

Trevo: An Exploratory Tool for Tree and Trait Data (Working title)

Pre-Paper Talk

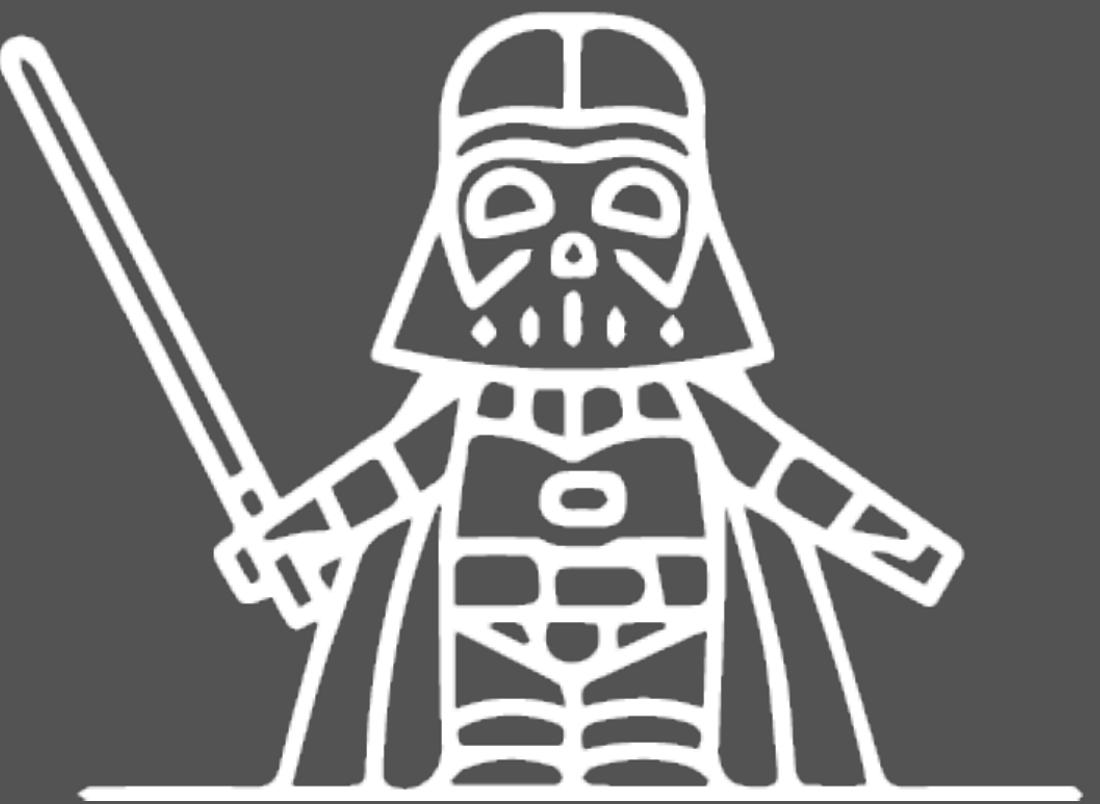


Background and Motivation

***What drives the diversity of life
on earth?***

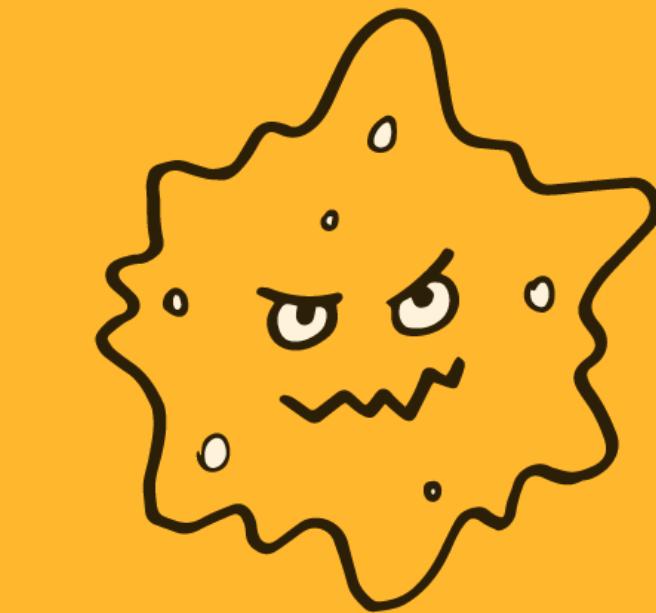
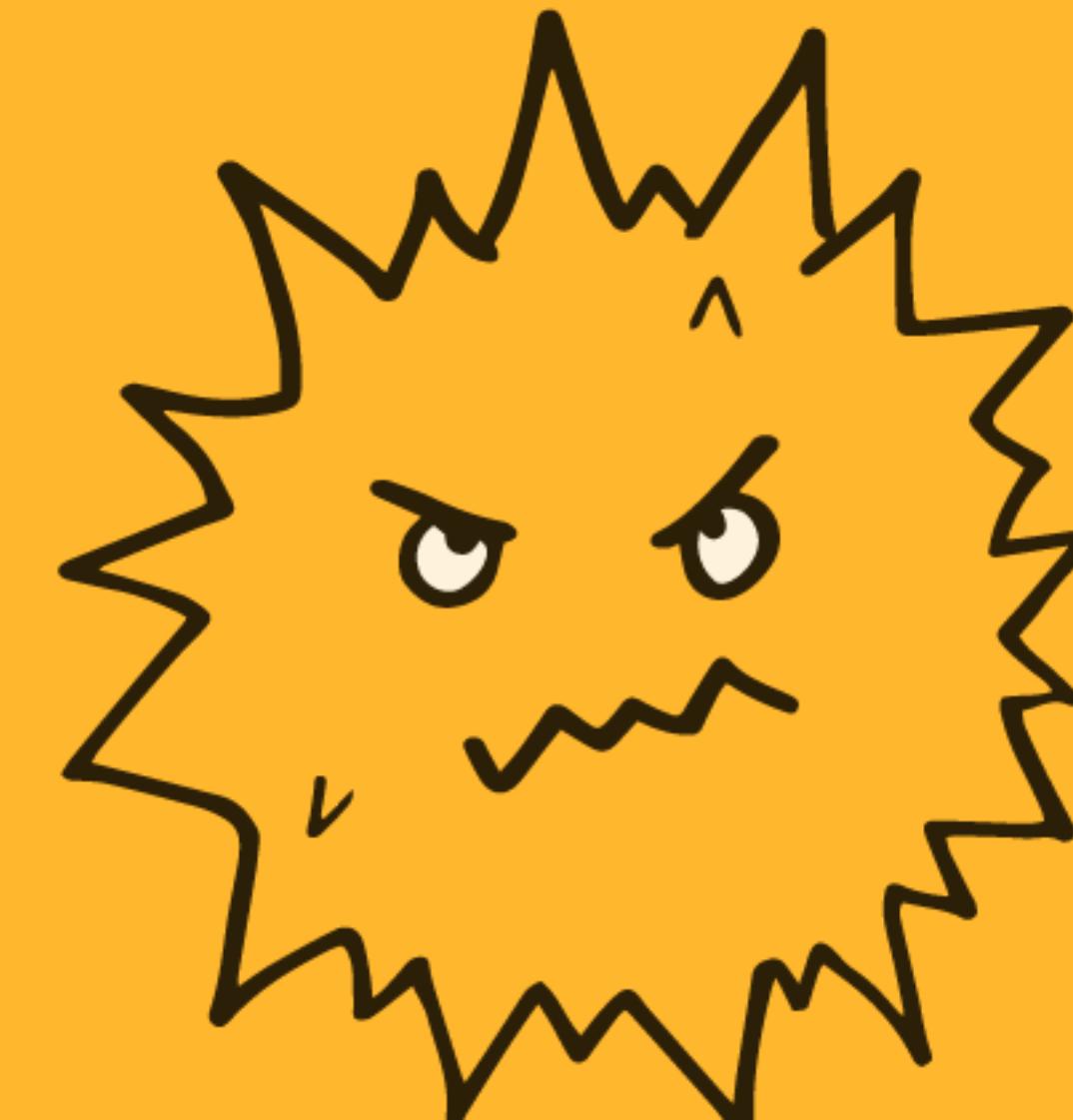
When and why do traits evolve?

***Do traits come from a common ancestor or a
result of other forces?***



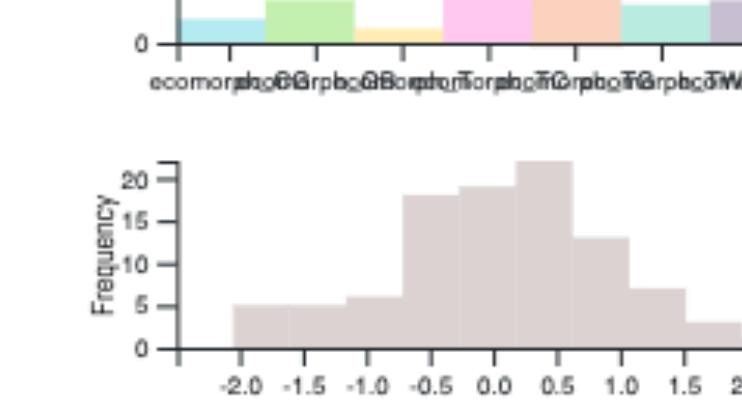
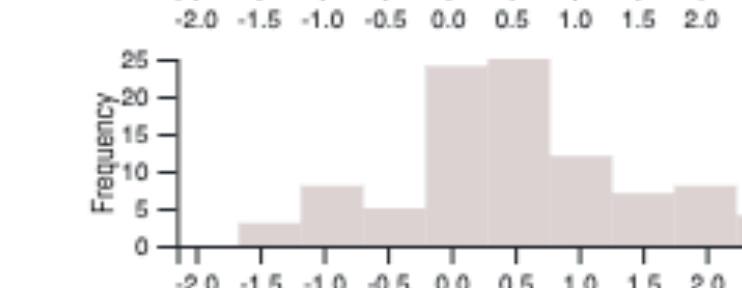
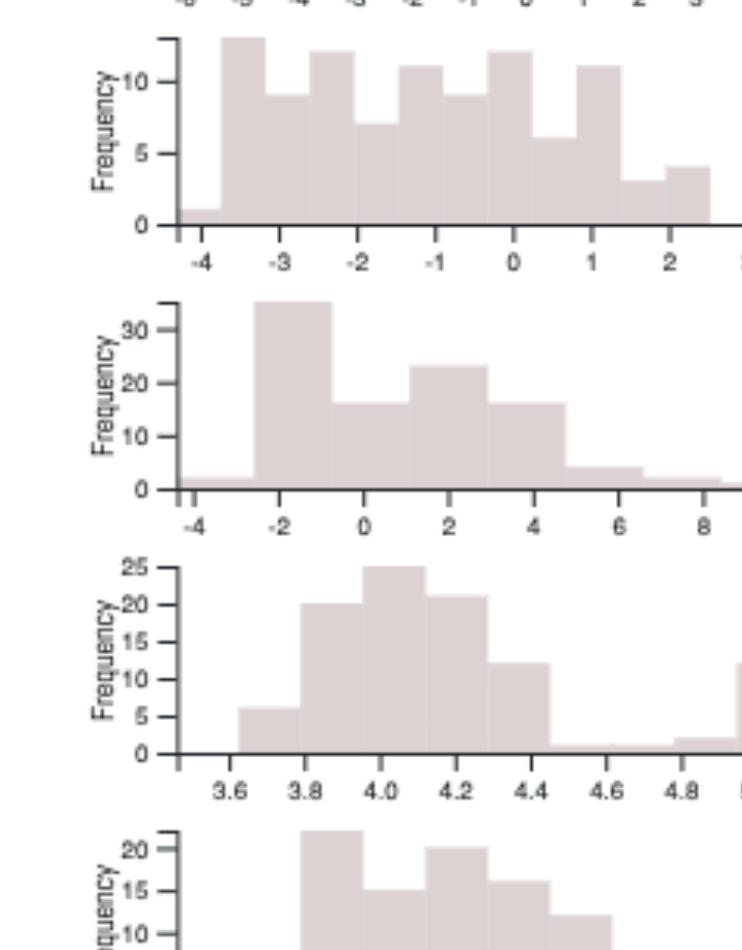
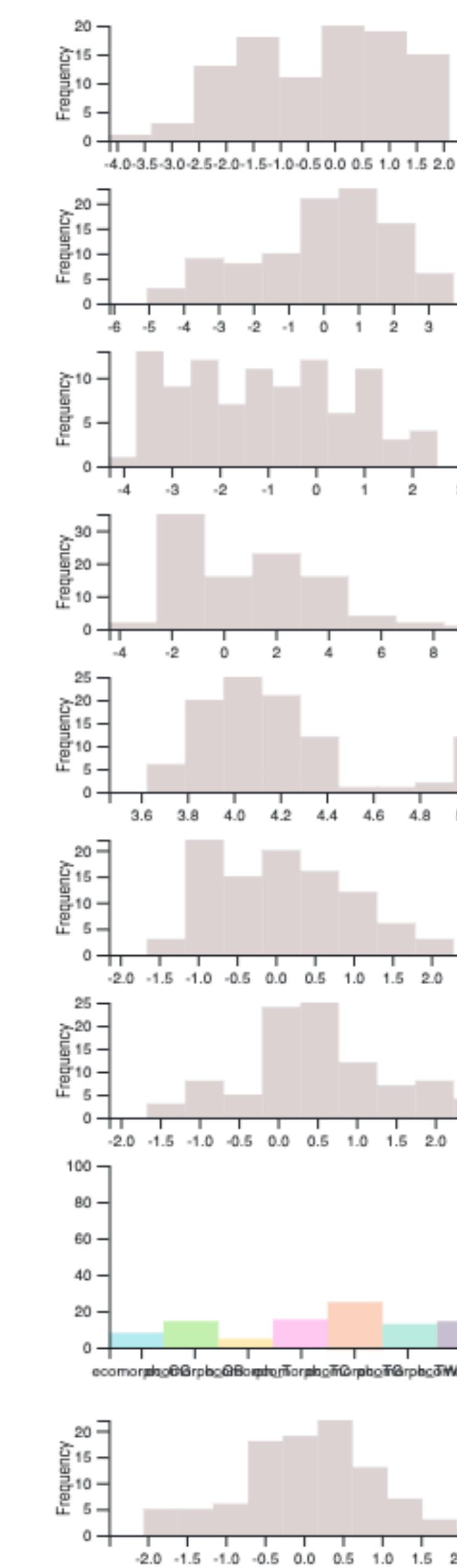
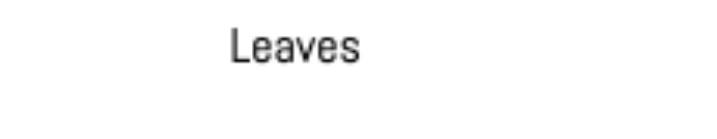
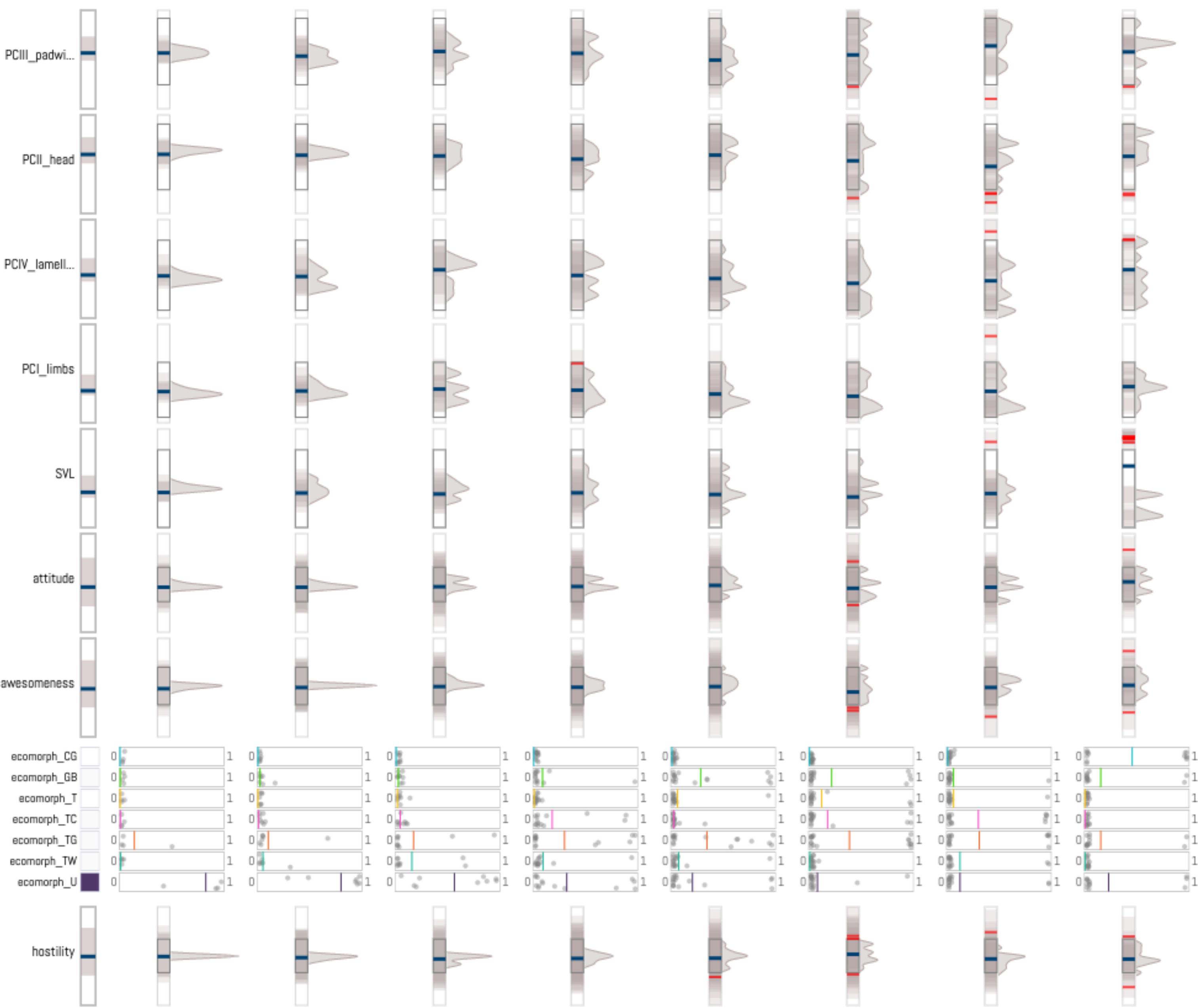
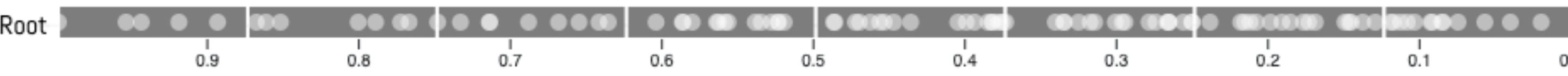
Study trees with traits (multivariate networks) that represent species' histories

- ***how animals adapt to their environments***
- ***the rate and mutation of virus strains, such as the coronavirus***



All island-anolis : 100 Paths

All island-anolis



COLOR BY TRAIT ▾ PHENGRAM



Contributions:

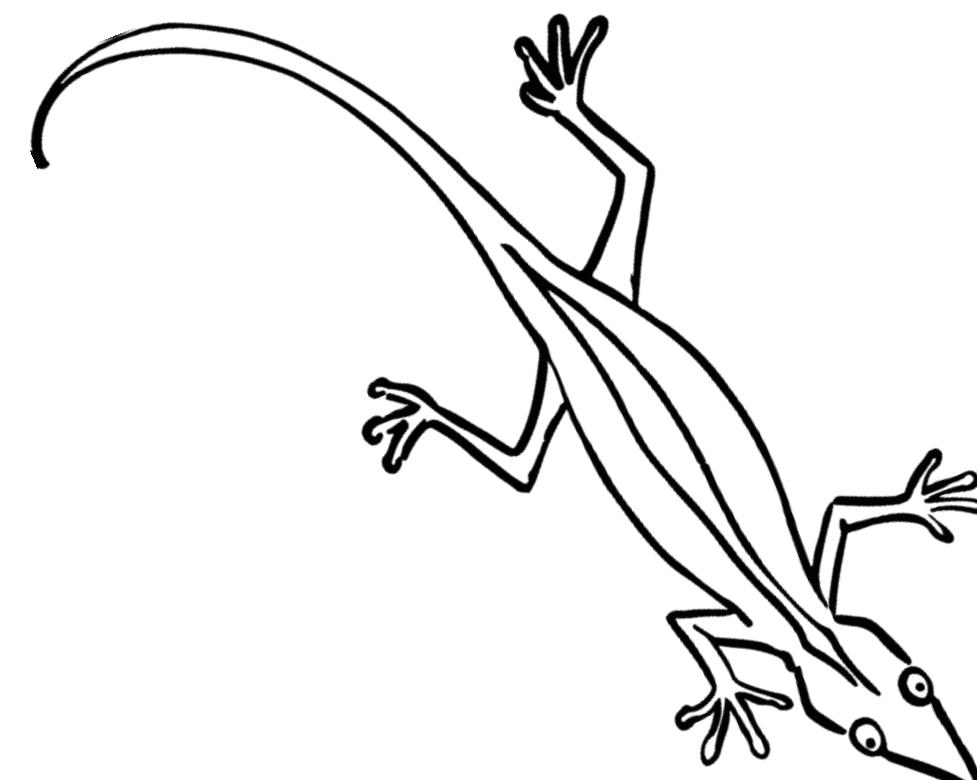
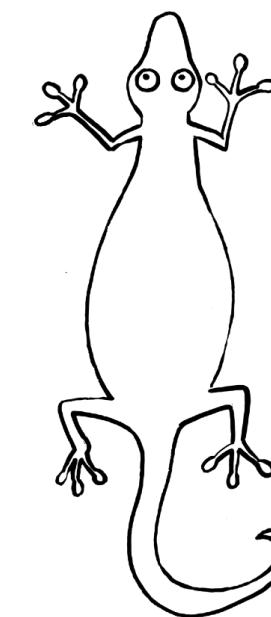
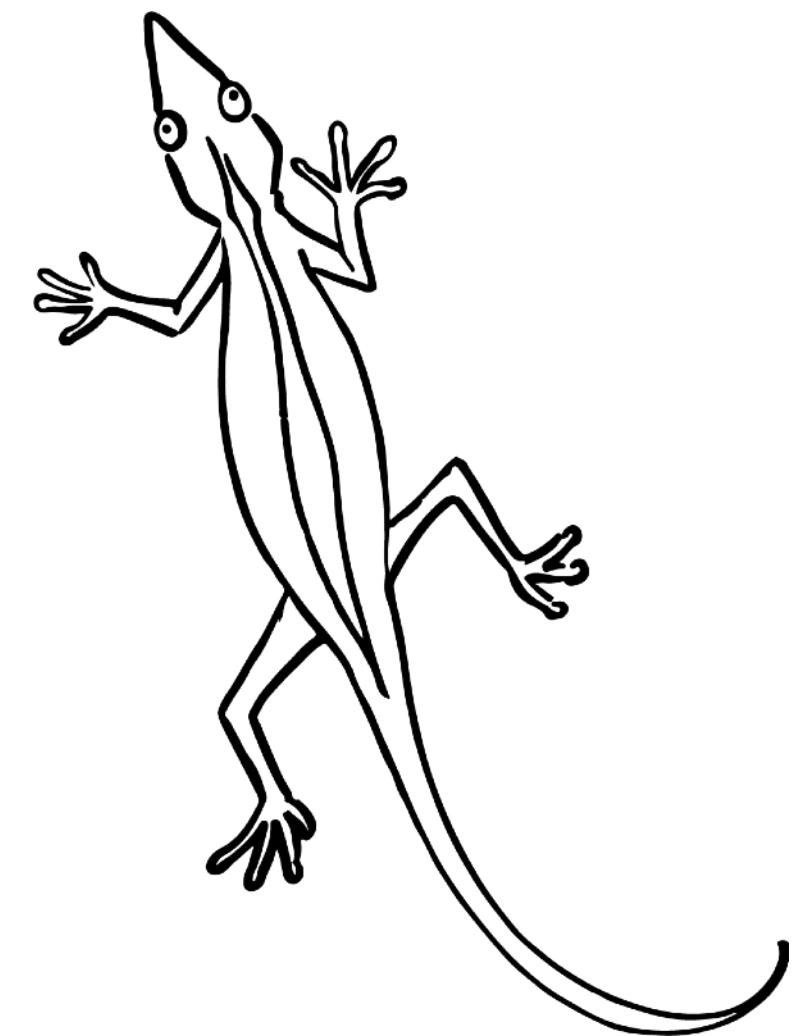
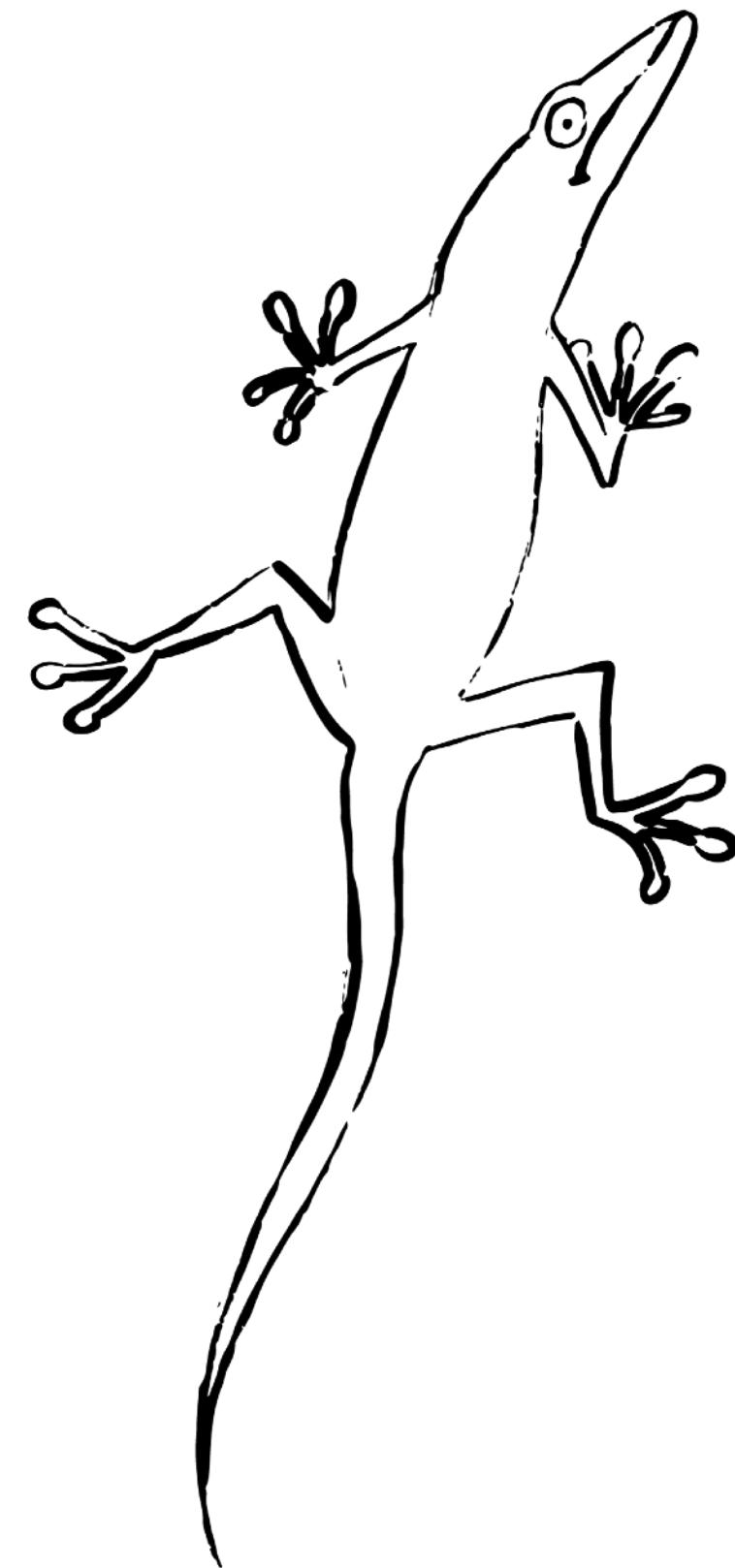
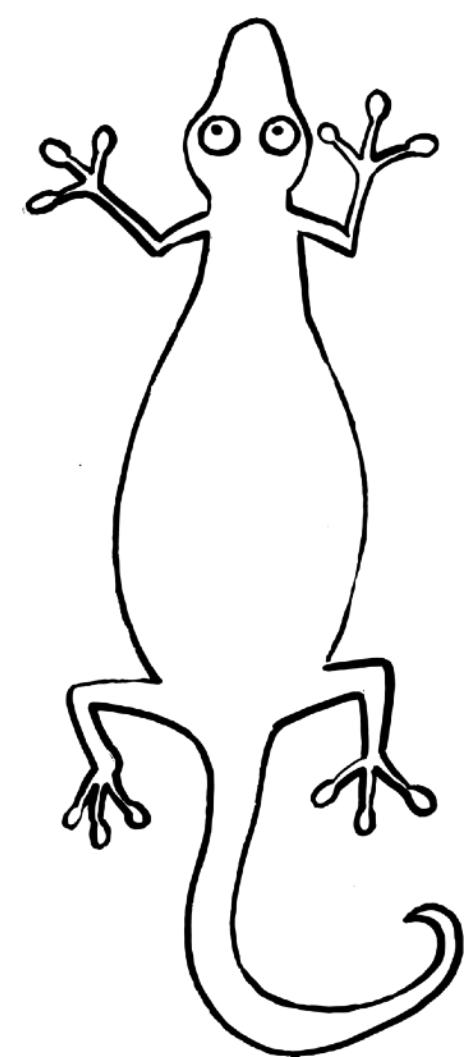
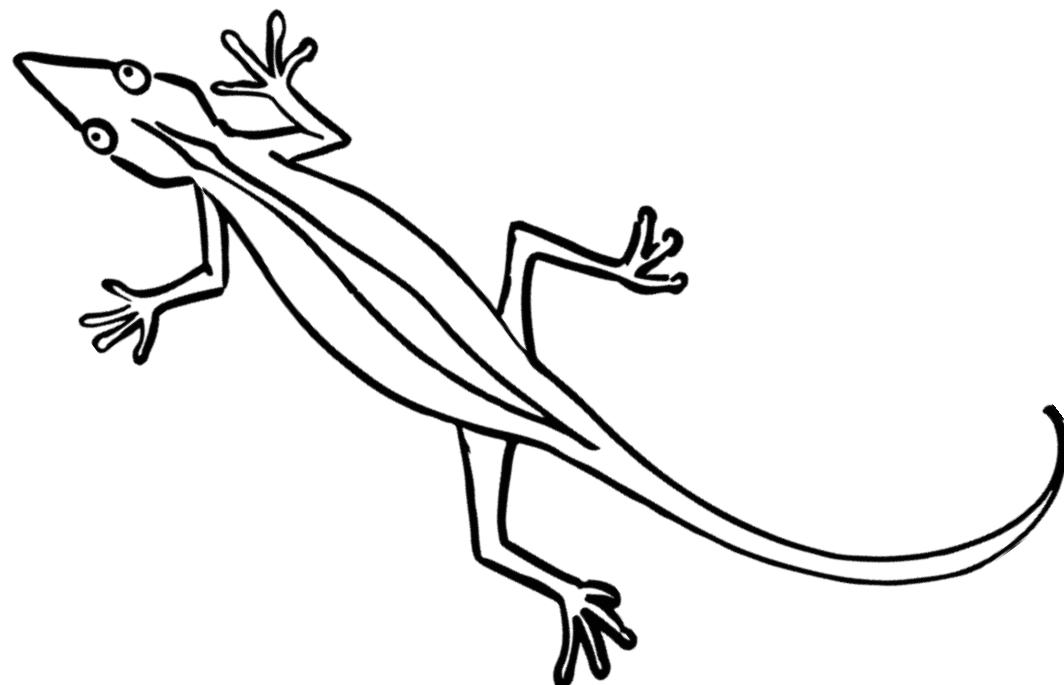
- **Domain Characterization of tasks.**
- **Novel visualization techniques to:**
 - *View multiple attributes under uncertainty in the context of topology*
 - *Identify patterns in paths along a tree based on attributes and topology*
- **Insights gained from experiments with criteria for rigor: Abundance, Transparency, and Reflexivity.**



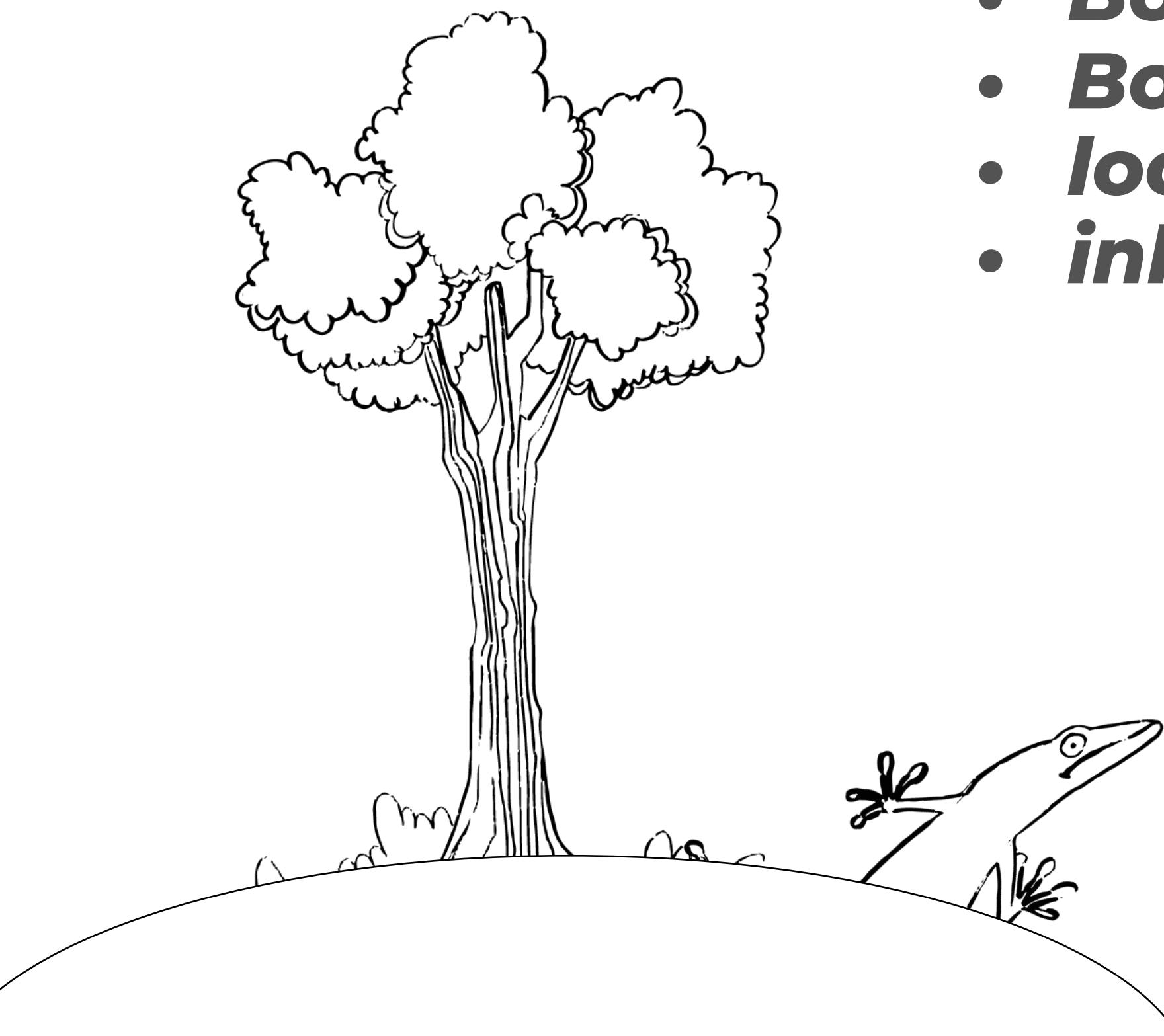
Task Characterization: Goal

***Explain why the living
world evolved the way it did***

***There are about 400 species
of anoles.***



Island A



Island B



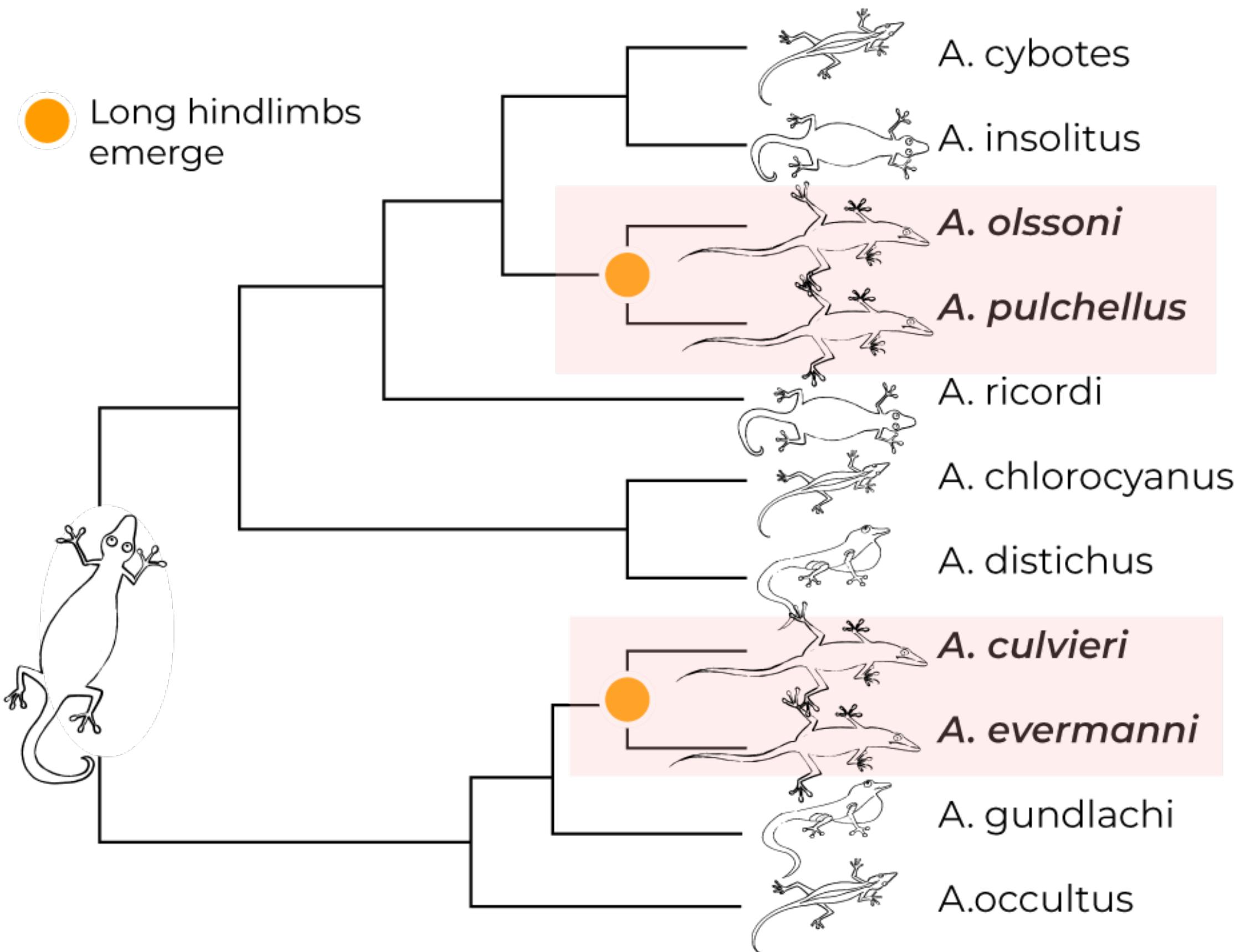
- ***Both have long hindlimbs,***
- ***Both are fast,***
- ***look similar,***
- ***inhabit similar environments.***



*Did I get my long legs
from our great, great,
great, grandmother?*

*Or did our legs come
from having to run
fast?*

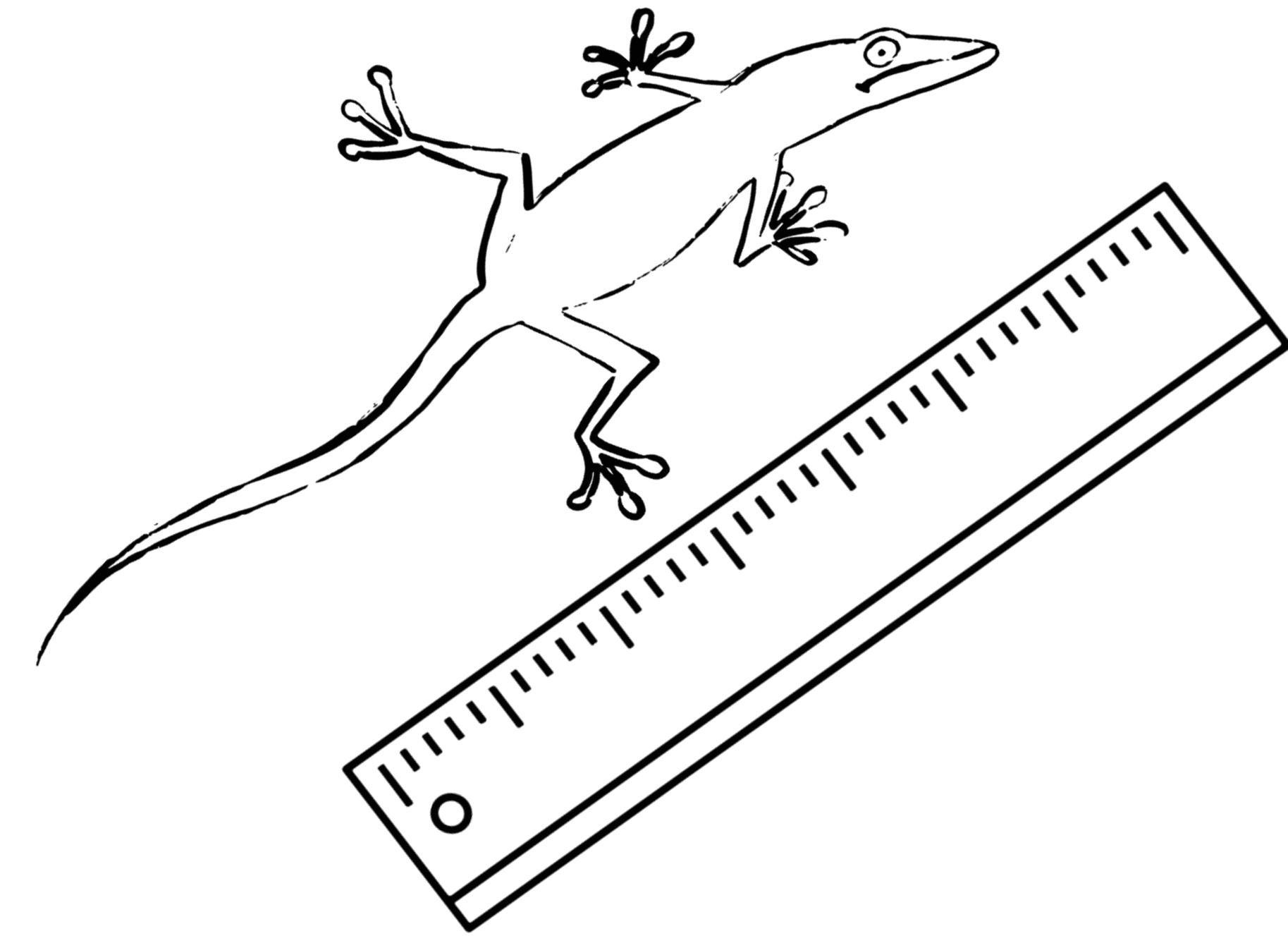
How do they determine this?



