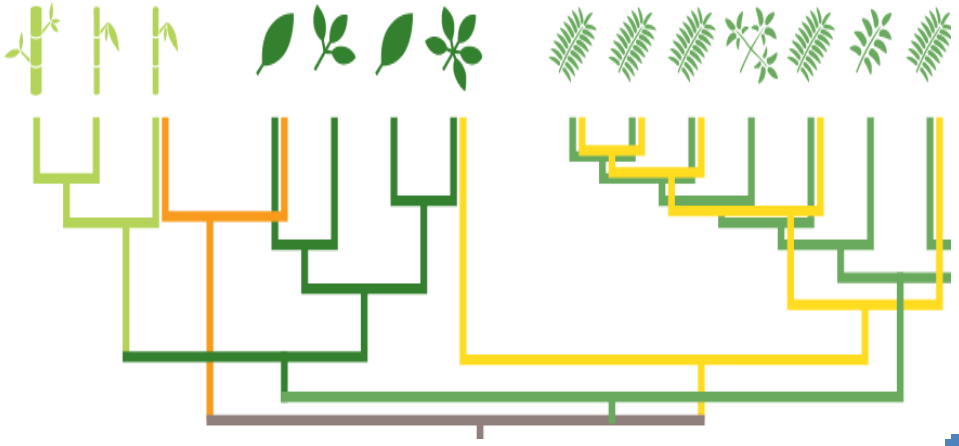


Connect to ecology/phylogeny

Evolution of plant functional traits

Images, Functional Traits, Phylogeny

- Correlation of plant functional traits
- Over- and underdispersion of traits:
  - Are traits phylogenetically constrained or broader properties of communities?
  - Implications for long-term health of communities



Summary

- Digitization enhances value of herbarium specimens – unexpected consequences
- Data aggregators (iDigBio, others) serve data and will increasingly provide access to analytical tools
- Herbarium specimens – locality records, images – provide rich data for range of biodiversity studies
- Linking data with phylogenies, other information holds particular promise
- Biodiversity data are heterogeneous, linkages difficult

iDigBio
Integrated Digitized Biocollections

Thank you!

www.idigbio.org

iDigBio Team & Collaborators

J. Fortes, A. Matsunaga, J. Miller,
C. Germain-Aubrey, B. Marchant

Florida Diversity Project

D. Soltis, R. Guralnick,
C. Germain-Aubrey, J. Allen, K. Neubig,
L. Majure, R. Abbott, M. Whitten,
B. Mishler, S. Laffan

BiotaPhy Project:

D. Soltis, J. Fortes, J. Beach, J. Soberon,
S. Smith, J. Cavner

Dimensions Project:

D. Soltis, J. Lichstein, S. Bohlman, M. Whitten,
J. Jantzen, S. Graves, K. Neubig

UF UNIVERSITY of FLORIDA

FLORIDA STATE UNIVERSITY 1851

FLORIDA MUSEUM OF NATURAL HISTORY

NSF

iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program (Cooperative Agreement EF-1115210). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. All images used with permission or are free from copyright.

Thank you!

www.idigbio.org

psoltis@flmnh.ufl.edu

facebook.com/iDigBio

twitter.com/iDigBio

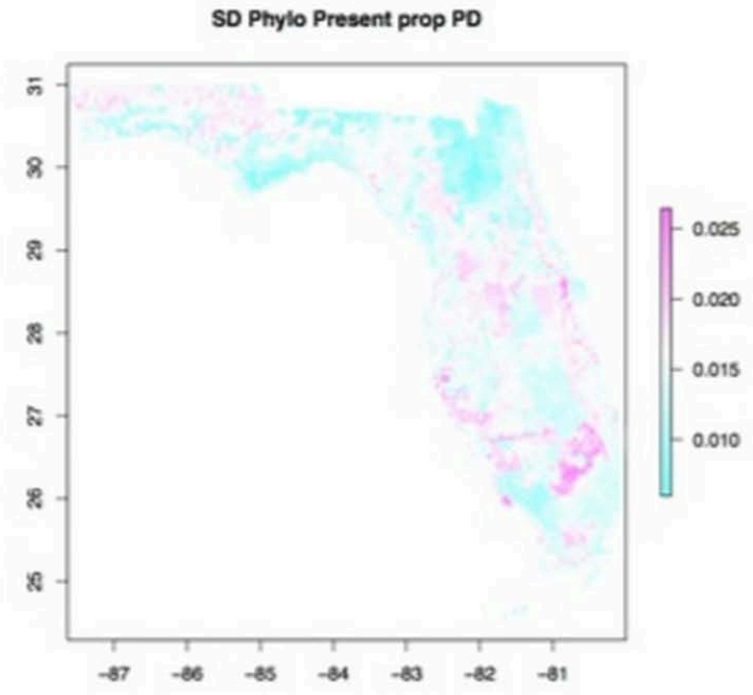
vimeo.com/idigbio

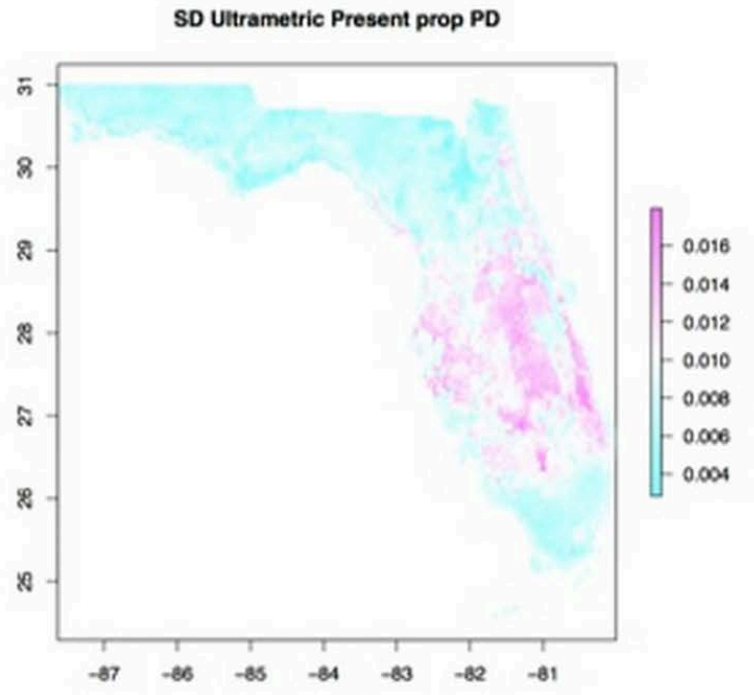
idigbio.org/rss-feed.xml

webcal://www.idigbio.org/events-calendar/export.ics

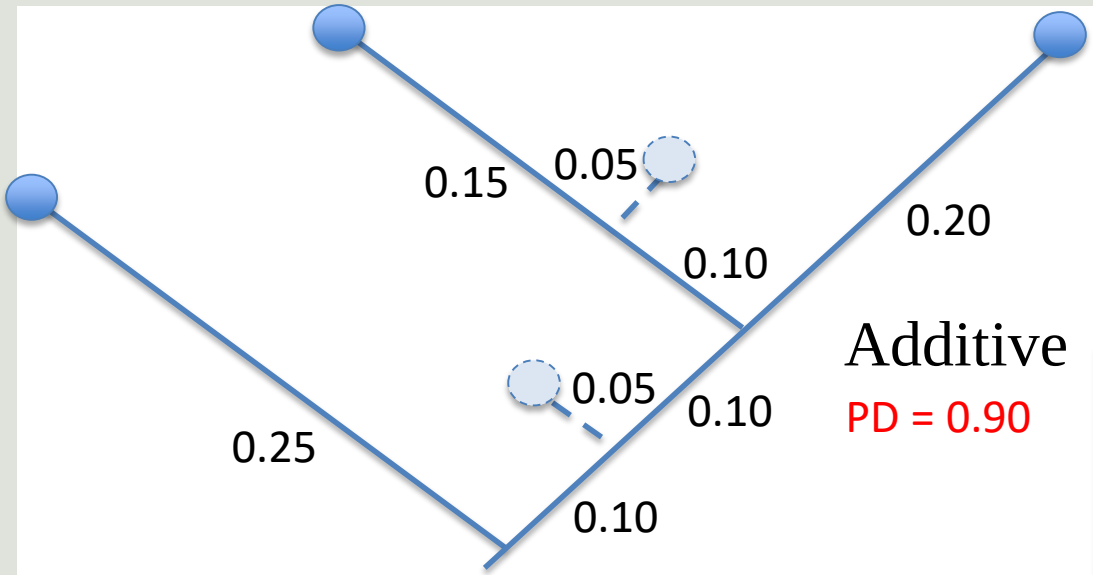
iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program (Cooperative Agreement EF-1115210). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. All images used with permission or are free from copyright.

Uncertainty





- Estimating uncertainty is essential
- Does it affect patterns of significance ?



