Using Specimens and Linked Data In Ecological and Evolutionary Research

Pamela S. Soltis
University of Florida

Collections: The Library of Life

~1400 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

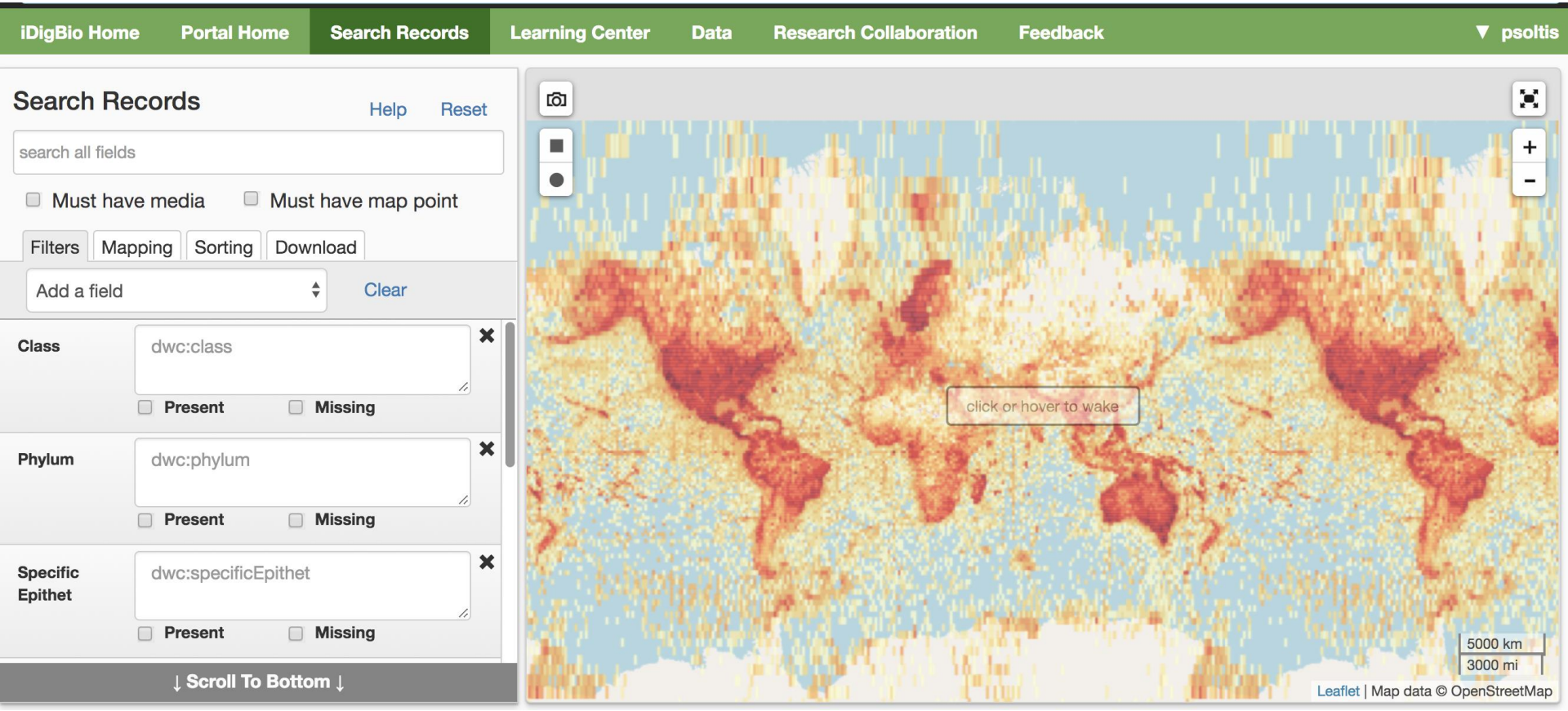
Most specimens locked away in cabinets, unavailable for general use.

DIGITIZATION!!!!

iDigBio
Integrated Digitized Biocollections

www.idigbio.org

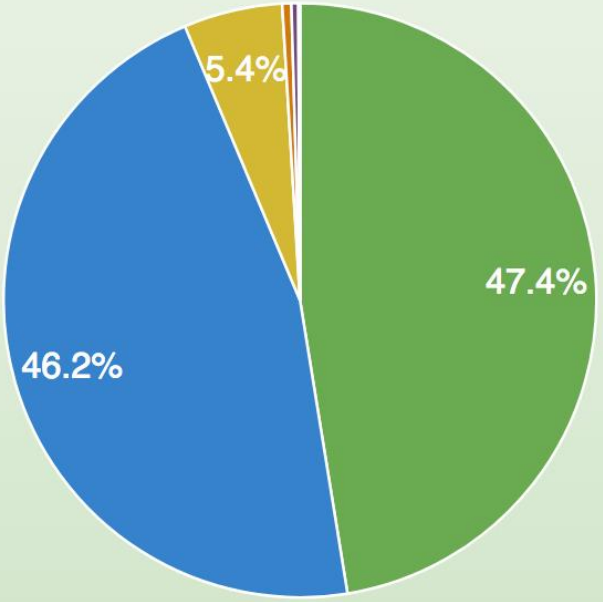
iDigBio: Enabling Biodiversity Research



Specimen Records in iDigBio

Specimen Records

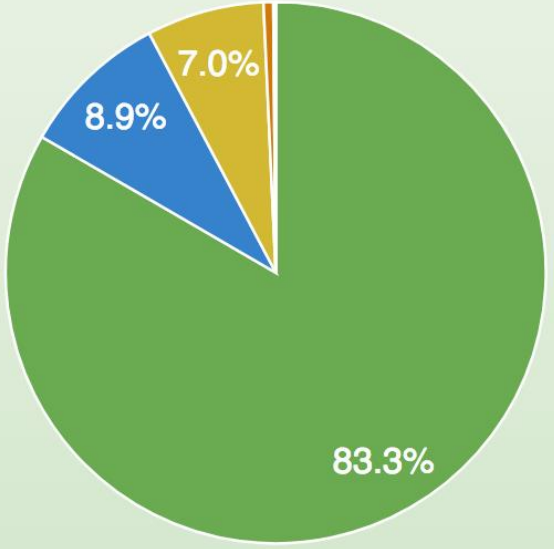
105,429,263



Plantae Animalia Fungi Chromista
other Protozoa Bacteria Eubacteria

Media Records

21,926,865



Plantae Animalia Fungi Chromista
other Bacteria Protozoa Eubacteria
Plantae,plantae

Click a pie wedge to jump to an advanced search for records with the given kingdom. Click a legend name to remove/add a pie wedge.

Papilio cresphontes – giant swallowtail

Specimen Records in iDigBio

Papilio cresphontes

iDigBio Integrated Digitized Biocollections

About iDigBio | Research | Tec

Take our 30-second survey

The U.S. National Science Foundation and iDigBio are required to collect information on use of digitized collections-based specimen data. Please help us meet this requirement every time you use this search portal. Sustainability of the national digitization effort depends on evidence of data use! **Maybe later**.

iDigBio Home | Portal Home | Search Records | Learning Center | Data | Research Collaboration | Feedback | ▼ psoltis

Search Records

Help Reset

search all fields

Must have media Must have map point

Filters | Mapping | Sorting | Download

Add a field Clear

Specific Epithet dwc:specificEpithet

Present Missing

Genus dwc:genus Add EOL Synonyms

click or hover to wake

Specimen Records in iDigBio

Papilio aristodemus

iDigBio Integrated Digitized Biocollections

About iDigBio | Research | Technical Inf

Take our 30-second survey

The U.S. National Science Foundation and iDigBio are required to collect information on use of digitized collections-based specimen data. Please help us meet this requirement every time you use this search portal. Sustainability of the national digitization effort depends on evidence of data use! **Maybe later**.

iDigBio Home | Portal Home | Search Records | Learning Center | Data | Research Collaboration | Feedback | ▼ psoltis

