



# BiotaPhy: An Overview

Doug Soltis  
University of Florida



**iDigBio**  
Integrated Digitized Biocollections



**BiotaPhy**



# Linking Heterogeneous Data: Connecting Specimens, Trees, Tools



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

J. Beach, D. Soltis, P. Soltis, J. Fortes,  
S. Smith, R. Folk

## RESOURCES:



### Lifemapper

- ecological niche modeling
- biodiversity and range analysis
- visualization



### Analytical Tools

- evolutionary models
- comparative methods
- visualization



### Open Tree of Life

- phylogenies
- taxonomy / names
- visualization



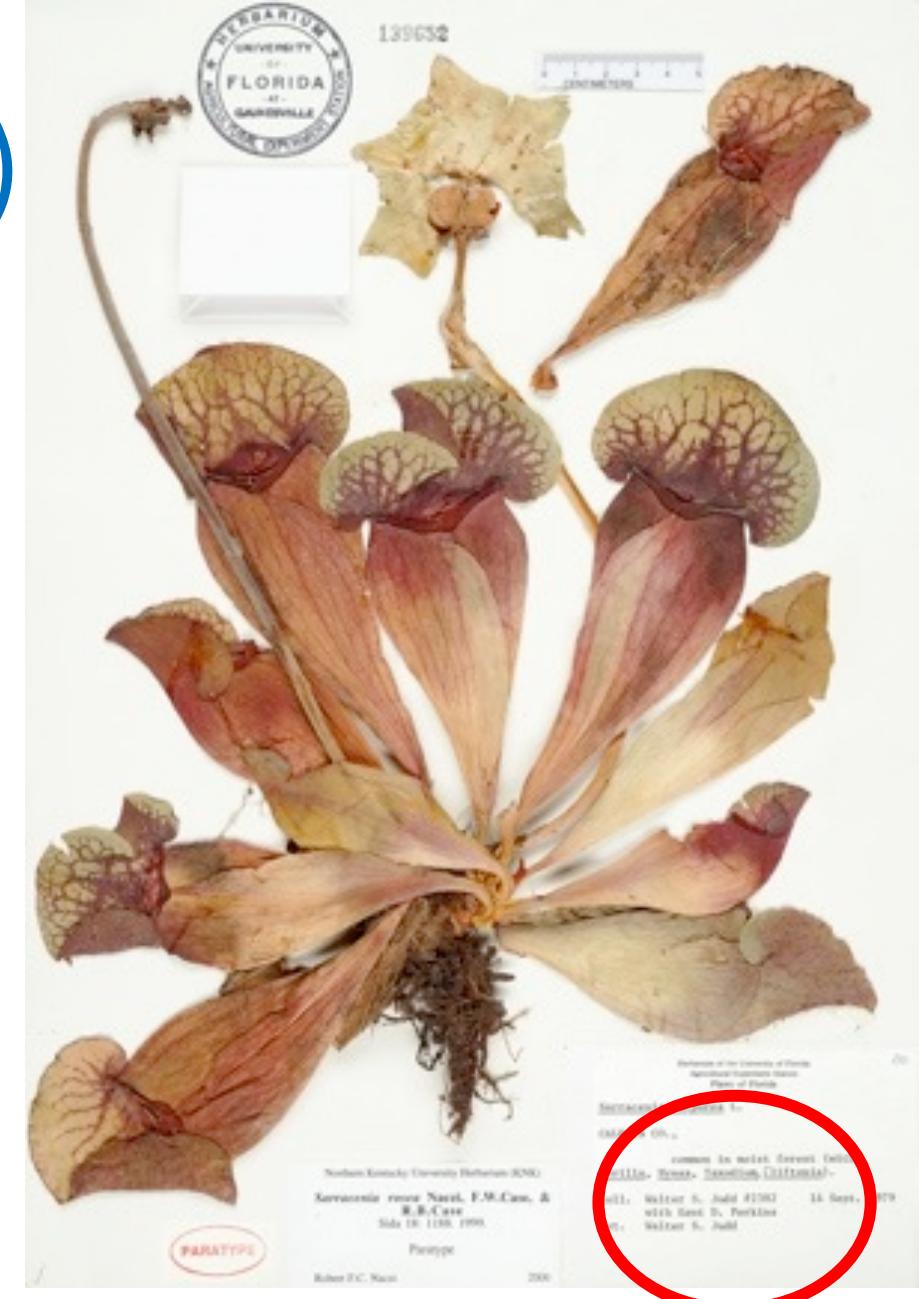
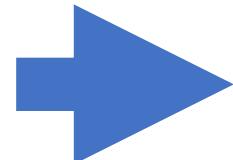
### iDigBio