Additive

PD = 0.90

additive >> ultrametric

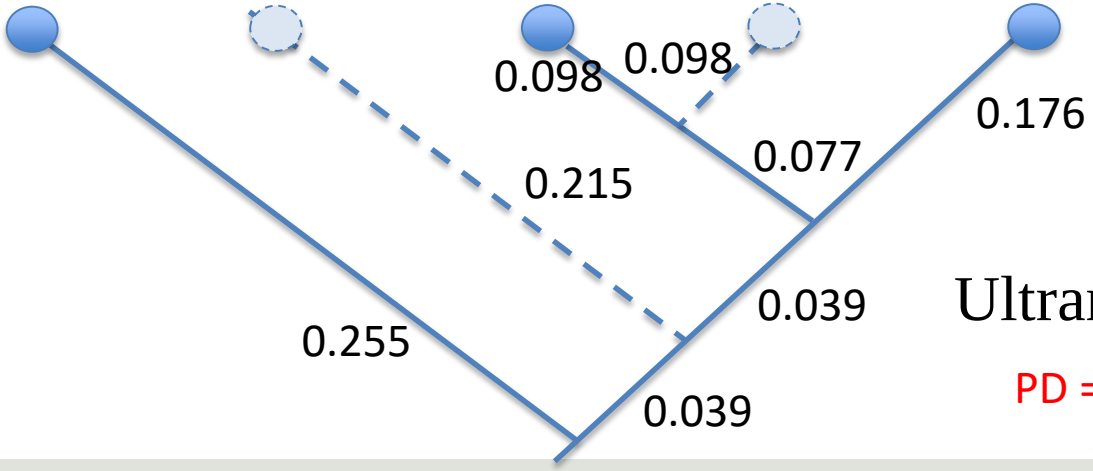
= terminal present in selected grid cell

= terminal missing in selected grid cell

= branch present in selected grid cell

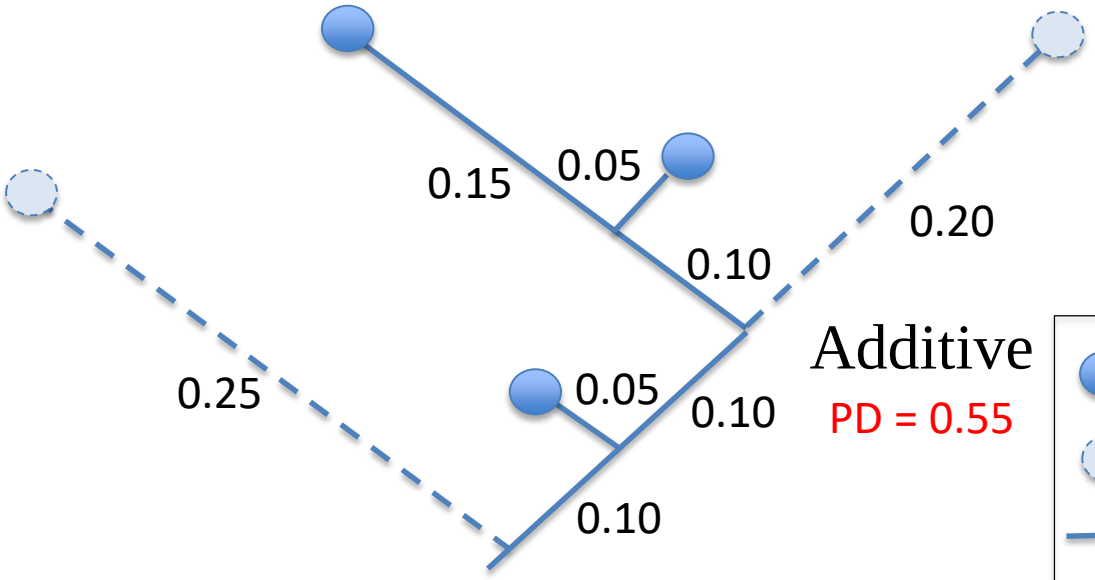
= branch missing in selected grid cell

All branches lengths are PD_P



Ultrametric

PD = 0.687



Additive

PD = 0.55

ultrametric ≈ additive

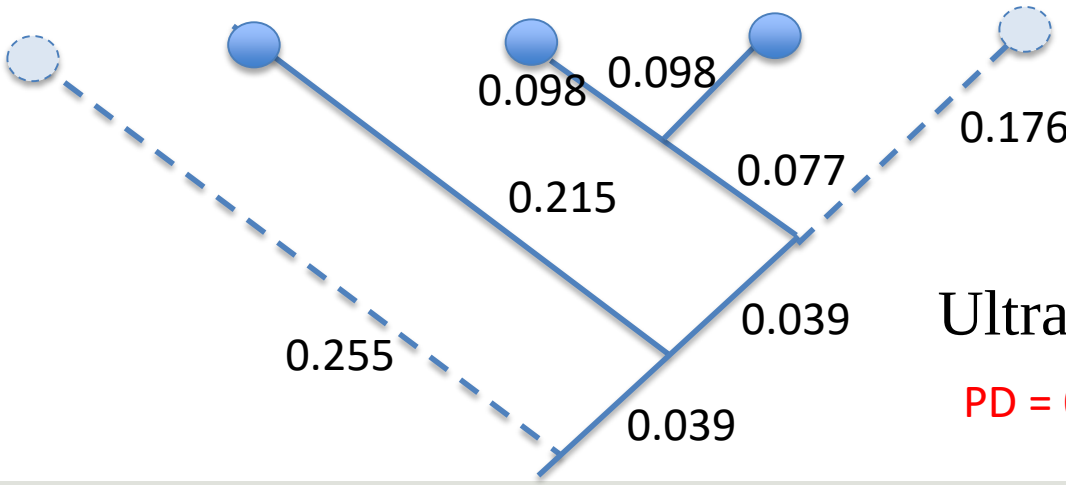
= terminal present in selected grid cell

= terminal missing in selected grid cell

= branch present in selected grid cell

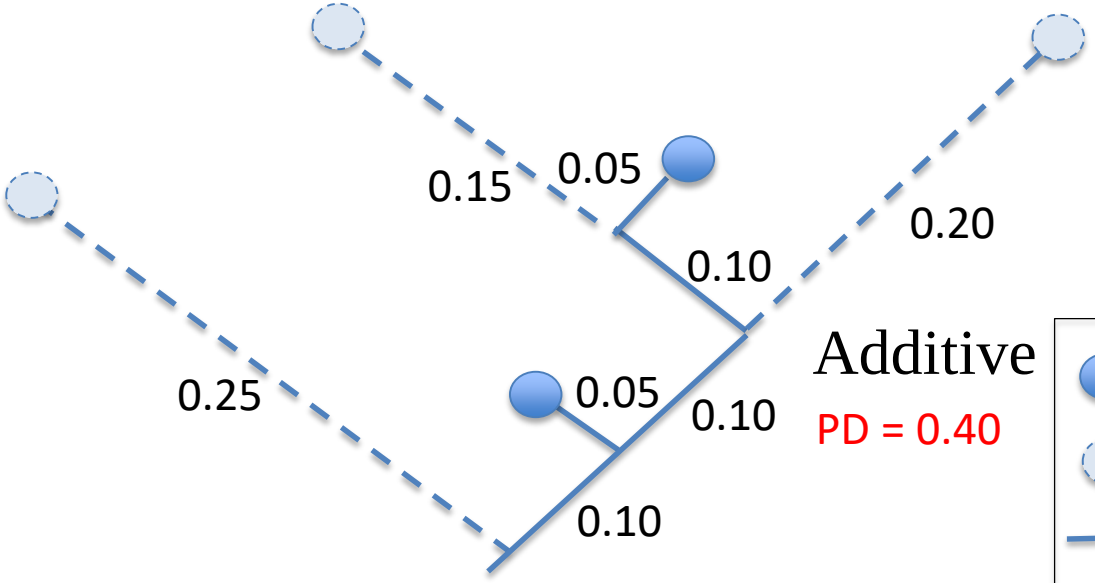
= branch missing in selected grid cell

All branches lengths are PD_P



Ultrametric

PD = 0.569



Additive

PD = 0.40

ultrametric > additive

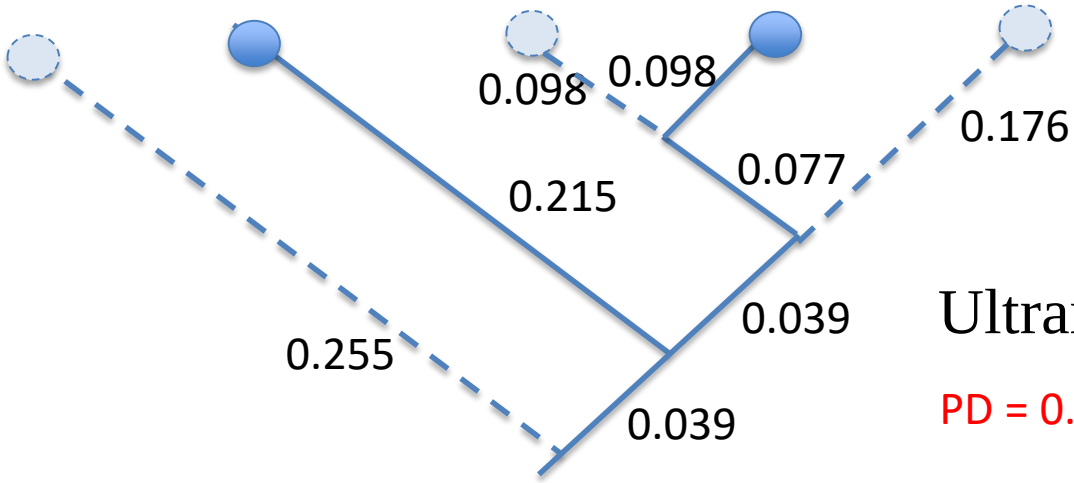
= terminal present in selected grid cell

