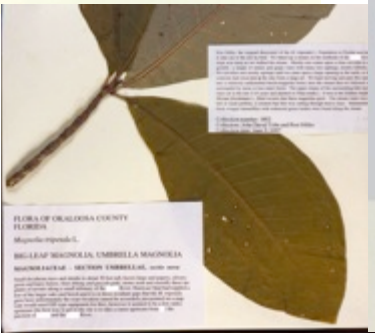
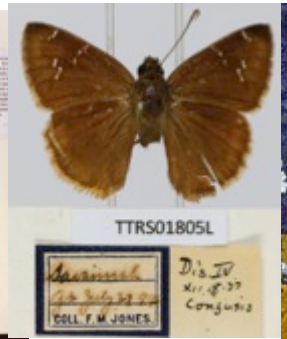
iDigBio
Integrated Digitized Biocollections

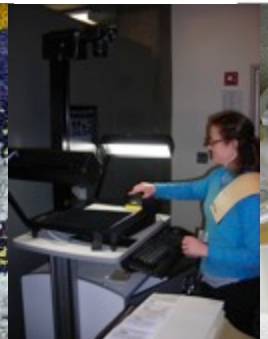
NSF

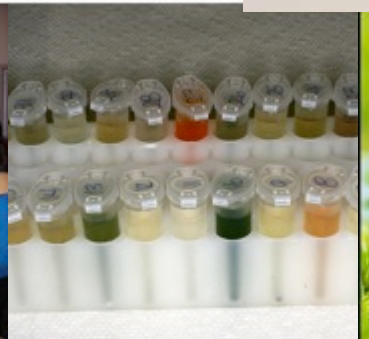
iDigBio Research Coordination and Scientific Community Outreach

Pamela S. Soltis
Co-PI & Director for Research









Components of iDigBio:

Digitization of Biodiversity Collections
Interactive, Integrative, Innovative

- Cyberinfrastructure
- Digitization
- Research
  - *Access to specimen data: Provide portal access to biodiversity data in a cloud-computing environment*
  - *Develop a computational environment to facilitate specimen-based integrative biodiversity research*
  - *Develop research workflows to anticipate computational needs and linkages*
- Education and Outreach

Components of iDigBio:

Digitization of Biodiversity Collections
Interactive, Integrative, Innovative

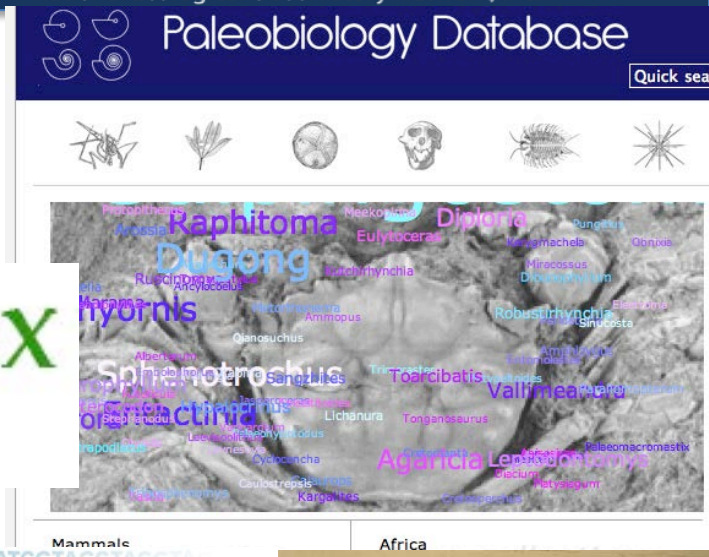
- Cyberinfrastructure
- Digitization
- Research
  - *Access to specimen data: Provide portal access to biodiversity data in a cloud-computing environment*
  - *Develop a computational environment to facilitate specimen-based integrative biodiversity research*
  - *Develop research workflows to anticipate computational needs and linkages*
- Education and Outreach

Plans for Year 2
(from last year's EAB meeting)