Search Records

Help Reset

search all fields

☐ Must have media ☐ Must have map point

Filters | Mapping | Sorting | Download

Add a field Clear

Scientific Name papilio aristodemus Add EOL Synonyms

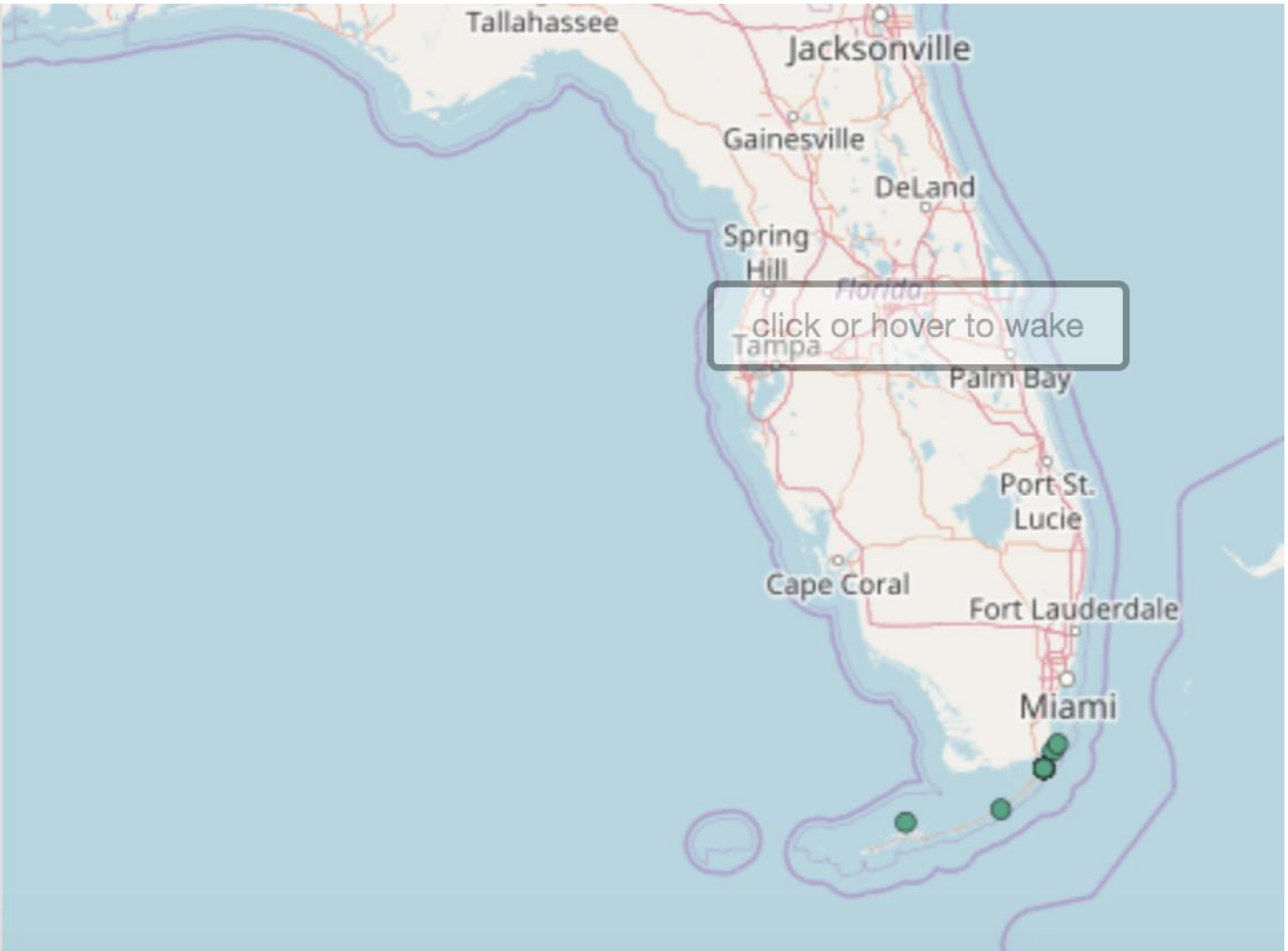
☐ Present ☐ Missing

Date Collected Start: yyyy-mm-dd End: yyyy-mm-dd

click or hover to wake

Specimen Records in iDigBio

Papilio aristodemus



Specimen Records in iDigBio

Papilio cresphontes

iDigBio Integrated Digitized Biocollections

About iDigBio | Research | Tec

Take our 30-second survey

The U.S. National Science Foundation and iDigBio are required to collect information on use of digitized collections-based specimen data. Please help us meet this requirement every time you use this search portal. Sustainability of the national digitization effort depends on evidence of data use! **Maybe later**.

iDigBio Home | Portal Home | Search Records | Learning Center | Data | Research Collaboration | Feedback | ▼ psoltis

Search Records

Help Reset

search all fields

Must have media Must have map point

Filters | Mapping | Sorting | Download

Add a field Clear

Specific Epithet dwc:specificEpithet

Present Missing

Genus dwc:genus Add EOL Synonyms

click or hover to wake

Specimen Records in iDigBio

Papilio cresphontes

Specimen Record

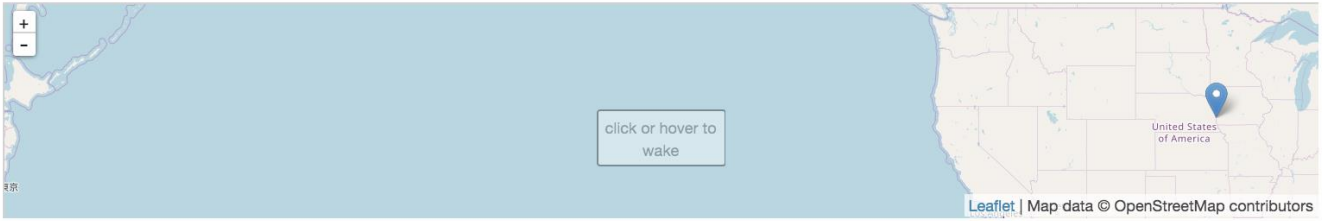
Animalia > Arthropoda > Insecta > Lepidoptera > Papilionidae

Papilio cresphontes Cr.

From Entomology Division, Yale Peabody Museum

Continent	North America
Country	United States
State/Province	Nebraska
County/Parish	Douglas County
Locality	Omaha
Latitude	41.2472
Longitude	-95.9557877

Institution Code	Ypm
Collection Code	Ent
Catalog Number	Ypm Ent 431144
Collected By	R A. Leussler
Date Collected	1909-06-19



Media

From Recordset

Entomology Division, Yale Peabody Museum

The systematic collections of the Yale Peabody Museum's Division of Entomology comprise over 1,000,000 curated specimens. Division holdings include important collections of Lepidoptera, arachnids, Orthoptera, Coleoptera, water beetles and midges, and specialty collections on evolutionary themes. Recent acquisitions also include historically important collections from other institutions. The Division also maintains a general entomological library of periodicals, books and reprints that includes coverage of arachnology donated from the Alexander Petrunkevitch Library.

Contacts

Name	Larry Gall	**Name**	Larry Gall	**Name**	Larry Gall
Role	Head, Computer Systems Office	**Role**	Head, Computer Systems Office	**Role**	Head, Computer Systems Office
Email	lawrence.gall@yale.edu	**Email**	lawrence.gall@yale.edu	**Email**	lawrence.gall@yale.edu
Name	Larry Gall				
Role	Head, Computer Systems Office				
Email	lawrence.gall@yale.edu				

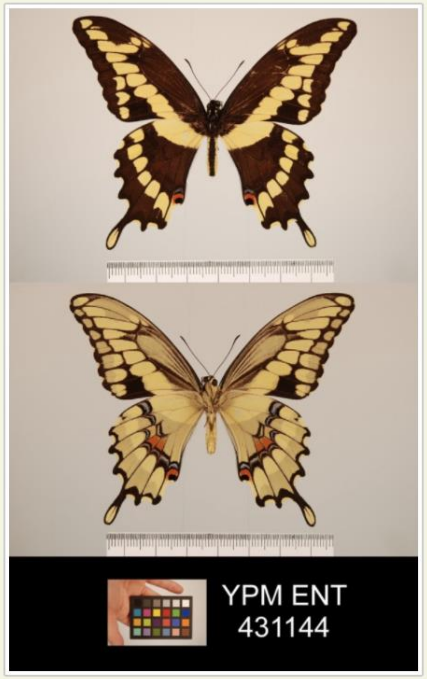
Specimen Records in iDigBio

Media Record

Animalia > Arthropoda > Insecta > Lepidoptera > Papilionidae

Papilio cresphontes Cr. view specimen record

From Entomology Division, Yale Peabody Museum



Media retrieved from:
http://deliver.odai.yale.edu/content/repository/YPM/id/36771/format/3

Open in browser

Download File

Papilio cresphontes

Specimen Records in iDigBio

Papilio cresphontes

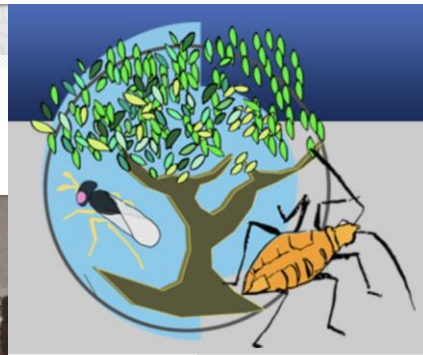
iDigBio
Integrated Digitized Biocollections

Thematic Collections Networks

ILLINOIS NATURAL HISTORY SURVEY
PRAIRIE RESEARCH INSTITUTE

invertnet

8,338 Slide Boxes
6,680 Vial Trays
3,448 Drawers
51,791 Total Images
186 Specimens Annotated

