Webinar Series

Data Use Skills
Featuring Data from Natural History Collections

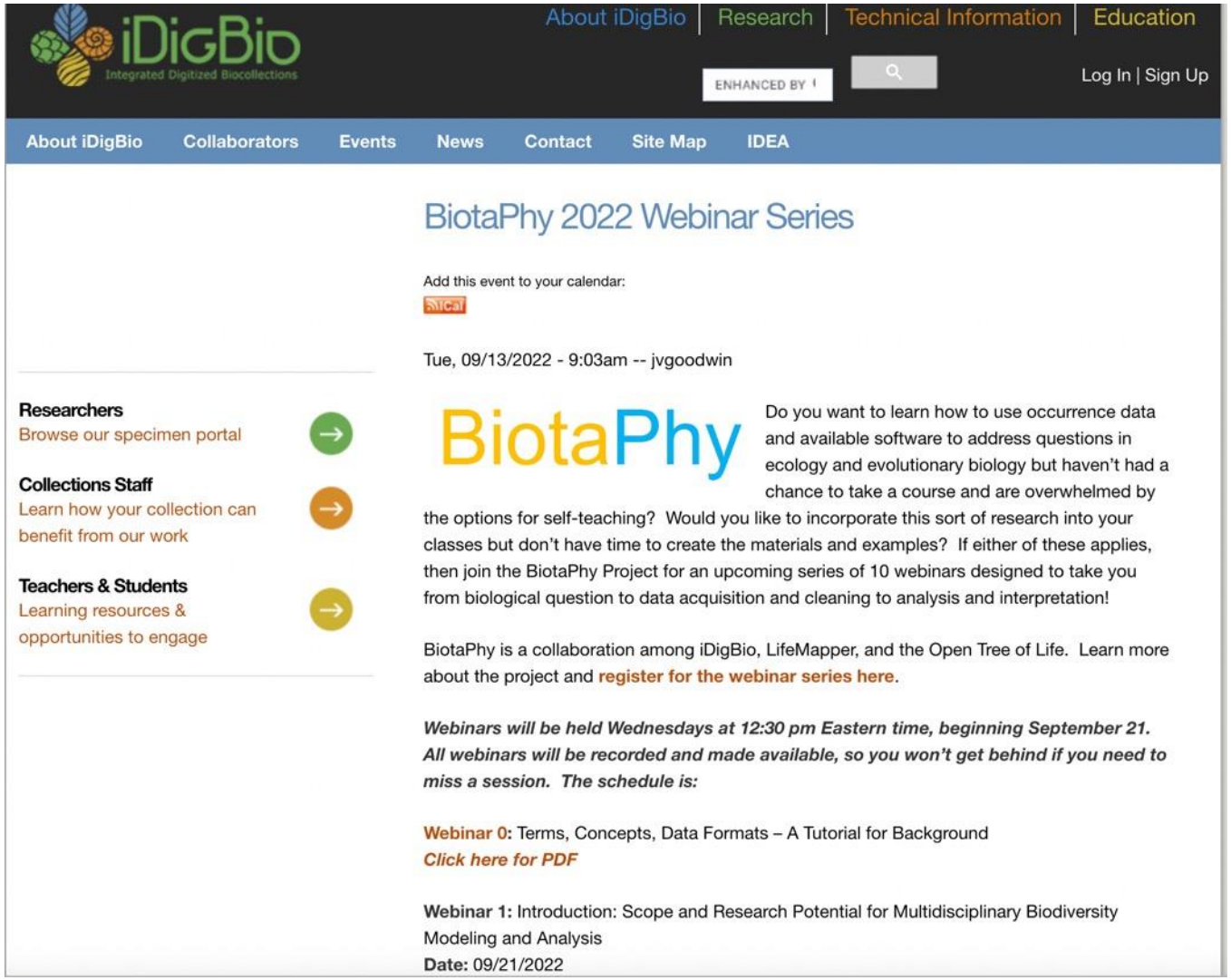
September 21-November 30, 2022

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

iDigBio:

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

iDigBio.org



Thank You

Maria Cortez
Aimee Stewart

Jill Goodwin
Gil Nelson

Webinar 7

Species Distribution Models: Interpretation and Ancestral Niche Reconstruction

Goals

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

Learning Objectives

Biological Objectives:

- **Assess species distribution models**
- **Interpret species distribution models**
- **Learn a method of ancestral niche reconstruction**

Technical Objectives :

- **Go over Maxent HTML outputs**
- **Format and requirements for ancestral niche reconstruction**

Webinar organization

1. SDM interpretation
2. Exploring Concepts: Ancestral niche
3. Practical

SDM Interpretation

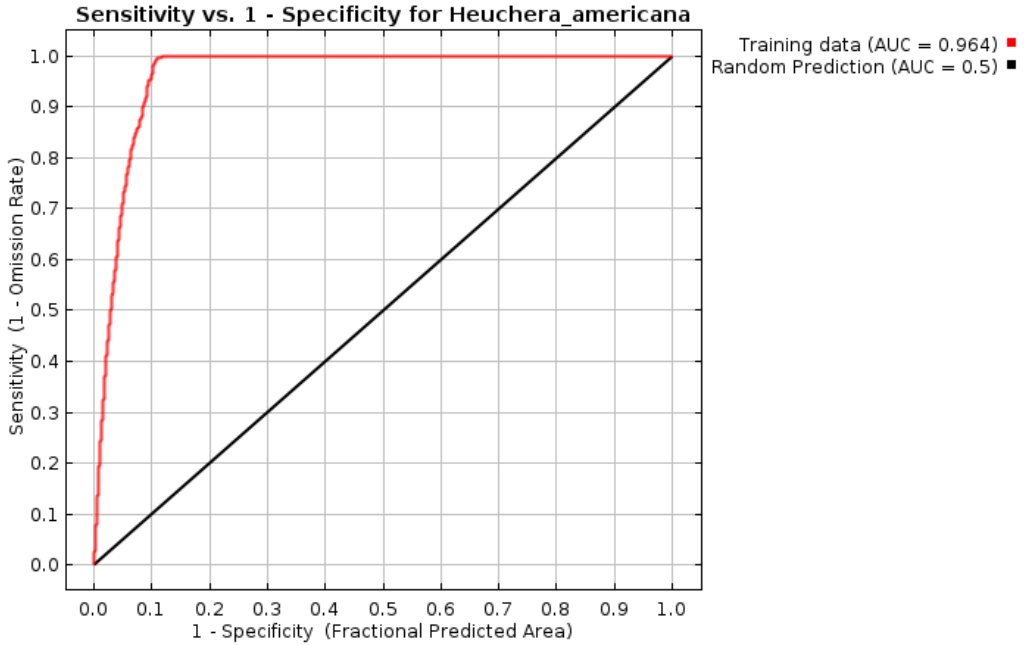
Assessment metrics

- How do you know the SDM (species distribution model) is telling you something meaningful?
- How do you know which environmental variables drive the model?