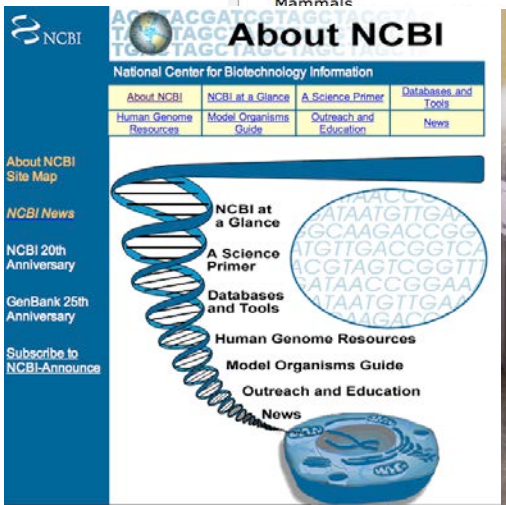
- Linked Data workshop (with NESCent)
- Foster additional workshops with research communities
- Symposia and training at national/regional meetings
- Develop network of genetic and tissue collections
- Engage post-docs and students in research

Linking Collections to…

- Ecology
- Paleontology
- Genomics
- Living Collections
- Other repositories



UTEX



Linking Collections to Ecology

- Through NEON

National Ecological Observatory Network

- Biological monitoring at sites across USA; collections
- Baseline for changes in species distribution and abundance over time

iDigBio
Integrated Digitized Biocollections

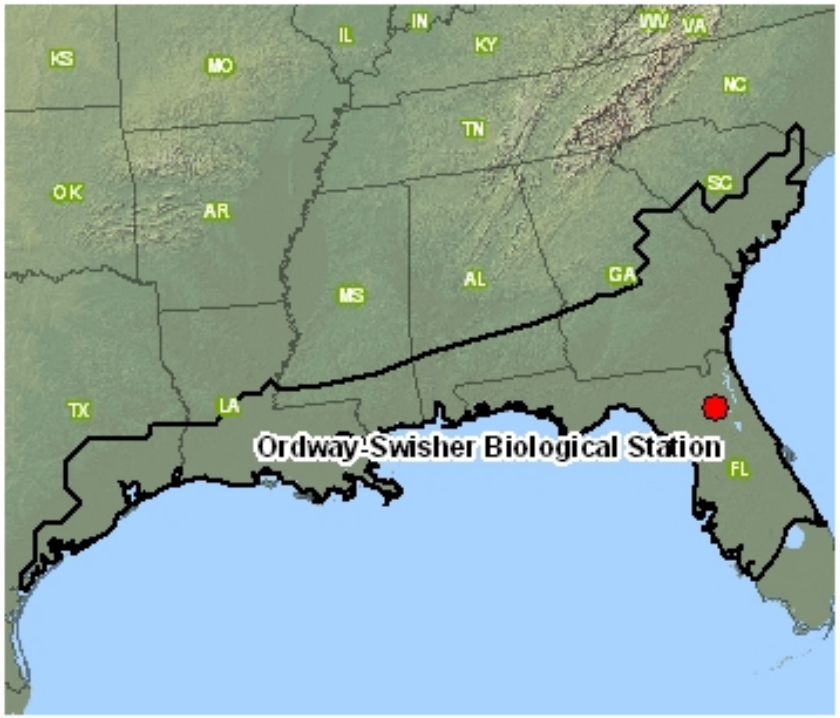
University of Florida
ORDWAY-SWISHER BIOLOGICAL STATION

topics of interest
- about osbs •
- research •
- education •
- conservation •
- NEON •
- ERT •
- support osbs •