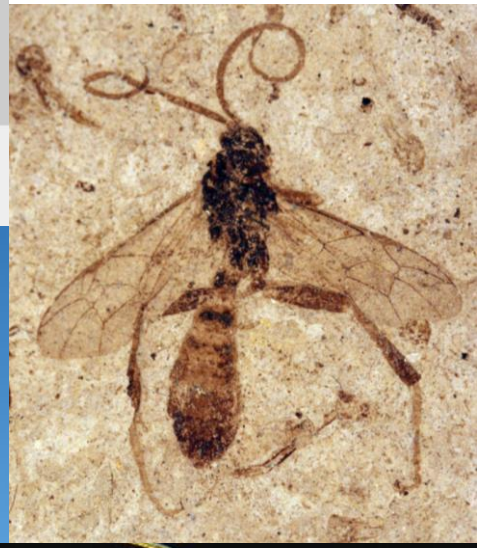


About
Partners and People
Products
Specimen Short

Tri-Trophic Thematic Collection Network

SCAN
Symbiota Collections of Arthropods Network

fossil insect COLLABORATIVE



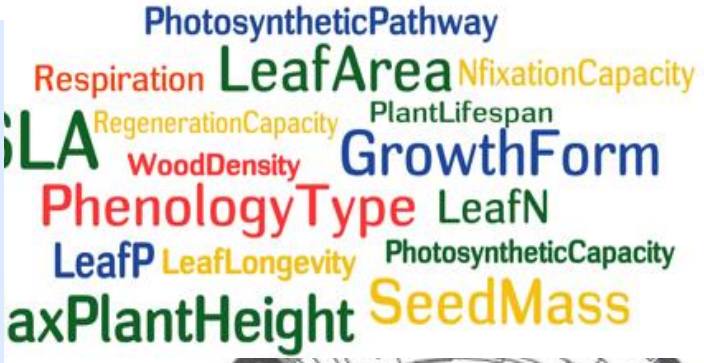
LEPNET: Lepidoptera of North America Network

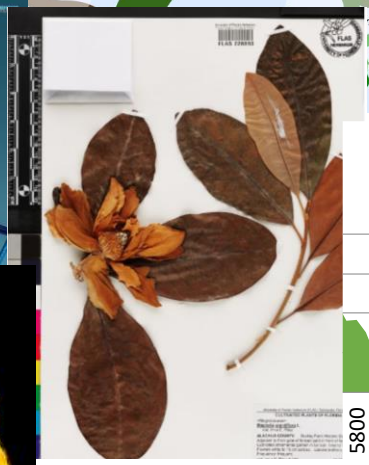
HOME | GENERAL INFORMATION | MEETINGS & TALKS | TECHNIQUES & PROTOCOLS | DATA | EDUCATION & OUTREACH | SCAN

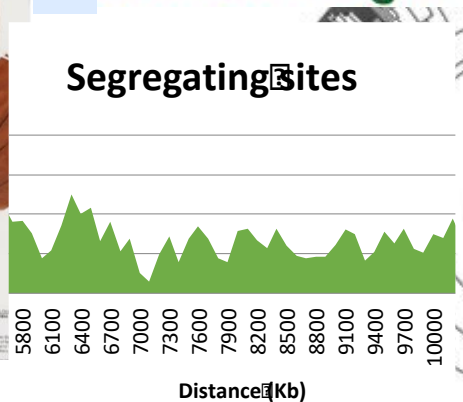
iDigBio
Integrated Digitized Biocollections

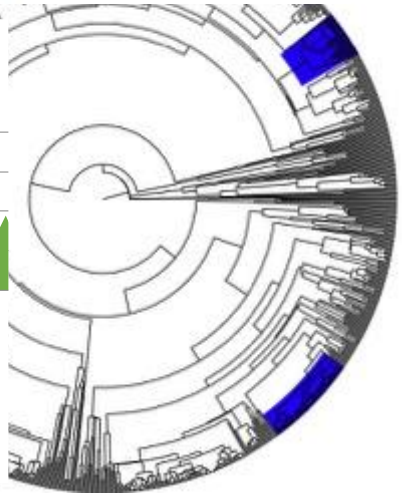
Linking Heterogeneous Data

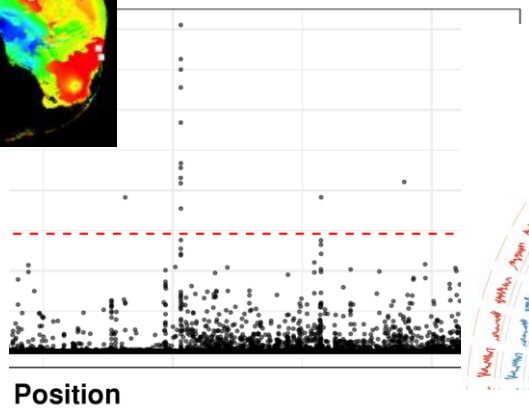
Genus	Species	Scientific name	Sample date	SLA dry mass (g)	leaf_area_cm2	SLA_cm2.g	LMA_g.m2	d13C(‰)
Chionanthu	virginicus	Chionanthus virginici	5/6/15	1.6	489.778	306.1	32.7	-33.17
Chionanthu	virginicus	Chionanthus virginici	5/6/15	2.36	462.018	195.8	51.1	-31.96
Chionanthu	virginicus	Chionanthus virginici	5/6/15	1.94	383.707	197.8	50.6	-32.69
Castanea	pumila	Castanea pumila	5/6/15	0.84				-32.91
Castanea	pumila	Castanea pumila	5/7/15	1.843	399.395	216.7	46.1	-32.08
Castanea	pumila	Castanea pumila	5/7/15	1.676	368.592	219.9	45.5	-30.83
Castanea	pumila	Castanea pumila	5/7/15	1.452	326.529	224.9	44.5	-31.42
Castanea	pumila	Castanea pumila	5/7/15	1.249	304.058	243.4	41.1	-30.76
Castanea	pumila	Castanea pumila	5/7/15	1.433	383.797	267.8	37.3	-33.07
Cartrema	americana	Cartrema americana	5/7/15	2.11	339.297	160.8	62.2	-30.79
Cartrema	americana	Cartrema americana	5/7/15	2.68	378.94	141.4	70.7	-31.52
Cartrema	americana	Cartrema americana	5/7/15	3.4	379.952	111.8	89.5	-30.74
Chionanthu	virginicus	Chionanthus virginici	5/7/15	2.866	393.473	137.3	72.8	-30.03
Diospyros	virginiana	Diospyros virginiana						
Quercus	laevis	Quercus laevis						
Cornus	florida	Cornus florida						
Cornus	florida	Cornus florida						
Cornus	florida	Cornus florida						
Cornus	florida	Cornus florida						

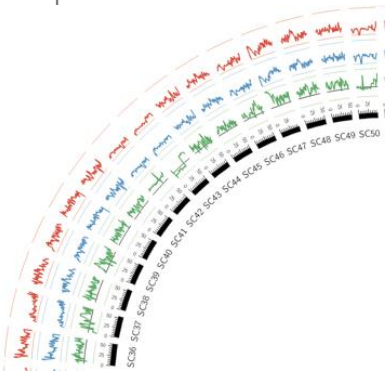










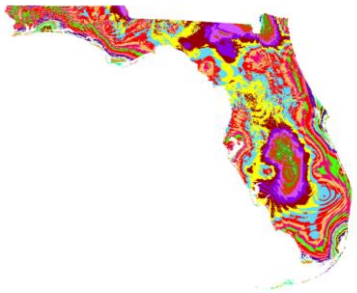


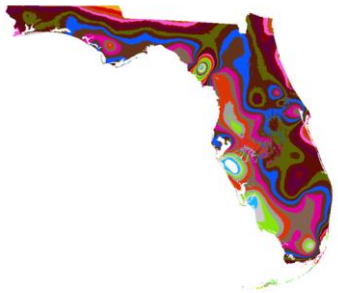
Modeling the Distribution of Florida Plants

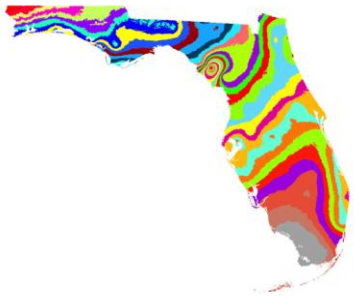
- Location information and environmental data
- Software to model the range of each species
- Project onto future climate conditions
- For Florida plants:
  - 1490 plant species (of 4100 species)
  - >511,000 georeferenced points (GPS)
  - Environmental features: temperature, precipitation, soil, etc.