- trait data
- specimen data / images
- fossil data / images

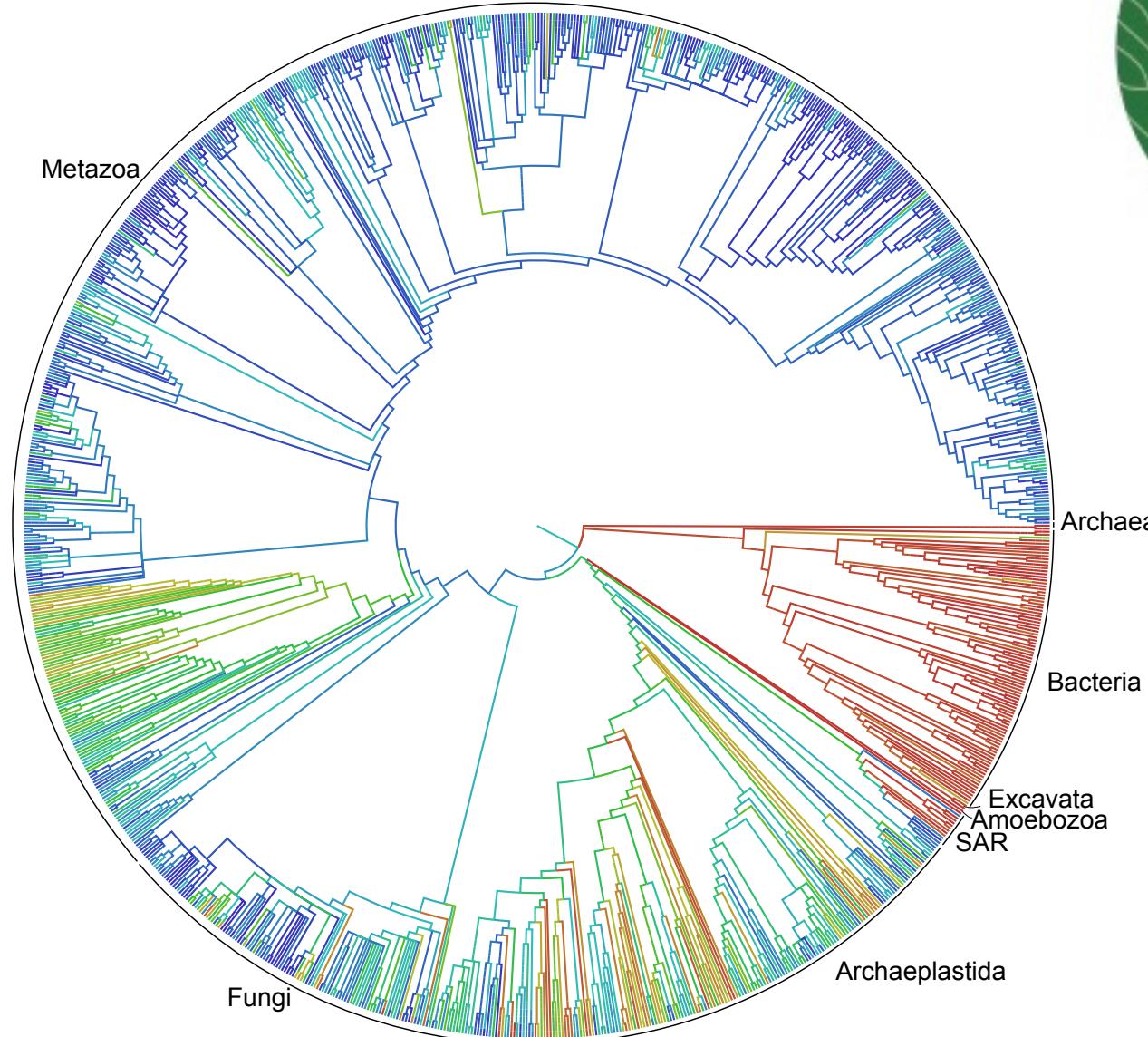


# Label Data (Occurrences)

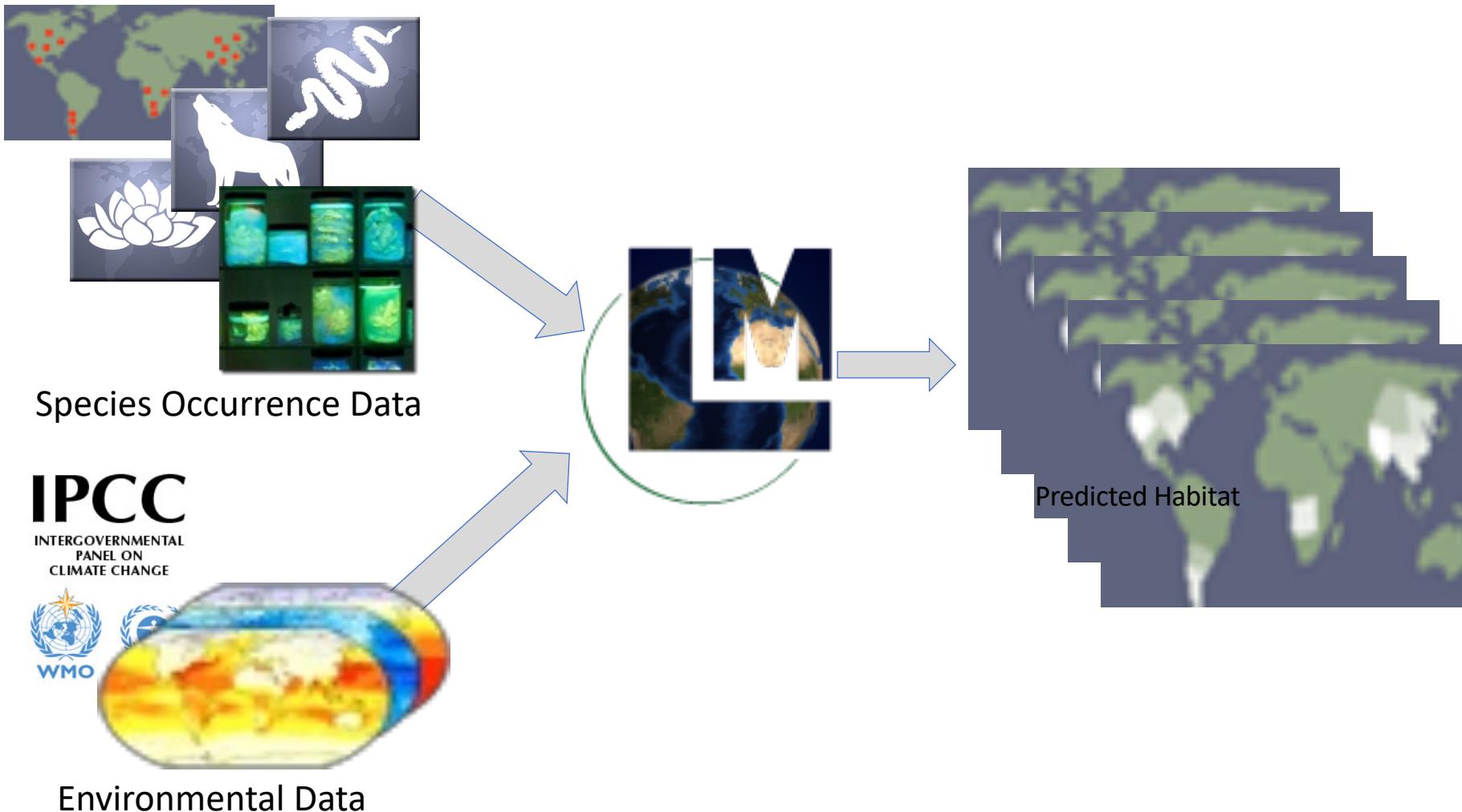
- Scientific name
- Date
- Collector
- Location – state, county, specific site, GPS coordinates
- Associated species