SDM Interpretation

Standard assessment metrics

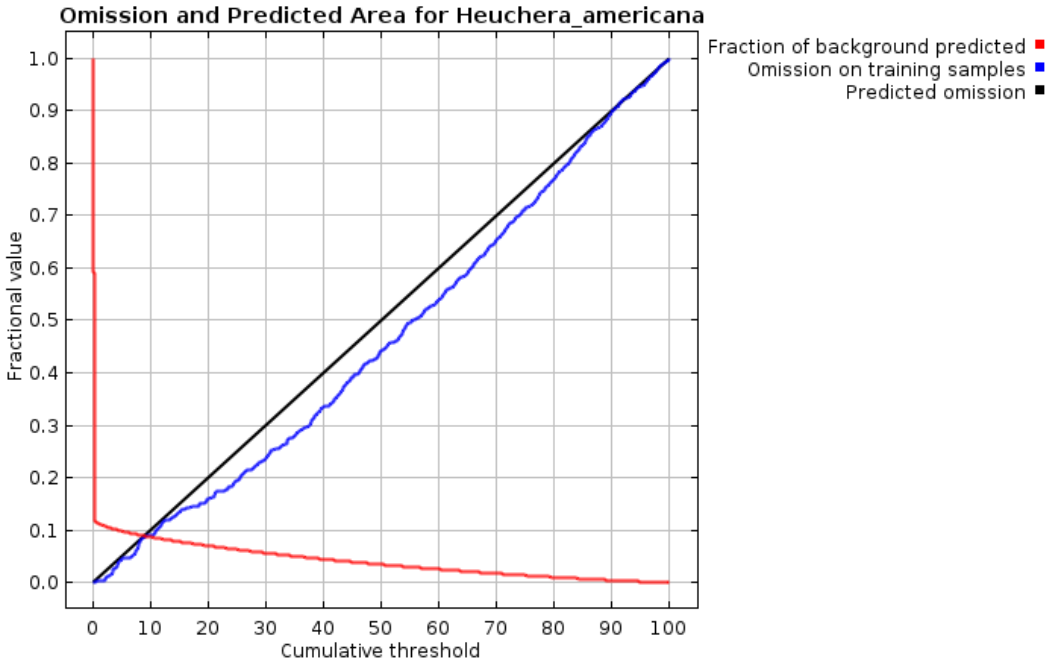
- AUC – Area Under the Curve
  - Plot sensitivity (true positive rate) vs. 1 – specificity (false positive rate)
  - Given a false positive classification threshold (x axis), predicted true positive rate
  - 1:1 line is random performance



SDM Interpretation

Standard assessment metrics

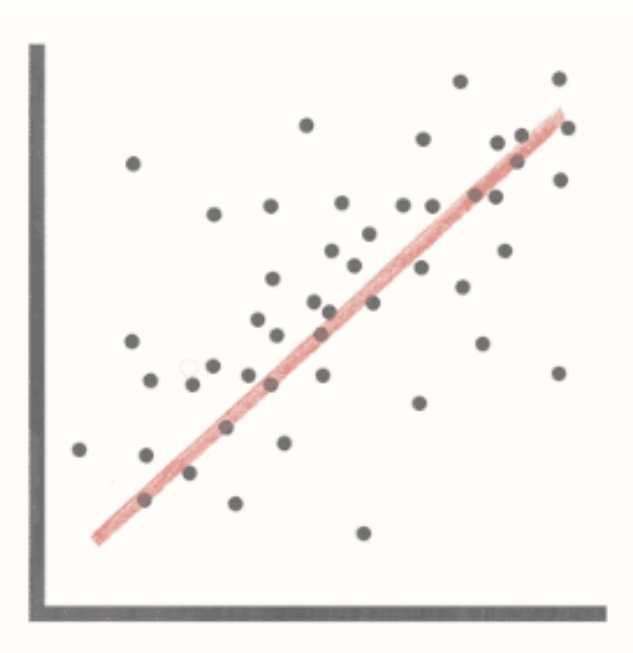
- Omission plot
  - Less commonly used
  - Omission means a sample is present but was not predicted by the model
  - 1:1 line is an ideal model – higher threshold, more omission



SDM Interpretation

Standard assessment metrics

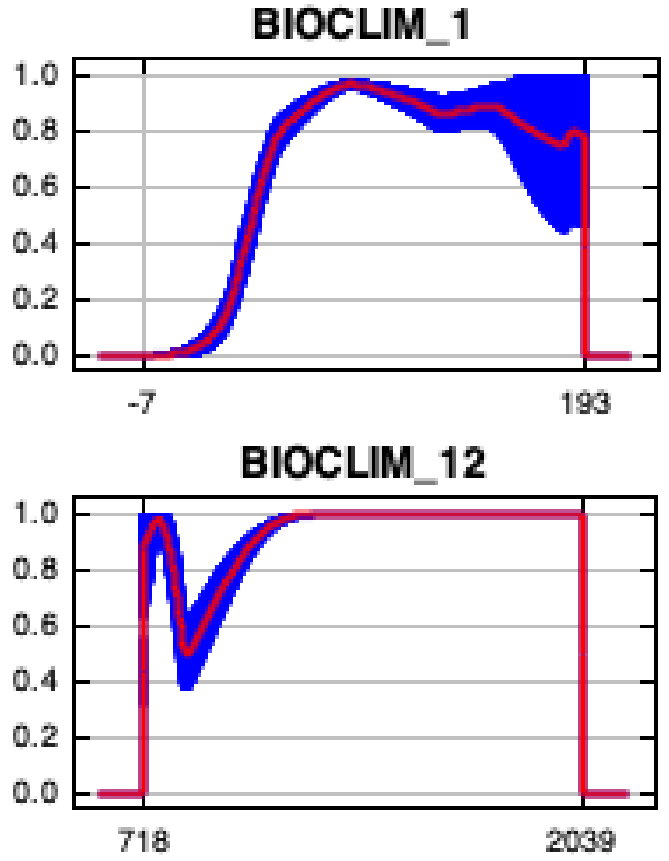
- Variable importance
  - Many ways to measure, but all are vulnerable to collinearity
  - Temperature and precipitation variables are highly correlated, so importance is not independent