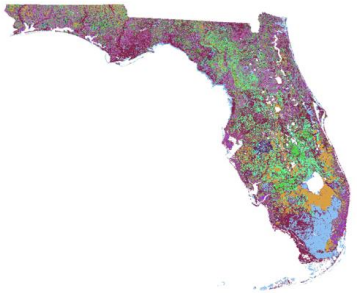
Charlotte Germain-Aubrey

Julie Allen



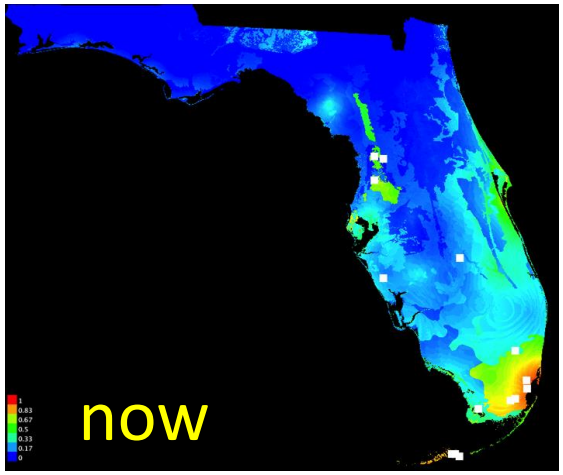


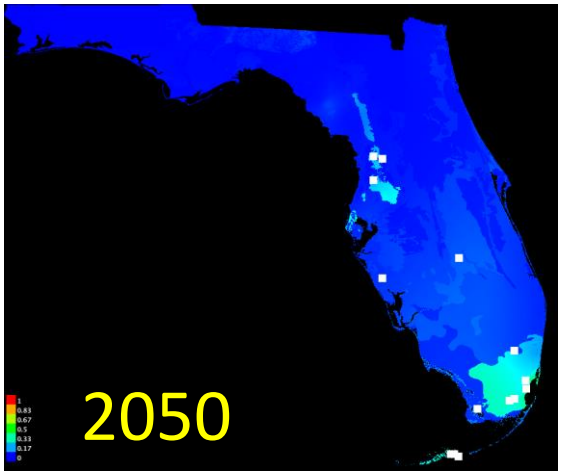




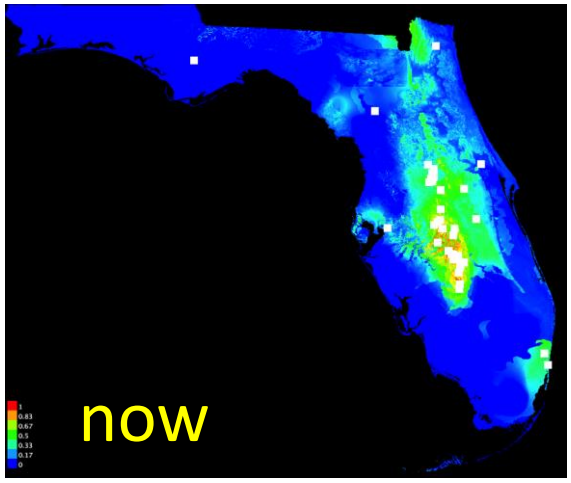
Responses to Climate Change: Winners & Losers

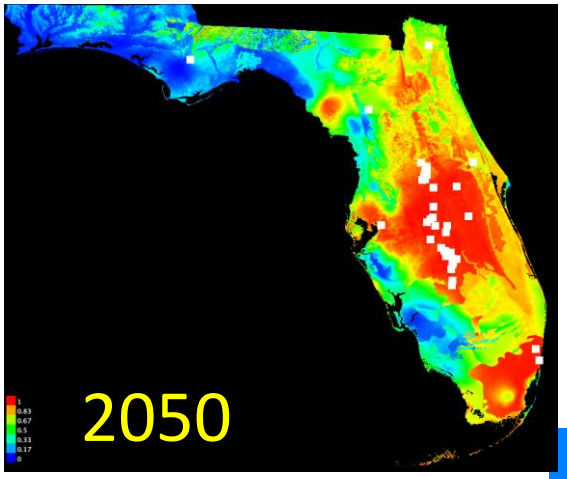
Abildgaardia ovata (flatspike sedge)



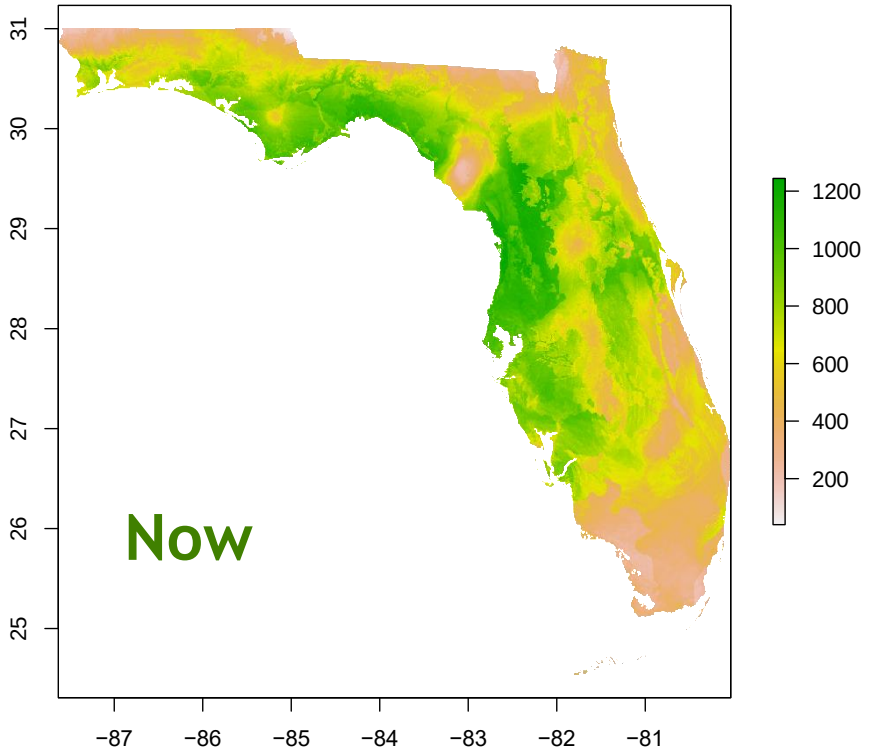


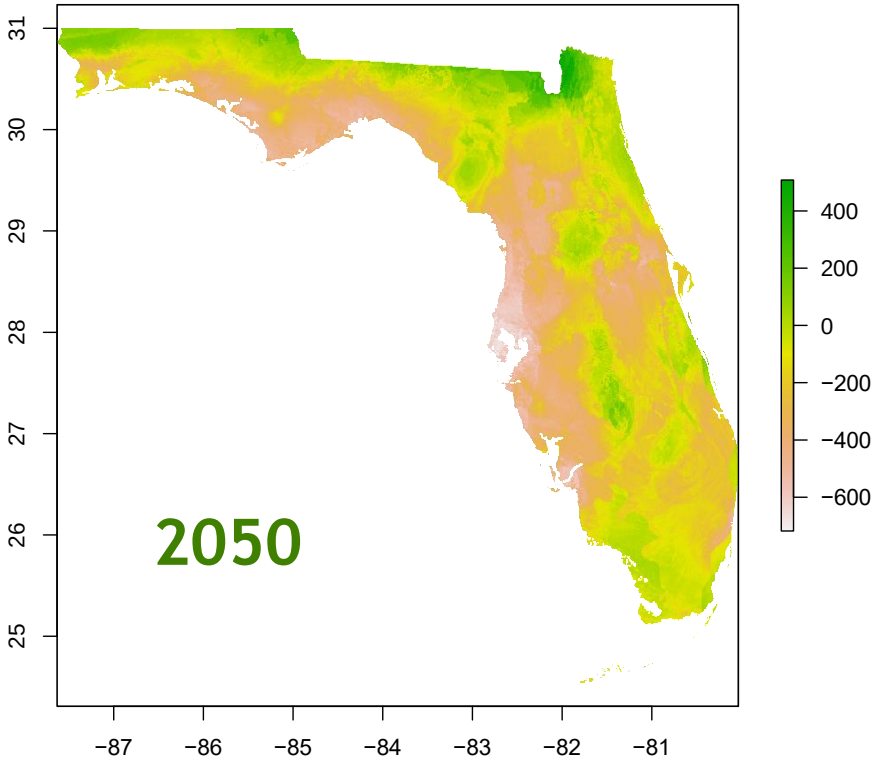
Prunus geniculata (scrub plum)