= terminal missing in selected grid cell

= branch present in selected grid cell

= branch missing in selected grid cell

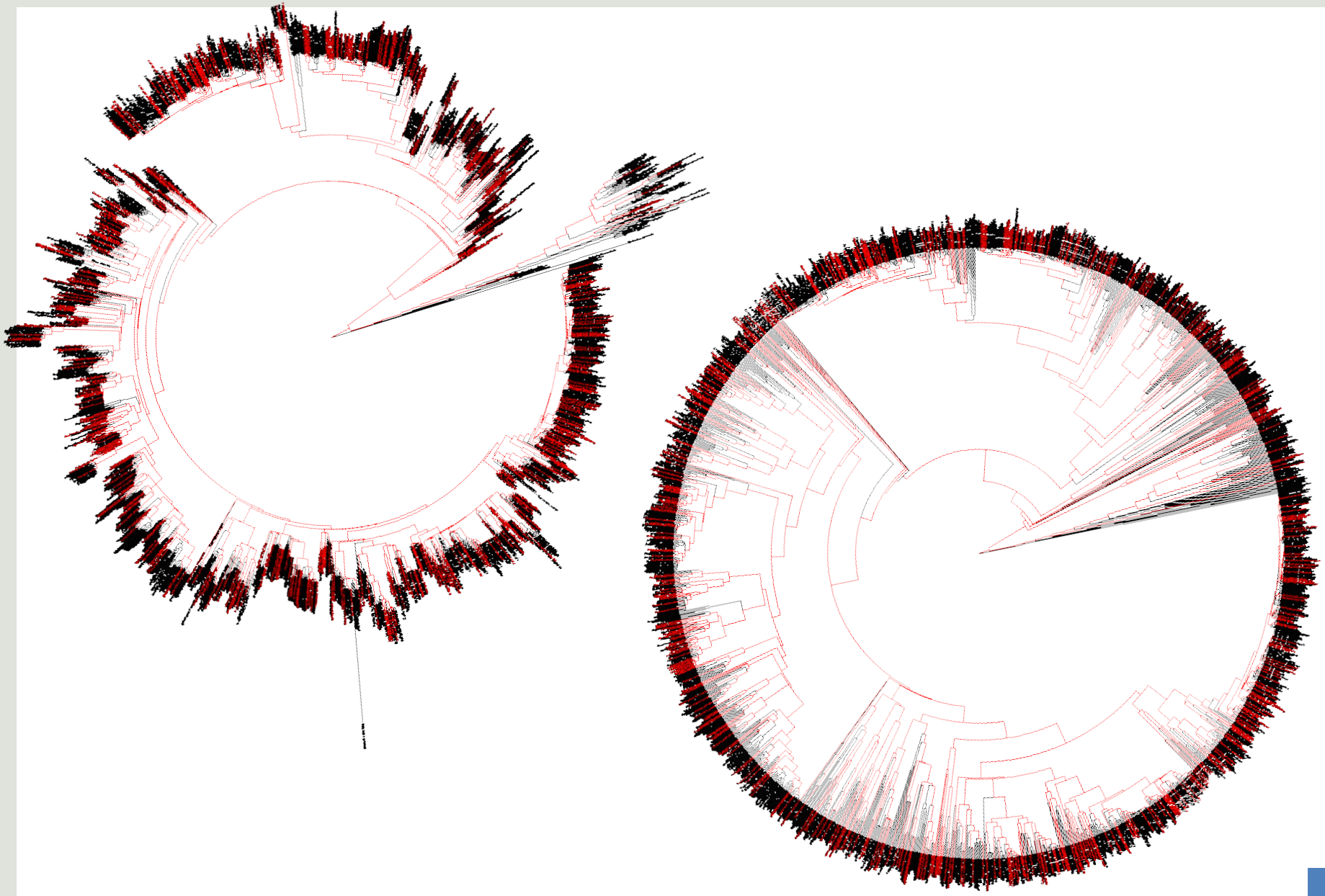
All branches lengths are PD_P



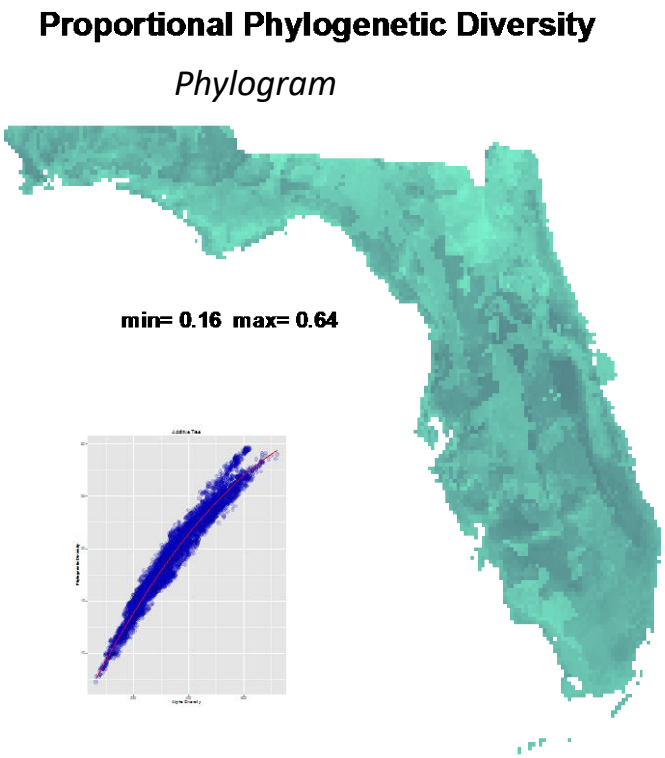
Ultrametric

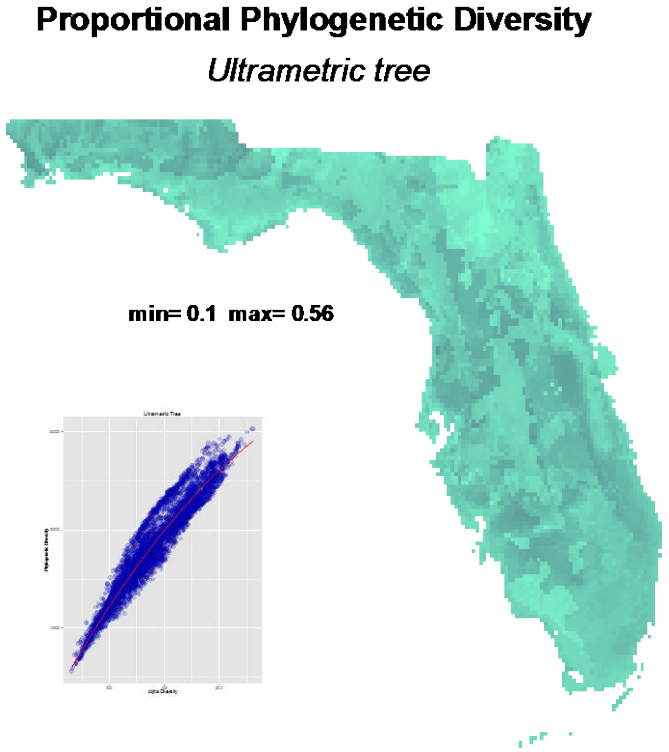
PD = 0.468

Phylogram vs. ultrametic tree | 1.c.



Phylogram vs. ultrametic tree | 1.c.