# Trees



# Lifemapper: Species Distribution Modeling



# A Project in Need of a Logo

Connectivity



# Connecting Specimens, Trees, Tools

## 5 Possible Workflows

### RESOURCES:



**Lm** Lifemapper  
 • ecological niche modeling  
 • biodiversity and range analysis  
 • visualization



**A** Analytical Tools  
 • evolutionary models  
 • comparative methods  
 • visualization

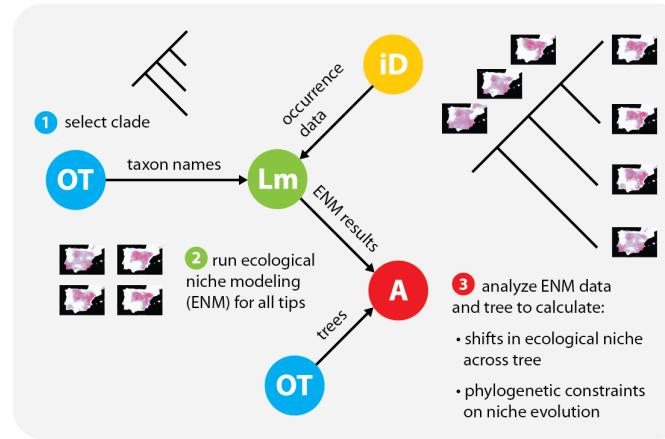


**OT** Open Tree of Life  
 • phylogenies  
 • taxonomy / names  
 • visualization



**iD** iDigBio  
 • trait data  
 • specimen data / images  
 • fossil data / images

### EXAMPLE WORKFLOWS:



### RESOURCES:



**Lm** Lifemapper  
 • ecological niche modeling  
 • biodiversity and range analysis  
 • visualization



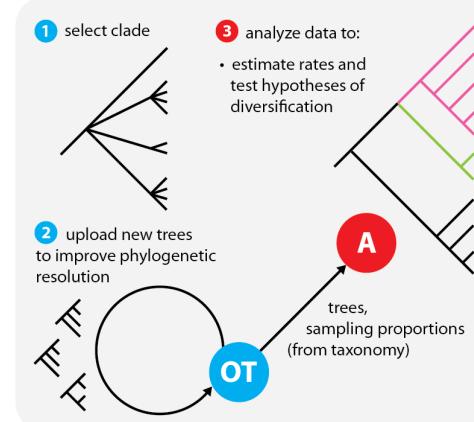
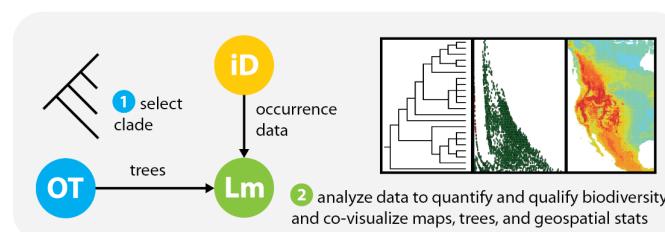
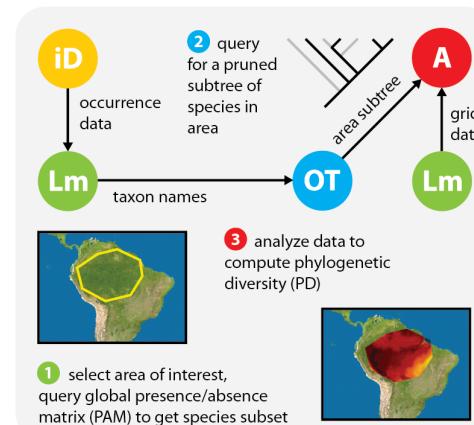
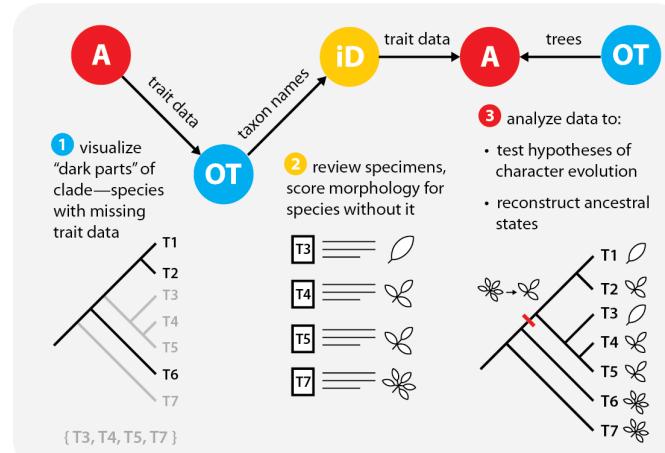
**A** Arbor  
 • evolutionary models  
 • comparative methods  
 • visualization



**OT** Open Tree of Life  
 • phylogenies  
 • taxonomy / names  
 • visualization



**iD** iDigBio  
 • trait data  
 • specimen data / images  
 • fossil data / images



# Ancestral Niche

## Connecting Trees, Specimens, Tools

### RESOURCES:



Lifemapper  
• ecological niche modeling  
• biodiversity and range analysis  
• visualization



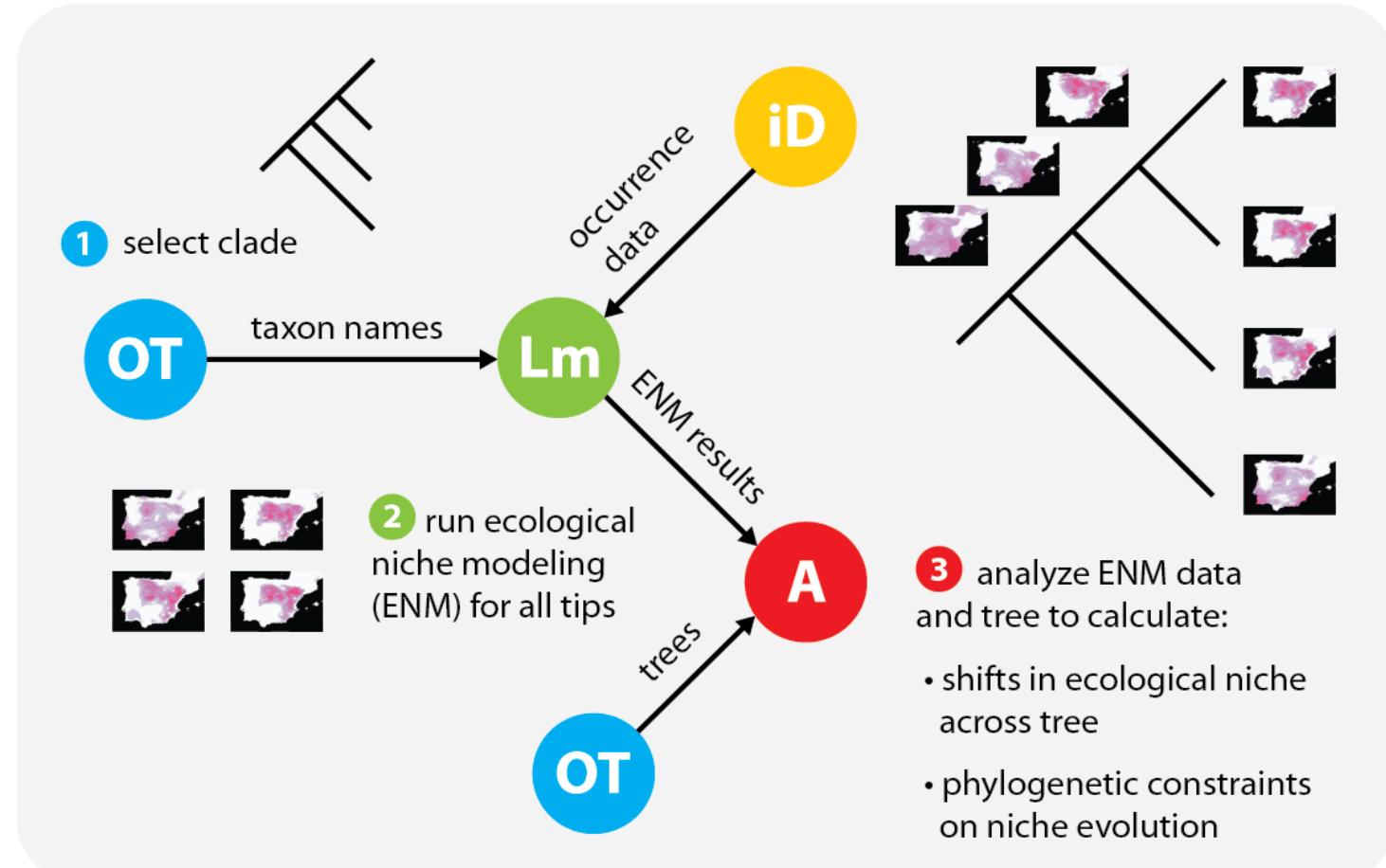
Analytical Tools  
• evolutionary models  
• comparative methods  
• visualization