Florida Plant Diversity

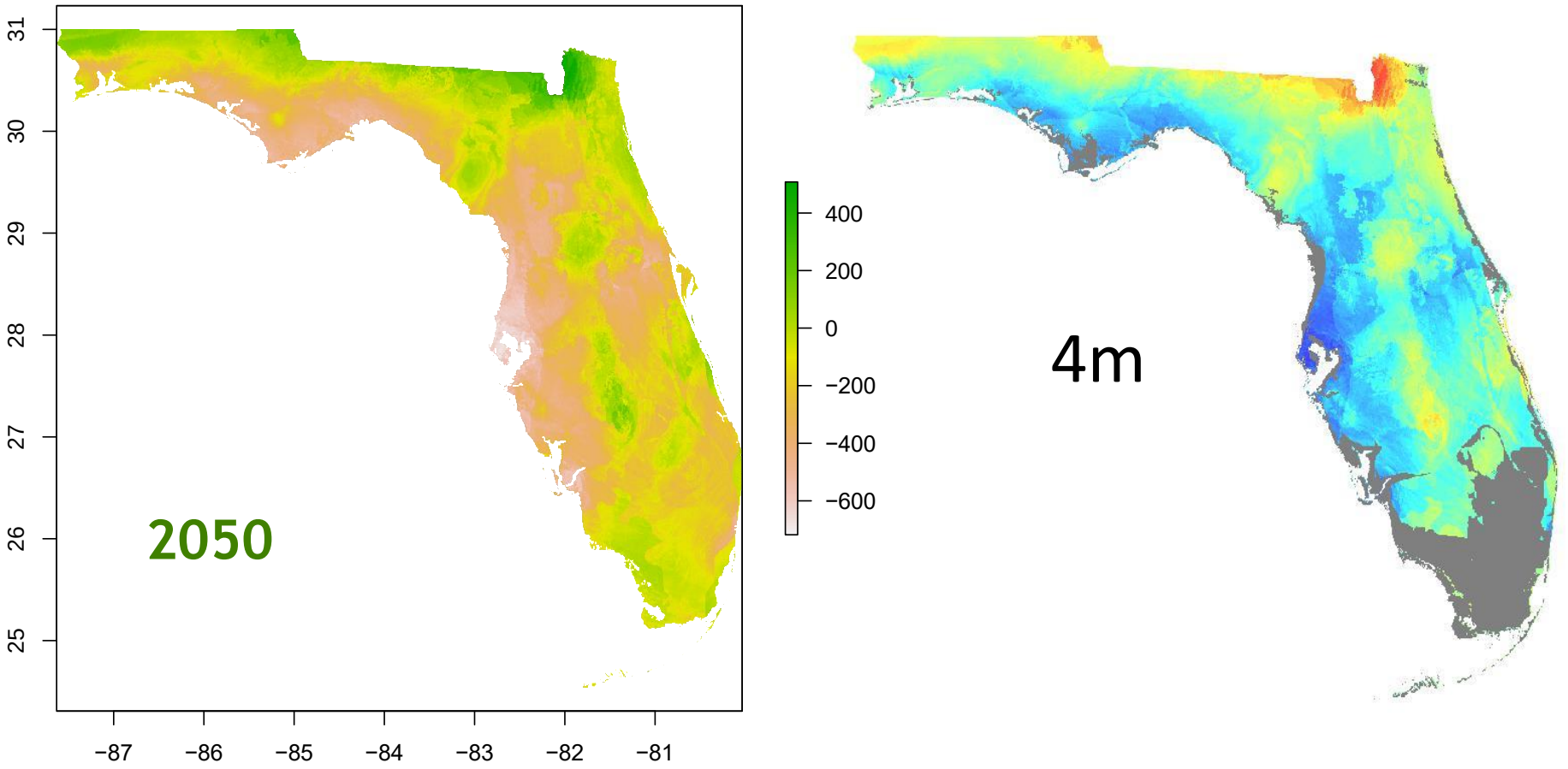




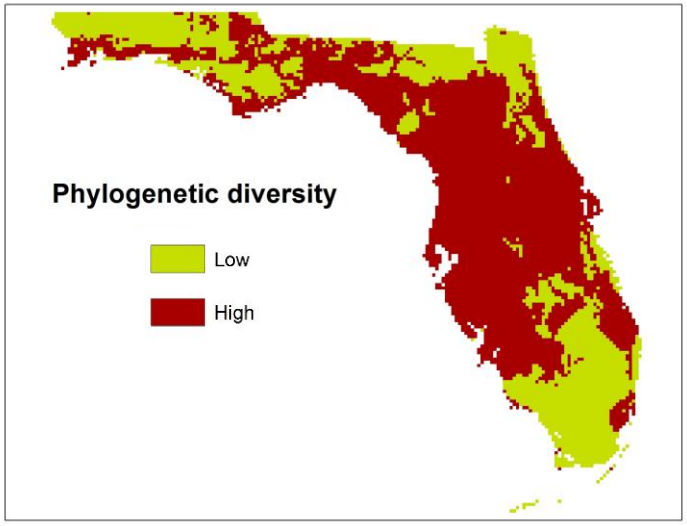
- High species diversity
- Low species diversity

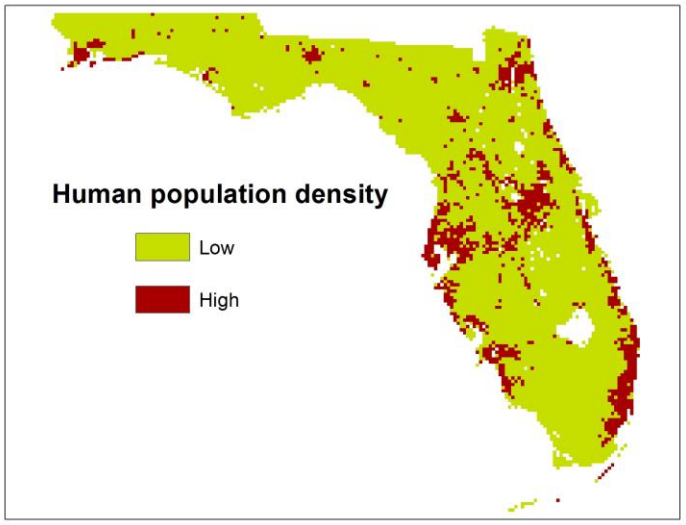
- Panhandle species: NORTH!
- Peninsula species: SOUTH!

Florida Plant Diversity: Sea Level Rise

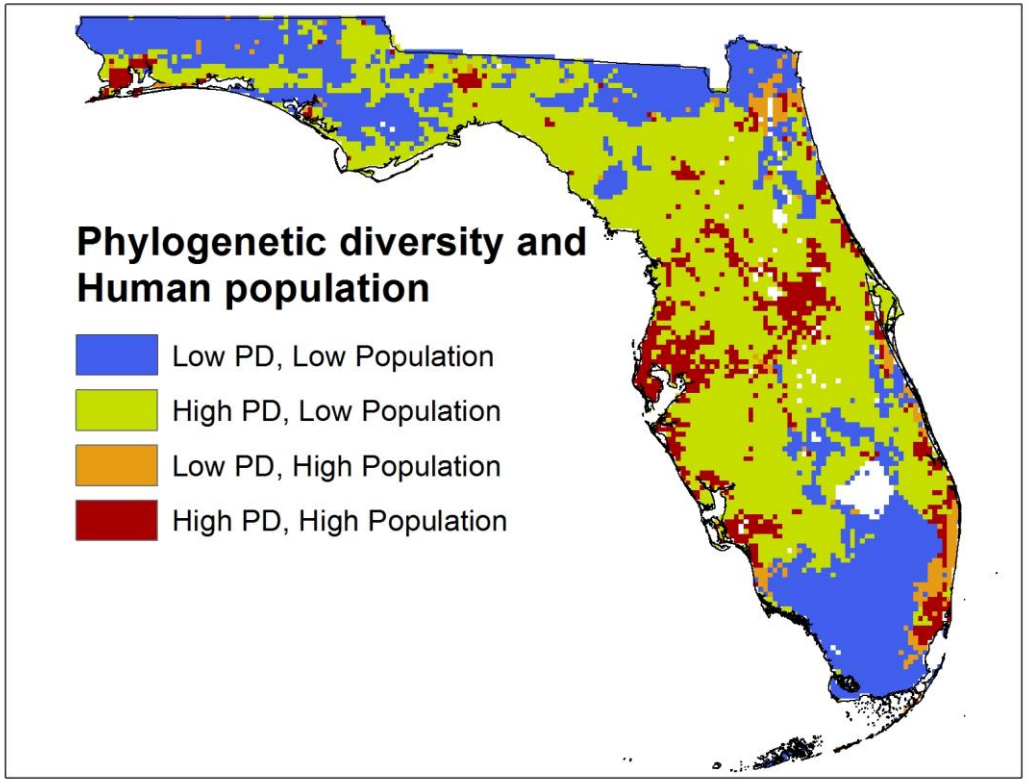


Florida Plant Diversity





Human population density



N. Barve

Phenological Asynchrony in a Changing Climate: Plant Bud Burst, Flowering, Fruiting, Insect Emergence, Bird Nesting & Migration, Mammal Nesting & Survival …