SDM Interpretation

Standard assessment metrics

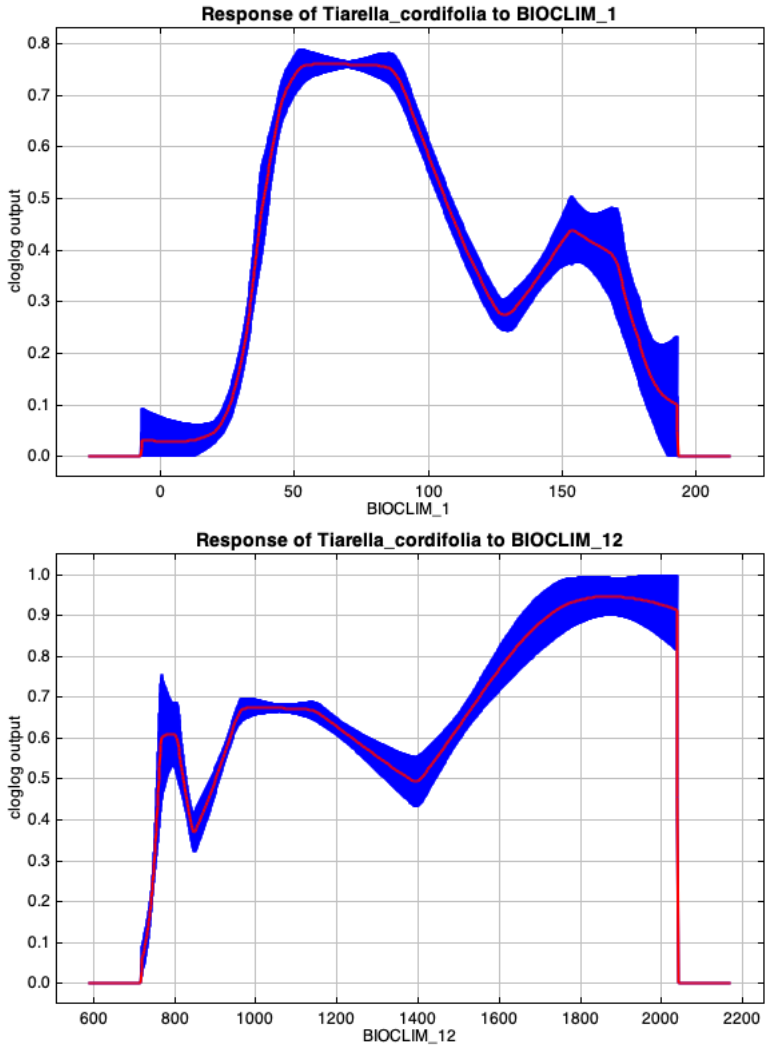
- Response curves
  - Shows what the model is actually doing – probability of presence expressed in terms of one variable
  - Often un-biological shapes – due to model's response to collinearity



SDM Interpretation

Standard assessment metrics

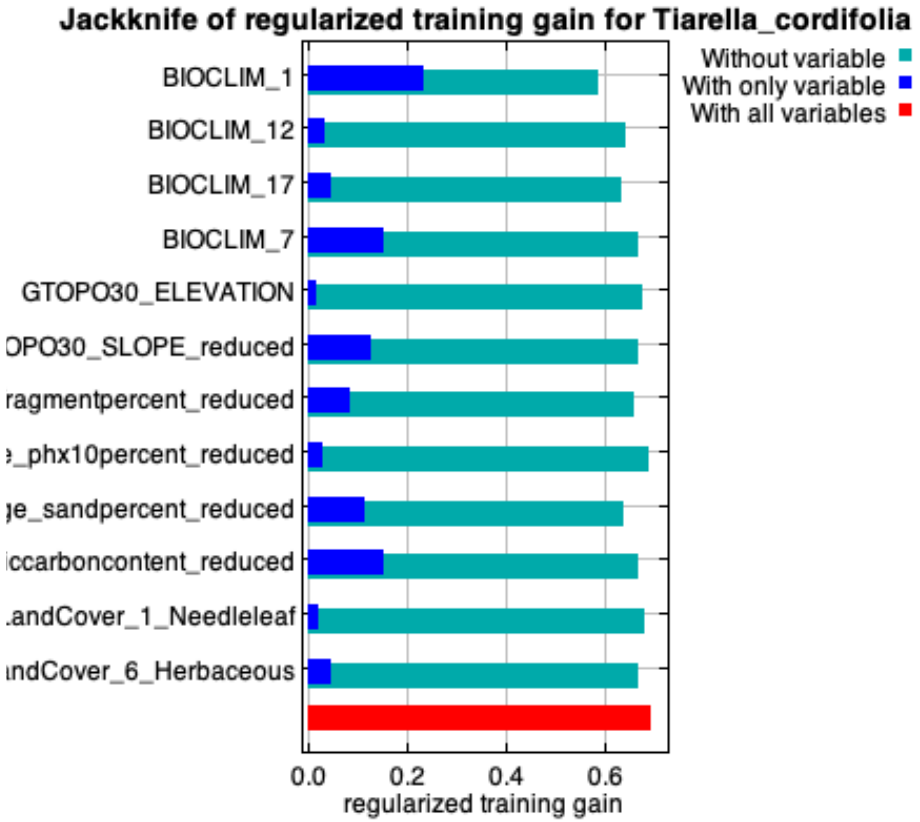
- Response curves
  - Maxent also reports response curves for a model with only one variable
  - Doesn't use full information in model



SDM Interpretation

Standard assessment metrics

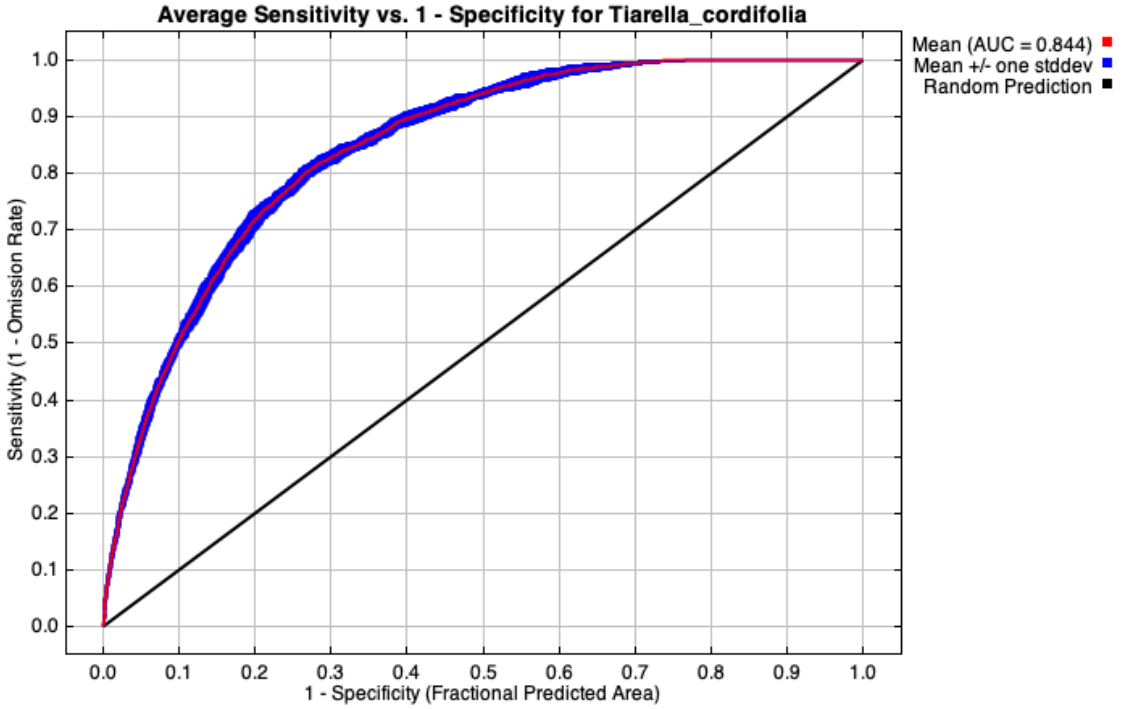
- Jackknife
  - Standard stats meaning – leave-one-out analysis
  - Most straightforward interpretation
  - Especially without variable (teal) – notice leaving variables out sometimes results in a lot of lost information, other times not