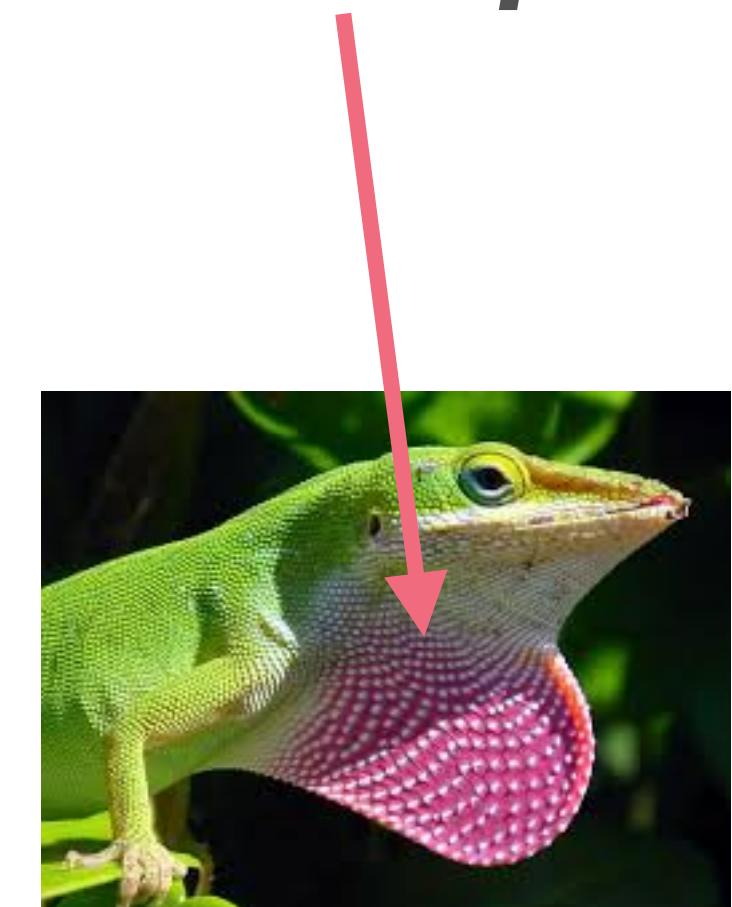
To understand the evolutionary process, they reconstruct histories of traits and map them to a tree.

Take samples of existing species

Continuous:
*Tail length,
Snout length,
Body size*

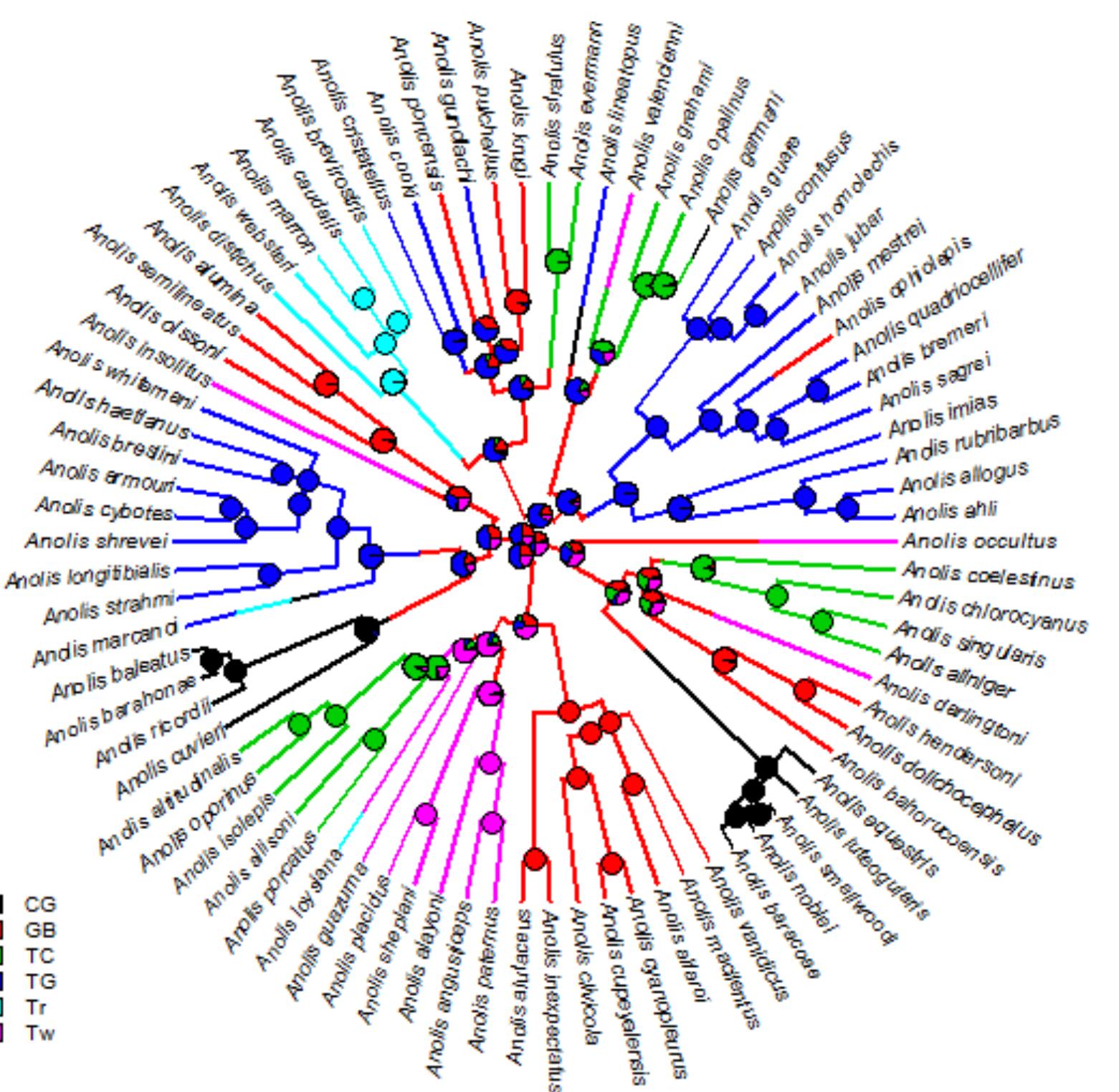


Discrete:
*Island of origin,
Do they lay eggs,
Do they have a
dewlap?*

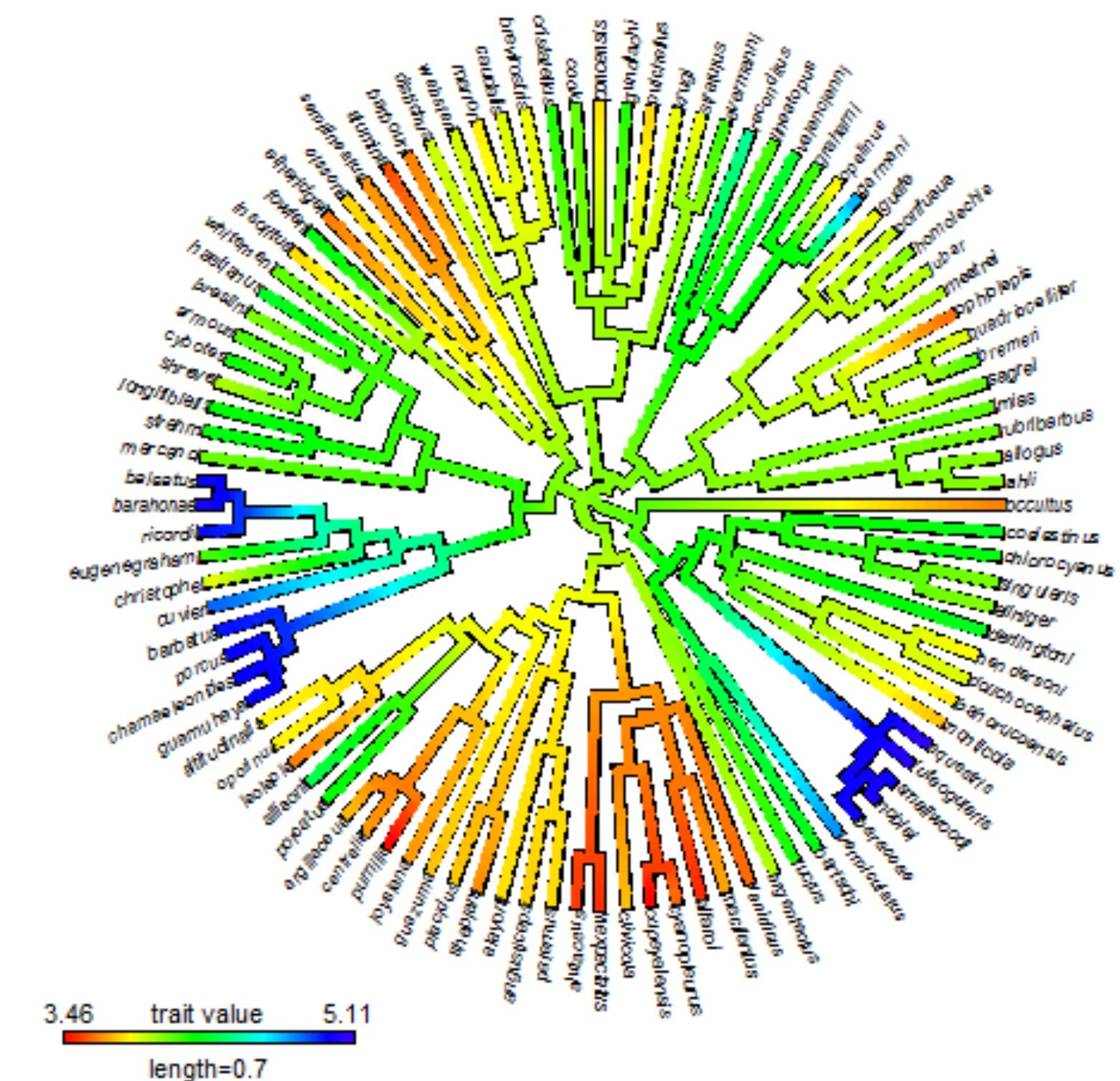


How they currently represent reconstructed traits

(these are from the Anole lizards)



Discrete trait: lizard group



Continuous trait: snout ventral length

Task Characterization:

Domain Requirements

- 1) *Understand how and when traits evolved.***
- 2) *Identify patterns in the trait data***

Task Characterization:

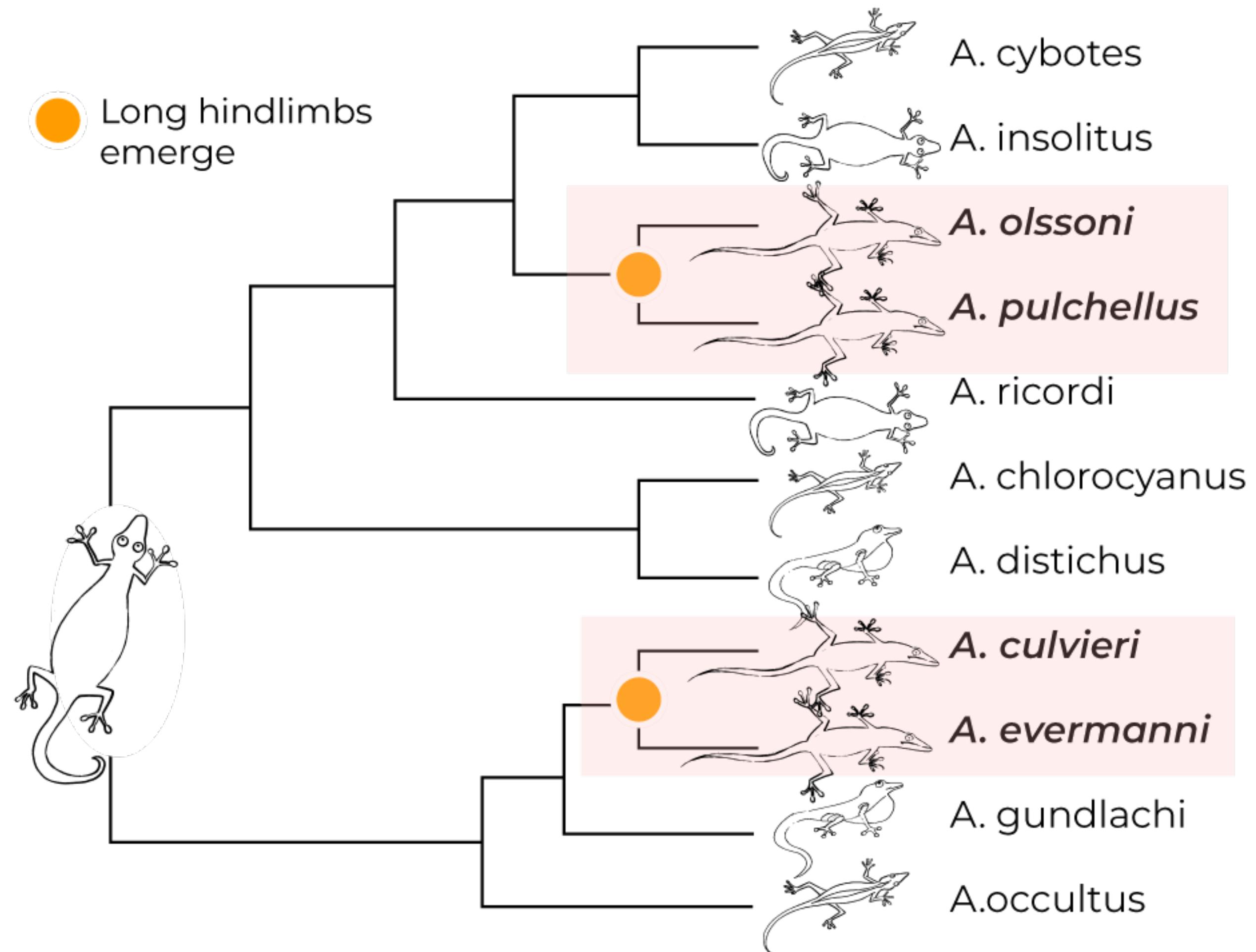
Understand how and when traits evolved

- Identify trait changes along species' histories
- Identify differences in distributions across subtrees
- Visualize the uncertainty of estimated trait values in species' histories.
- Identify evolutionary outliers.

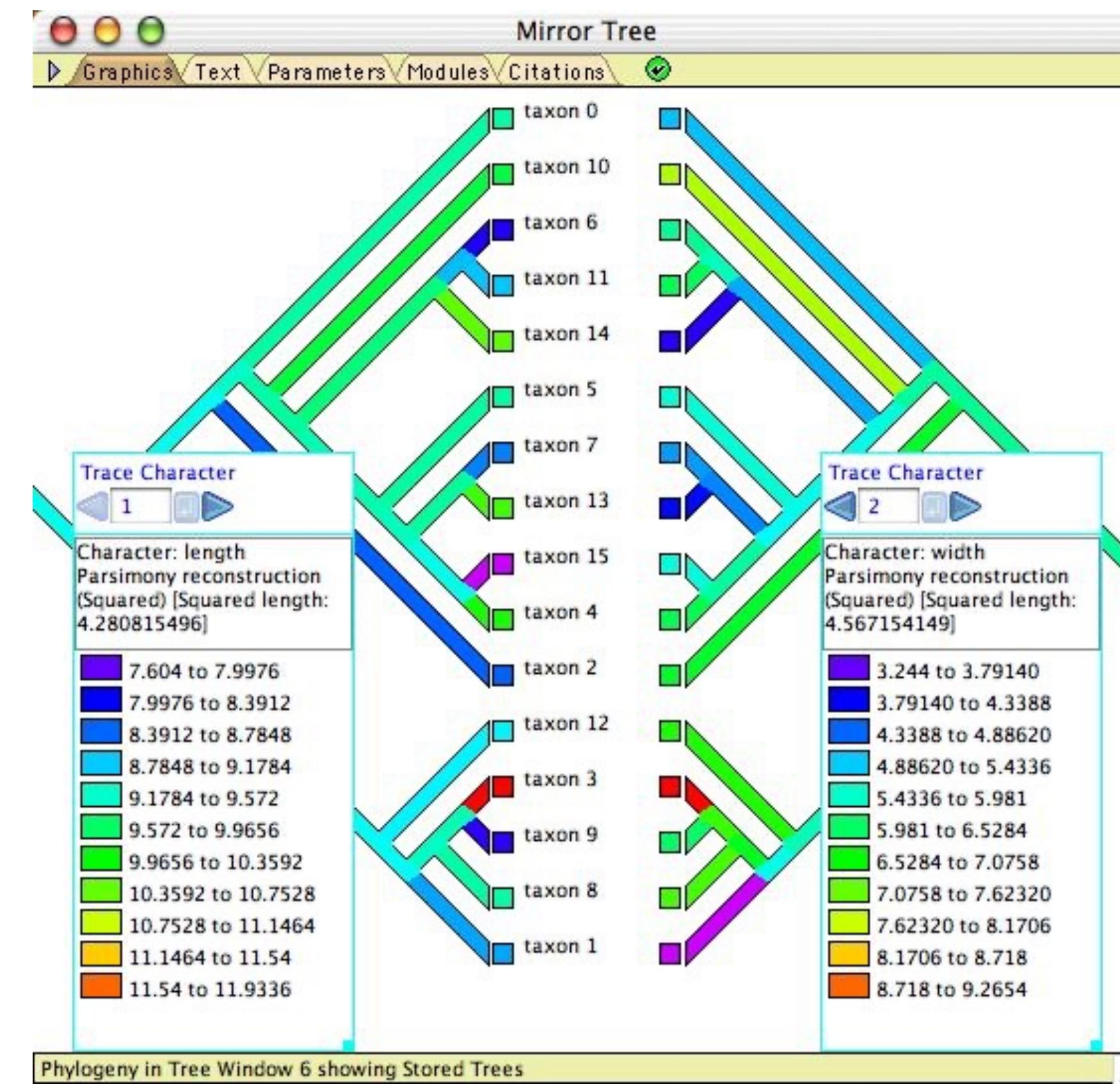
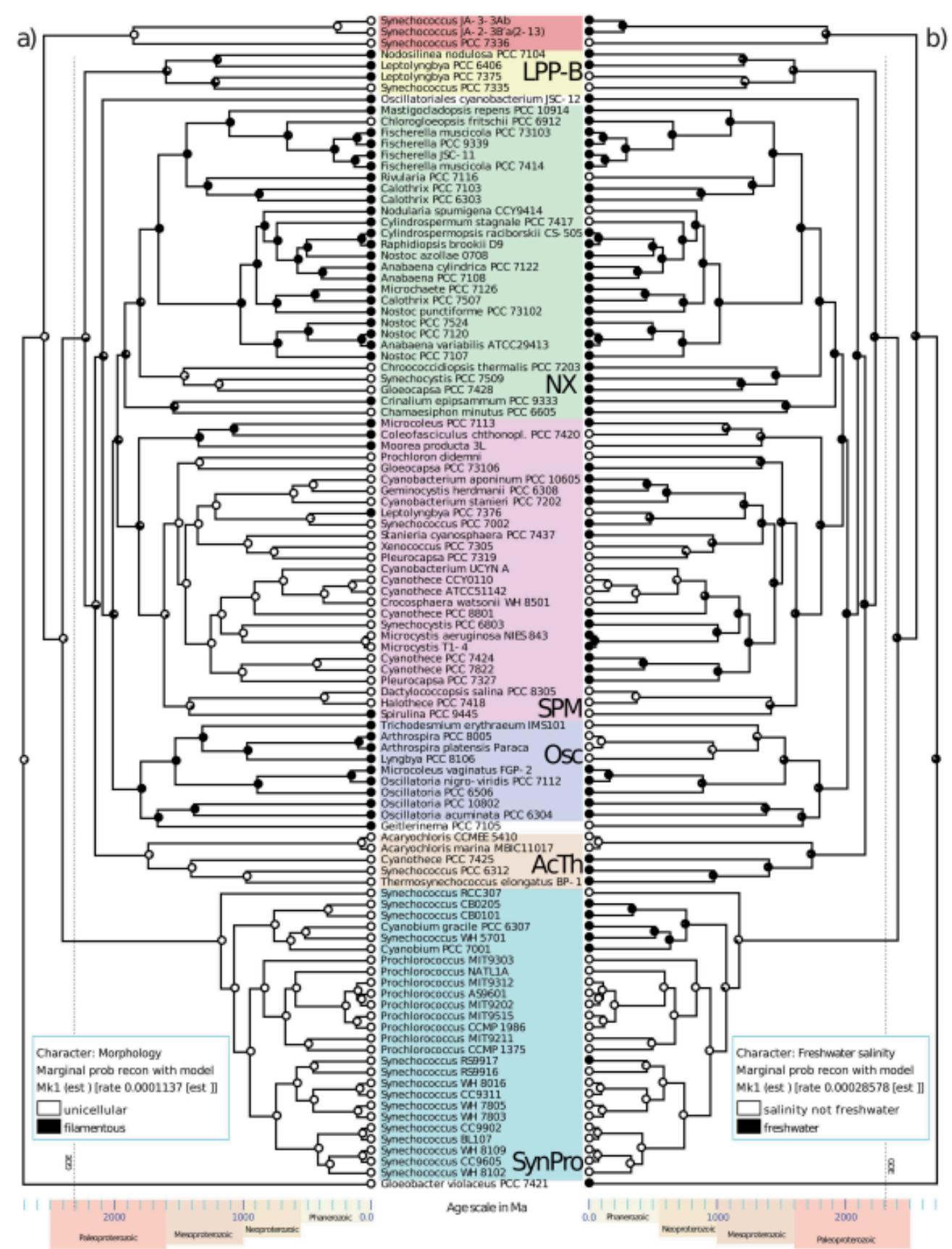
Task Characterization:

Identify patterns in the trait data

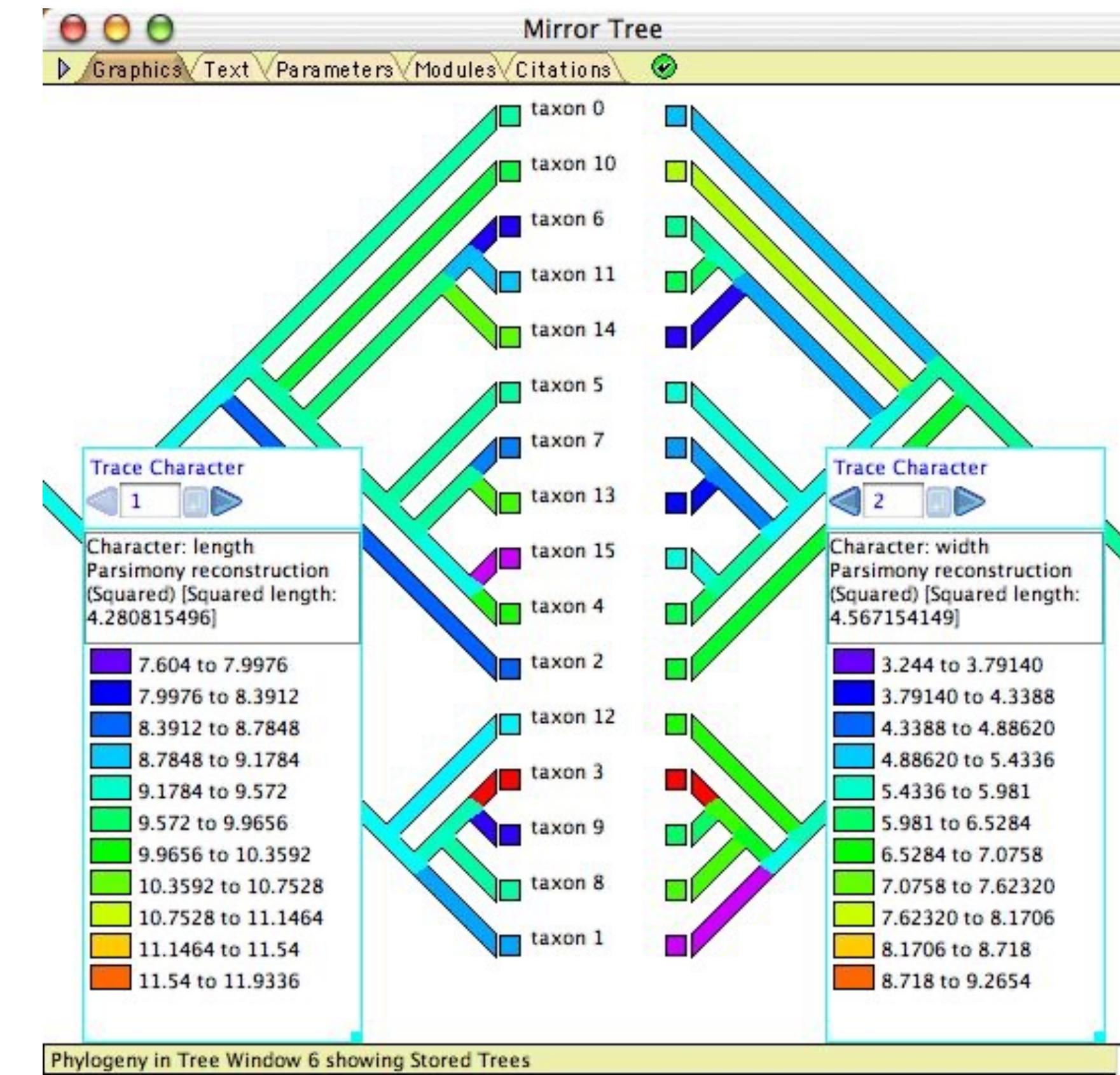
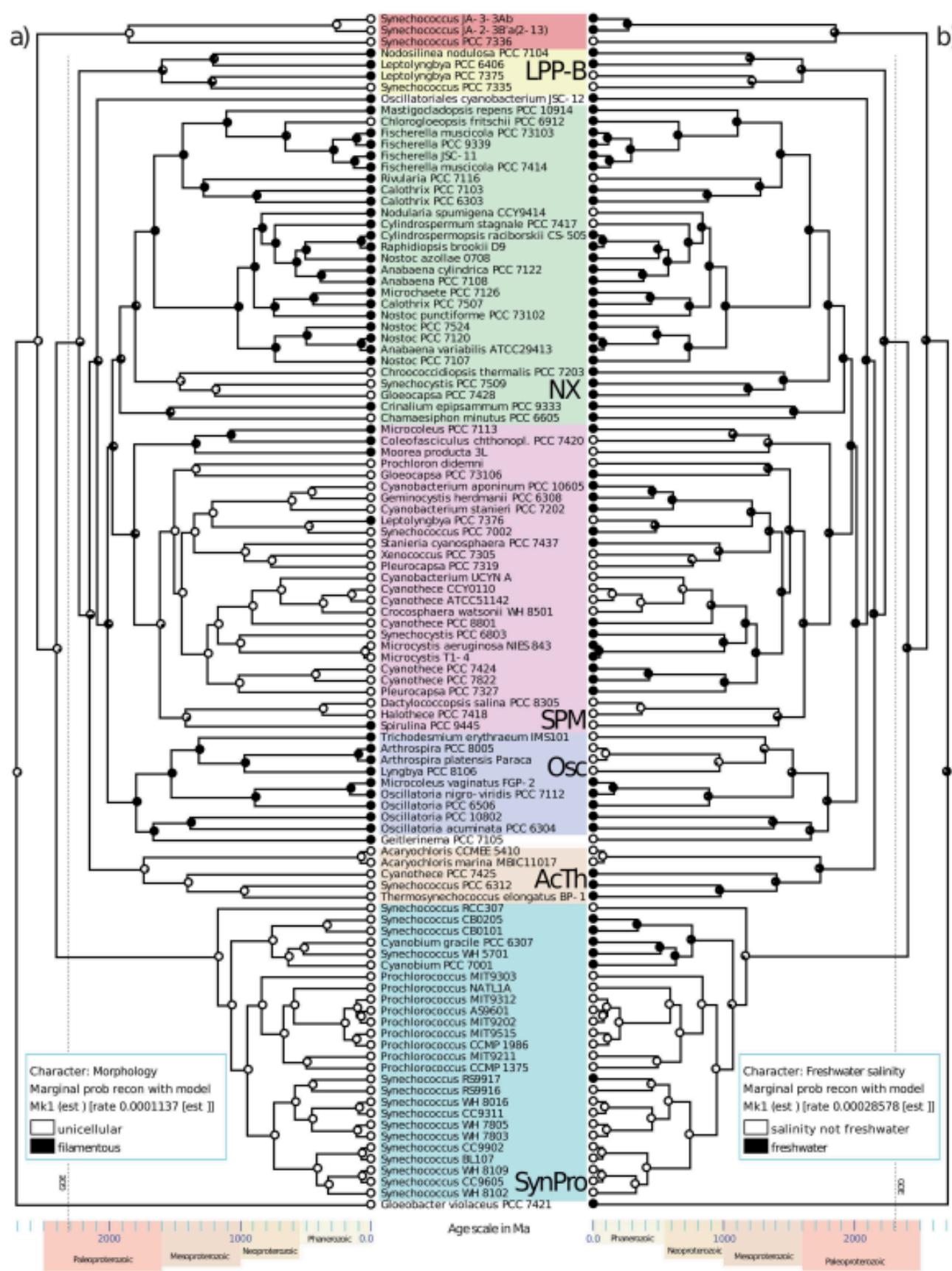
- identify nodes that are separated in the network but share a common state.
- Compare path histories for multiple attributes at once.
- Identify patterns between multiple paths



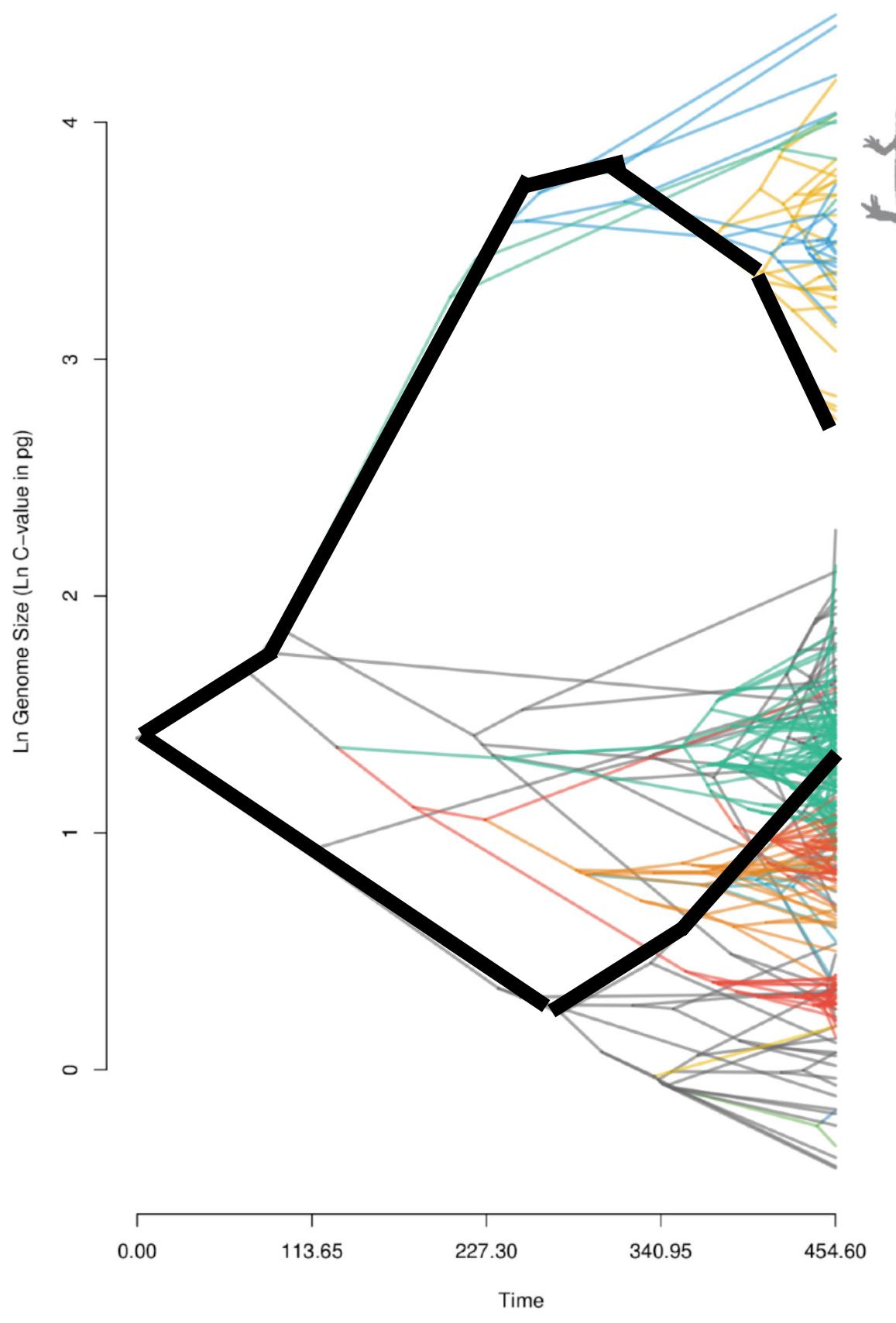
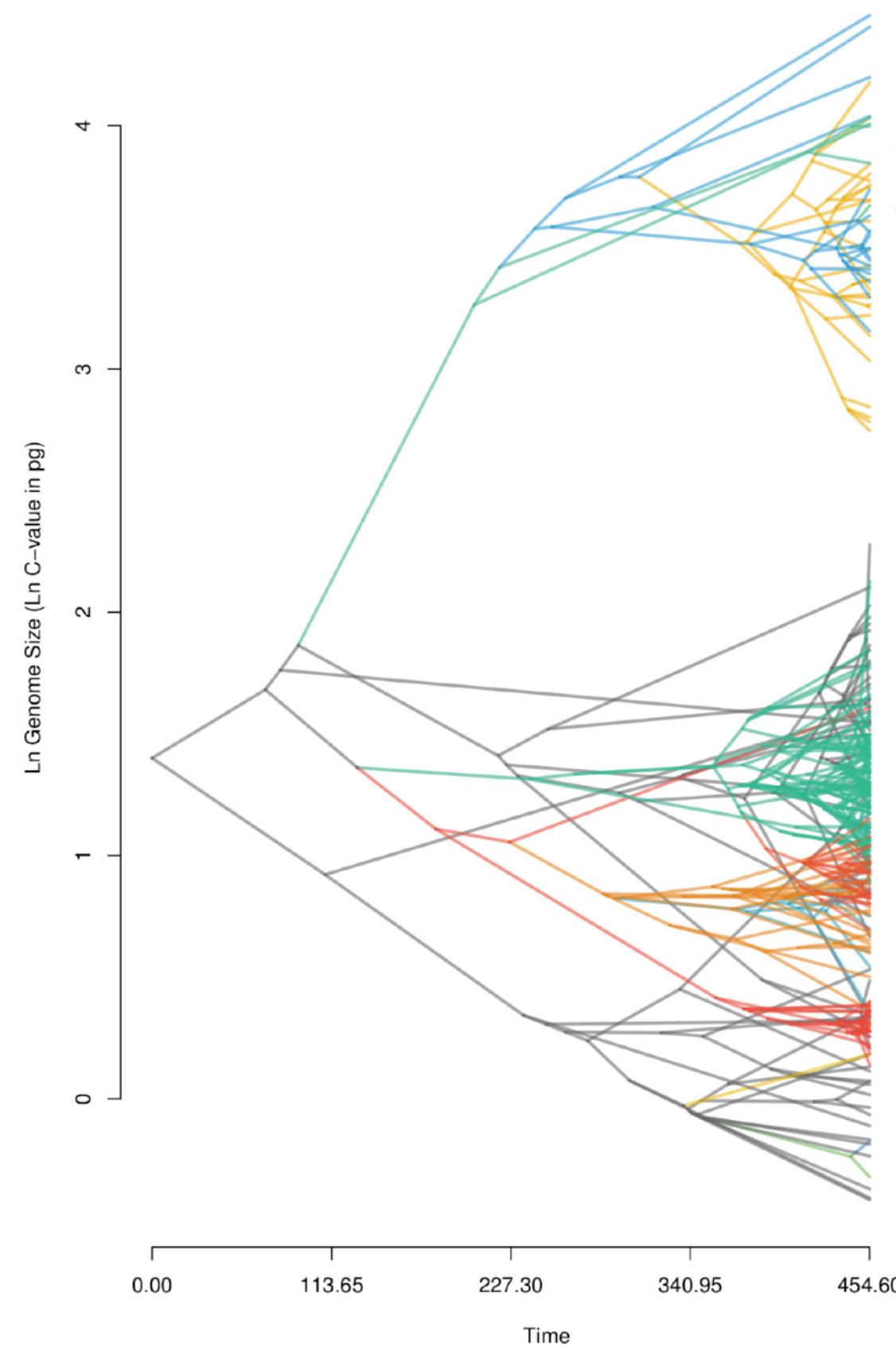
**Often compare multiple traits.
Currently do this by mirroring trees with different attributes.**



**“If you have one continuous trait you can do things.
If you have two - ok. If you have three or four or five,
there is nothing really sufficient”**



View trends in trait change in the context of the topology

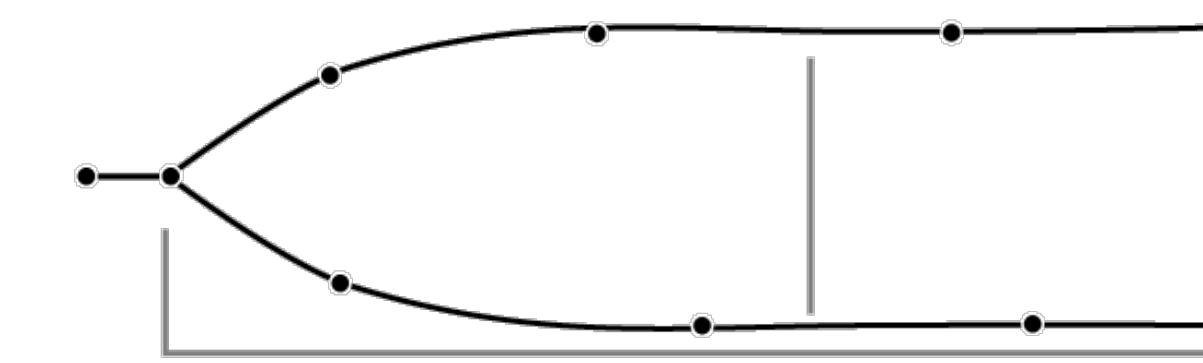
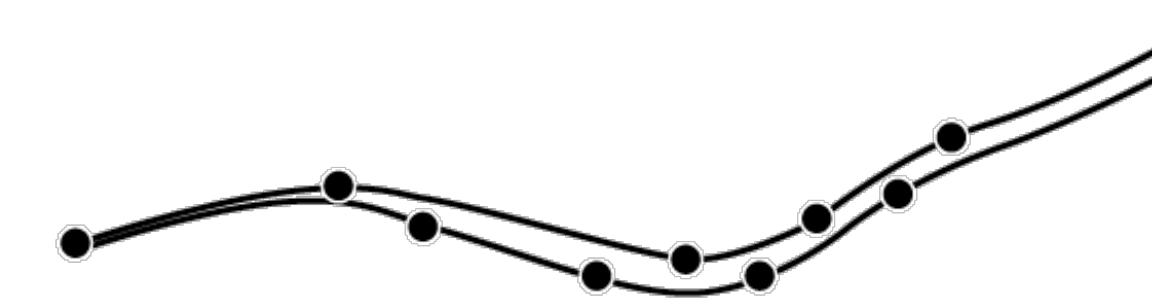
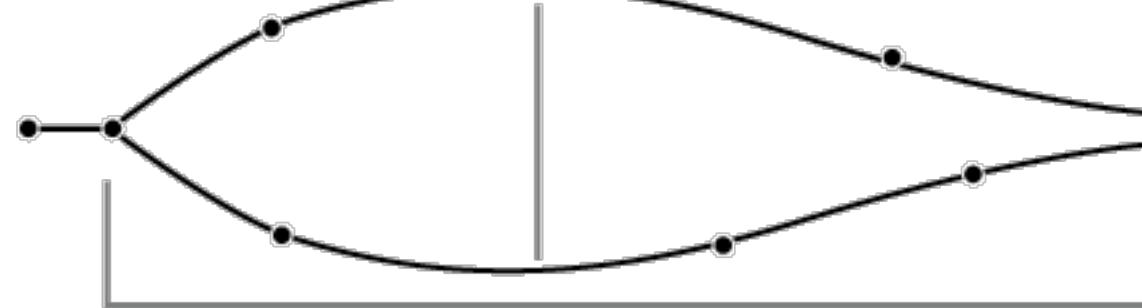


Patterns that indicate evolutionary events can be hard to see

Trends obscured

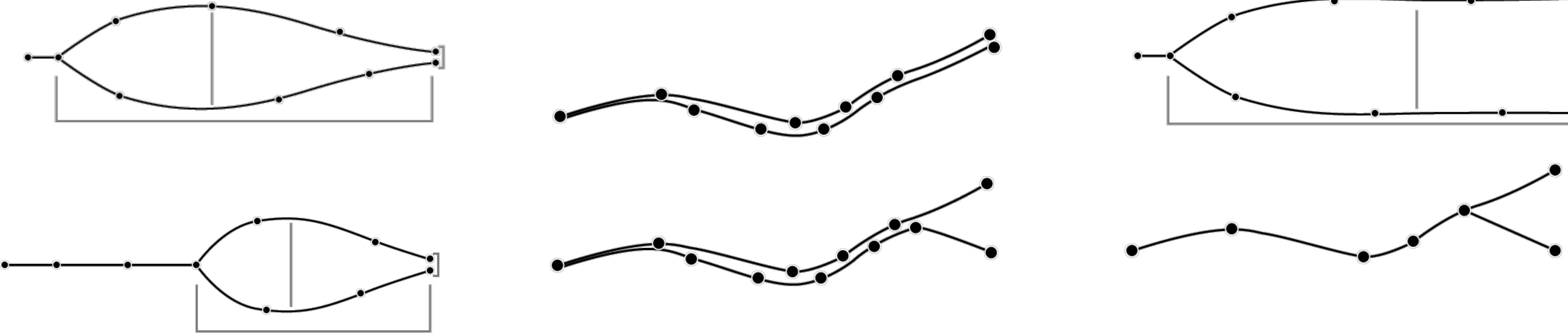
Hard to isolate individual pairs

Known patterns of interest:



How do you explain these?