Bartomeus et al. 2011

http://pheno-mismatch.org/

http://budburst.org/phenology_whyphenology

Codistribution of plant bugs (Miridae; Insecta: Hemiptera) and their host plants

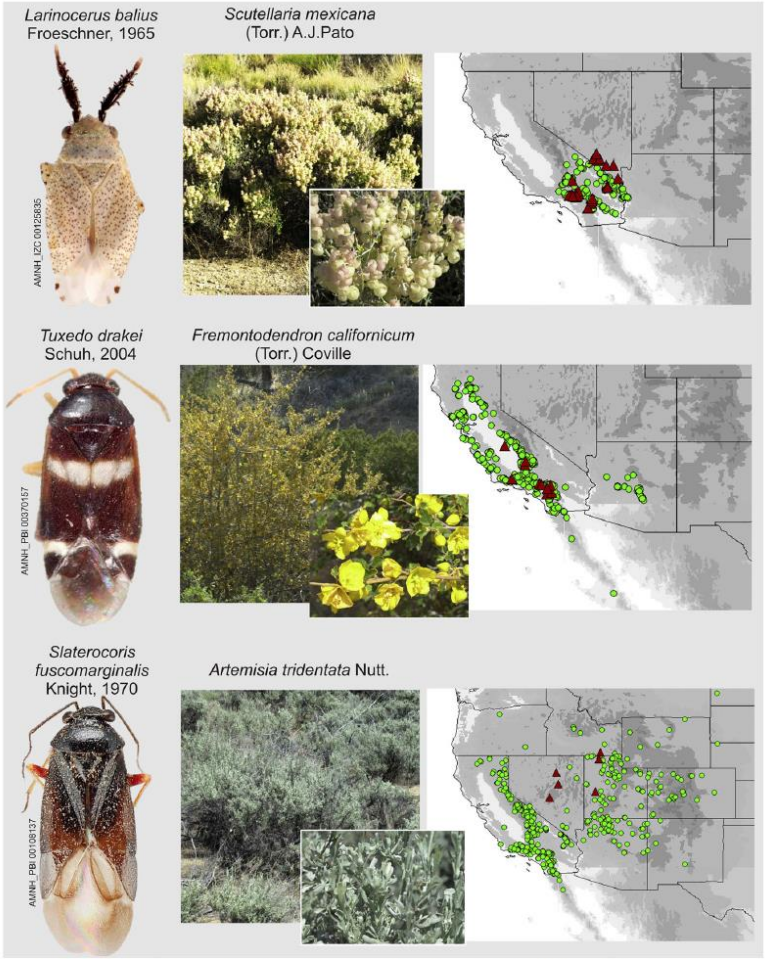


Fig. 1. Three species of Nearctic Miridae in the subfamilies Phylinae and Orthotylinae, their host plants and distribution ranges (plant bug species: maroon triangle; plant species: green circle), illustrating the relatively smaller distribution ranges of these plant bugs compared to their host plant.

Plant bugs: 1339 species, 61,016 records

Plants: 313 species, 196,012 records

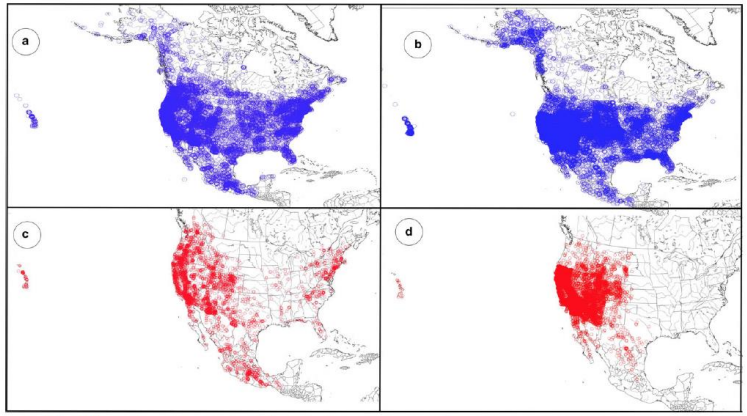
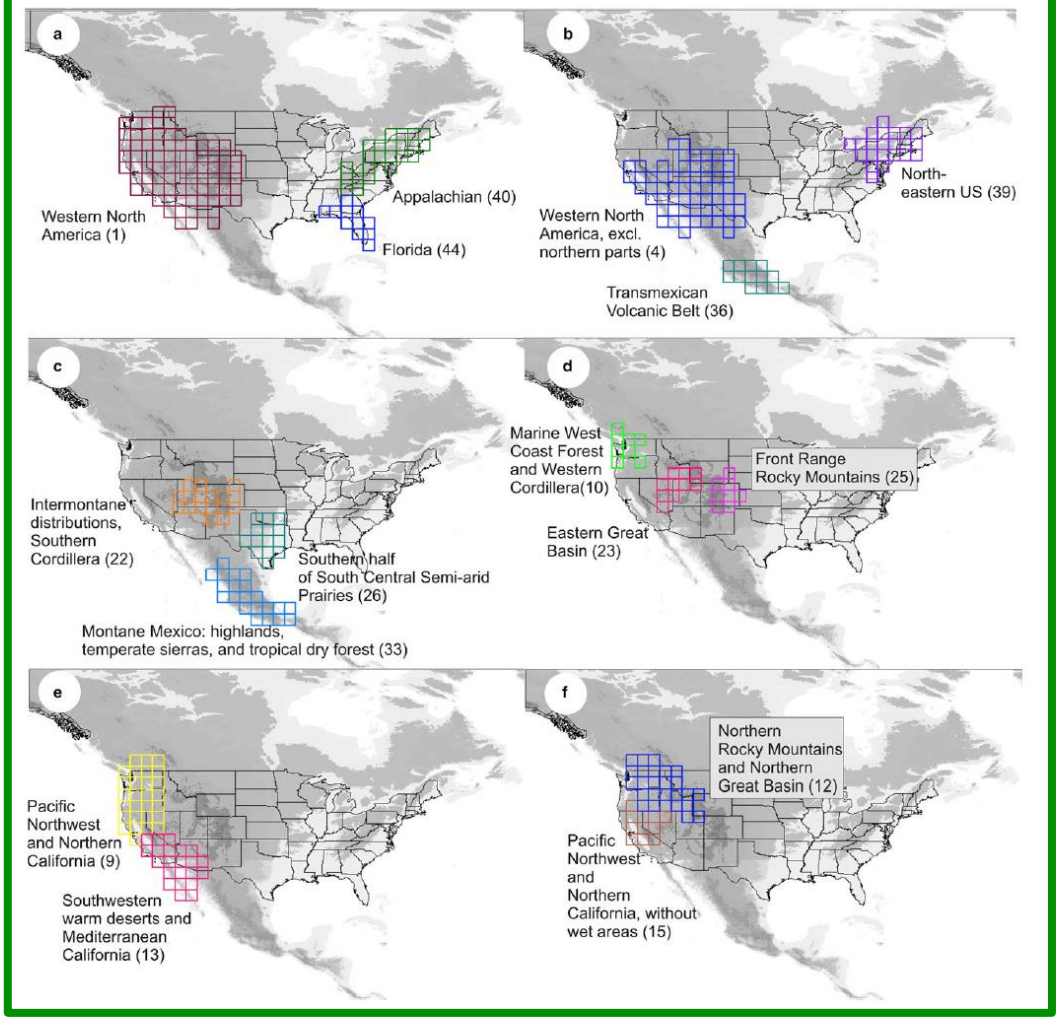


Fig. 2. Maps illustrating all unique localities included in the analysis, documenting dense sampling and localities of scoring species: (a) all unique localities for the default Miridae data set; (b) all unique localities for the plant data set; (c) unique localities of scoring plant bug species; and (d) unique localities of scoring plant species.