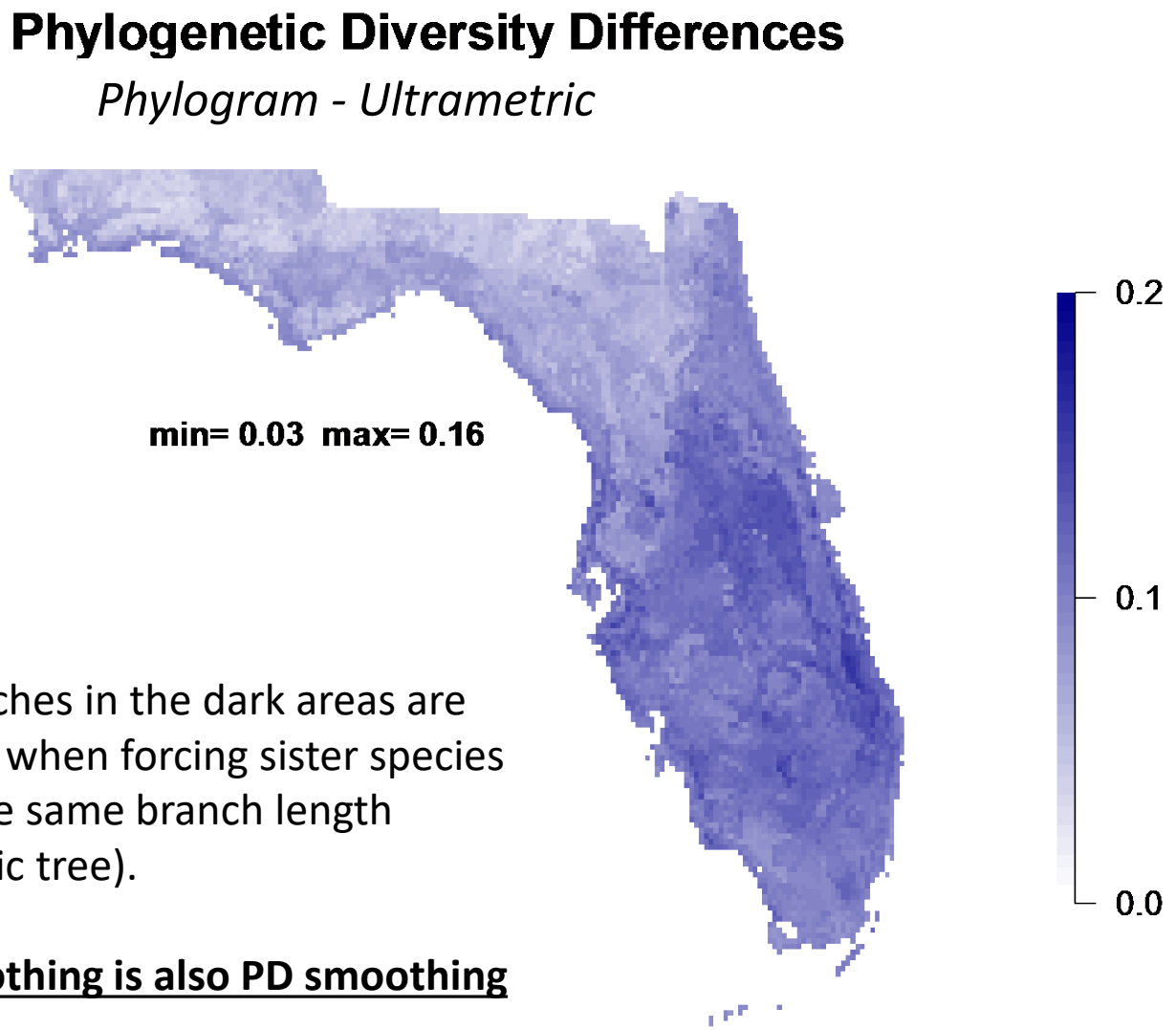


- Genetic diversity
- Accumulation of evolutionary change

- Time diversity
- Amount of evolutionary history

- More easily comparable to other methods/studies

Phylogram vs. ultrametic tree | 1.c.



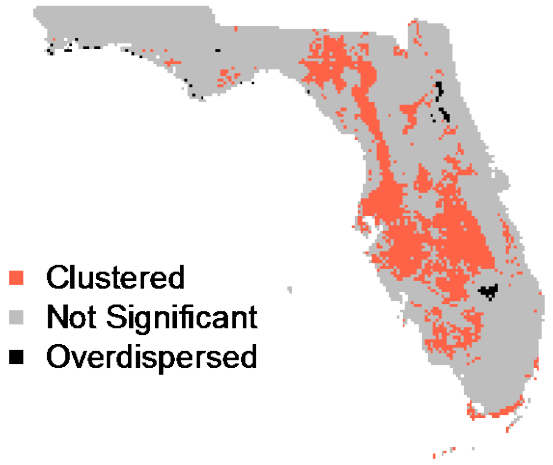
Long branches in the dark areas are shortened when forcing sister species to have the same branch length (ultrametric tree).

Rate smoothing is also PD smoothing

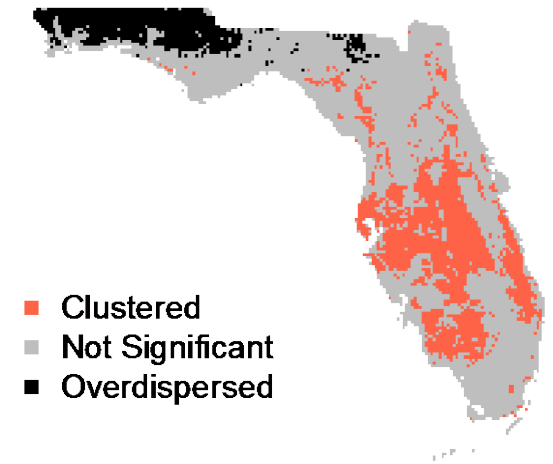
Phylogenetic Diversity | 2.

RAxML

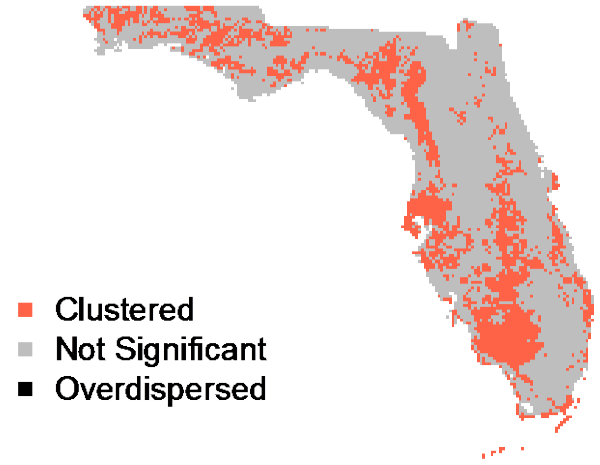
Phylogram/no branch lengths

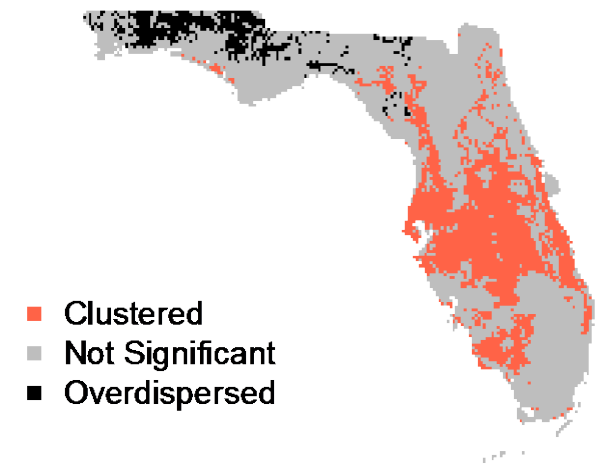


Ultrametric/dated



PhyloMatic





EPA ecoregions

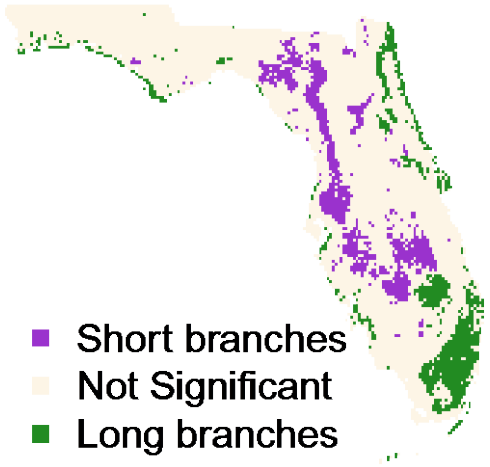
Regions with >50% of pixels significant for over- or under- dispersal

EPA ecoregion	PhyloMatic	RaxML
Southern Pine Plains and Hills	over	over
Dougherty Plain	over	over
Tifton Upland	-	over
Tallahasee Hills/Valdosta Limesink	-	-
Southeastern Floodplains and Low Terraces	-	over
Gulf Coast Flatwoods	-	-
Southwestern Florida Flatwoods	under	under
Central Florida Ridges and Uplands	under	-
Eastern Florida Flatwoods	under	-
Okefenokee Plains	-	-
Sea Island Flatwoods	-	-
Okefenokee Swamp	-	-
Floodplains and Low Terraces	-	-
Sea Islands/Coastal Marsh	-	-
Gulf Barrier Islands and Coastal Marshes	-	-
Big Bend Coastal Marsh	-	-
Everglades	-	-
Big Cypress	under	under
Miami Ridge/Atlantic Coastal Strip	-	-
Southern Coast and Islands	-	-

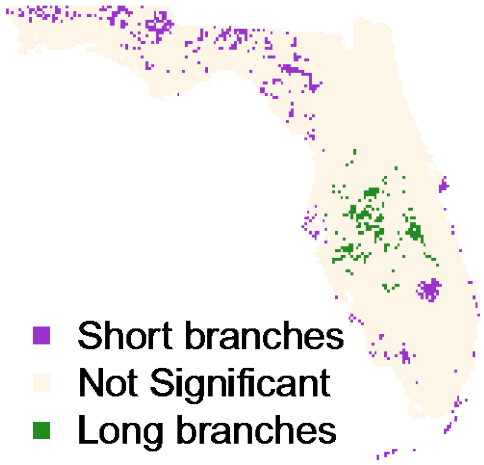
Relative Phylogenetic Diversity | 2.