• home \ neon

national ecological observatory network, inc

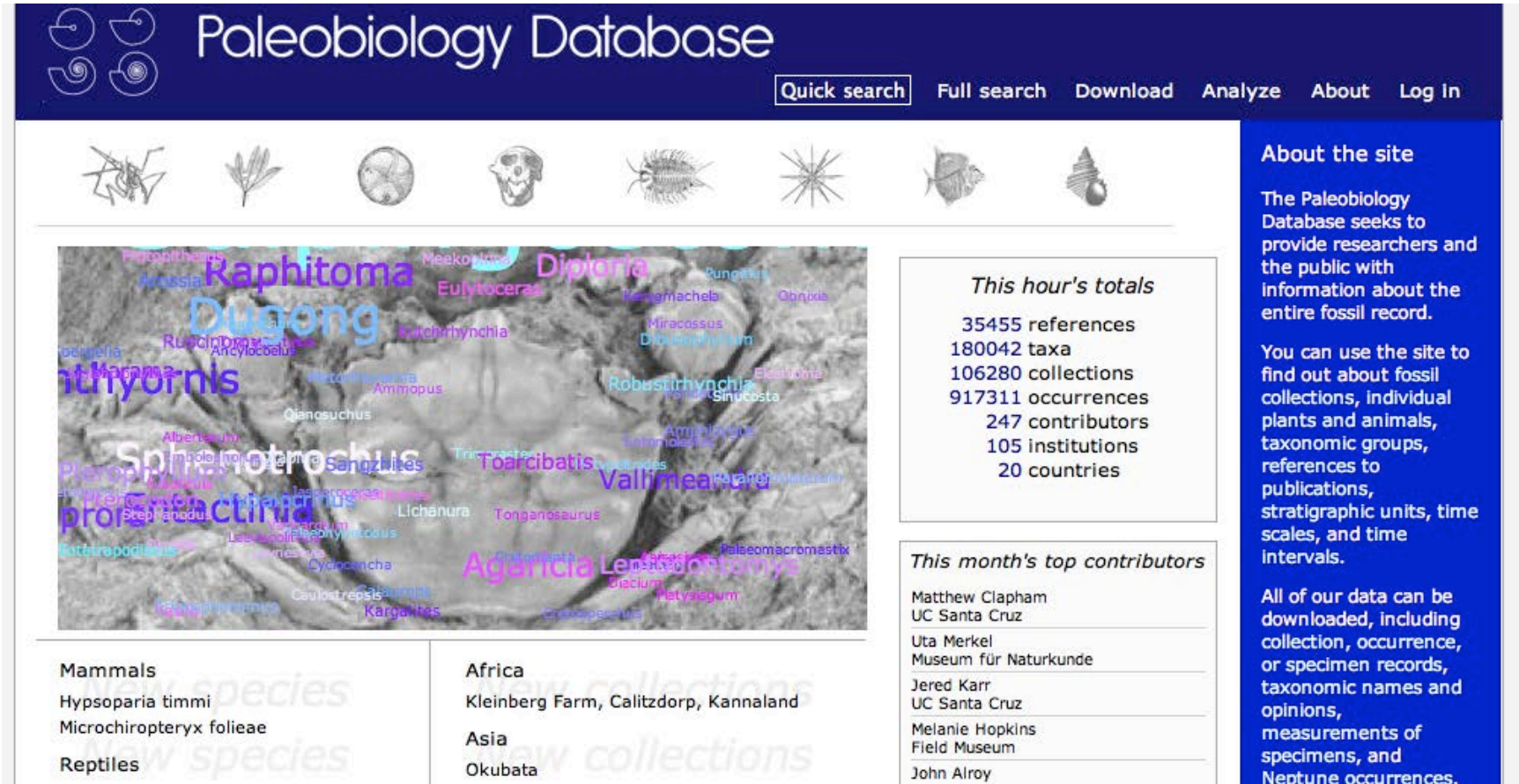
neon
National Ecological Observatory Network, Inc.



NEON site at Ordway-Swisher Biological Station (Photo Credit: NEON)

Linking Collections to Paleobiology

- Paleobiology Database
  - (http://paleodb.org/cgi-bin/bridge.pl)

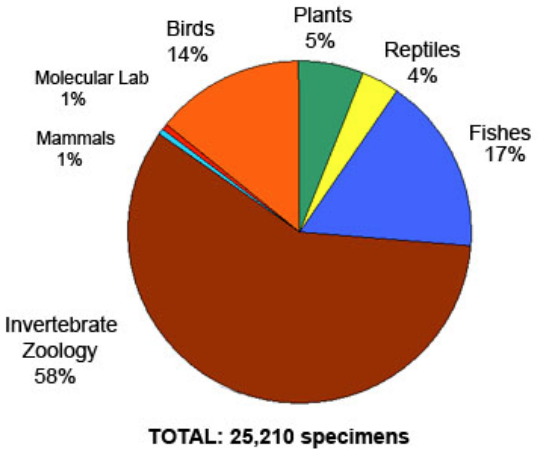


Linking Collections to Paleobiology

- Paleocollections Workshop, April, 2012
  - 30 Participants (outside iDigBio)
  - Topics
    - Georeferencing
    - Digitization, OCR, NLP
    - Grand challenges and research priorities
  - Paper in progress
- Paleocollections Working Group established
- Funding of paleo-focused TCN
  - **Digitizing Fossils to Enable New Syntheses in Biogeography –** Creating a PALEONICHES TCN (B. Lieberman, University of Kansas)