Weirauch et al. 2017 *Cladistics*

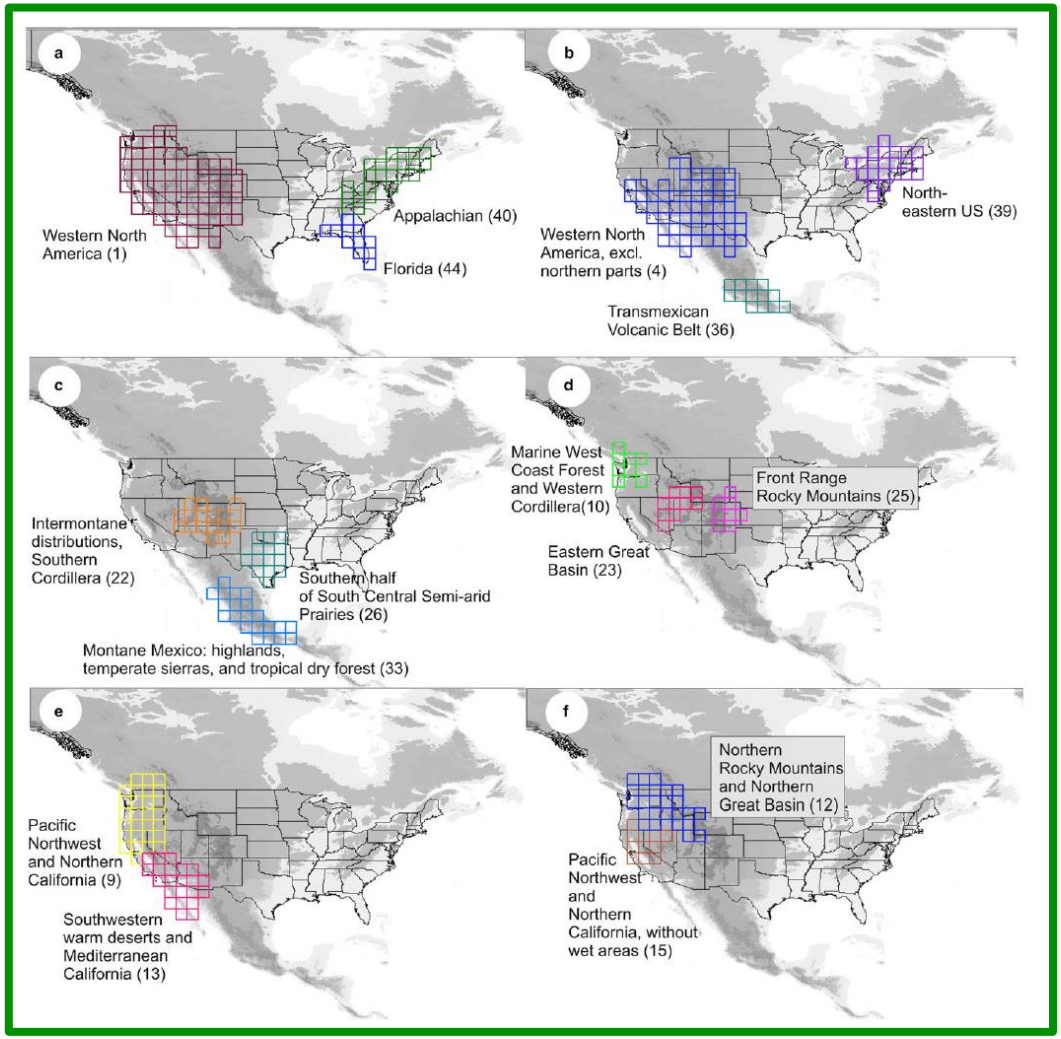
Codistribution of plant bugs (Miridae; Insecta: Hemiptera) and their host plants

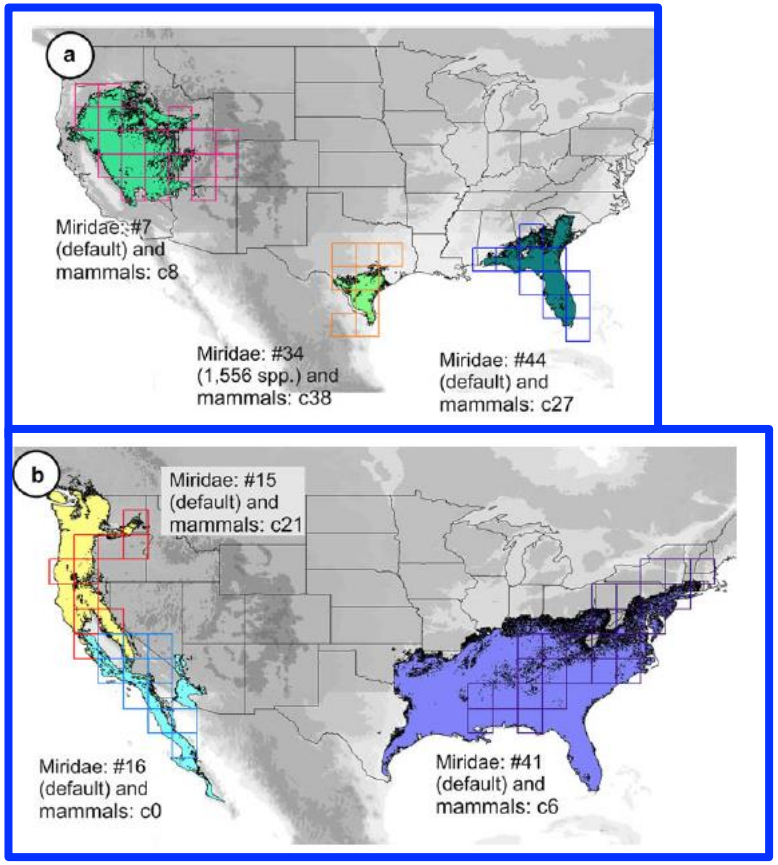


- Shown: 16 Areas of Endemism (AoE) for Miridae
- In all: 45 AoE for Miridae,10 AoE for plants
- Smaller AoE for Miridae than host plants, although generally congruent

Weirauch et al. 2017 *Cladistics*

Codistribution of plant bugs (Miridae; Insecta: Hemiptera) and their host plants





AoE congruent with mammals, too

Weirauch et al. 2017 *Cladistics*

Specimen Records in iDigBio

Papilio cresphontes

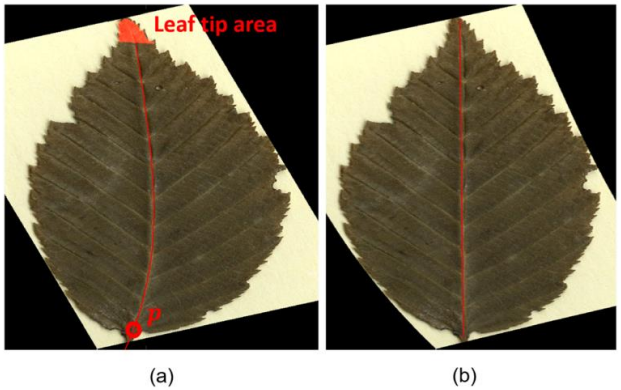
Traits from Labels, Images, & Text

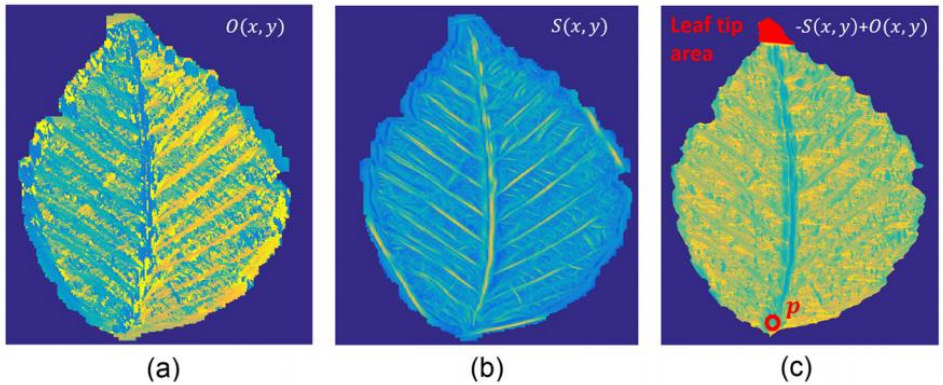
Machine Learning: Herbarium specimens

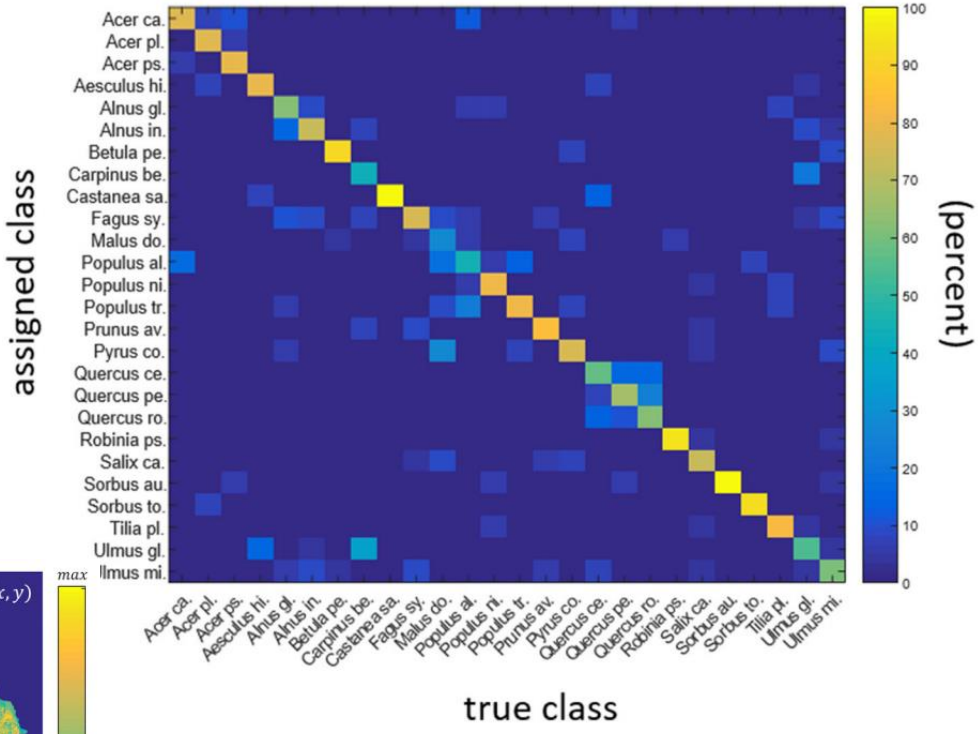
Classifying German trees to species

Leaf shape, venation

85% accuracy





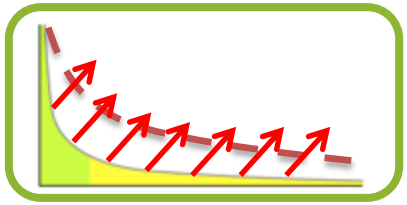


Unger et al. 2016

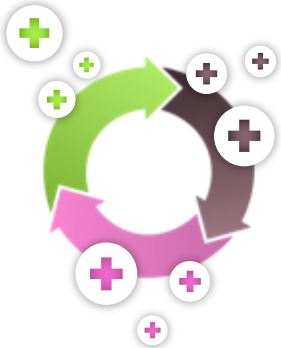
Pl@ntNet =

Free web & mobile app dedicated to plant identification and gathering of botanical observations; extending to specimens, phenology