SDM Interpretation

Standard assessment metrics

- Replication
  - Randomly divide test and training data multiple times
  - Maxent allows multiple strategies
  - Allows for uncertainty estimation



SDM Interpretation

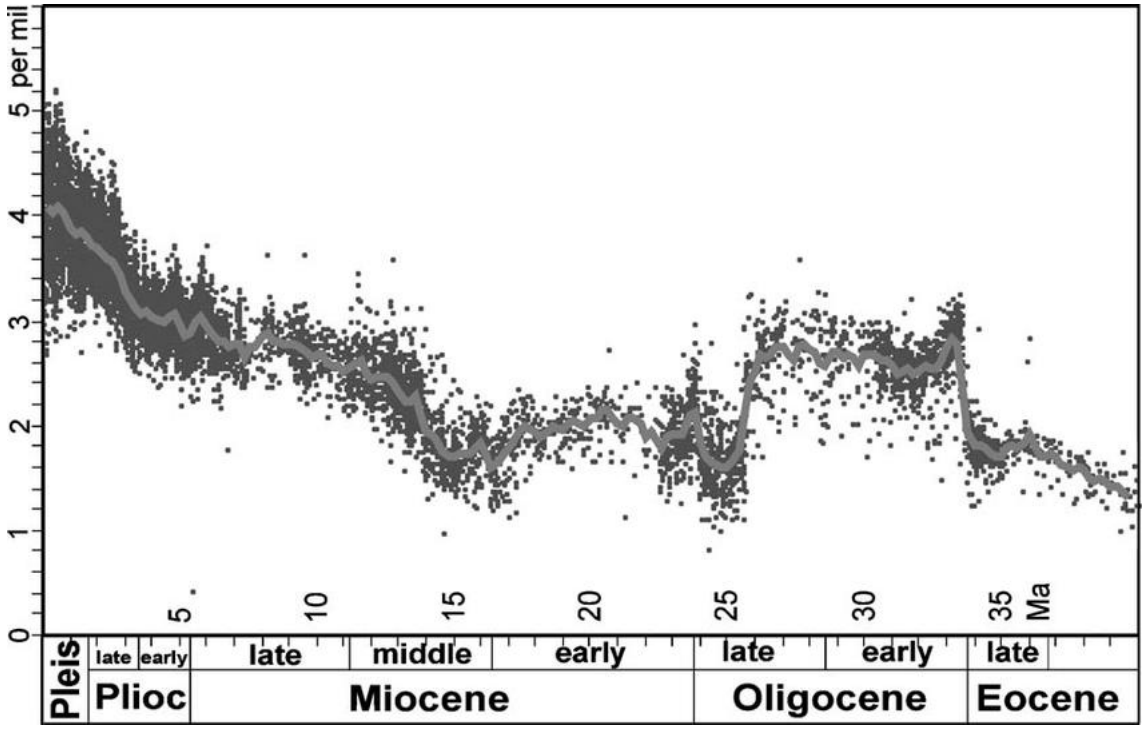
Summary

- Be careful interpreting in terms of physiology
- Use multiple means of assessment
- Use replication
- Focus on the question — extrapolation and niche limits are harder to estimate, centroid is easy

Exploring Concepts: Ancestral niche

Going back in time

- Hindcasting: Project model predictions onto past climate
- Our predictions could be very wrong if niche has evolved through time
- Our predictions are irrelevant if modern species did not exist back then



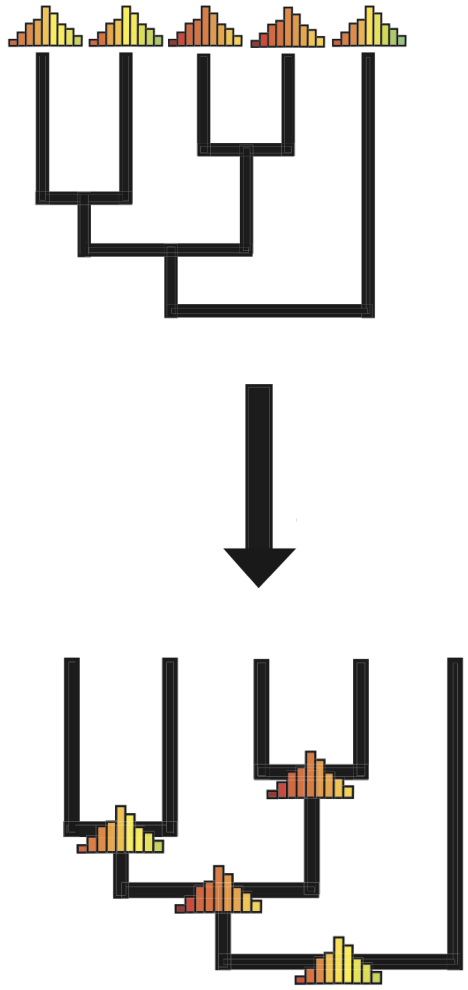
Exploring Concepts: Ancestral niche

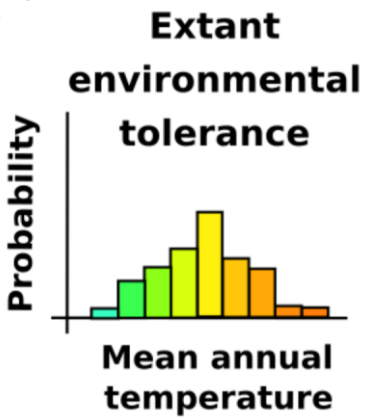
Ancestral niche

- Reconstruction of the niche of a hypothetical ancestor

Scenario

- Question about the past, too far in the past to hindcast a current species' distribution

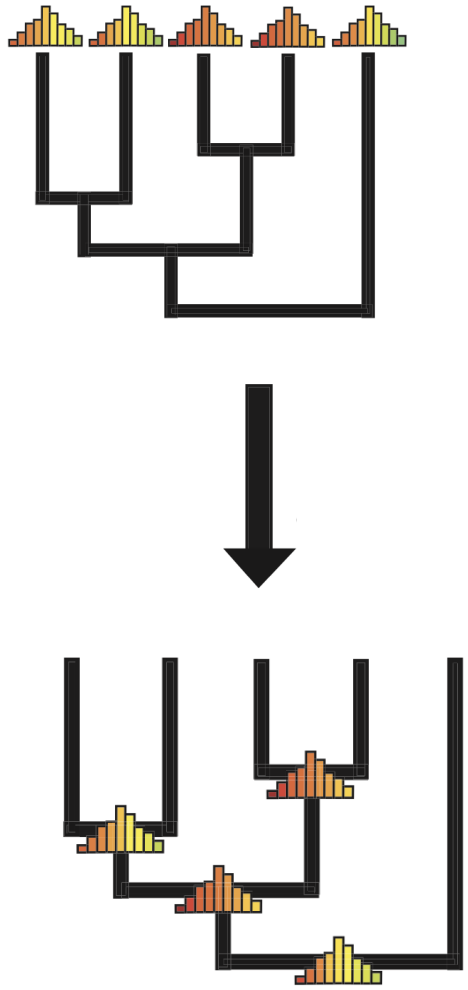


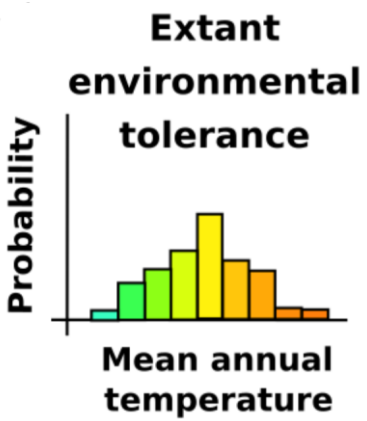


Exploring Concepts: Ancestral niche

Ancestral niche

- Uses SDM outputs directly in the form of environmental responses
- Similar but also dissimilar to standard character reconstruction
- Focus is on variance in niche ("niche breadth"), not just the most likely value ("centroid")