Linking Collections to Genomics

- National network of tissue and genetic resources



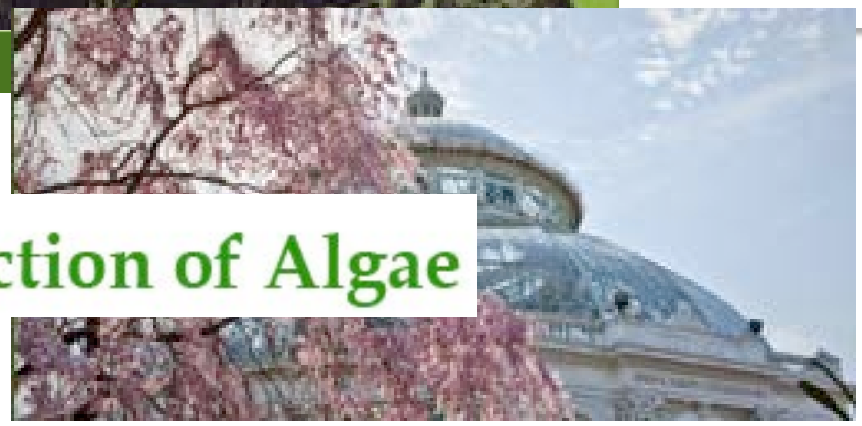
FLMNH GRR: the Genetic Resources Repository
Florida Museum of Natural History

Linking Collections to Genomics

- Participated in NSF-sponsored workshop on DNA banks at Missouri Botanical Garden, Jan. 2013
- Assembling a file of national DNA and tissue banks
- Developing a common web portal to these collections (available April 2013)
- DNA resources ultimately connected to specimen data in iDigBio

Linking to Living Collections

- Botanical gardens, zoos, culture collections



Linking to Living Collections

- Botanical gardens, zoos, culture collections
  - Discussions with botanical gardens about database structures: living collections vs. herbaria

New Partnerships for Research

- iPlant Collaborative
  - Taxonomic Names Resolution Service
  - Workflows
- AVAToL – Assembling, Visualizing, and Analyzing the Tree of Life
  - Workshop at NSF, May 2013

Research Working Group

- Summit II research break-out group
- Community input on potential research questions
  - Research-oriented discussions with TCNs to identify computational needs to address the questions they posed, assuming data are available
- What data sources and computational software is needed?

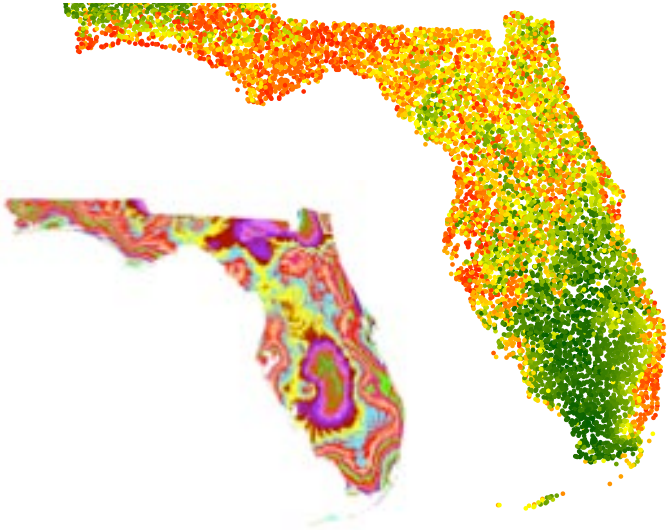
Components of iDigBio:

Digitization of Biodiversity Collections
Interactive, Integrative, Innovative

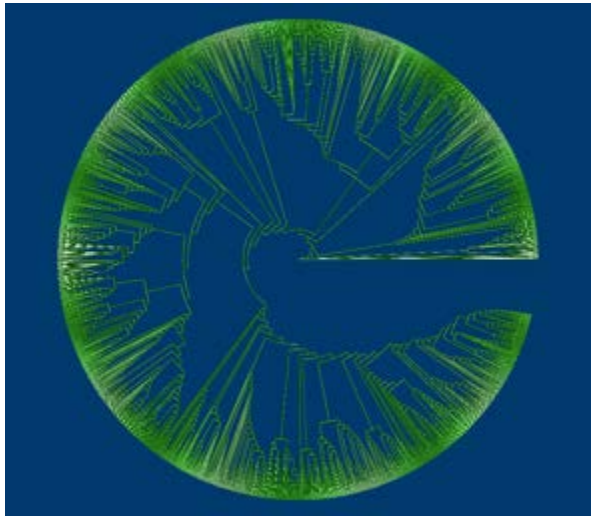
- Cyberinfrastructure
- Digitization
- Research
  - *Access to specimen data: Provide portal access to biodiversity data in a cloud-computing environment*
  - *Develop a computational environment to facilitate specimen-based integrative biodiversity research*
  - *Develop research workflows to anticipate computational needs and linkages*
- Education and Outreach

Florida Plant Diversity in a Changing Climate

Integrating specimen data, climate change models, and phylogeny



+

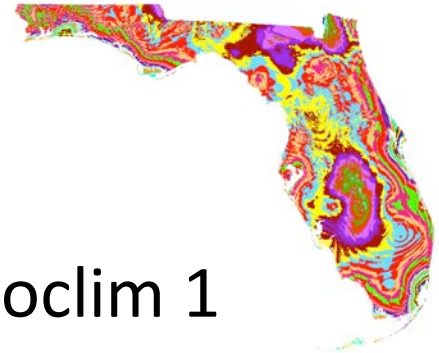


2609 species (of ~4200)
all included in phylogeny

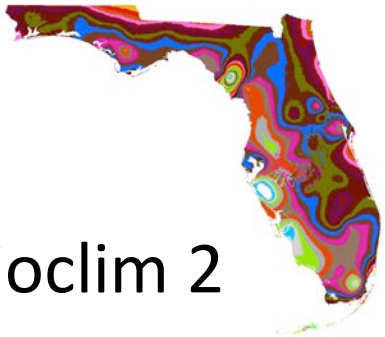
Phylogenetic tree, 2609 species
GenBank, new (1000 spp)