Pl@ntNet is based on an innovative crowdsourcing workflow



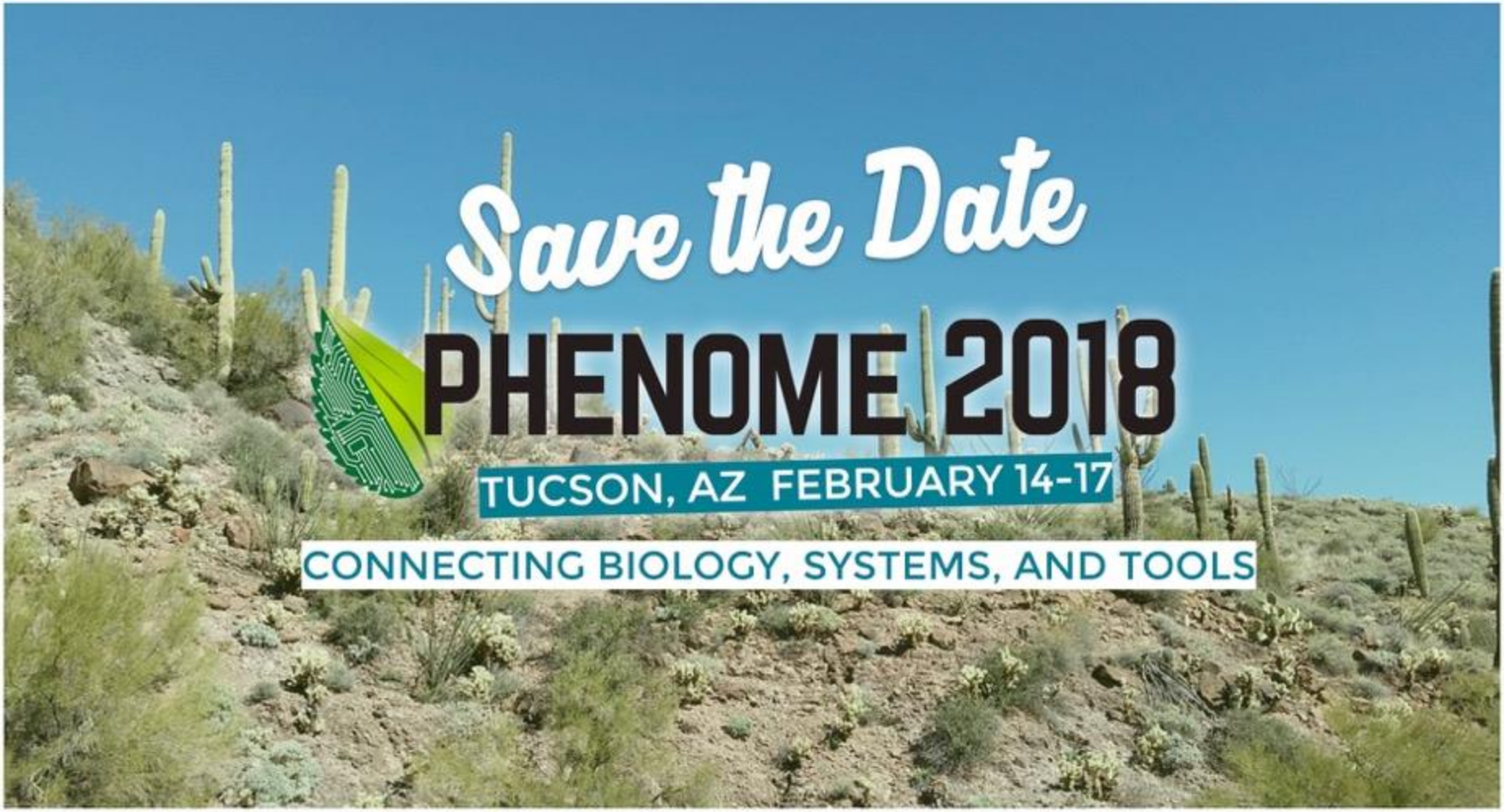
BIODIV. DATA



MACHINE LEARNING

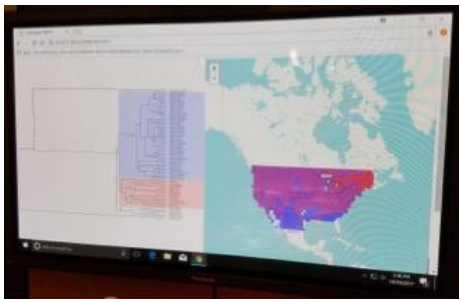
NATURE OBSERVERS

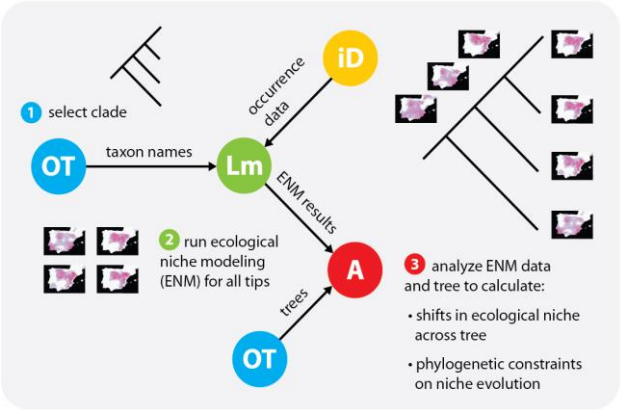
iDigBio
Integrated Digitized Biocollections



iDigBio: Enabling Biodiversity Research Through Tool Development & Collaboration

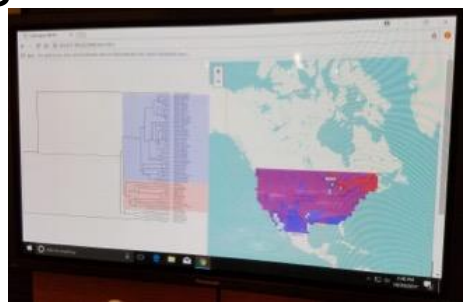
- GUODA: Global Unified Open Data Access
- Jupyter: Python & R notebook infrastructure
- FreshData: Notifications of new data records
- Effechecka: Taxonomic Checklist Generator
- Taxonomic Name Resolution Service (iPlant)
- Links, interfaces, & workflows - Open Tree of Life and Lifemapper
- USVH as a model – linking resources

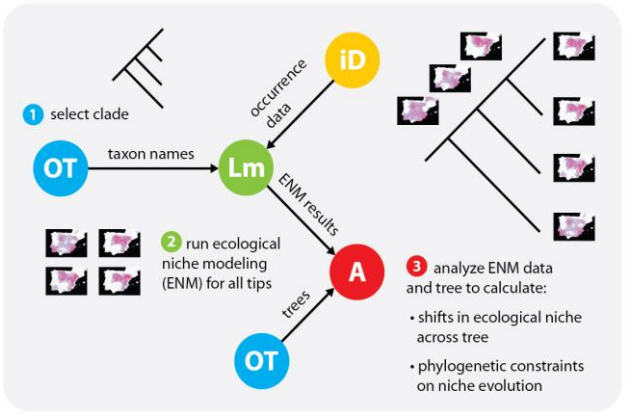




iDigBio: Enabling Biodiversity Research Through Tool Development & Collaboration

- GUODA: Global Unified Open Data Access
- Jupyter: Python & R notebook infrastructure
- FreshData: Notifications of new data records
- Effechecka: Taxonomic Checklist Generator
- Taxonomic Name Resolution Service (iPlant)
- Links, interfaces, & workflows - Open Tree of Life and Lifemapper
- USVH as a model – linking resources
- Training: Data & Software Carpentries, digitization, software use





iDigBio
Integrated Digitized Biocollections

Thank you!

idigbio.org/wiki

facebook.com/iDigBio

twitter.com/iDigBio

vimeo.com/iDigBio

idigbio.org/rss-feed.xml

idigbio.org/events-calendar/export.ics

iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program. 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.