Known patterns of interest:



How do you explain these?

Visualization:

- ***View multiple attributes under uncertainty in the context of topology***
- ***Identify patterns in paths along a tree based on attributes and topology***

DATA: ISLAND-ANOLIS ▾

SUMMARY VIEW ▾

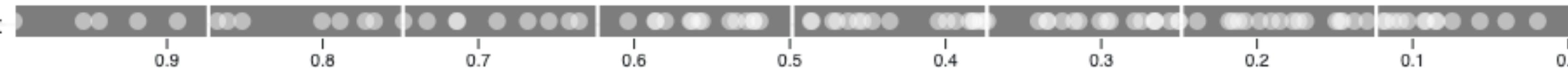
SHOWN ATTRIBUTES ▾

SHOWN SUBTREES ▾

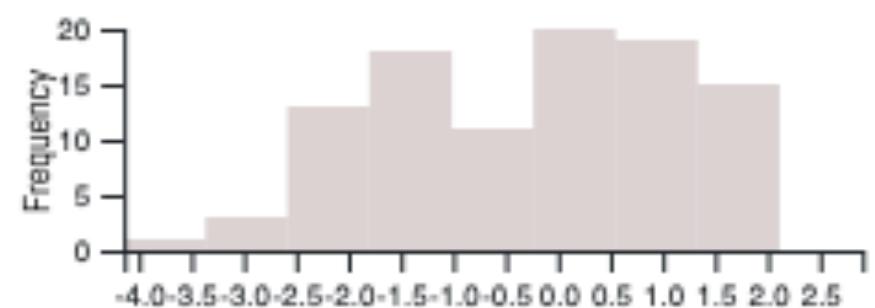
LOG SCALE

All island-anolis : 100 Paths

Root



Leaves



PCIII_padwi...

PCII_head

PCIV_lamell...

PCI_limbs

SVL

attitude

awesomeness

View multiple attributes under uncertainty in the context of topology

ecomorph

ecomorph(CG)

ecomorph(GB)

ecomorph(T)

0

1

0

1

1

0

0

1

1

0

0

1

1

0

0

1

1

0

1

0

1

0

1

0

1

0

1

0

1

0

1

0

1

0

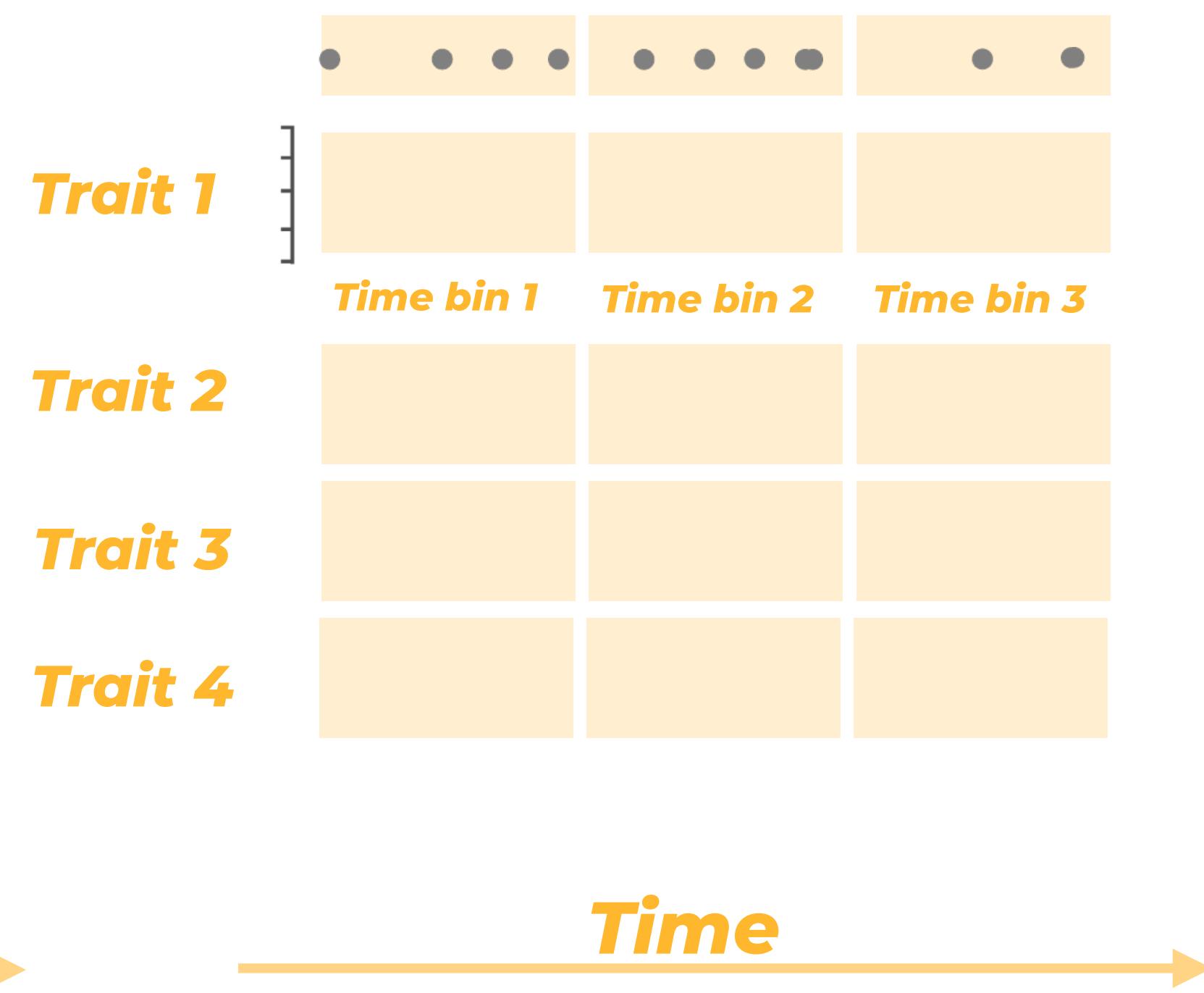
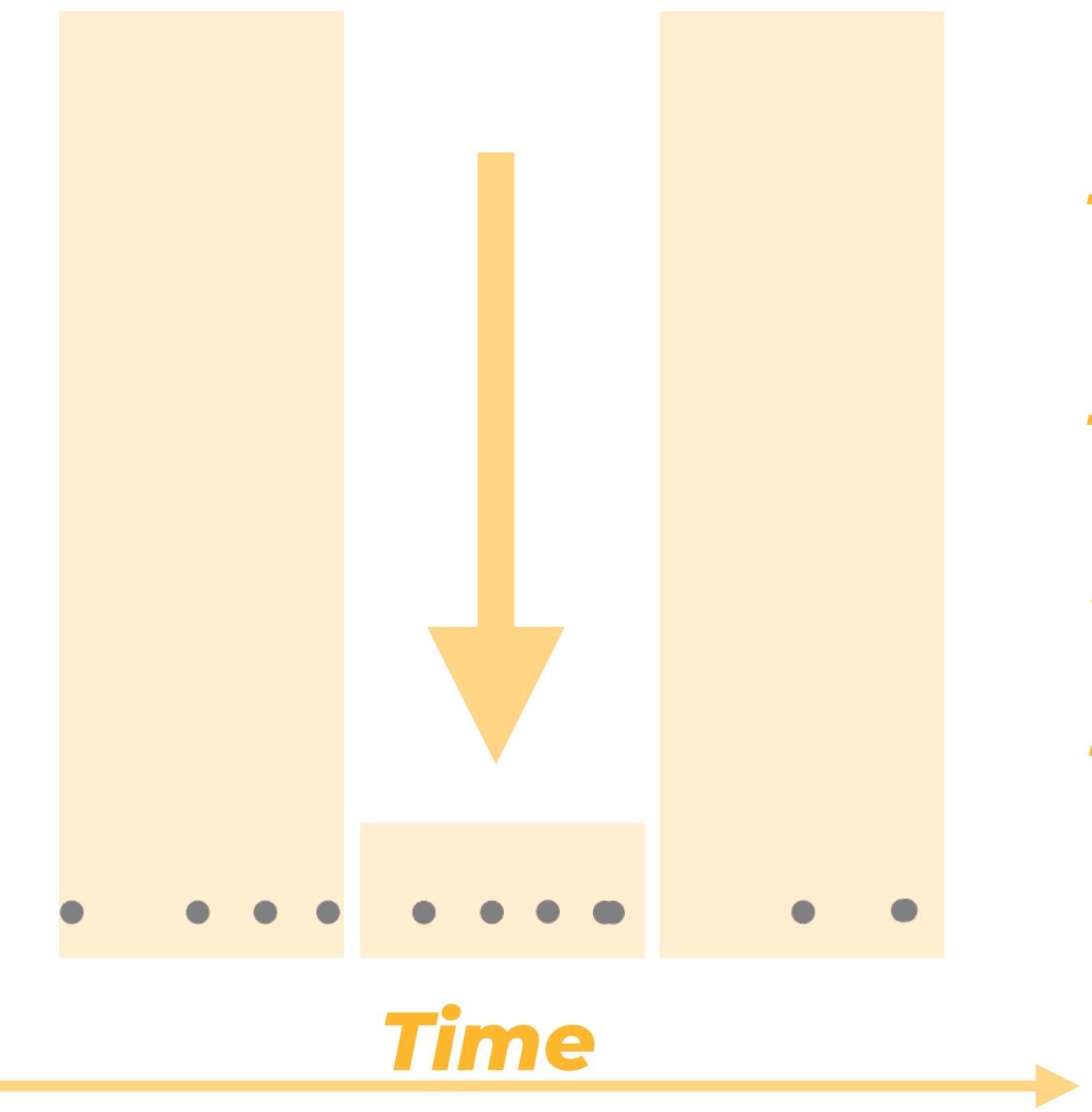
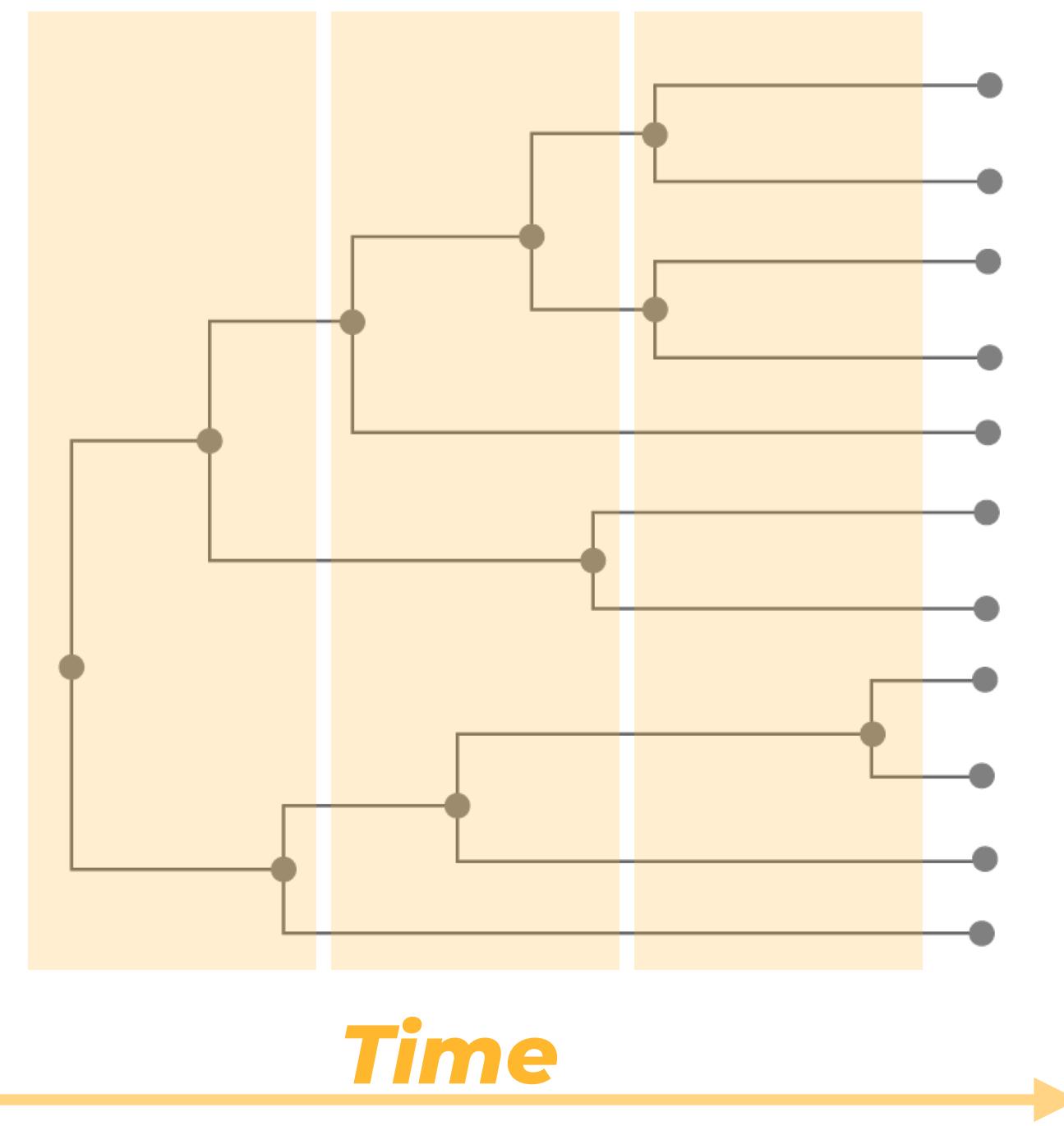
1