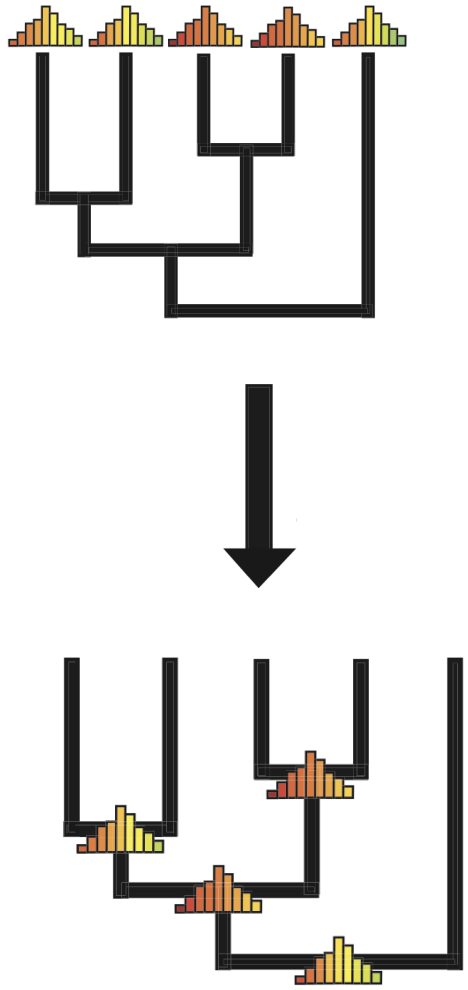


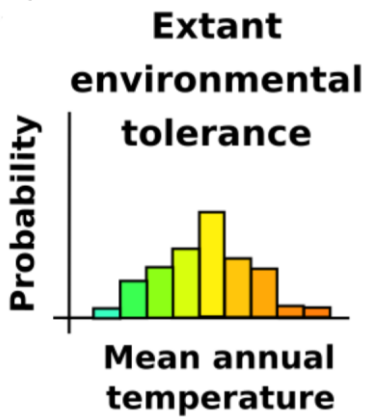
Exploring Concepts: Ancestral niche

Ancestral niche

- Multiple methods

- Available packages:
  - phyloclim (R package, available on CRAN)
  - ambitus (available at https://github.com/ryanafolk/ambitus)
  - Histogram method (discussed later today)

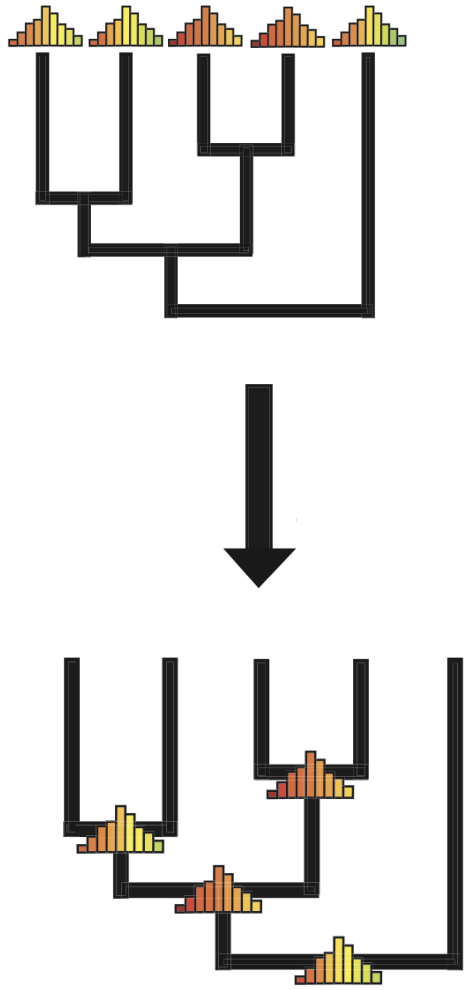


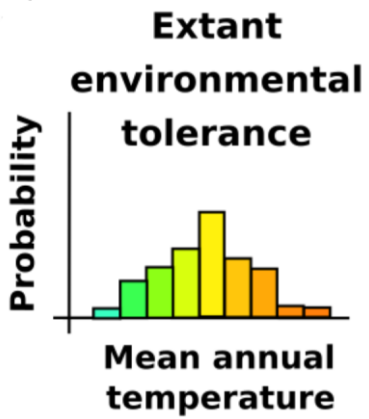


Exploring Concepts: Ancestral niche

Different modes of thinking

- E-space: The area of the environment a species occupies
  - Only need a tree and SDMs, many applications
- G-space: The area of the Earth a species occupies
  - Need paleoclimate data too