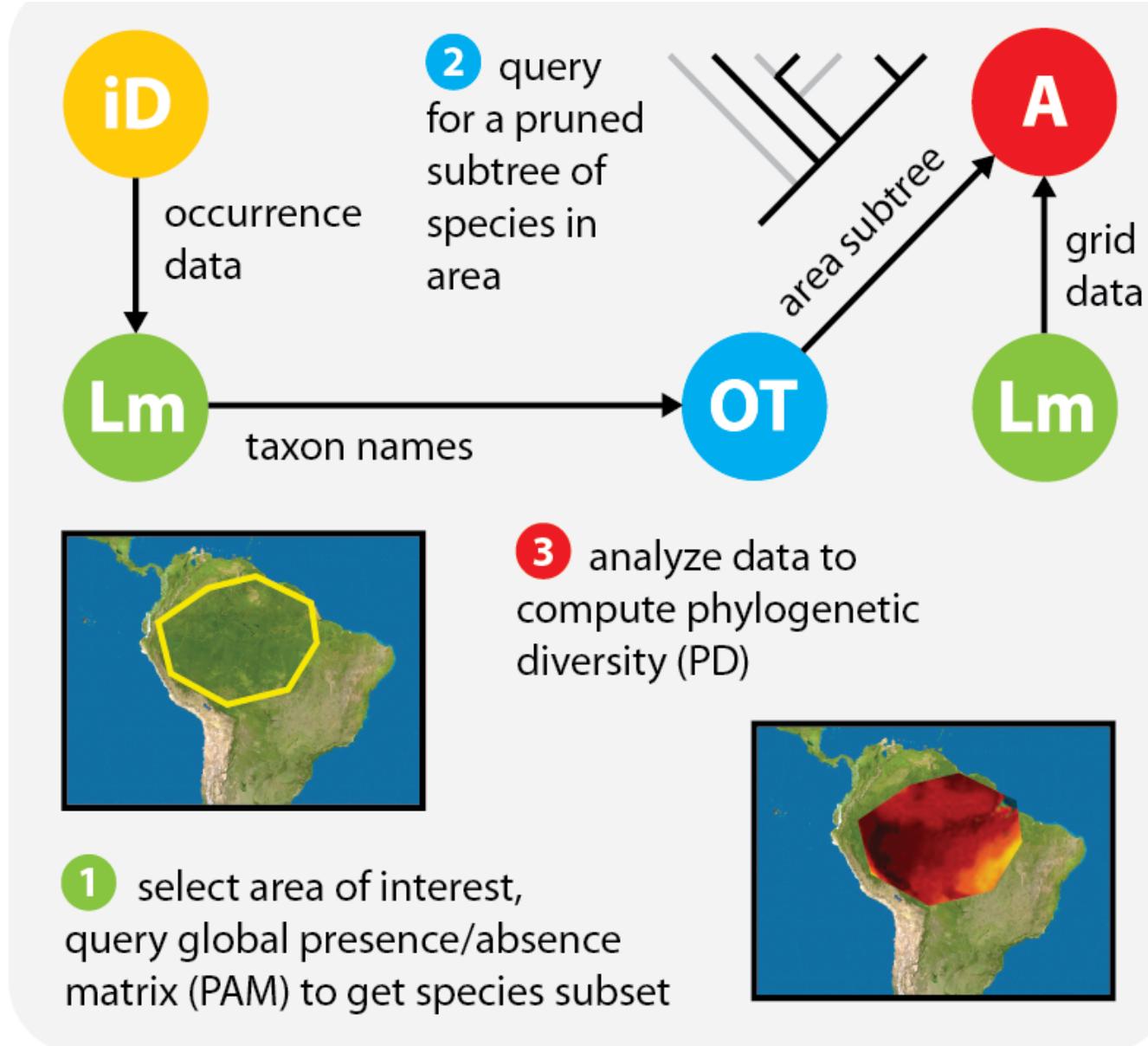
Open Tree of Life  
• phylogenies  
• taxonomy / names  
• visualization



iDigBio  
• trait data  
• specimen data / images  
• fossil data / images



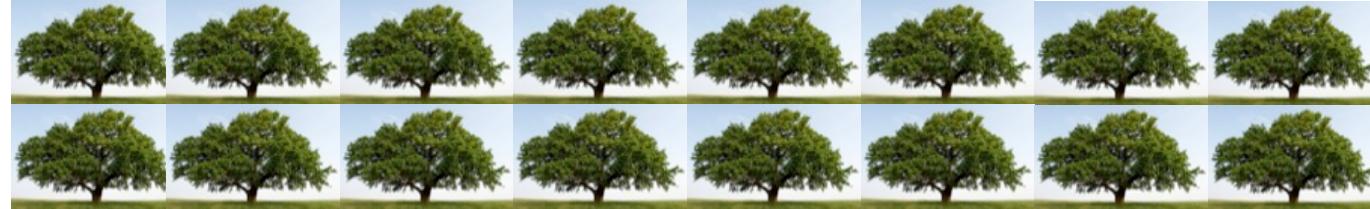
# Spatial Phylogenetics



# Phylogenetic Diversity (PD):

*How much of the Tree of Life is present in an area?*

Oaks



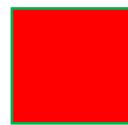
Vs.