0

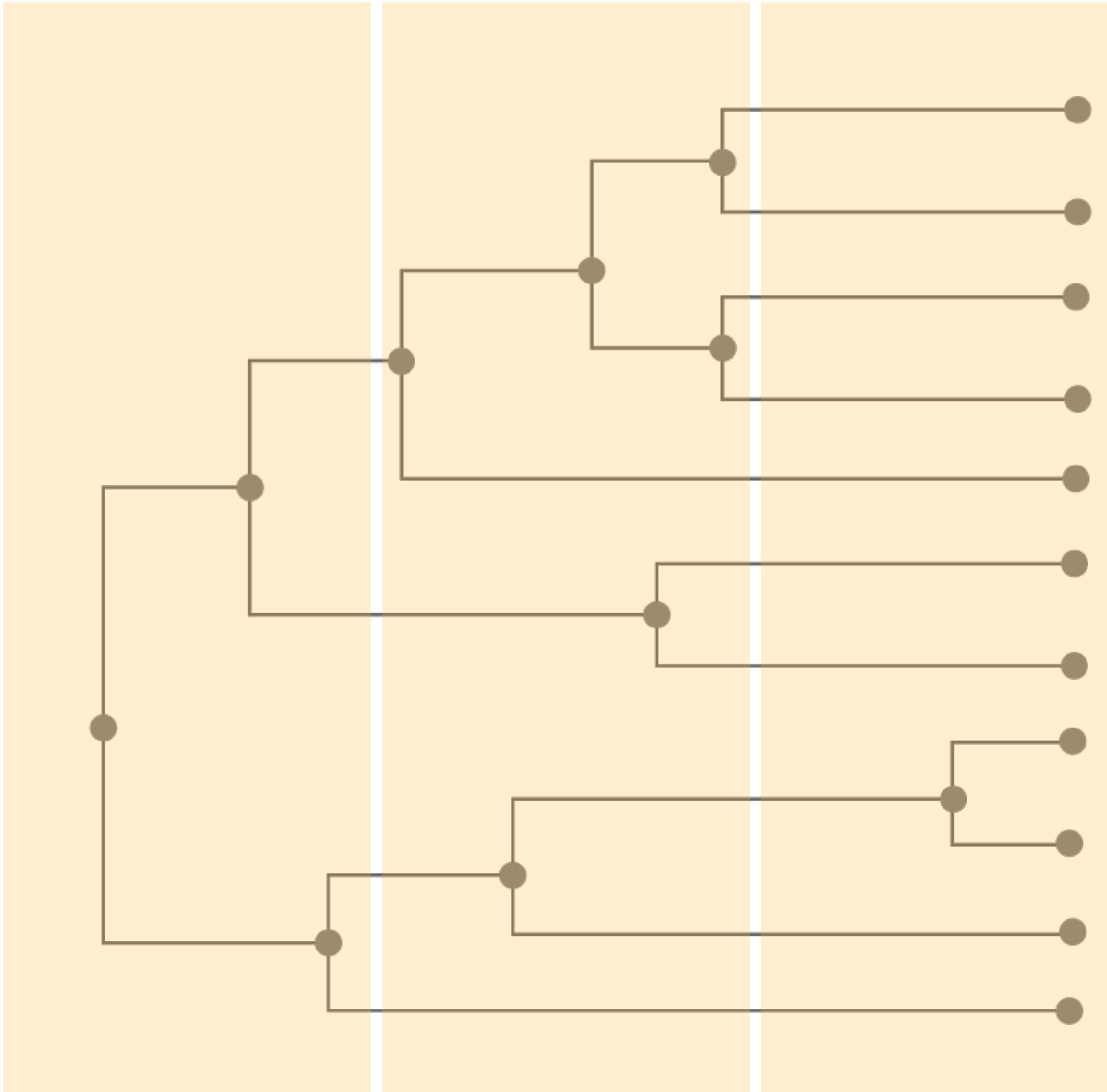
1

0

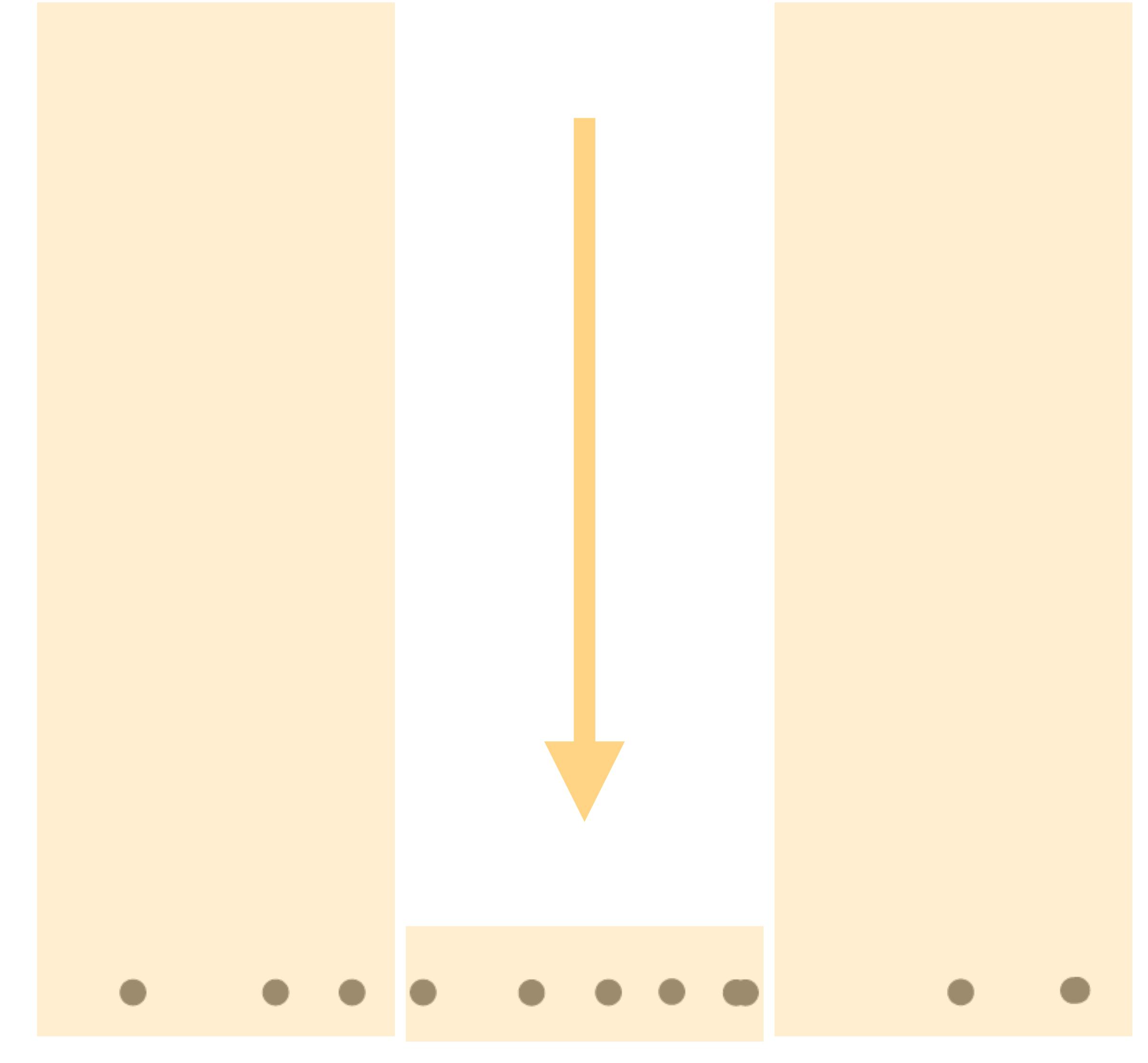




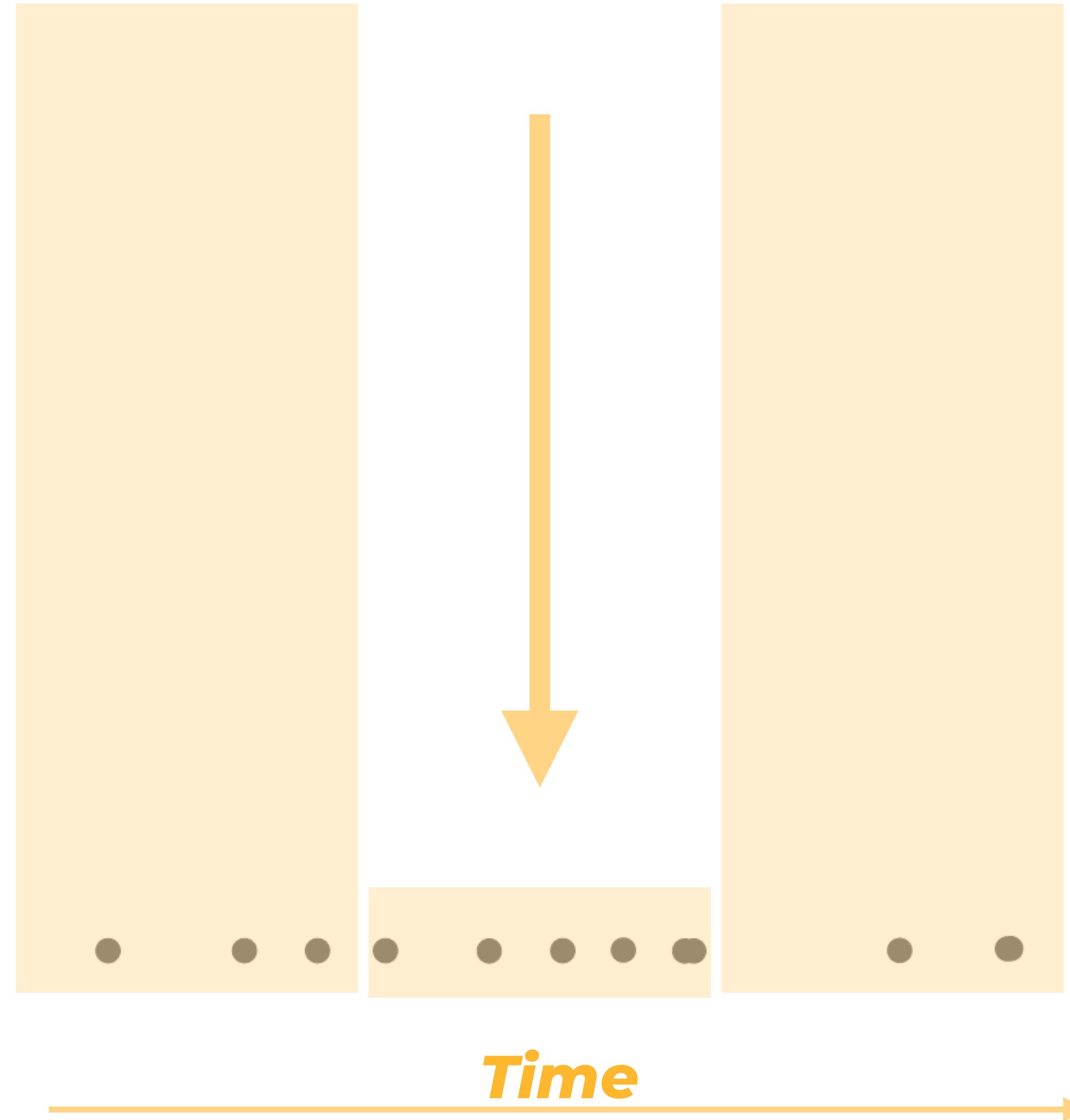
A.



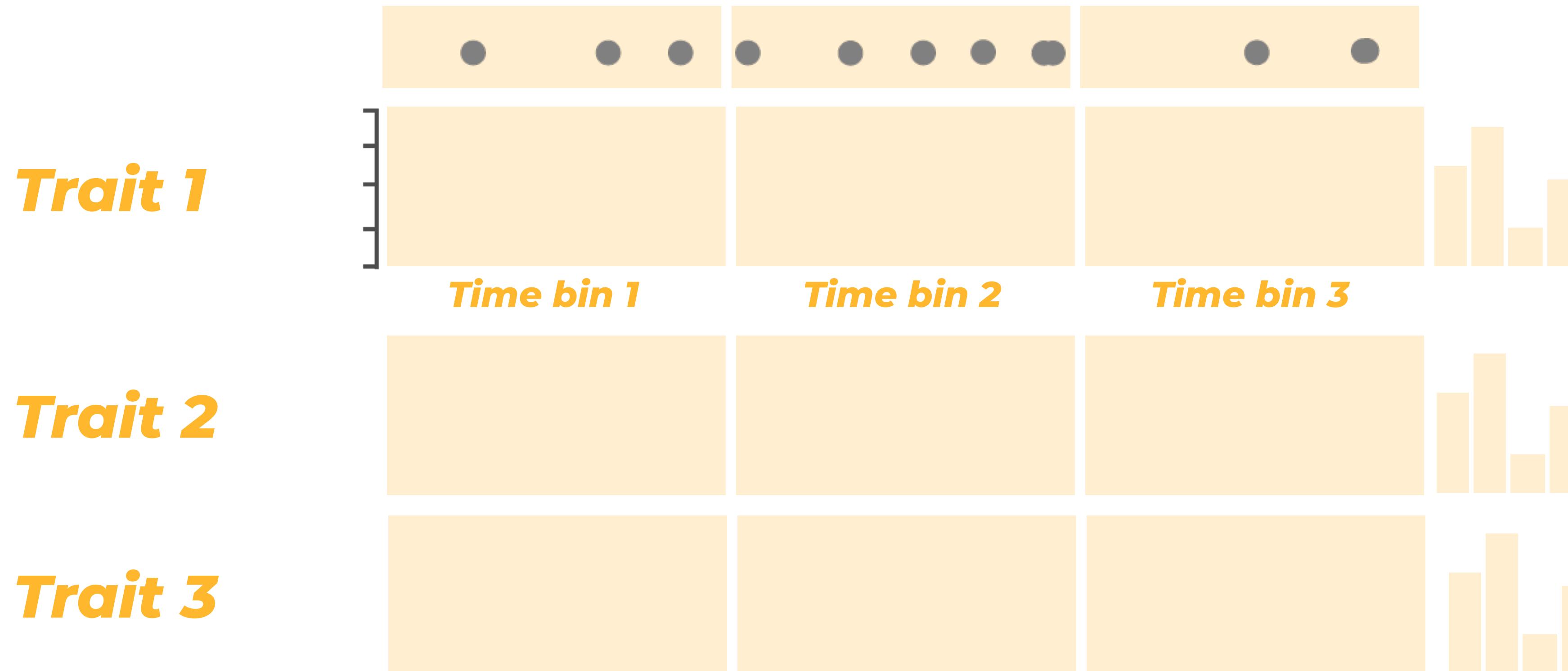
B.



Squished the tree down

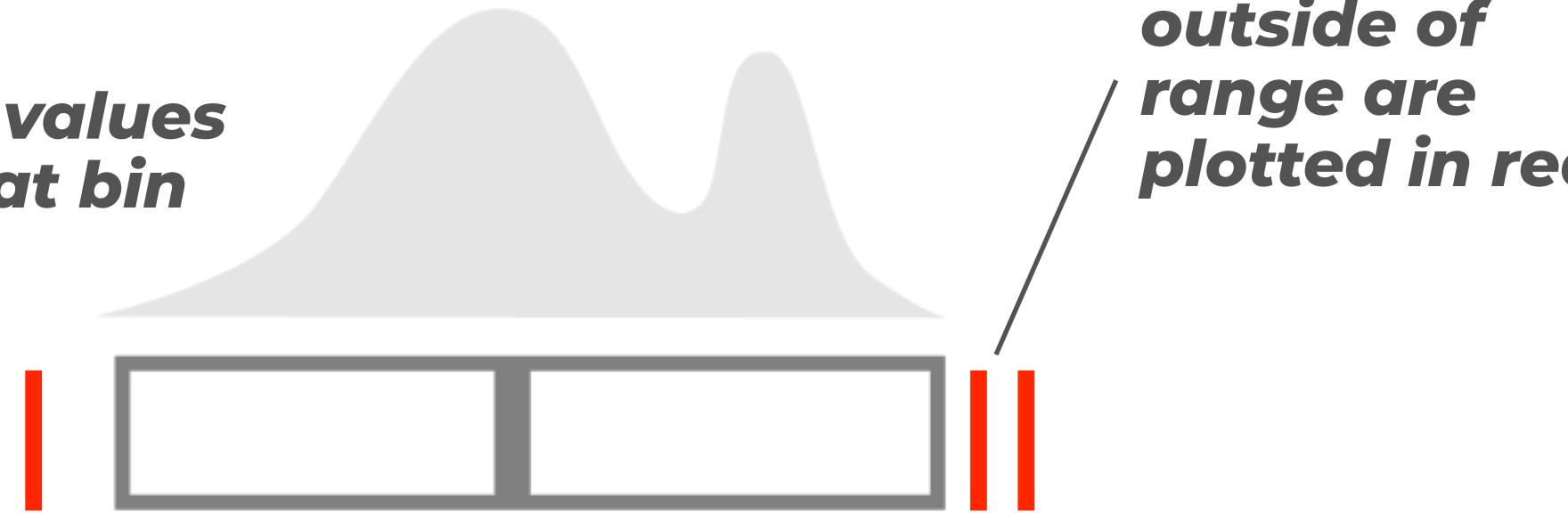


Each trait gets a row

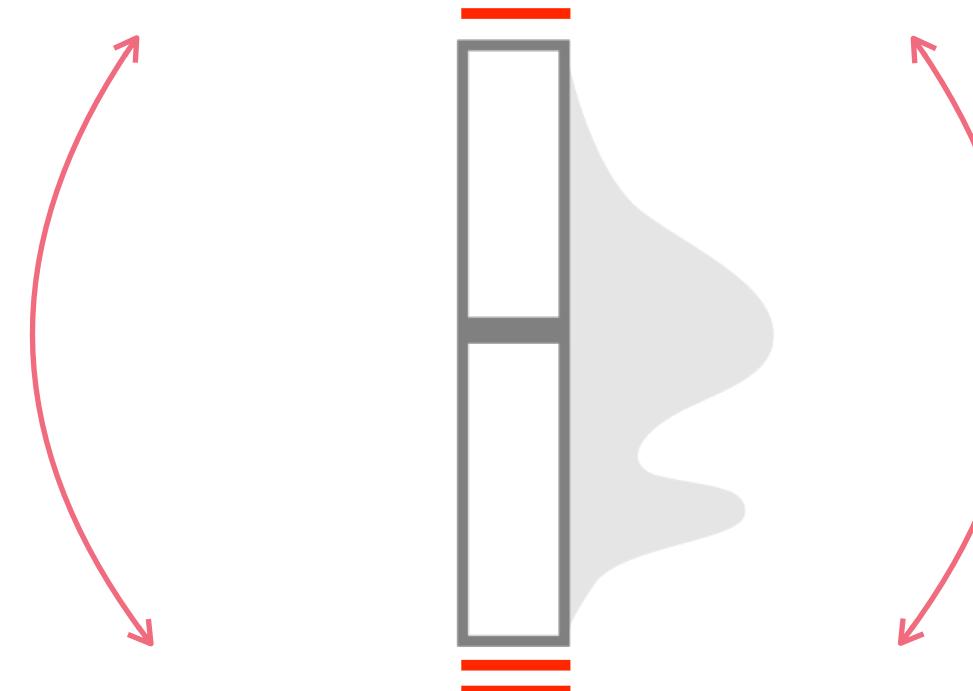


Continuous

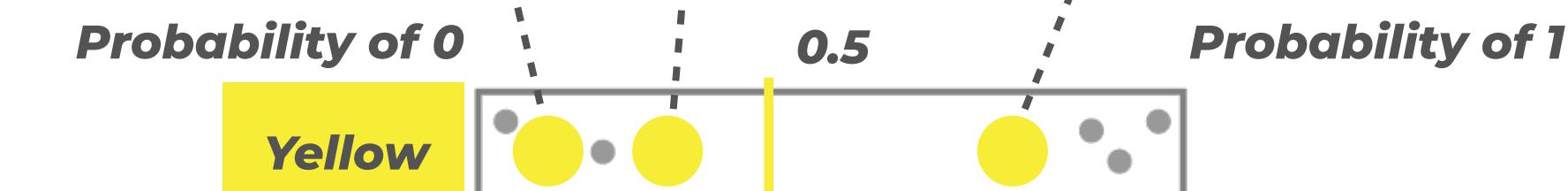
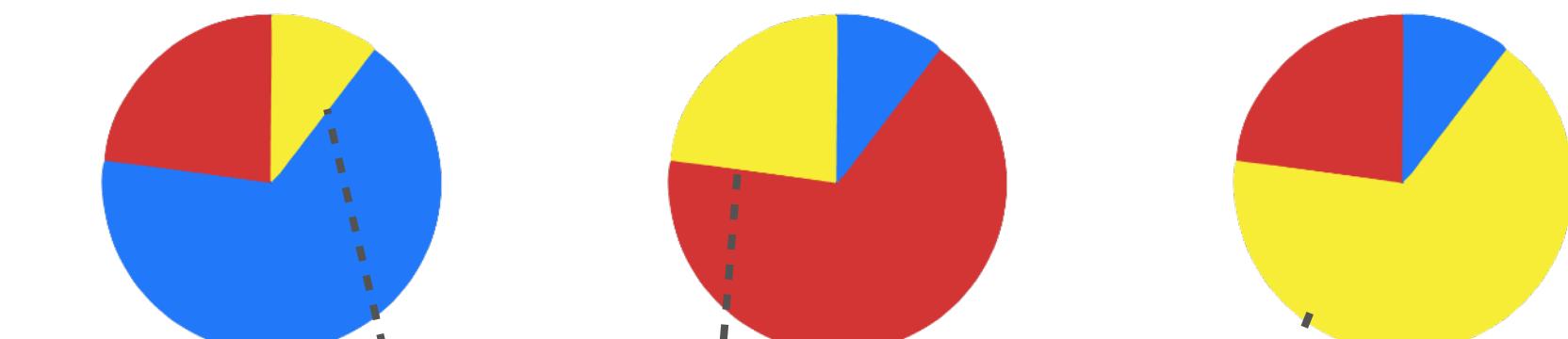
Distribution of values
for nodes in that bin



Mean and normal
distribution range



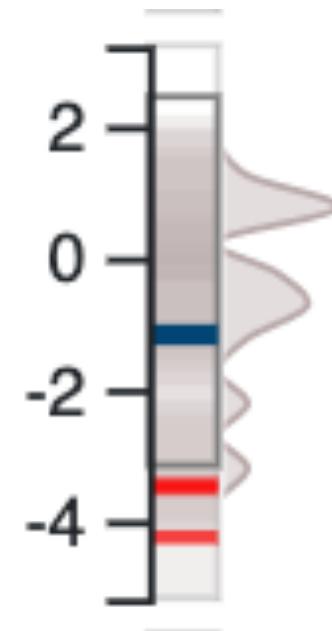
Discrete



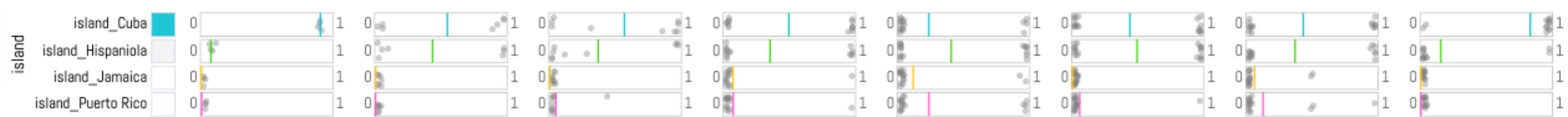
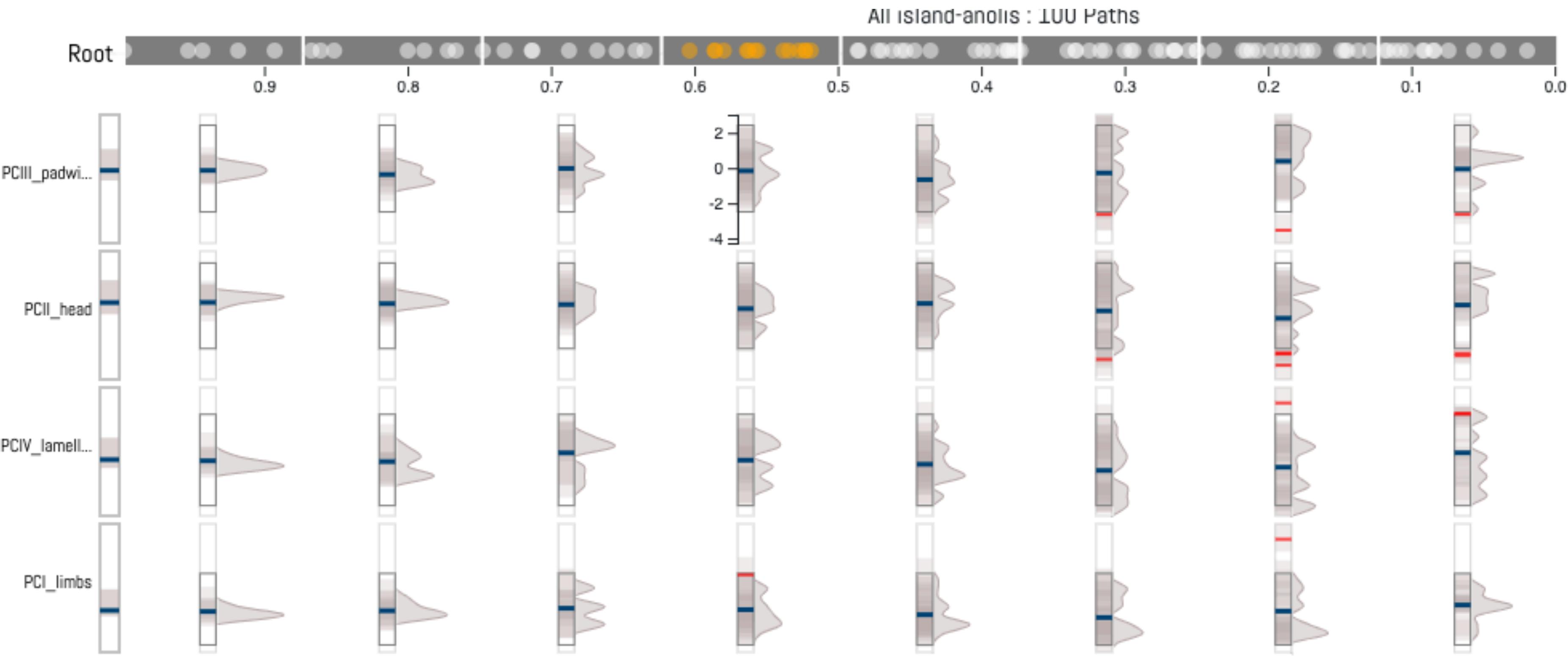
Internal node
probabilities that they
belong to that state
plotted as dots