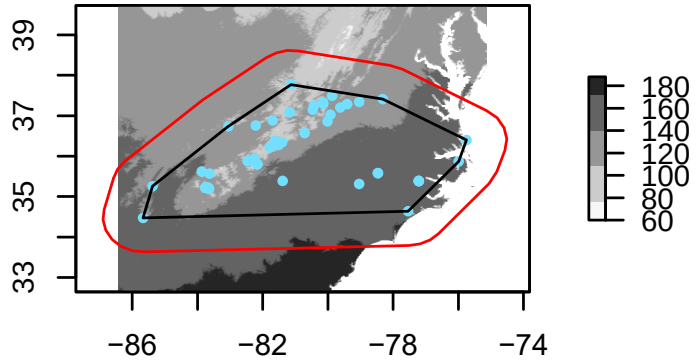




Exploring Concepts: MaxEnt

Projection

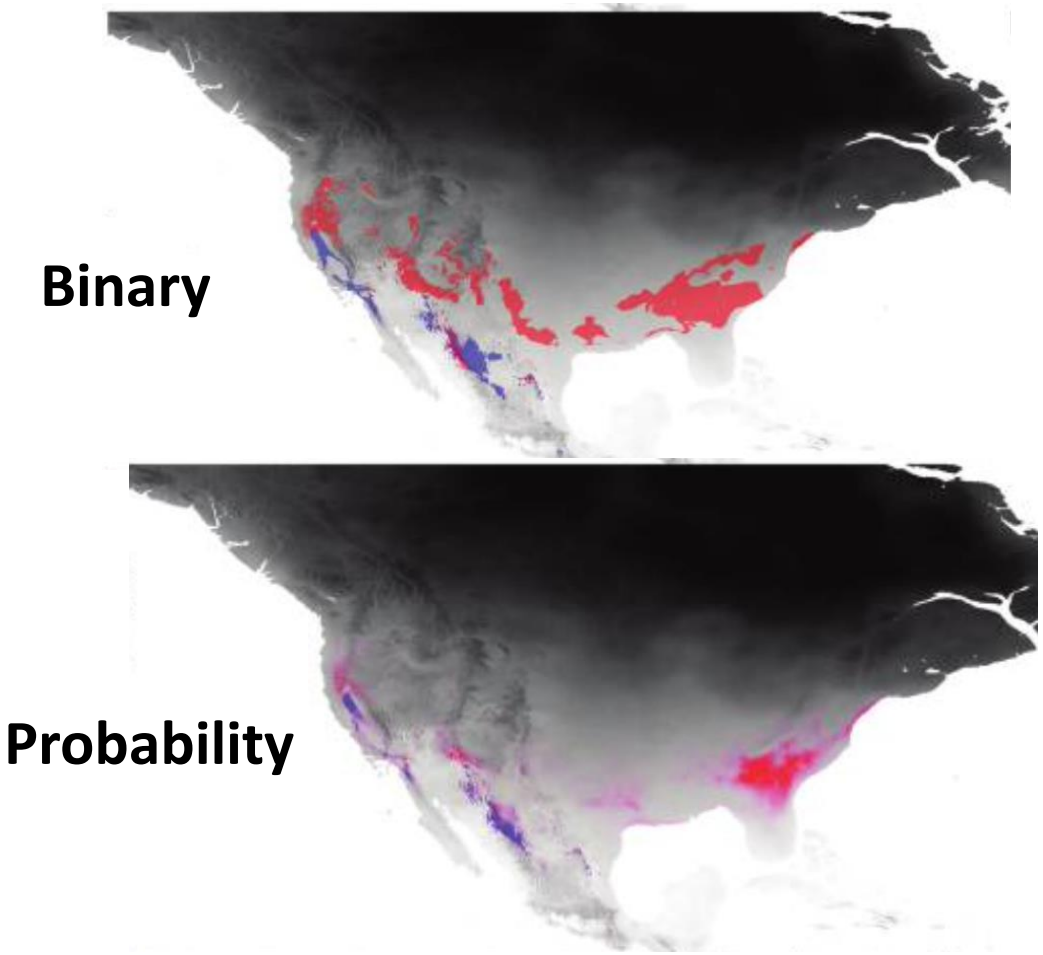
- Model states occurrence probability in terms of environmental conditions, which need not be those that obtain at present
- With the assumption that the relationship between environment and occurrence holds true through time, can predict for any time



Exploring Concepts: Ancestral niche

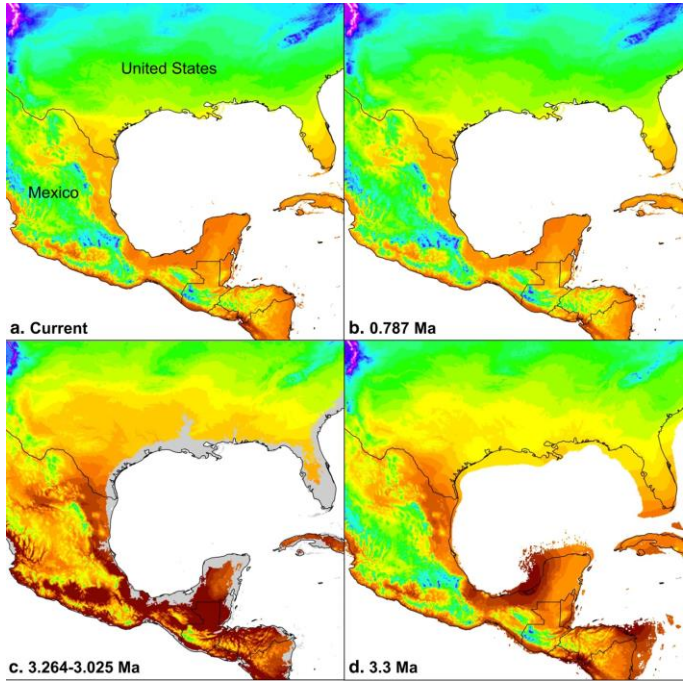
Projecting ancestral reconstruction

- Apply reconstructed niche probability to paleoclimatic data
- Three types of statements:
  - Probability
  - Binary (yes or no)
  - Distance (how far from ideal habitat)



Data sources

PaleoClim



Pleistocene: late-Holocene, Meghalayan (4.2-0.3 ka), v1.0

Pleistocene: mid-Holocene, Northgrippian (8.326-4.2 ka), v1.0

Pleistocene: early-Holocene, Greenlandian (11.7-8.326 ka), v1.0

Pleistocene: Younger Dryas Stadial (12.9-11.7 ka), v1.0

Pleistocene: Bølling-Allerød (14.7-12.9 ka), v1.0

Pleistocene: Heinrich Stadial 1 (17.0-14.7 ka), v1.0

Pleistocene: Last Interglacial (ca. 130 ka), v1.0

Pleistocene: MIS19 (ca. 787 ka), v1.0*

Pliocene: mid-Pliocene warm period (3.205 Ma), v1.0*

Pliocene: M2 (ca. 3.3 Ma), v1.0*

SDMs: Before we start …

We will be living outside of Docker today – github repositories run directly

SDMs: Before we start …

To look at Maxent diagnostics download for the first time OR update the tutorials repository.

Initial download:

git clone https://github.com/biotaphy/tutorials

Update tutorial:

cd tutorials

git pull

SDMs: Before we start …

Identifying climatic drivers of hybridization in Heuchereae (Saxifragaceae)

R.A. Folk*†, M.L. Gaynor*, N.J. Engle-Wrye, B.C. O'Meara, P.S. Soltis, D.E. Soltis, R.P. Guralnick, S.A. Smith, C.J. Grady, Y. Okuyama

* co-first authors

† author for correspondence, rfolk@biology.msstate.edu

"histogram method"

In revision

SDMs: Before we start ...

In downloaded repository navigate to

tutorials/data/input/heuchera_rfolk_sdm/

Click on .html file to get navigable output

SDMs: Before we start ...