RAxML

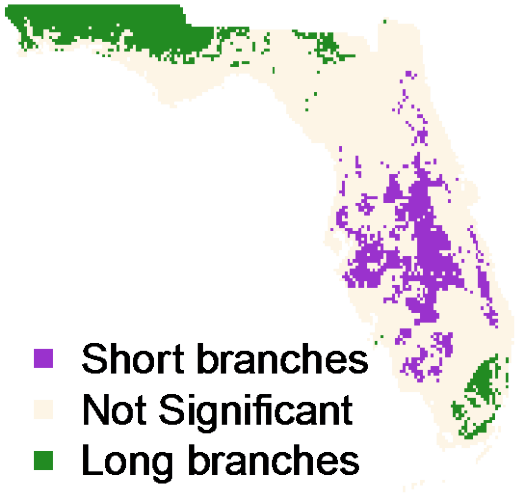
Phylogram/no branch lengths

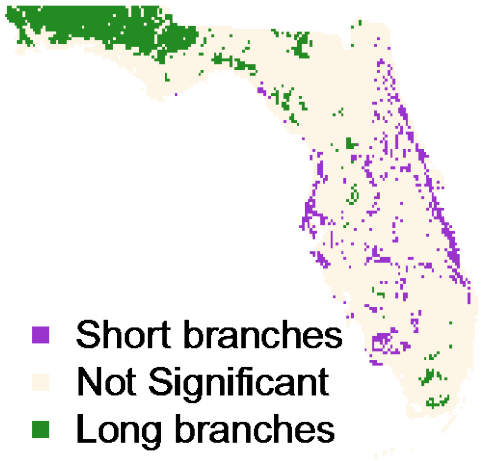


Phylomatic

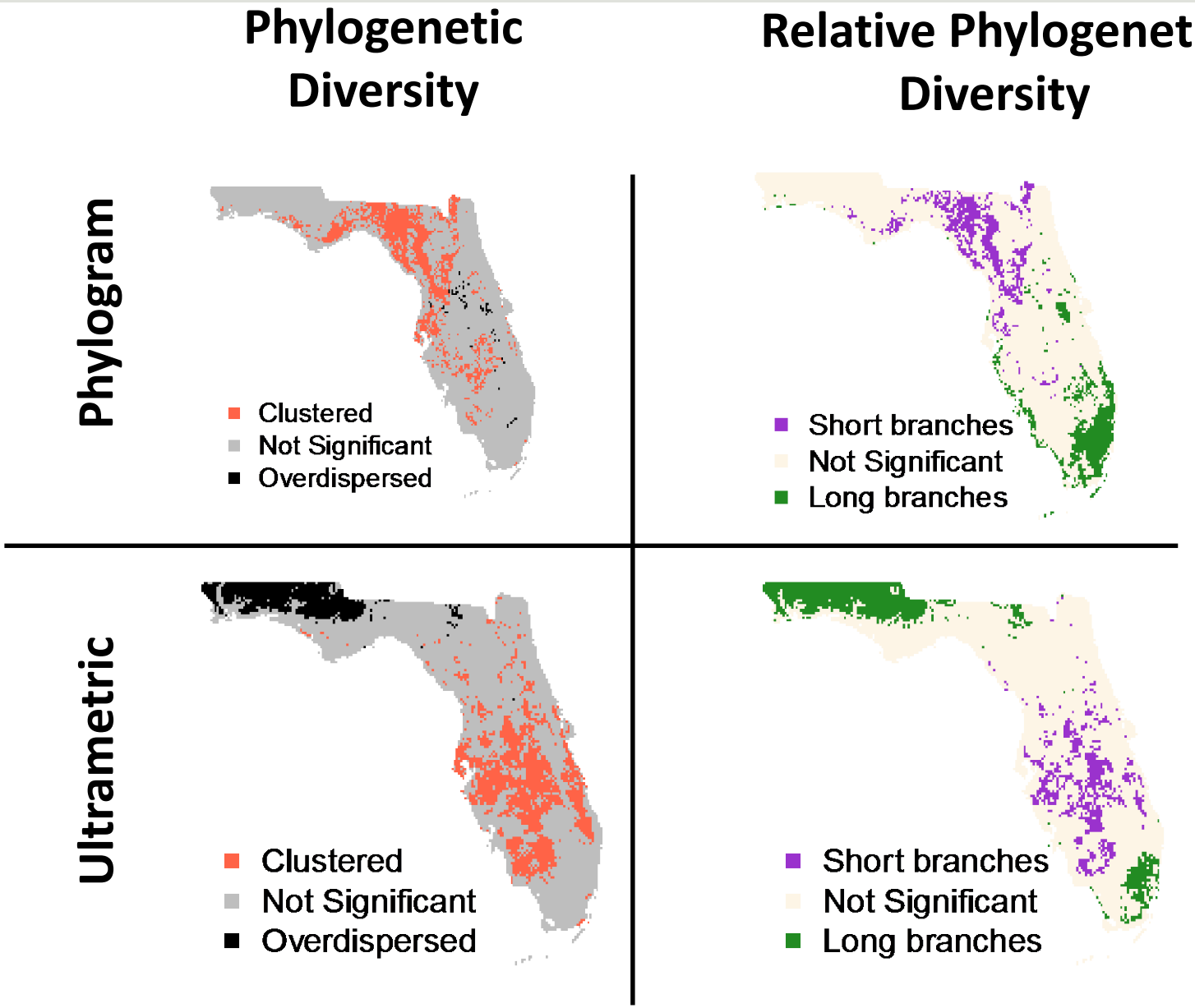


Ultrametric/dated



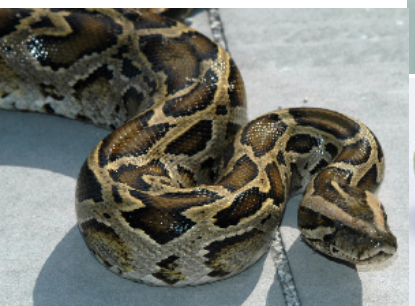


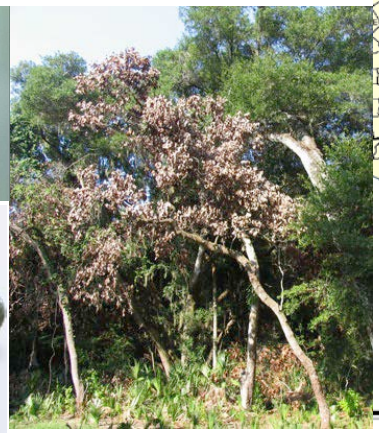
Phylodiversity | 3.

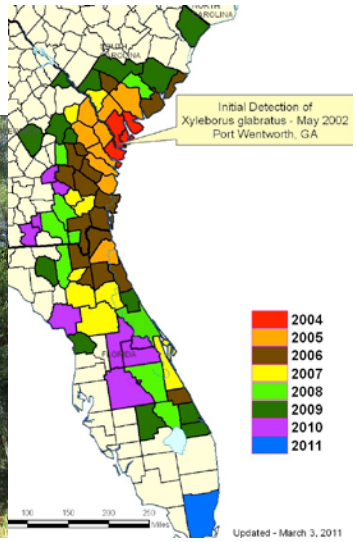


Invasive Species: Ecological & Economic Impacts

- Where have invasives been introduced, and how quickly are they spreading?
- What is the pattern of spread, and do patterns covary with other species?
- How does climate change affect the spread of invasives?
- Can we predict future invasions?