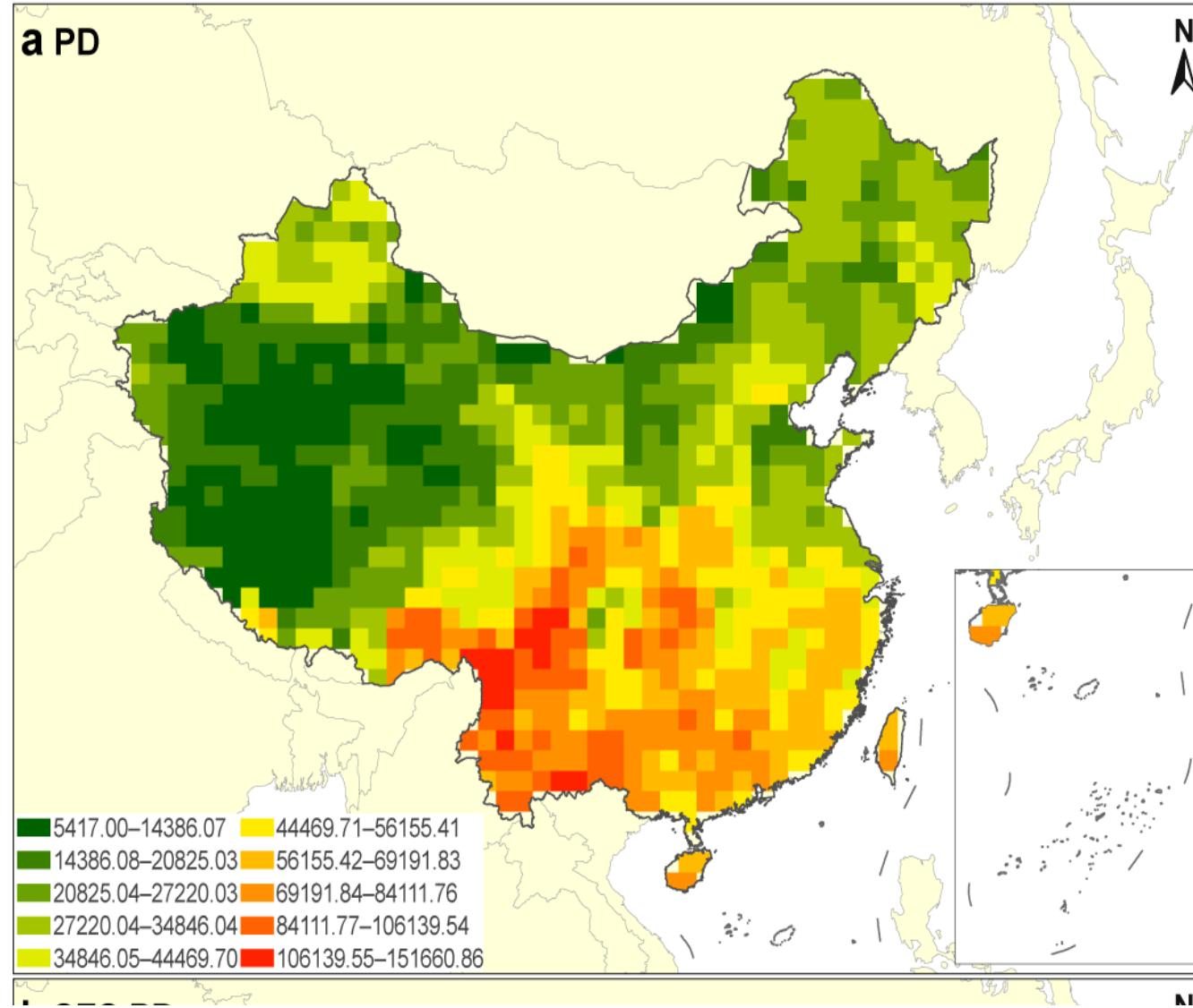
# PD: Genera of China Angiosperms



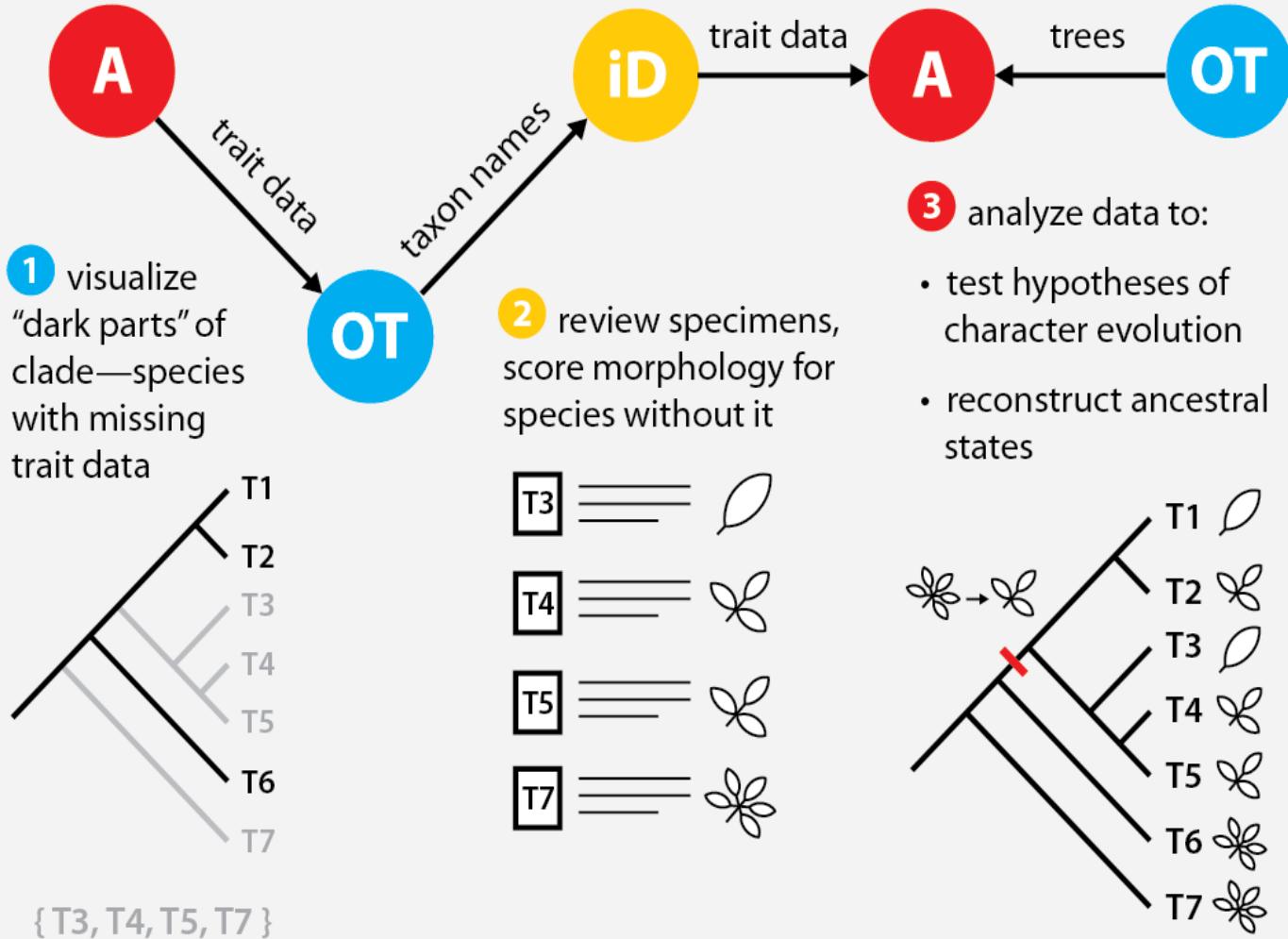
Low PD



High PD



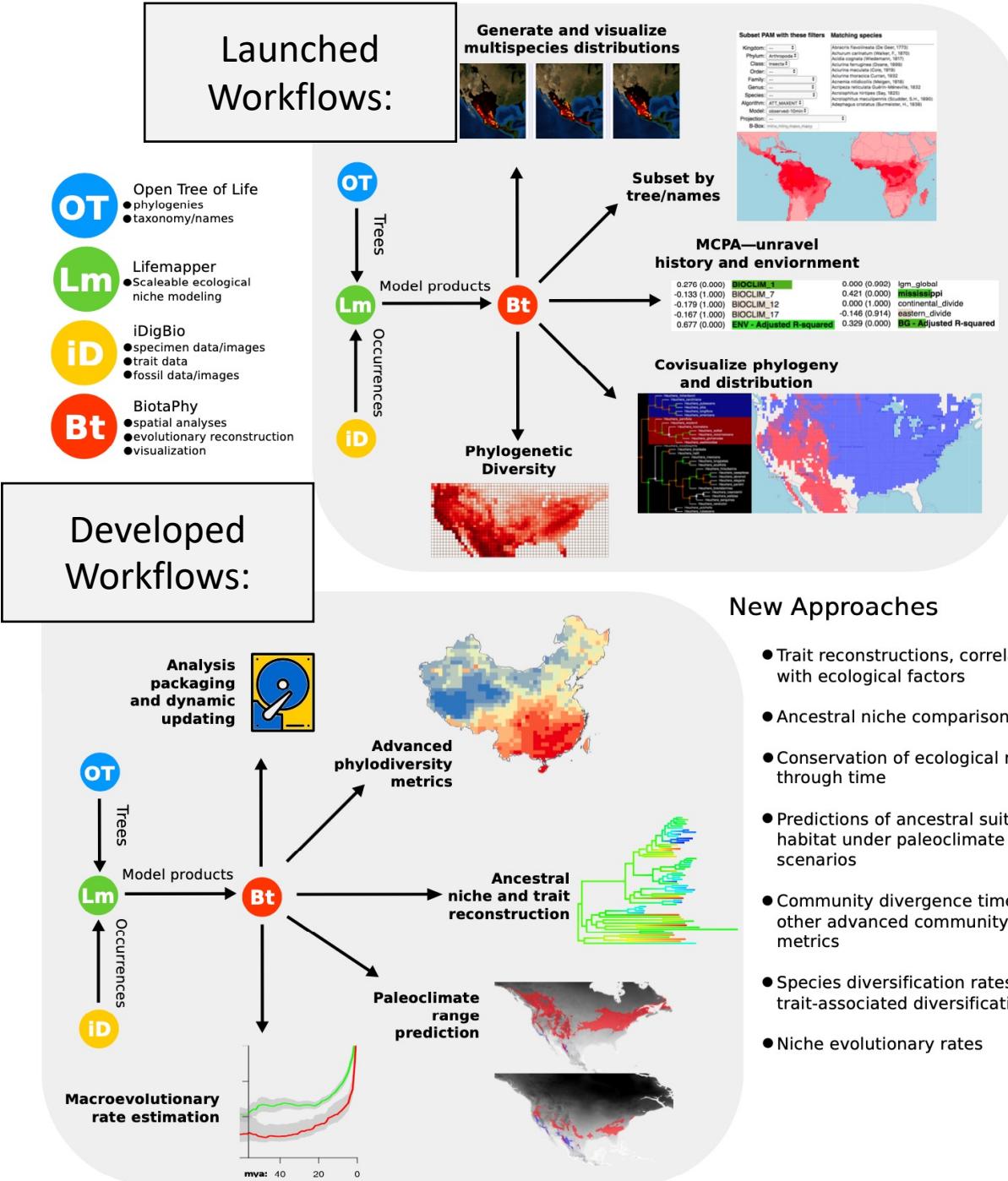
# Trait Evolution



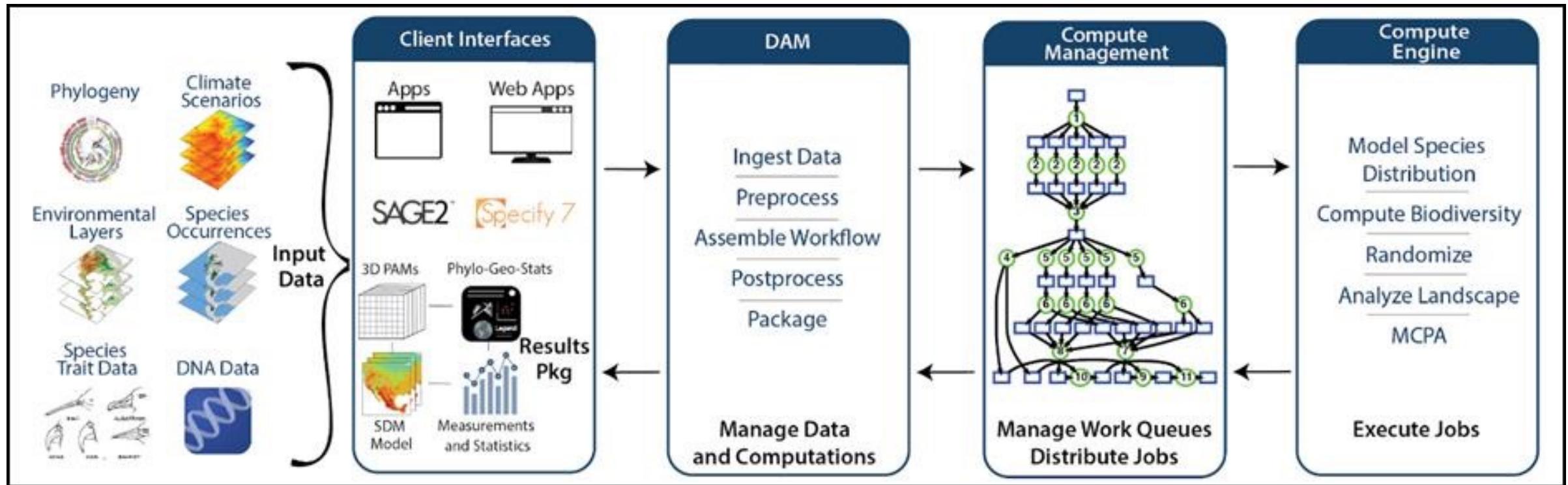
# BiotaPhy 2: Goals

- BiotaPhy Platform: community biodiversity gateway for data- intensive science
- Increase accessibility and use of Open Tree of Life, iDigBio, and Lifemapper
- A sustainable, long-term, biodiversity community infrastructure
- Workshops