C. Germain-Aubrey, J. Allen, K. Neubig, L. Majure, R. Abbott, G. Burleigh, D. Soltis

Ecological Niche Modeling

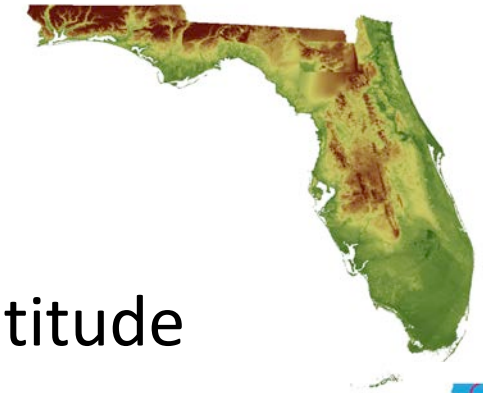


Bioclim 1



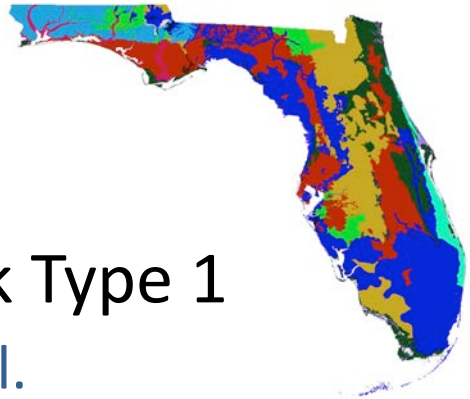
Bioclim 2

Bioclim 3

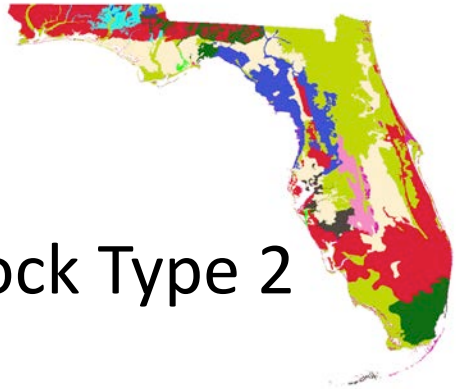


Altitude

Land Cover



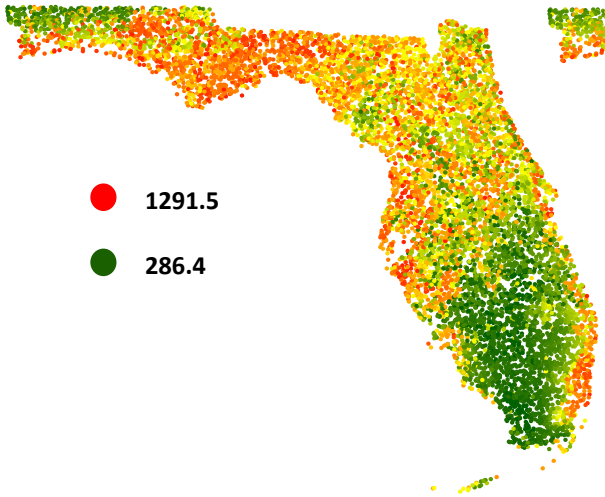
Rock Type 1



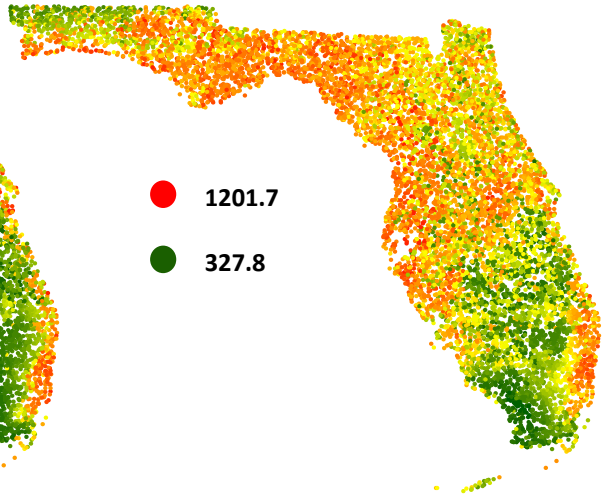
Rock Type 2

C. Germain-Aubrey et al.

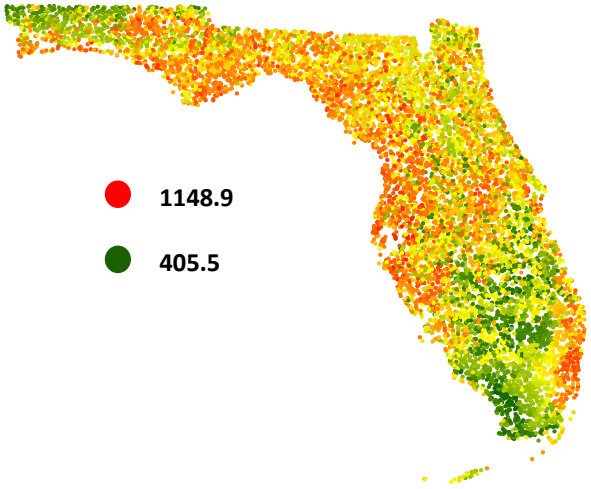
Species Indices: Present, 2050, 2080



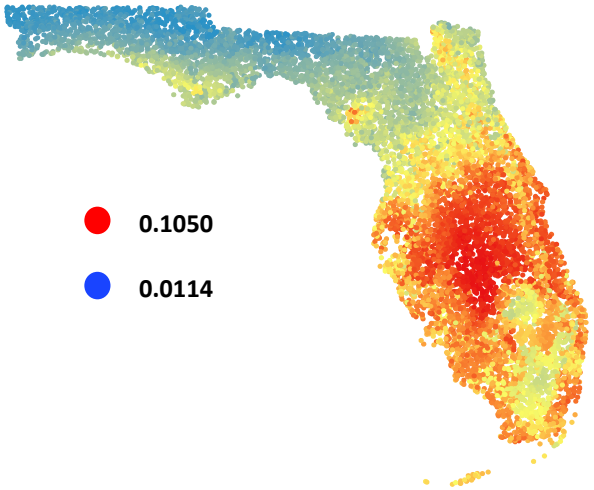
Present Species Index

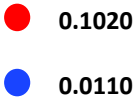