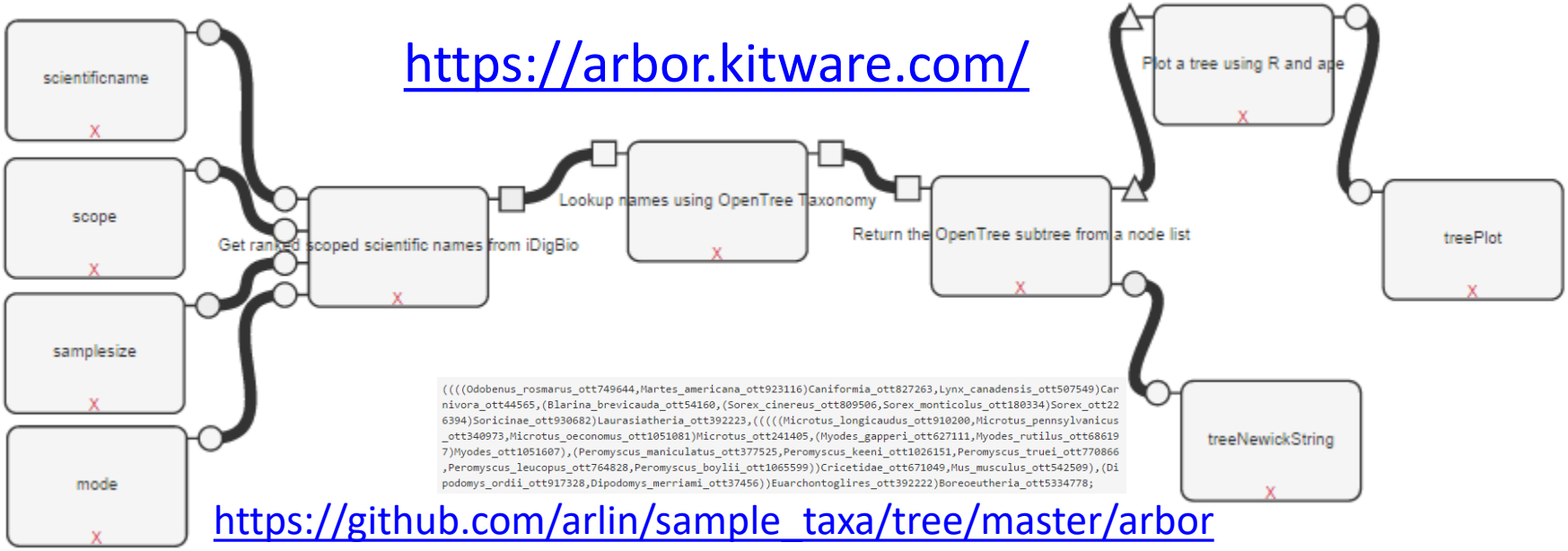
Topics of Networks

- Flowering time and climate change in New England
- Use of bryophytes and lichens as indicators of climate change across North America
- Invasive species of the Great Lakes region
- Plants, insect herbivores, parasitoids
- Flora of the SE US and response to climate change
- Response of paleo marine communities to environmental change

15 TCNS and collaborating institutions: ~250 institutions in 50 states

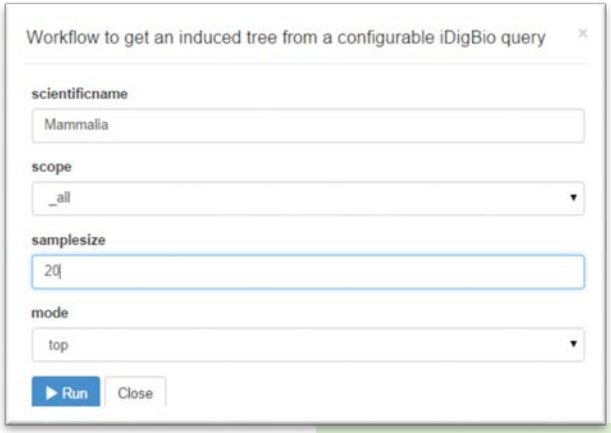
Arbor, OpenTree, and iDigBio

https://arbor.kitware.com/



((((Odobenus_rosmarus_ott749644,Martes_americana_ott923116)Caniformia_ott827263,Lynx_canadensis_ott507549)Carnivora_ott44565,(Blarina_brevicauda_ott54160,(Sorex_cinereus_ott809506,Sorex_monticolus_ott180334)Sorex_ott226394)Soricinae_ott930682)Laurasiatheria_ott392223,(((((Microtus_longicaudus_ott910200,Microtus_pennsylvanicus_ott340973,Microtus_oeconomus_ott1051081)Microtus_ott241405,(Myodes_gapperi_ott627111,Myodes_rutilus_ott686197)Myodes_ott1051607),(Peromyscus_maniculatus_ott377525,Peromyscus_keeni_ott1026151,Peromyscus_truei_ott770866,Peromyscus_leucopus_ott764828,Peromyscus_boylii_ott1065599))Cricetidae_ott671049,Mus_musculus_ott542509),(Dipodomys_ordii_ott917328,Dipodomys_merriami_ott37456))Euarchontoglires_ott392222)Boreoeutheria_ott5334778;

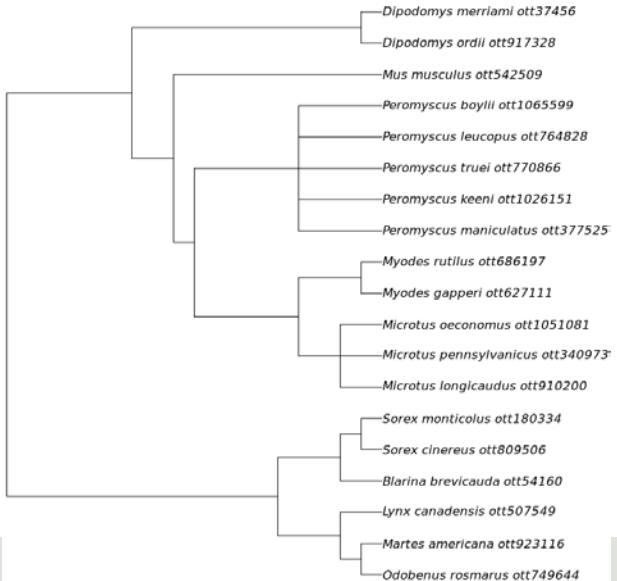
https://github.com/arlin/sample_taxa/tree/master/arbor




```
{
    "query": {
        "query_string" : {
            "default_field" : "order",
            "query" : "rodentia"
        }
    },
    "aggregations": {
        "my_agg": {
            "terms": {
                "field": "scientificname",
                "size": 100
            }
        }
    }
}
```

Success! Produced the following outputs:

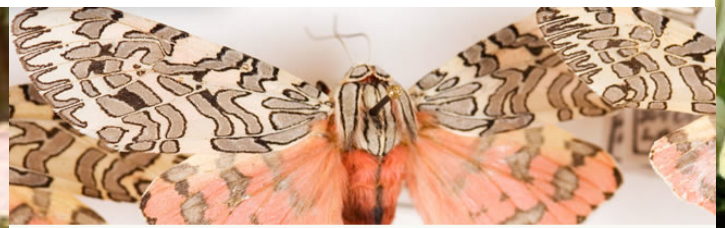
- Workflow to get an induced tree from a configurable iDigBio query treeNewickString [string]
- Workflow to get an induced tree from a configurable iDigBio query treePlot [image]