# BiotaPhy 2 Workflows



# BiotaPhy: Architecture

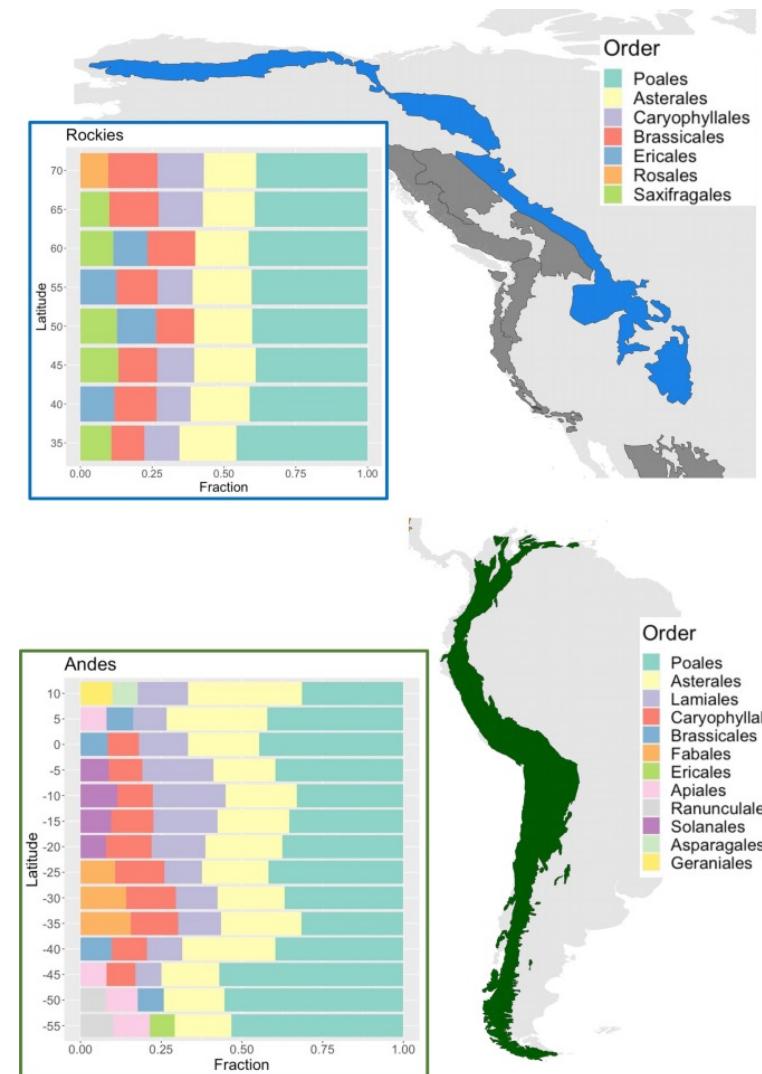
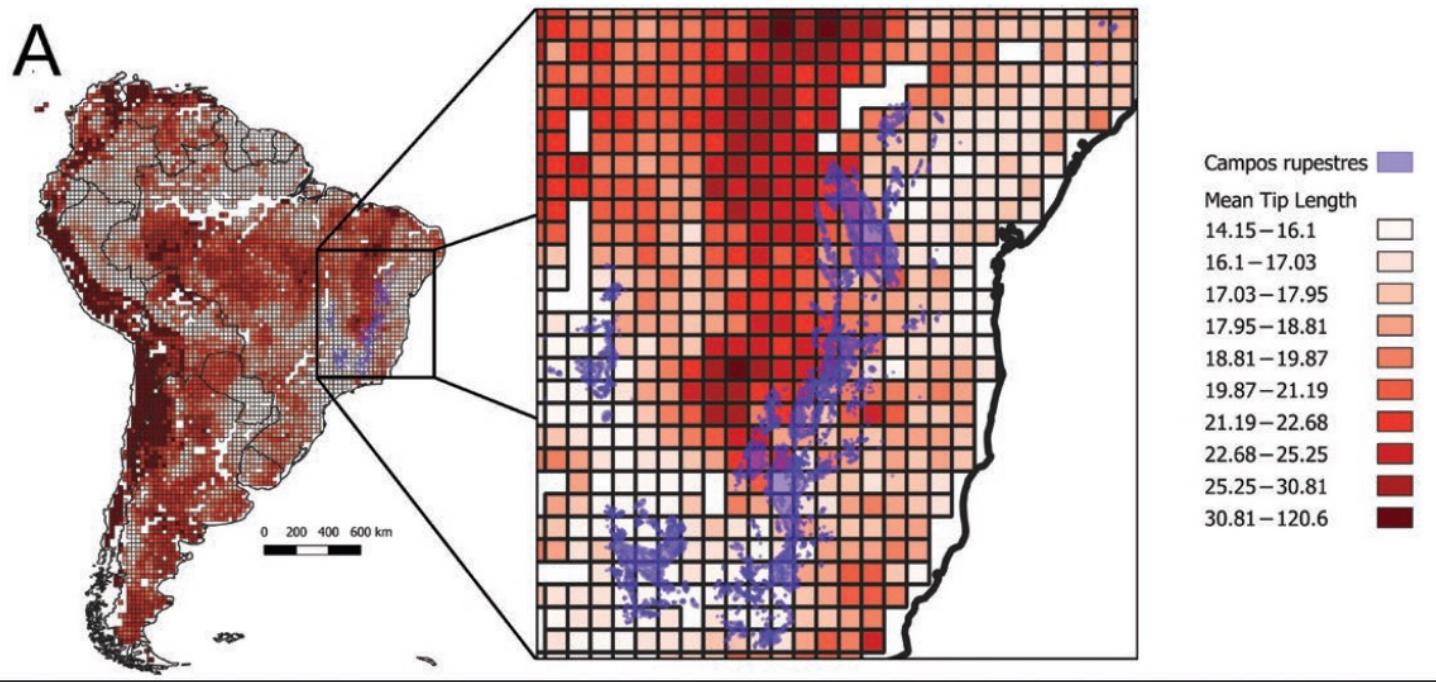




# BiotaPhy: Big Project Applications

CJ Grady

SEED PLANT COMMUNITY AGE AND OCBILS 15



# USING BIOTAPHY TO INVESTIGATE BIODIVERSITY OF ALPINE ECOSYSTEMS

*Hector Fox*  
*Botany Workshops*  
*Summer 2021*



Figueroa et al. 2022. Contrasting patterns of phylogenetic diversity and alpine specialization across the alpine flora of the American mountain range system. *Alpine Botany* **132**, 107–122  
<https://doi.org/10.1007/s00035-021-00261-y>

De Souza Cortez et al. 2020. Is the age of plant communities predicted by the age, stability and soil composition of the underlying landscapes? An investigation of OCBILs. *Biological Journal of the Linnean Society* 133: 297-316.