Context for binned internal nodes and timeframe

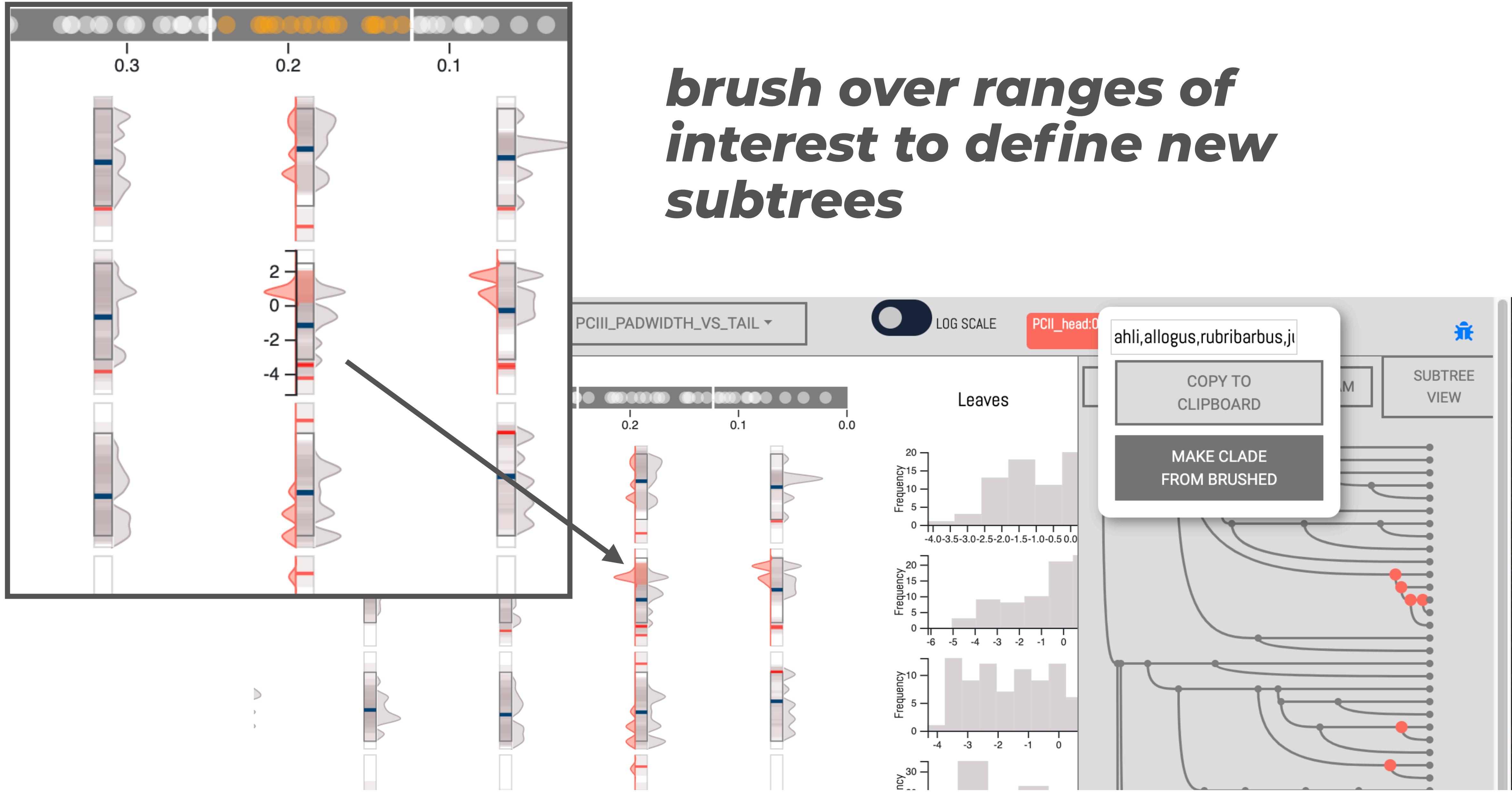
Continuous traits such as tail length and limb length



Discrete traits such as island



brush over ranges of interest to define new subtrees



DATA: ISLAND-ANOLIS ▾

SUMMARY VIEW ▾

SHOWN ATTRIBUTES ▾

SHOWN SUBTREES ▾

TRAIT: PCIII_PADWIDTH_VS_TAIL ▾

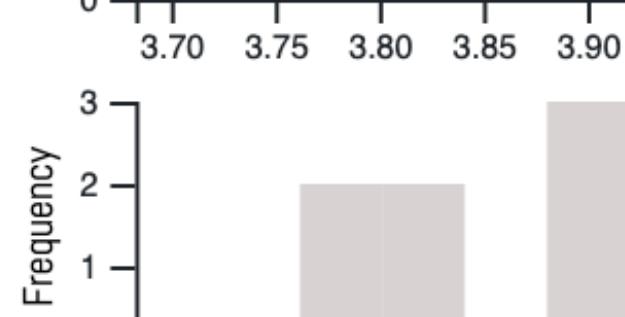
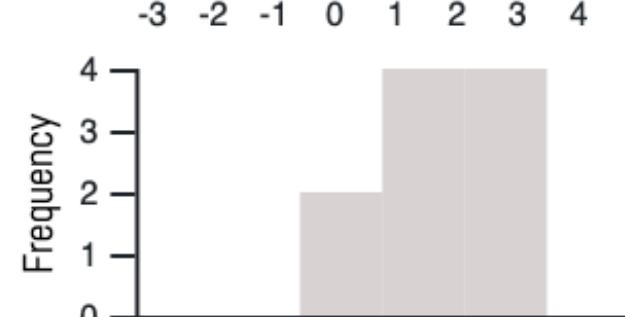
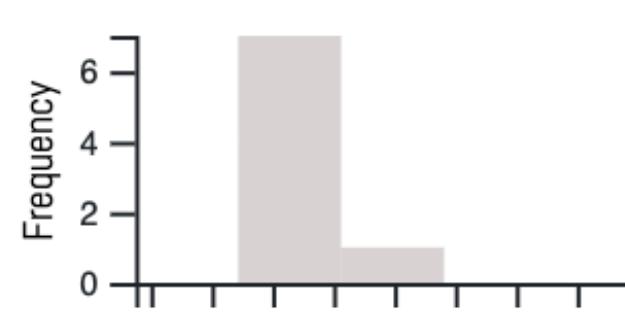
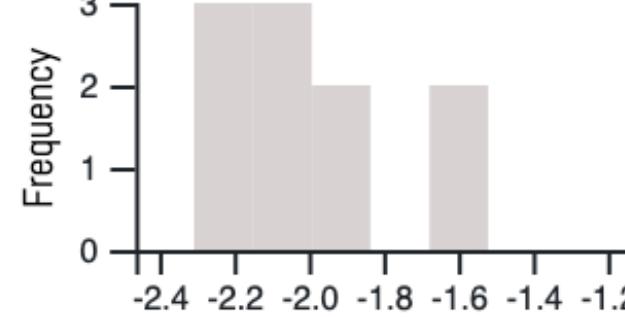
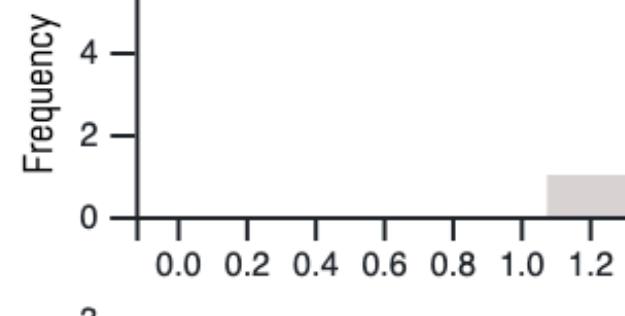
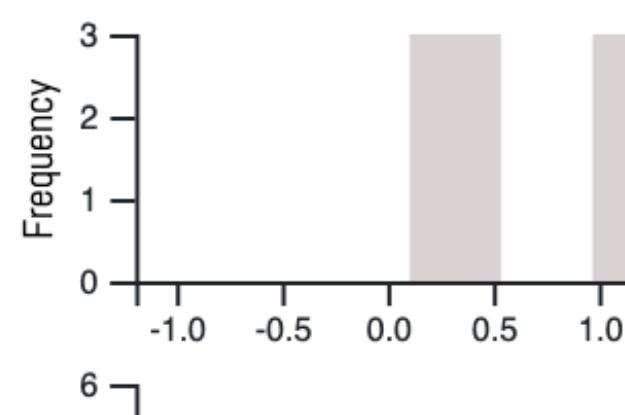
LOG SCALE

Group-Brush-10 : 10 Paths

Root



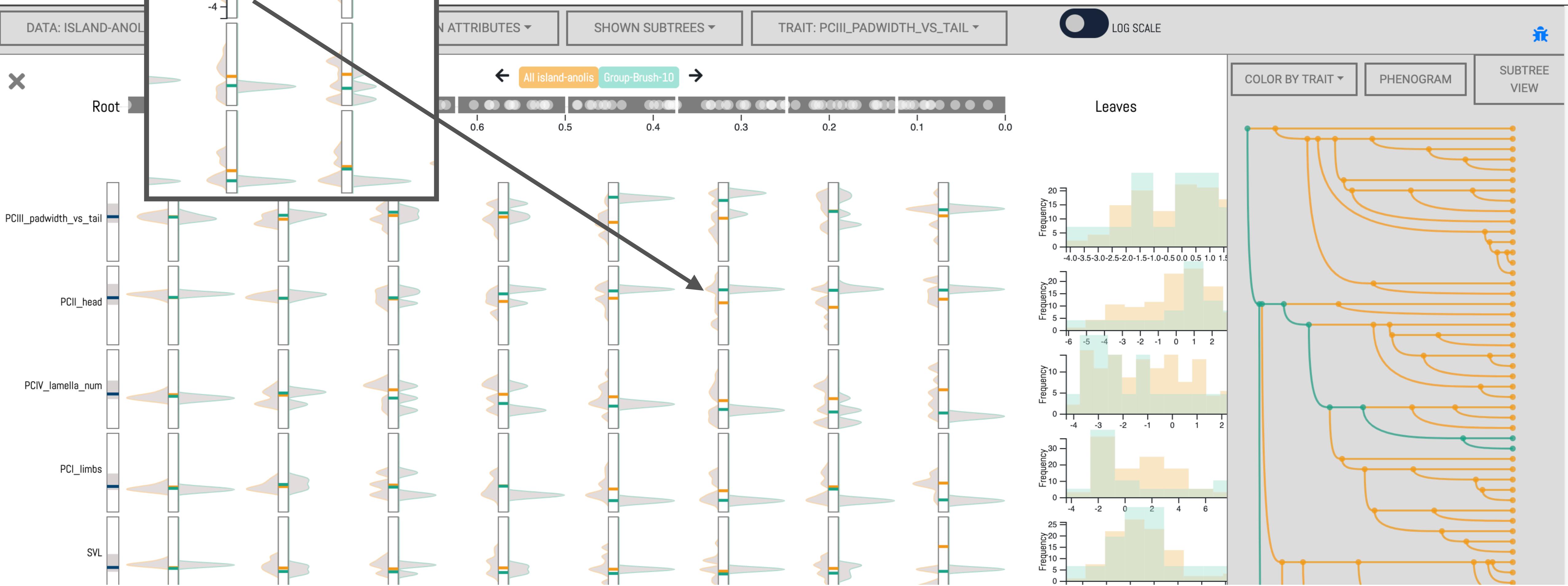
Leaves



**ADD TO
COMPARISON**

Group-Brush-10

Overlay trait value distributions between subtrees



DATA: ISLAND-ANOLIS

PAIR VIEW

SHOWN ATTRIBUTES

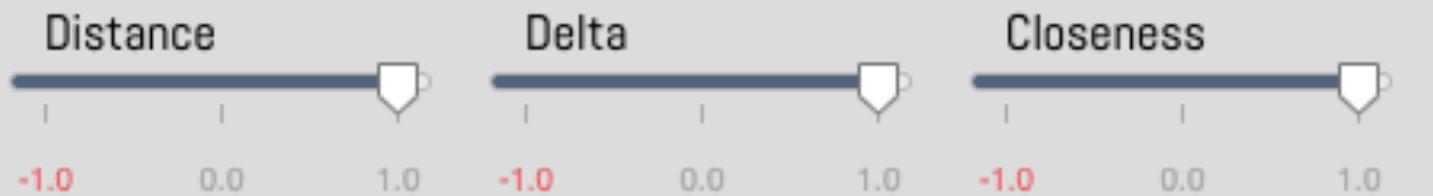
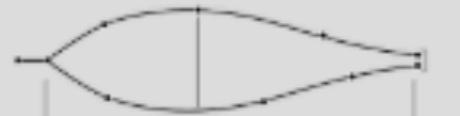
SHOWN SUBTREES

TRAIT: PCIII_PADWIDTH_VS_TAIL

LOG SCALE



CONVERGENCE (ANCIENT)



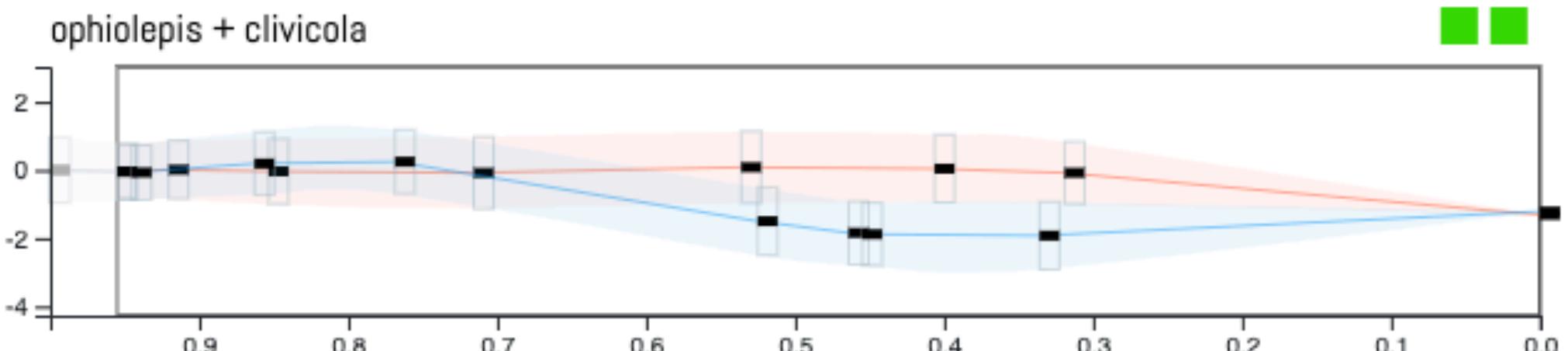
Trait: PCIII_padwidth_vs_tail
Pairs Shown: Top 1%
Num of Pairs: 50