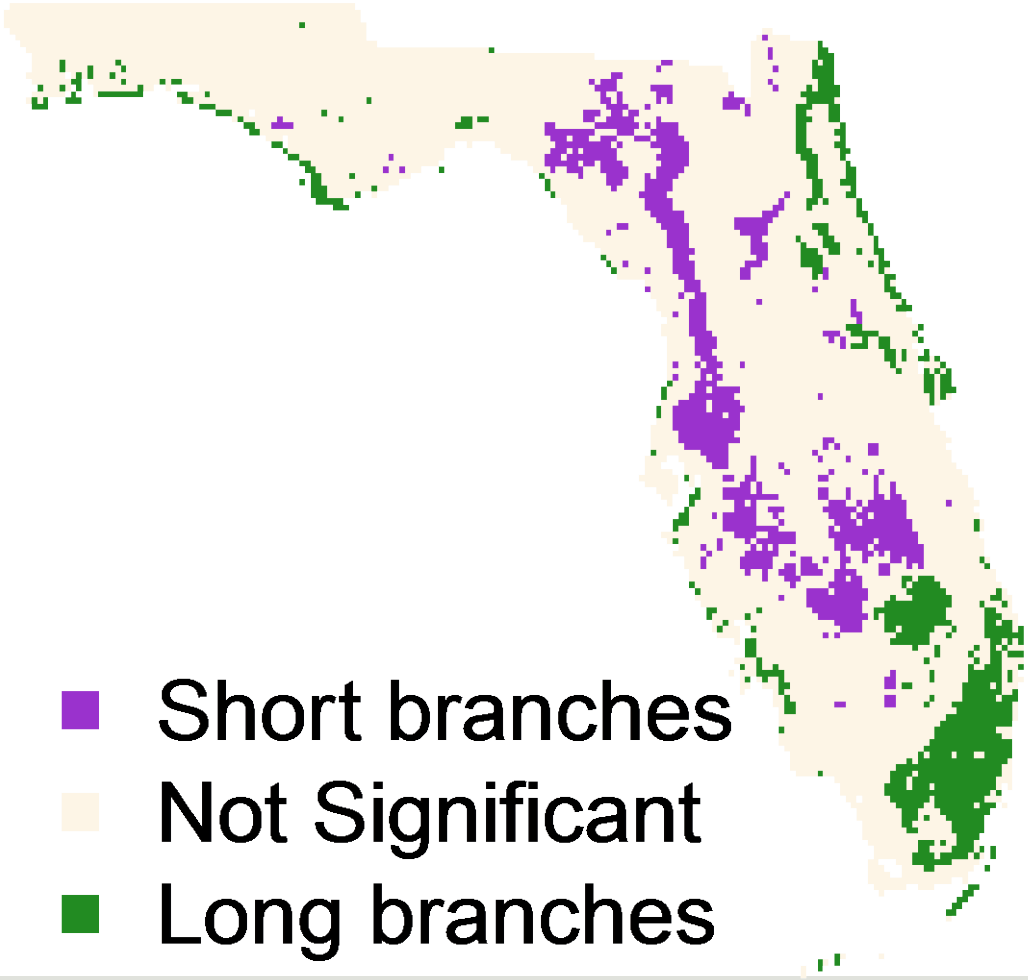
A. Matsunaga

What Can We Do with Specimen Data?

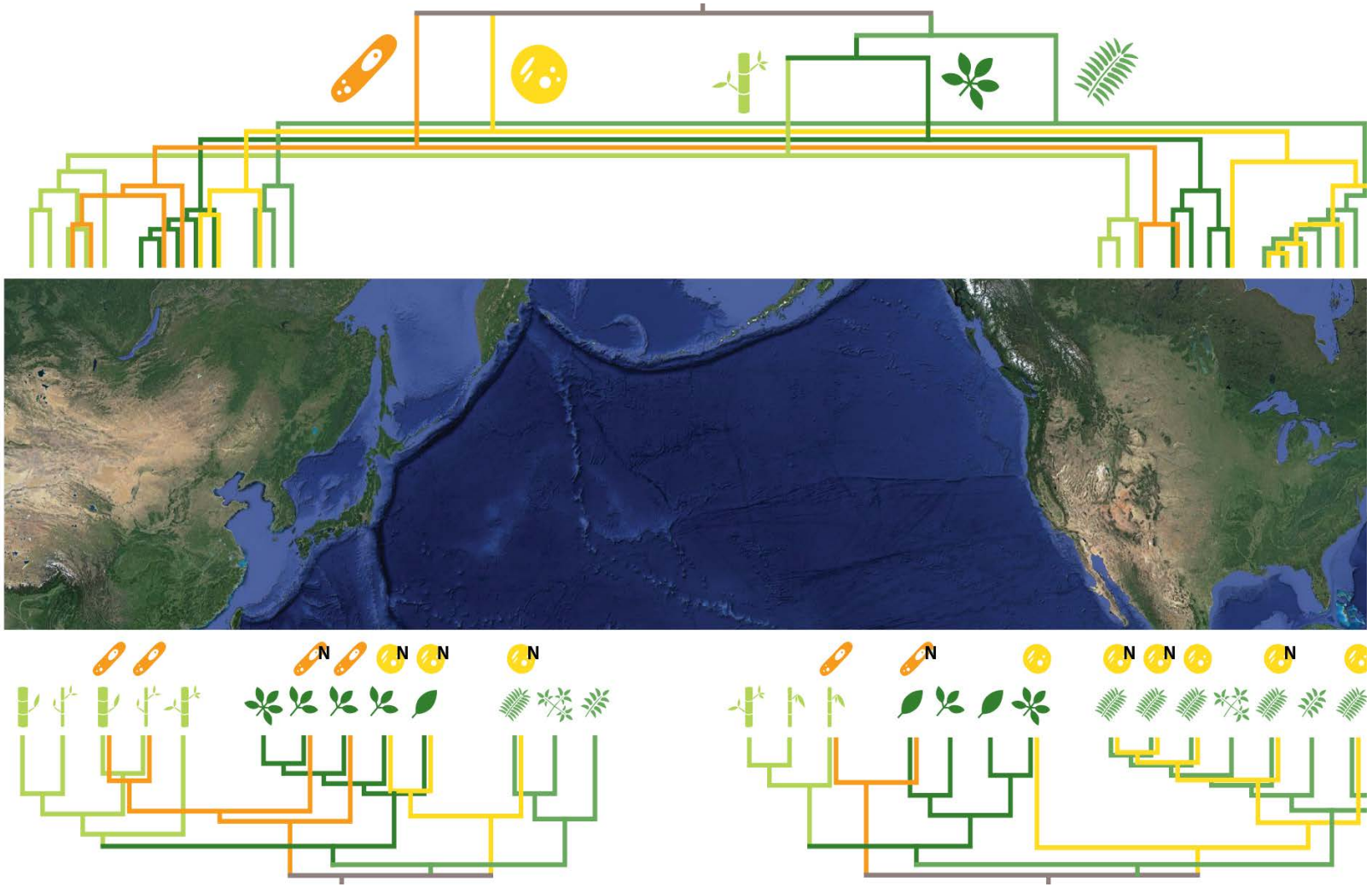
- Monitor shifts in biodiversity through time
- Track invasive species
- Ecological Niche Modeling, climate change
- Track phenological shifts
- Integrate with evolutionary history
- Past movements and climate change
- Landscape genetics

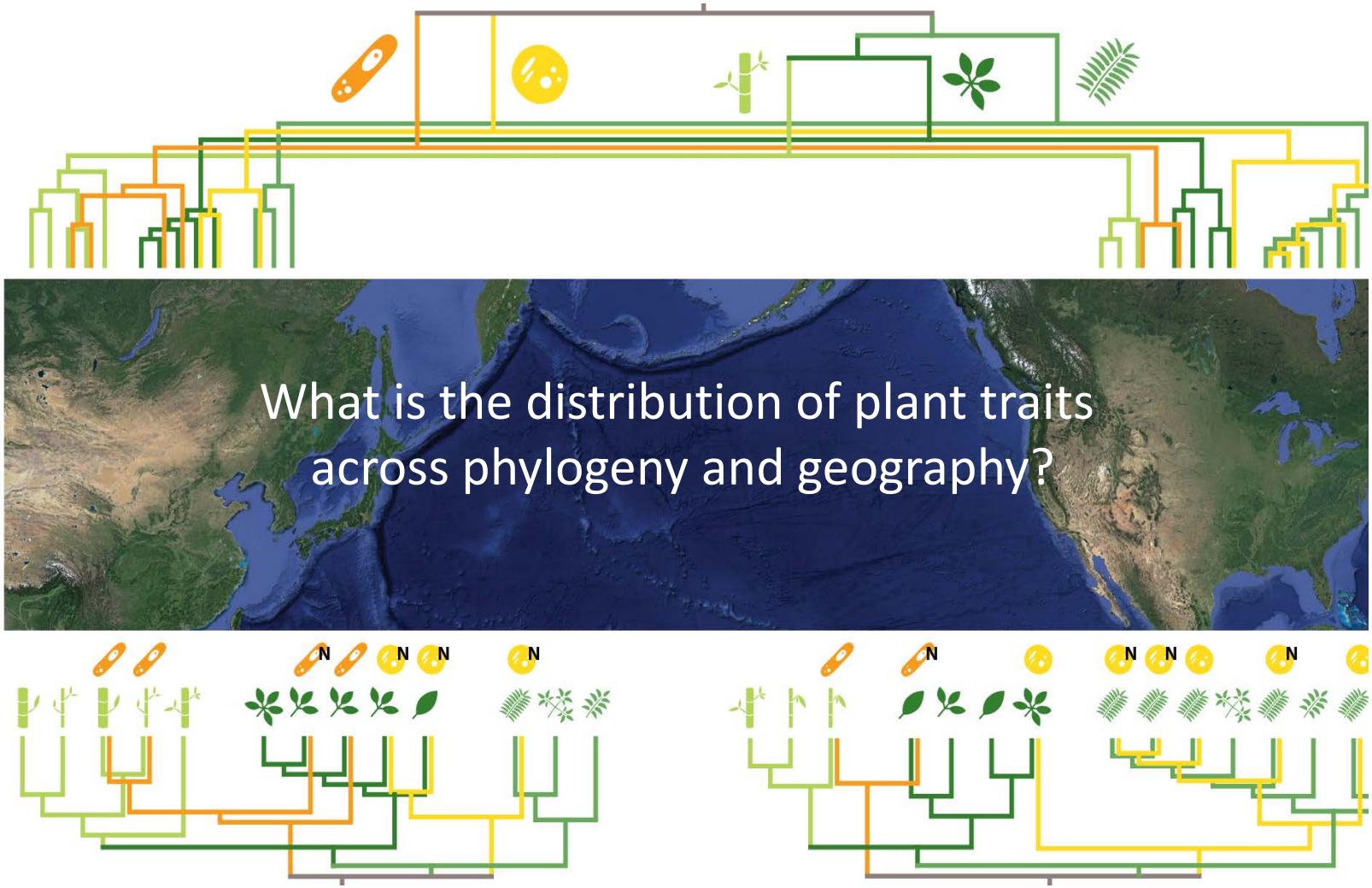
Relative Phylogenetic Diversity - angiosperms: PD/PD with all branches being equal