In downloaded repository navigate to

tutorials/data/input/heuchera_rfolk_sdm/

Click on .html file to get navigable output

→ browser

Ancestral niche: Data preparation

Initial download:

git clone https://github.com/ryanafolk/heuchera_ancestral_niche

For your reference: https://github.com/ryanafolk/pno_calc

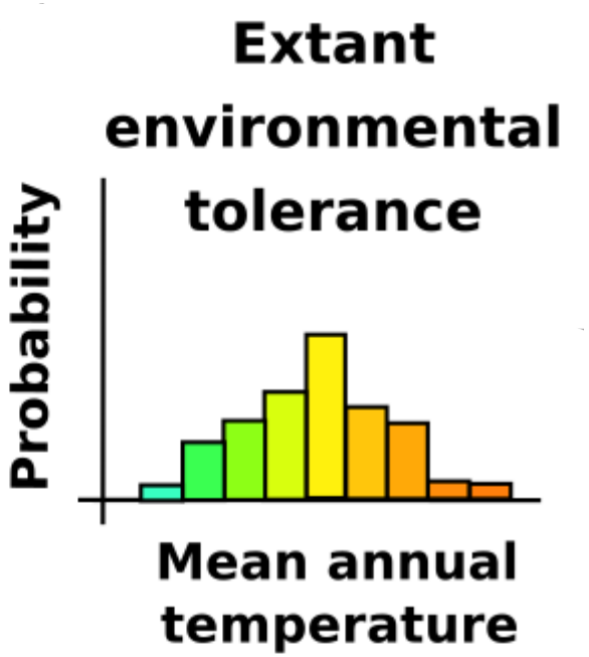
Ancestral niche: Data preparation

"PNO" (Predicted Niche Occupancy"):

environmental variable as predictor,

model suitability as response

	A	B	C	D	E	F	G	H	I	J
Formula Bar		-251	-240.08	-229.16	-218.24	-207.32	-196.4	-185.48	-174.56	-163.64
2	Bensoniella_oregona	0	0	0	0	0	0	0	0	0
3	Conimitella_williamsii	0	0	0	0	0	0	0	0	0
4	Darmera_peltata	0	0	0	0	0	0	0	0	0
5	Elmera_racemosa	0	0	0	0	0	0	0	0	0
6	Heuchera_abramsii	0	0	0	0	0	0	0	0	0
7	Heuchera_acutifolia	0	0	0	0	0	0	0	0	0
8	Heuchera_alba	0	0	0	0	0	0	0	0	0
9	Heuchera_americana	0	0	0	0	0	0	0	0	0
10	Heuchera_bracteata	0	0	0	0	0	0	0	0	0
11	Heuchera_caroliniana	0	0	0	0	0	0	0	0	0
12	Heuchera_chlorantha	0	0	0	0	0	0	0	0	0
13	Heuchera_cylindrica	0	0	0	0	0	0	0	0	0
14	Heuchera_eastwoodiae	0	0	0	0	0	0	0	0	0
15	Heuchera_elegans	0	0	0	0	0	0	0	0	0
16	Heuchera_glabra	0	0	0	1.52E-07	1.39E-06	6.09E-07	1.60E-06	2.96E-06	1.47E-06
17	Heuchera_glomerulata	0	0	0	0	0	0	0	0	0
18	Heuchera_grossulariifolia	0	0	0	0	0	0	0	0	0
19	Heuchera_hallii	0	0	0	0	0	0	0	0	0
20	Heuchera_longiflora	0	0	0	0	0	0	0	0	0
21	Heuchera_longipetala	0	0	0	0	0	0	0	0	0
22	Heuchera_maxima	0	0	0	0	0	0	0	0	0
23	Heuchera_merriamii	0	0	0	0	0	0	0	0	0



Ancestral niche: Data preparation

Tree: Newick format

```
(((Rodgersia_pinnata:0.103156,Darmera_peltata:0.103156):0.053155,((((Tiarella_trifoliata:0.048043,(Tiarella_polyphyll
```

Paleoclimatic data: See paper in review, contact me with questions about format

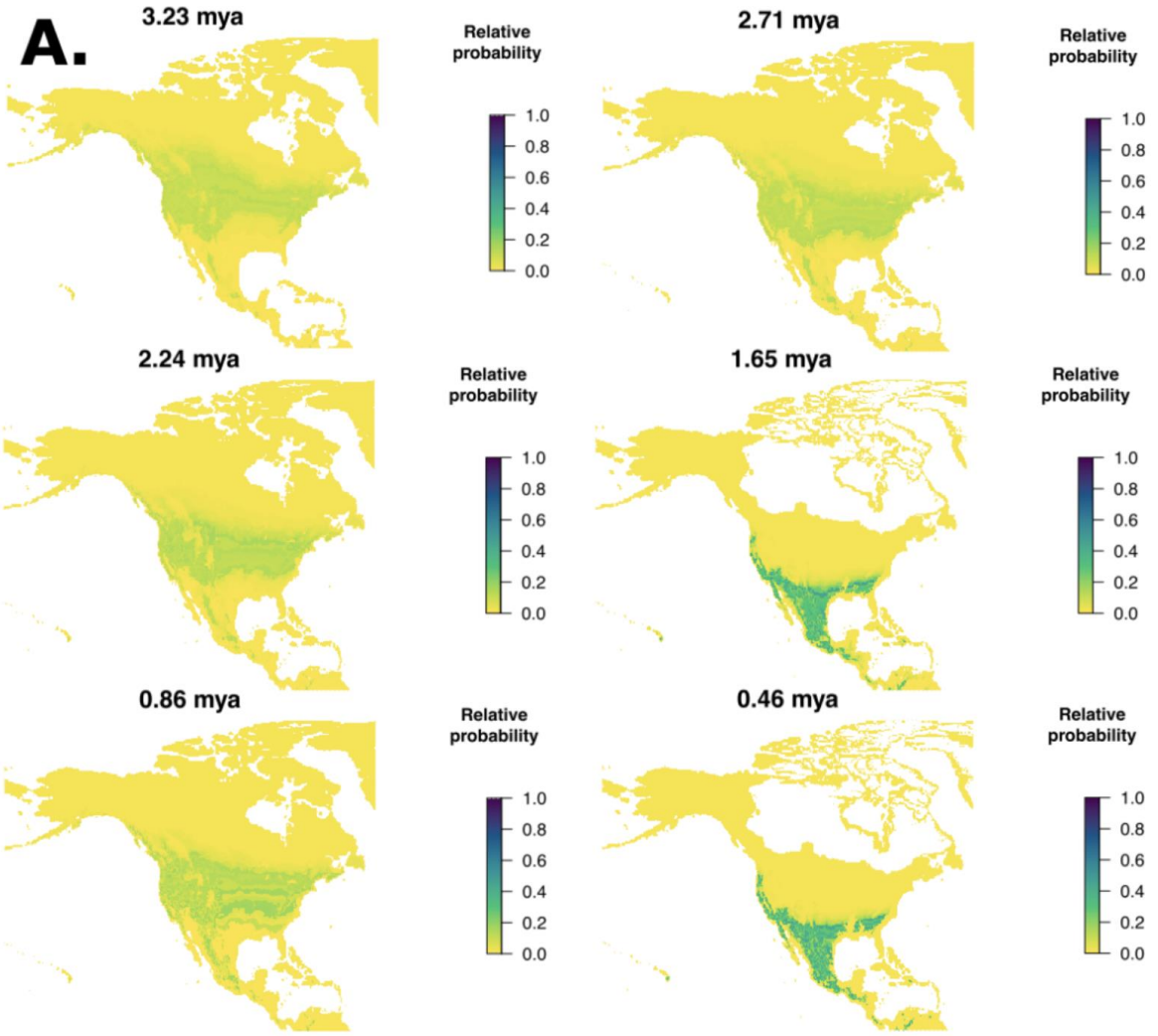
Optional: Phylogenetic uncertainty (.tsv)