# USING BIOTAPHY TO INVESTIGATE PLANT COMMUNITIES IN OCBILs

**Maria Beatriz Cortez, Ryan Folk, CJ Grady, Jonathan  
Spoelhof, Stephen Smith, Douglas Soltis and Pamela Soltis.**

Old Climatically stable Infertile Landscapes

# Correlating chromosome and climate evolution



Jonathan Spoelhof



# **Webinar Series**

## **Data Use Skills**

**Featuring Data from Natural History Collections**

**September 21-November 30, 2022**

<https://www.idigbio.org/content/biotaphy-2022-webinar-series>

# iDigBio:

<https://www.idigbio.org/content/biotaphy-2022-webinar-series>

# iDigBio.org



The screenshot shows the iDigBio website with a dark header. The header includes the iDigBio logo (Integrated Digitized Biocollections) and navigation links: About iDigBio, Research, Technical Information, Education, Log In, and Sign Up. A search bar is also present. Below the header, there's a banner for the "BiotaPhy 2022 Webinar Series". The banner includes a link to add the event to a calendar, the date (Tue, 09/13/2022 - 9:03am), and the organizer (jvgoodwin). To the left, there are three sections for Researchers, Collections Staff, and Teachers & Students, each with a descriptive text and a colored circular arrow icon. To the right, there's a large "BiotaPhy" heading, a description of the project, and information about the webinar series schedule and registration.

**BiotaPhy 2022 Webinar Series**

Add this event to your calendar: [iCal](#)

Tue, 09/13/2022 - 9:03am -- jvgoodwin

**Researchers**  
Browse our specimen portal

**Collections Staff**  
Learn how your collection can benefit from our work

**Teachers & Students**  
Learning resources & opportunities to engage

**BiotaPhy**

Do you want to learn how to use occurrence data and available software to address questions in ecology and evolutionary biology but haven't had a chance to take a course and are overwhelmed by the options for self-teaching? Would you like to incorporate this sort of research into your classes but don't have time to create the materials and examples? If either of these applies, then join the BiotaPhy Project for an upcoming series of 10 webinars designed to take you from biological question to data acquisition and cleaning to analysis and interpretation!

BiotaPhy is a collaboration among iDigBio, LifeMapper, and the Open Tree of Life. Learn more about the project and [register for the webinar series here](#).

*Webinars will be held Wednesdays at 12:30 pm Eastern time, beginning September 21. All webinars will be recorded and made available, so you won't get behind if you need to miss a session. The schedule is:*

**Webinar 0:** Terms, Concepts, Data Formats – A Tutorial for Background  
[Click here for PDF](#)

**Webinar 1:** Introduction: Scope and Research Potential for Multidisciplinary Biodiversity Modeling and Analysis  
**Date:** 09/21/2022

**Webinar 0:** Terms, Concepts, Data Formats – A Tutorial for Background  
[Click here for PDF](#)

**Webinar 1:** Introduction: Scope and Research Potential for Multidisciplinary Biodiversity Modeling and Analysis  
**Date:** 9/21/2022

[Recording 1](#) [Recording 2](#)  
[Click here for PDF](#)

**Webinar 2:** Resolving Nomenclature: Making Appropriate Taxonomic Choices  
**Date:** 09/28/2022  
[Recording](#)  
[Click here for PDF](#)

**Webinar 3:** Clean Your Dirty Data  
**Date:** 10/05/2022  
[Recording](#)  
[Click here for PDF](#)

**Webinar 4:** Georeferencing with GEOLocate  
**Date:** 10/12/2022  
[Recording](#)  
[Click here for PDF](#)

**Webinar 5:** Big Data Munging (a.k.a Splitting and Merging Occurrence Data by Taxa from Multiple Sources)

**Date:** 10/19/2022

[Recording](#)  
[Click here for PDF](#)

**Webinar 6:** Species Distribution Modeling 1  
**Date:** 10/26/2022  
[Recording](#)  
[Click here for PDF](#)

**Webinar 7:** Species Distribution Modeling 2  
**Date:** 11/02/2022  
[Recording](#)  
[Click here for PDF](#)

**Webinar 8:** Introducing Presence-Absence Matrices for Large Scale Analyses  
**Date:** 11/9/2022  
[Recording](#)  
[Click here for PDF](#)

**Webinar 9:** Phylogenetic Diversity: Integrating Phylogenies with Species and Biogeographic Data  
**Date:** 11/16/2022

# Webinar 0

**Terms, Concepts, Data Formats –  
A Tutorial for Background**

**Introduce the technical terminology that  
will be used throughout the webinars**

## Webinar 1

**Introduction: Scope and Research  
Potential for Multidisciplinary  
Biodiversity Modeling and Analysis**

**Introduce biological and computational  
workflow concepts for integrating:  
biological specimen (species occurrence)  
data, phylogenetic trees, and multi-species  
distribution models!**

## Webinar 2

# Resolving Nomenclature: Making Appropriate Taxonomic Choices

**Learn why and how to treat nomenclatural  
data that will be used in multiple  
biodiversity analysis**

## Webinar 3

# Clean Your Dirty Data

**Introduce the importance of cleaning  
data before processing biodiversity  
analyses and practice a few ways to do so**

## Webinar 4

# Georeferencing with GEOLocate

## **How to use locality information in specimen records to obtain georeferences (i.e., latitude and longitude)**

## Webinar 5

**Big Data Munging: Finding, Acquiring,  
and Preparing Species Occurrence Data  
and Tree Data**

**Learn how to find, manipulate, combine  
and use occurrence and tree data in  
biodiversity analyses**

## Webinar 6

# Species Distribution Models: What are they? How to Create One?

**Create Species Distribution Models using Maxent  
using occurrence data with the minimum number  
of points defined in the configuration file or the  
Rare Species Model algorithm for data without  
the required minimum number of points.**

## Webinar 7

# Species Distribution Models: Interpretation and Ancestral Niche Reconstruction

**Learn how to interpret species distribution  
models build using Maxent, and how to perform  
ancestral niche reconstruction**

## Webinar 8

# Introducing Presence-Absence Matrices for Large Scale Analyses

**Learn how to calculate basic diversity  
statistics and how to utilize a Presence and  
Absence Matrix (PAM) to optimize  
calculations involving multiple taxa**

## Webinar 9

# Phylogenetic diversity: Integrating Phylogenies with Species and Biogeographic Data

**Integrate phylogenetic trees and species  
distribution models to understand modern  
day diversity patterns using a species PAM.**