2050 Species Index

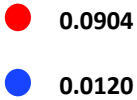


2080 Species Index





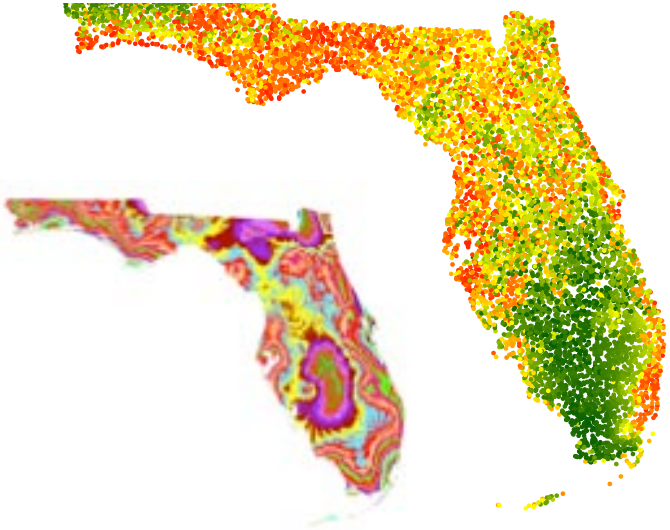
mic Ratio



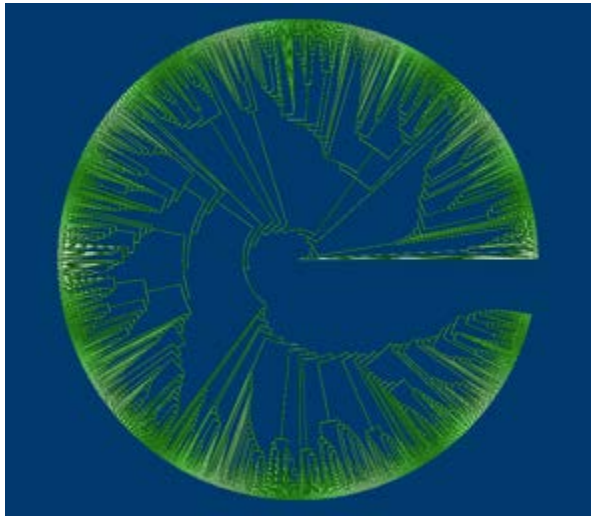
2080 Endemic Ratio

Florida Plant Diversity in a Changing Climate

Integrating specimen data, climate change models, and phylogeny



+



2609 species (of ~4200)
all included in phylogeny

Phylogenetic tree, 2609 species
GenBank, new (1000 spp)

C. Germain-Aubrey, J. Allen, K. Neubig, L. Majure, R. Abbott, G. Burleigh, D. Soltis

Florida Plant Diversity in a Changing Climate

Computational & Data Requirements:

access to georeferenced specimen data

access to GenBank

access to Bioclim information

tools for georeferencing

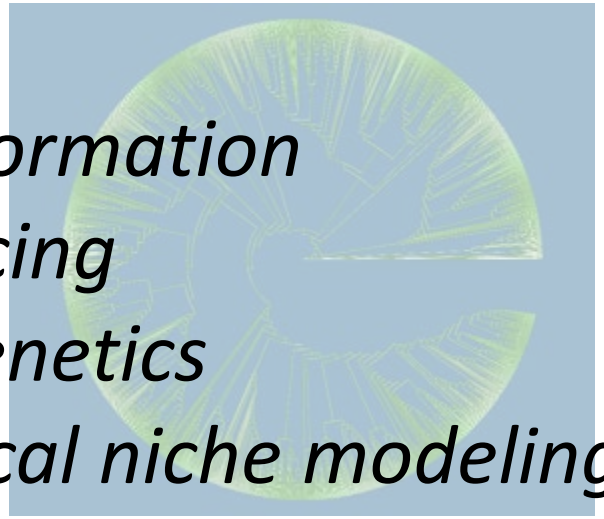
software for phylogenetics

software for ecological niche modeling

ability to incorporate custom scripts

Formal development of workflow

application to FLMNH fish & Lepidoptera



C. Germain-Aubrey, J. Allen, K. Neubig, L. Majure, R. Abbott, G. Burleigh, D. Soltis

Scientific Community Outreach

- Year 2 activities
  - Symposium and workshop at Botany 2012
  - Digitization workshops (Gil Nelson & Deb Paul)
  - DNA Banking workshops
    - Missouri Bot. Gard., Harvard, Smithsonian
  - SPNHC Symposium (*Diverse Uses for Natural History Collections*)
  - Planning for Year 3

Scientific Community Outreach

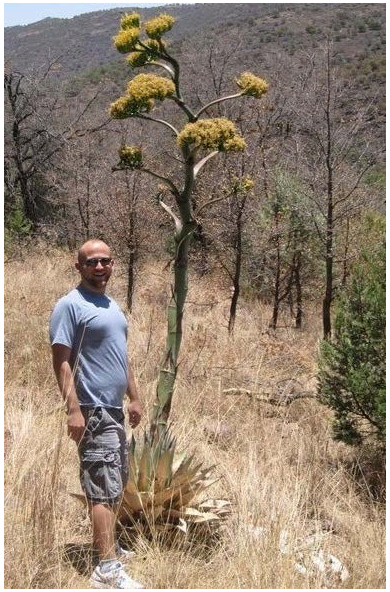
- Year 3 activities
  - 3 symposia at Botany 2013
    - ***Symposium:*** **Public Participation in Scientific Research: Emerging Resources for Botany Organizers: Austin Mast, Sarah Newman**
    - ***Symposium*****: Herbarium Digitization for Research, Teaching, and the Public Organizer: Eric Ribbens et al.**
    - ***Symposium*****: Broadening Participation - Recruiting and Retaining Outstanding Scientists in the Botanical Sciences Organizers: Anna Monfils, Ann Sakai**
  - Additional digitization workshops (Gil Nelson & Deb Paul)
  - Additional training workshops

Student & Post-doc Training

- Post-doc Charlotte Germain-Aubrey
- Graduate RAs Grant Godden and Ryan Moraski

Florida Plants
Fish, Lepidoptera

DNA Banking
Outreach



Georeferencing

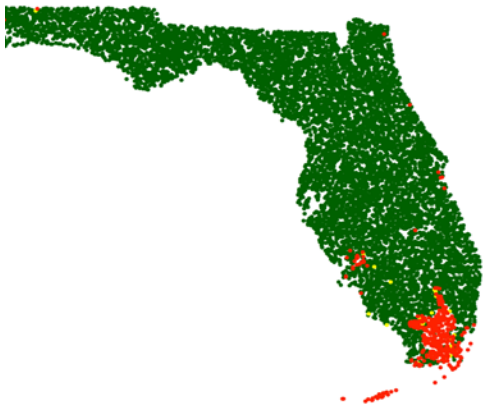
Plans for Year 3

- Engage Research Working Group
- Develop detailed workflows for research
- More explicit integration with other data sources
- Linked Data workshop (with NESCent)
- Develop international collaborations to enhance research opportunities
- Foster additional workshops with research communities
- Symposia and training at national/regional meetings
- Continue to engage post-docs and students in research

iDigBio
Integrated Digitized Biocollections

Thank you!

Species Movement: Past, Present, and Future



0.6545

04

0.3548

0.0001

Present-2080

Research Activities for Year 2

- Continuing discussions with other data sources
- Working toward an integrated national network of genetic resources
  - Survey of genetic collections to assess needs, capabilities, etc.
  - workshop in January
- Development of sample research workflows using specimen data to enable development of a computational environment for integrated research
- NESCent-hosted workshop on linked data – spring