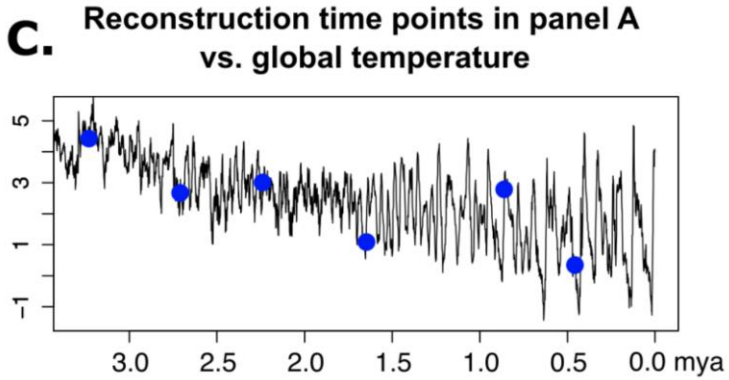
Gen	t_n91	t_n92	t_n93	t_n94	t_n95	t_n96	t_n97	t_n98	t_n99
1	0.5197541	0.4388296	0.2114912	0.1259142	0.1238122	0.0967981
2	0.5197541	0.4388296	0.2114912	0.1259142	0.1226340	0.0983163
3	0.5197541	0.4388296	0.2114912	0.1259142	0.1226340	0.0975708
4	0.5197541	0.4388296	0.2114912	0.1259142	0.1226340	0.0975708
5	0.5197541	0.4388296	0.2114912	0.1259142	0.1229654	0.0975708
6	0.5197541	0.4388296	0.2114912	0.1259142	0.1229654	0.0975708
7	0.5197541	0.4388296	0.2017740	0.1259142	0.1229654	0.0990773
8	0.5197541	0.4388296	0.2017740	0.1259142	0.1221857	0.0990773
9	0.5197541	0.4388296	0.2017740	0.1259142	0.1221857	0.0990773
10	0.5197541	0.4388296	0.2017740	0.1259142	0.1221857	0.0990773
11	0.5197541	0.4388296	0.1932840	0.1246685	0.1221857	0.0990773
12	0.5197541	0.4388296	0.1870675	0.1246685	0.1221857	0.0990773
13	0.5197541	0.4388296	0.1870675	0.1253817	0.1221857	0.0990773
14	0.5197541	0.4388296	0.1783095	0.1253817	0.1221857	0.0990773
15	0.5197541	0.4388296	0.1783095	0.1253817	0.1221857	0.0990773
16	0.5197541	0.4388296	0.1783095	0.1253817	0.1221857	0.0990773
17	0.5197541	0.4388296	0.1783095	0.1253817	0.1221857	0.0990773
18	0.5197541	0.4388296	0.1776911	0.1253817	0.1221857	0.0990773
19	0.5197541	0.4388296	0.1776911	0.1253817	0.1221857	0.0990773
20	0.5197541	0.4388296	0.1776911	0.1253817	0.1221857	0.0990773
21	0.5197541	0.4552106	0.1747089	0.1253817	0.1225292	0.1001253
22	0.5381017	0.4736103	0.1747089	0.1253817	0.1225292	0.1001253
23	0.5381017	0.4736103	0.1703725	0.1253817	0.1225292	0.1001253
24	0.5381017	0.4736103	0.1703725	0.1253817	0.1225292	0.1001253
25	0.5381017	0.4736103	0.1655152	0.1253817	0.1225292	0.1001253
26	0.5381017	0.4736103	0.1583839	0.1253817	0.1225292	0.1001253
27	0.5381017	0.4736103	0.1583839	0.1253817	0.1225292	0.1001253
28	0.5381017	0.4736103	0.1583839	0.1253817	0.1225941	0.1001253
29	0.5381017	0.4736103	0.1502337	0.1253817	0.1225941	0.1001253

→ terminal

Ancestral niche: Data preparation

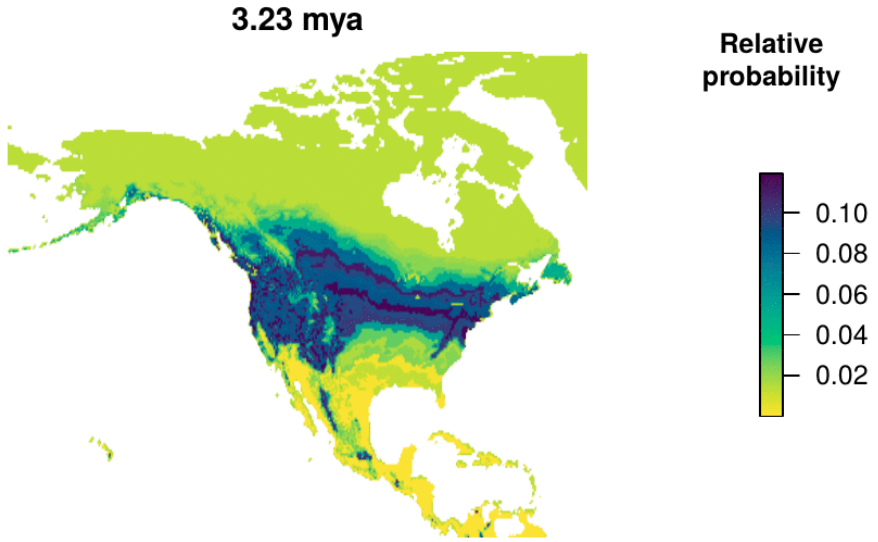
Output

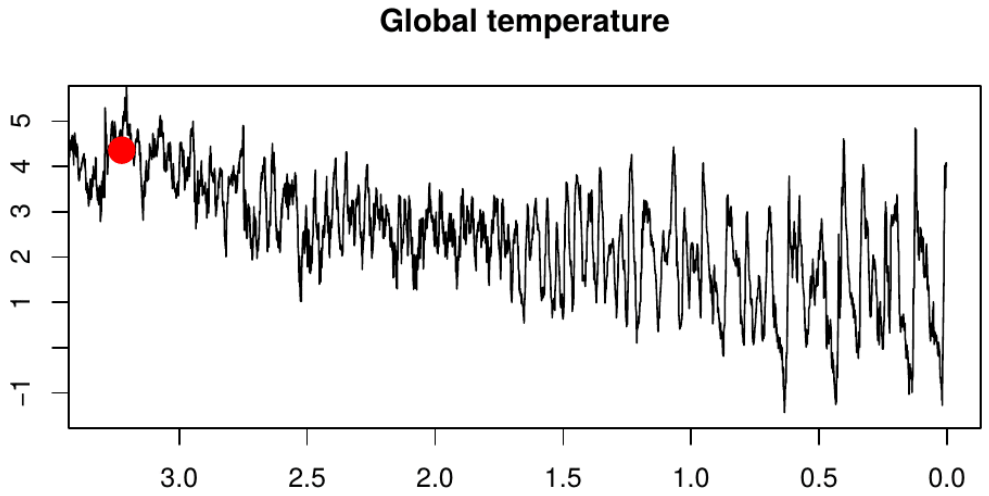




Ancestral niche: Data preparation

Output





Session Summary, Q&A and Discussion

- SDM interpretation
- SDM assessment
- Concepts behind ancestral niche reconstruction
- Use of “histogram method”

Session Summary, Q&A and Discussion

Any questions??

Please use the chat to write your question!