Integrating Phylogenetics and Plant Trait Data

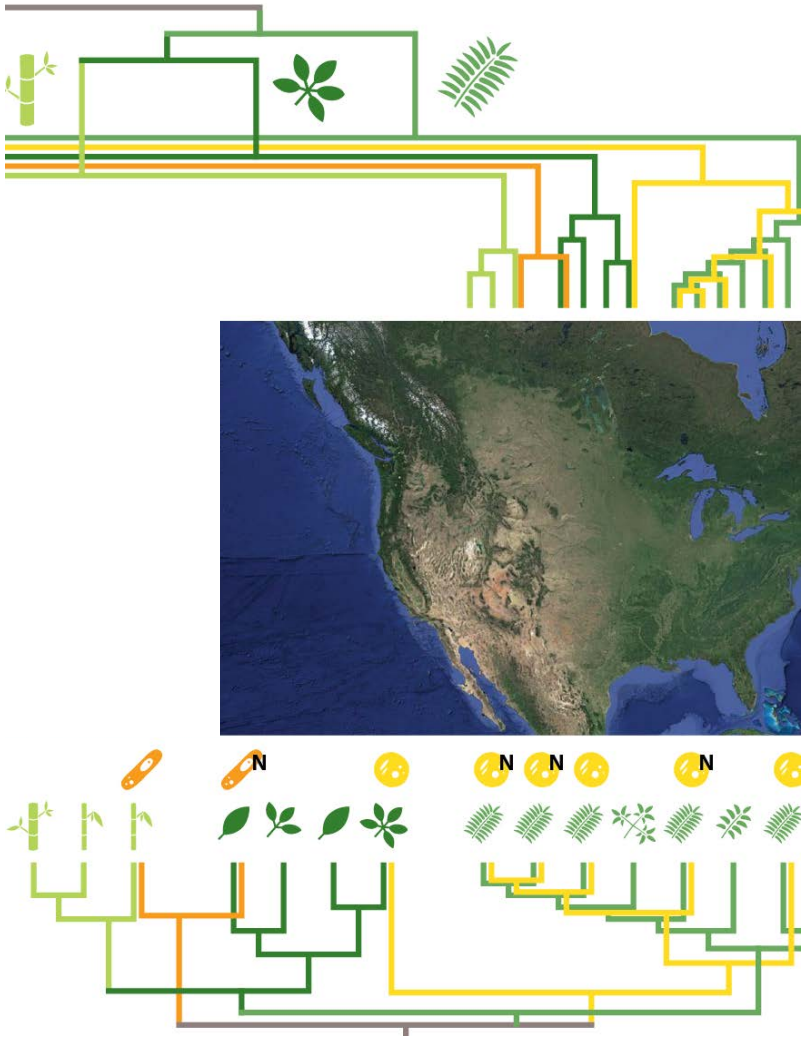


Integrating Phylogenetics and Plant Trait Data

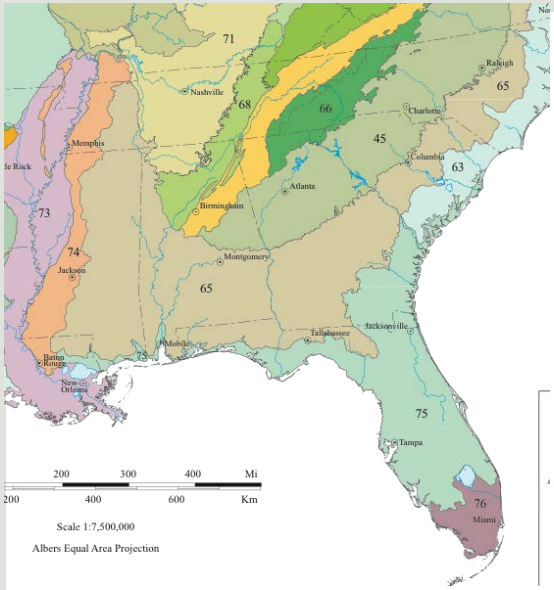


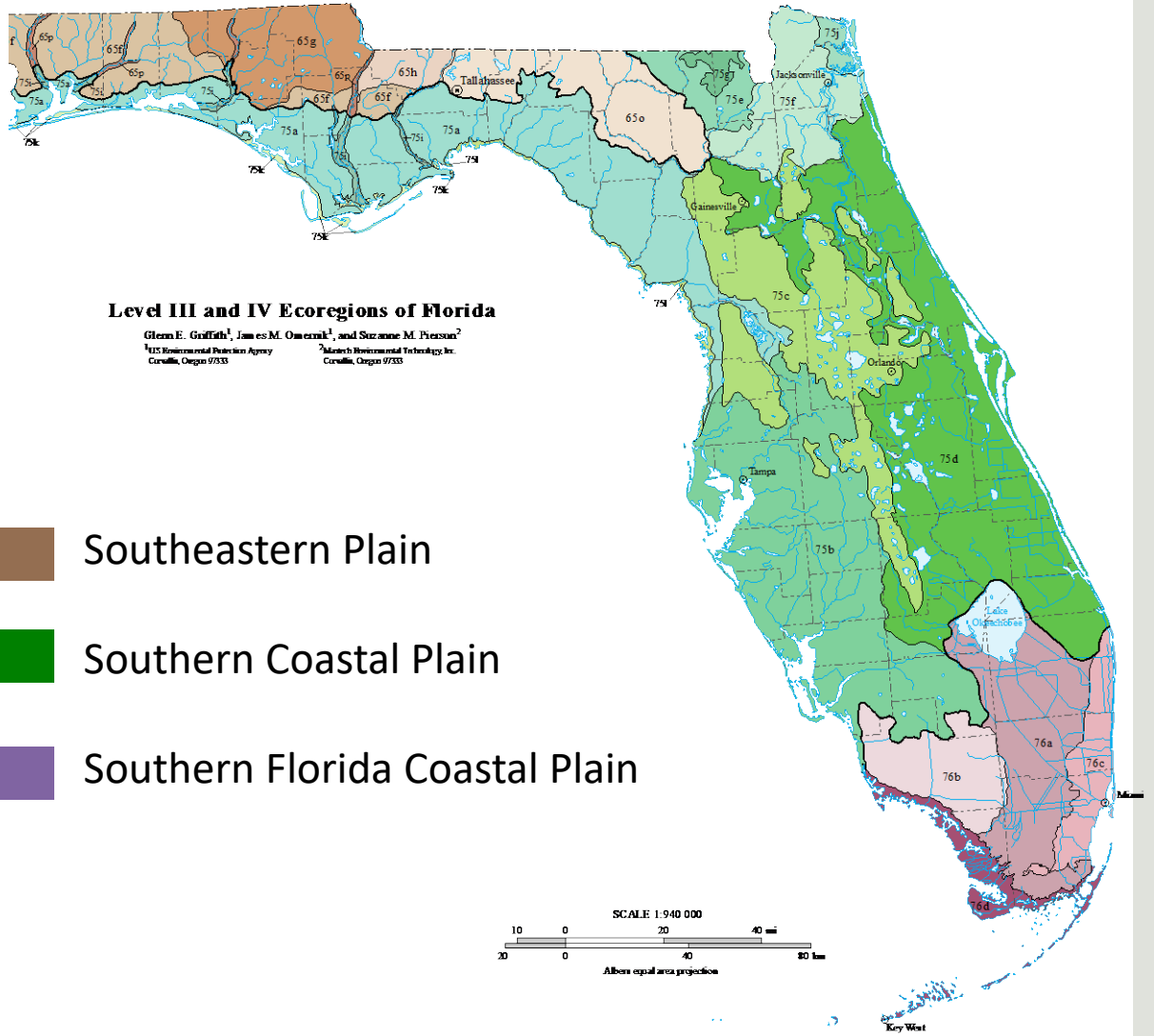
Integrating Phylogenetics and Plant Trait Data

- Community-level phylogenies
  - Comparisons of phylogenetic diversity across spatial scales
  - Correlation of plant functional traits
  - Over- and underdispersion of traits:
    - Are traits phylogenetically constrained or broader properties of communities?
    - Implications for long-term health of communities



Florida – EPA Regions





Southeastern Plain

Southern Coastal Plain

Southern Florida Coastal Plain

~ 4,100 plant species

Ziziphus celata

Gitzendanner et al.