ECOMORPH

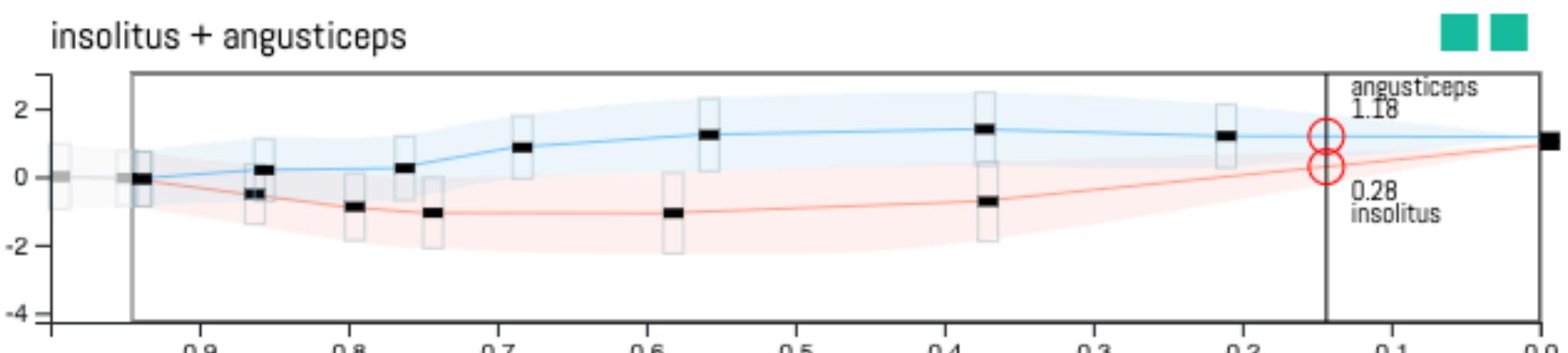
COLOR BY TRAIT

VIEW PHYLOGENY

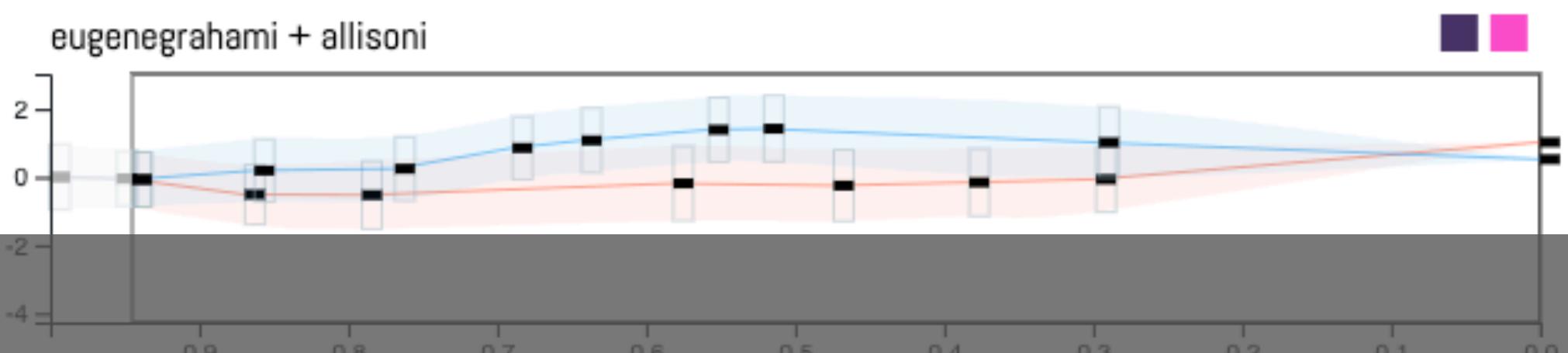
SUBTREE VIEW 33



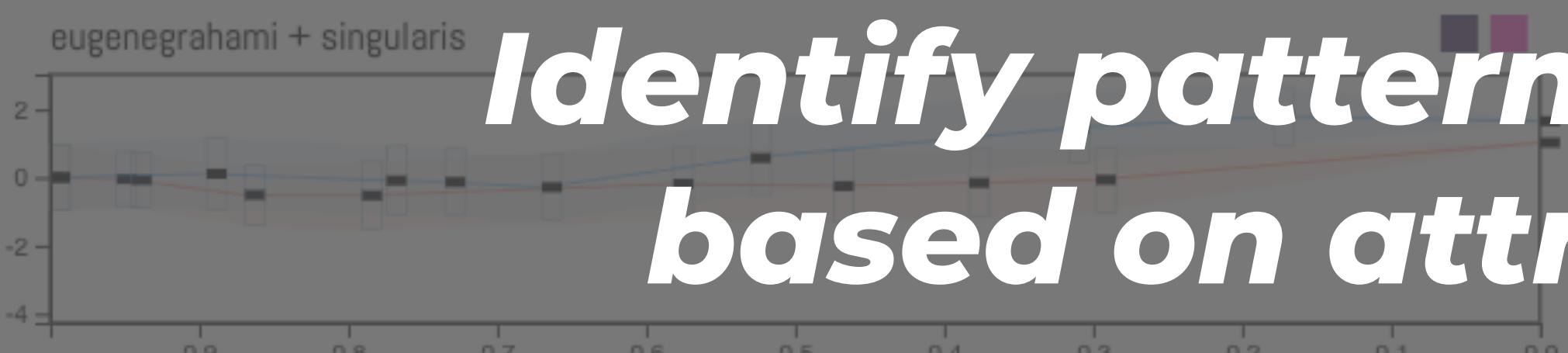
Ranked Top 1% in Other Traits



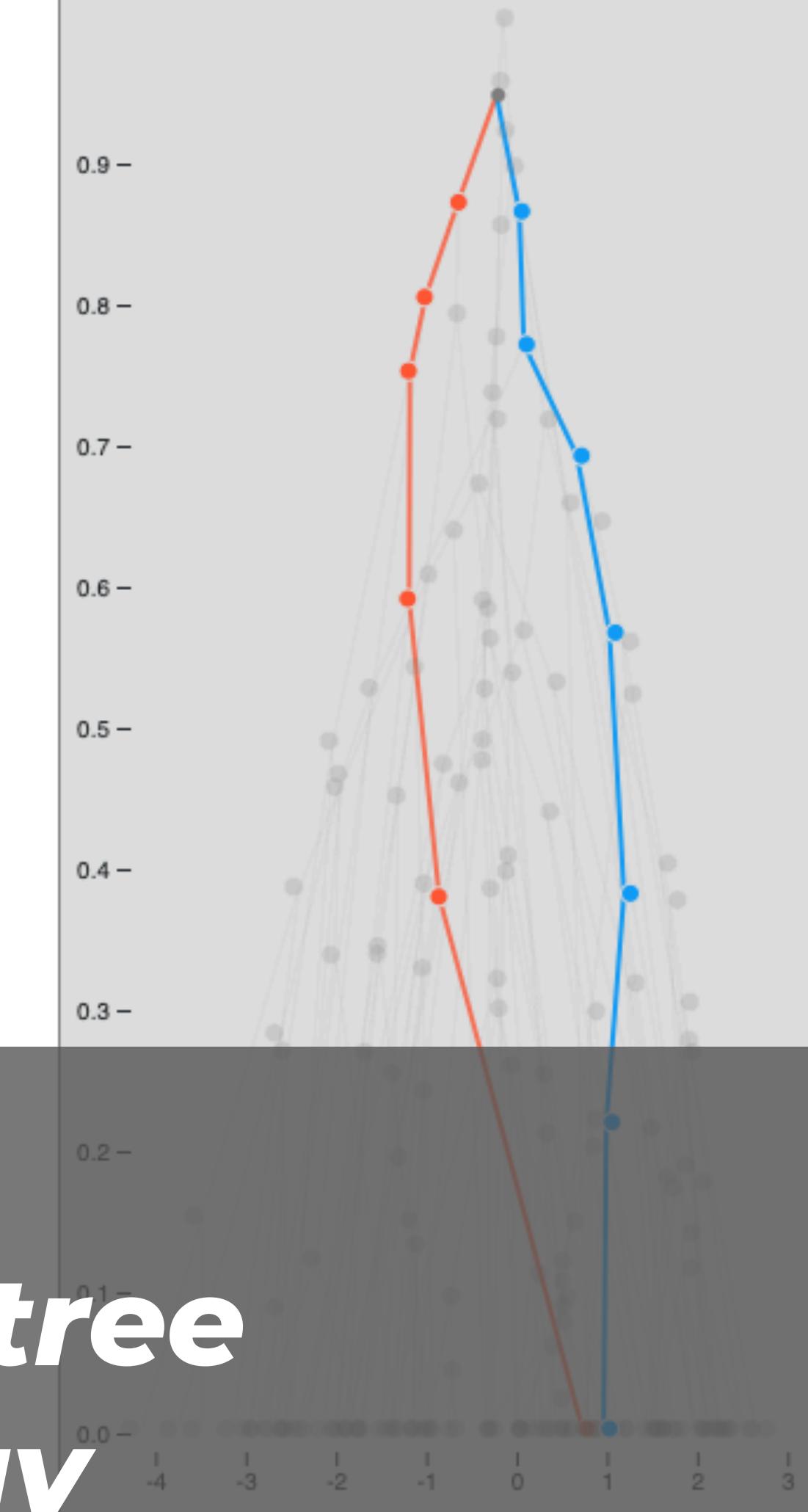
Ranked Top 1% in Other Traits



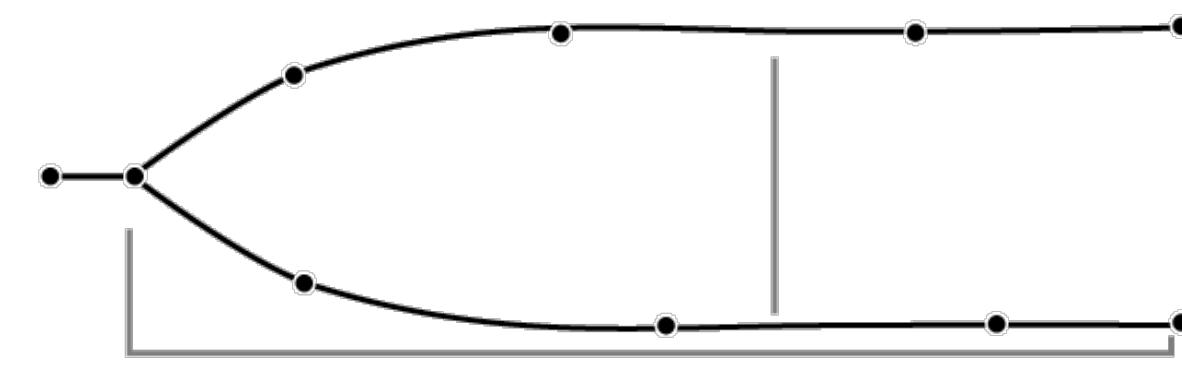
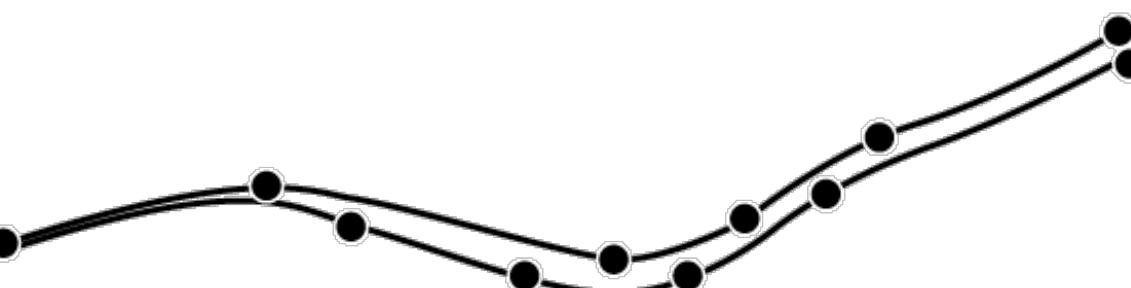
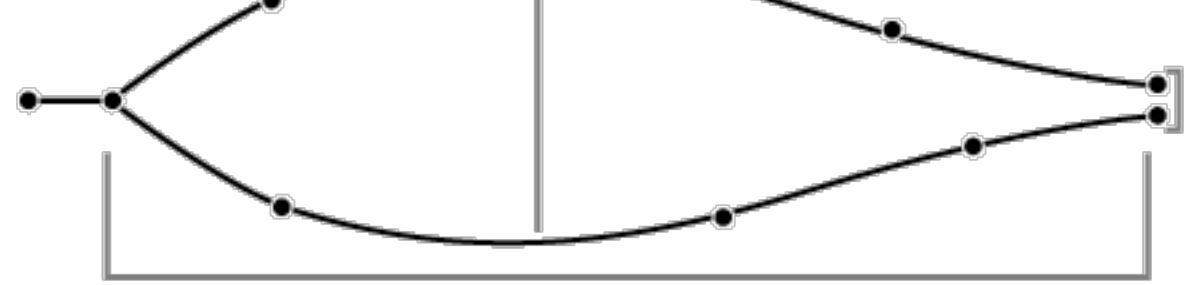
Ranked Top 1% in Other Traits



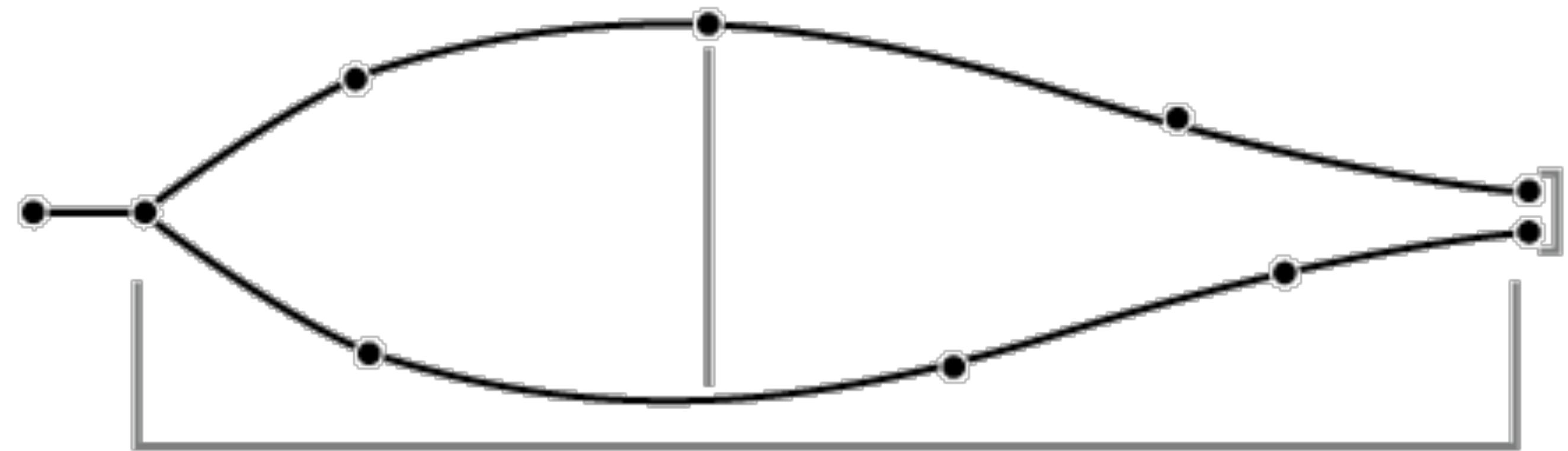
Ranked Top 1% in Other Traits



Identify patterns in paths along a tree based on attributes and topology



***Max difference
after split but
before tips***



***Value
difference
at tips***

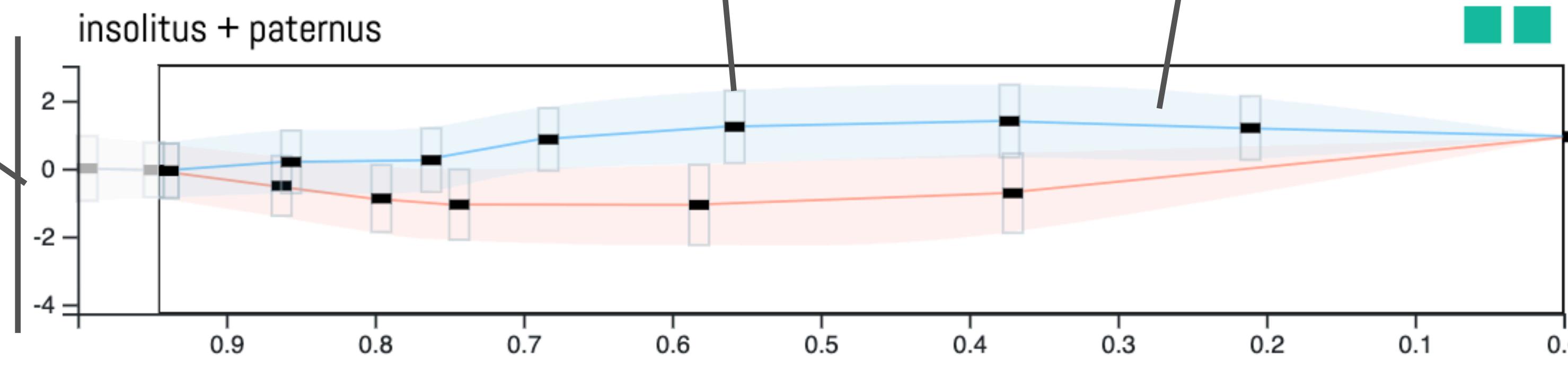
***Distance
from split***

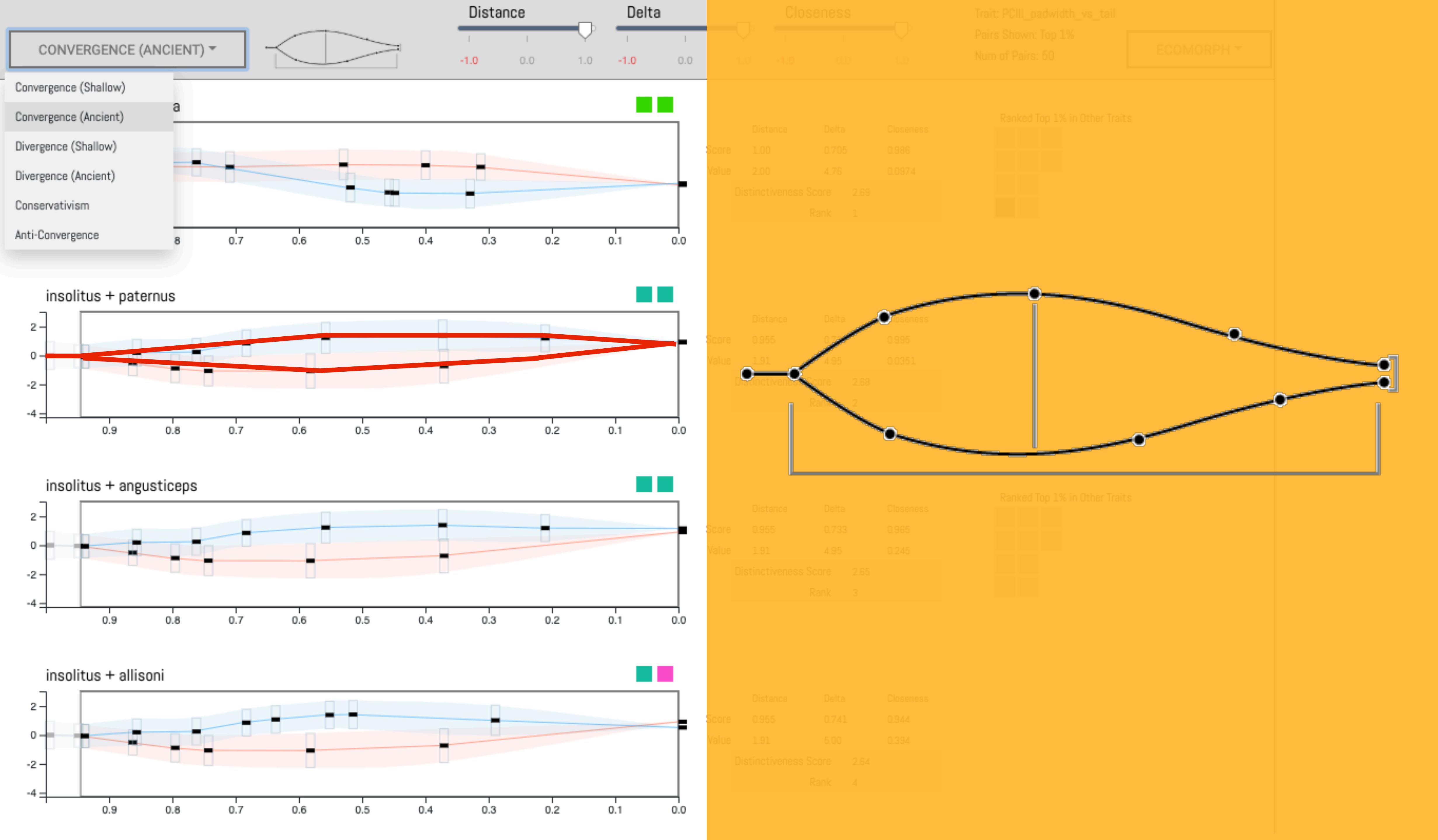
**Trait value
range**

**Internal nodes
and 95%
confidence
intervals**

**Confidence
intervals**

Species nodes







Process



4-13-19

Started background work, correspondence with Harmon Lab

5/13/19 - 7/31/19

Field study in Moscow

10/18/19

2/3/20 - 2/6/20

2nd visit to Moscow

3rd visit to Moscow

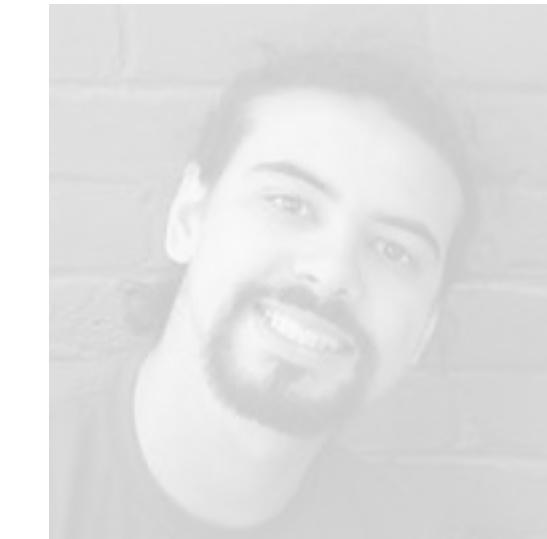
Ran a workshop
5/24/19

Conducted interviews, became familiar with domain space

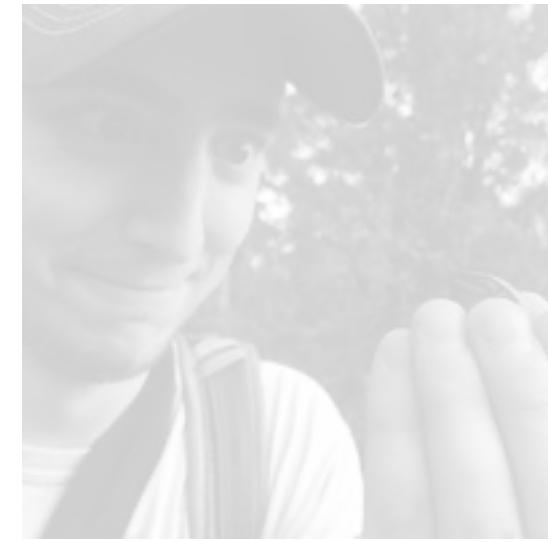
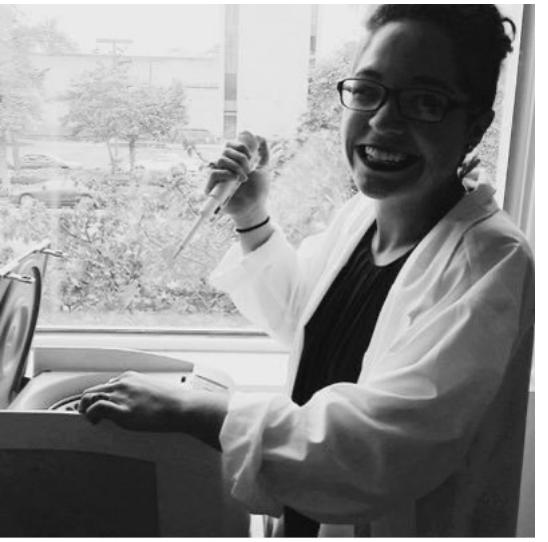
We are here



Primary collaborators



Secondary collaborators



6 Criteria for Establishing Rigor

INFORMED,
REFLEXIVE,
ABUNDANT,
PLAUSIBLE,
RESONANT, and
TRANSPARENT.

6 Criteria for Establishing Rigor

“Leave it to design study researchers to decide, and argue for how best to achieve these criteria - given the specific people, data, and context involved in the study”

*Experimented with
methods to establish Rigor*

*Focus on **REFLEXIVE, ABUNDANT, TRANSPARENT***

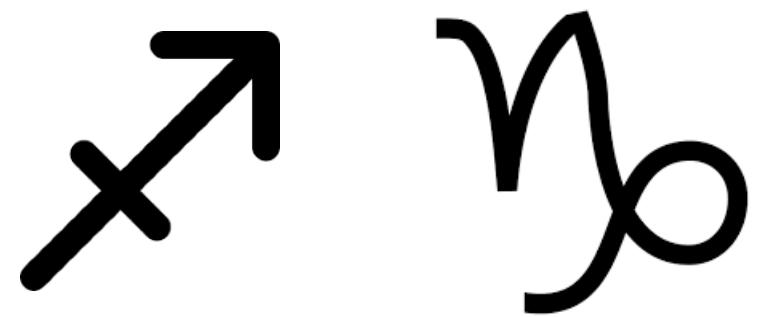
REFLEXIVE:
“is explicit and thoughtful self-awareness of a researcher’s own role in a study,”
(Meyer & Dykes, 2019)

REFLEXIVE: *Embedding myself in the Harmon lab and becoming one of herd.*



Took regular reflexive notes through the process.

ABUNDANT:
“A design study with abundance has rich details; many voices, datasets, contexts, and designs; and significant time in the field,”
(Meyer & Dykes, 2019)



*(I was sagittarius sun,
Capricorn rising if you're
curious)*

**“significant time
in the field”**

**Embedded myself in
the lab for 3 months,
became one of the
herd.**

12/05/2019	Sketchbook	Sent out emails to get datasets from 4 other labs in Utah Wed Dec 11 2019 00:00:00 GMT-0700 (Mountain Standard Time) Tags: correspondence, ,
12/10/2019	Revisited discrete summary view	
12/11/2019	Sent out emails to get datasets from 4 other labs in Utah, Brainstormed names, made sain	
12/17/2019	Meeting with Miriah	
12/20/2019	Adding multiple traits to ranking system	
12/27/2019	Added Histogram for multiple traits in ranking	
12/28/2019	Updated Discrete Summary View	
12/30/2019	Showed Alex new discrete view. He suggested Jitter and circles Instead of bars to handle m	
01/02/2020	Checking in Email to Austin and Luke , Checking in Email to Austin and Luke	
01/04/2020	Email Feedback from Austin	
01/15/2020	Got Salamander data from Austin, Meeting with Austin skype	

Dear [XXXX],

My name is Jen and I am a PhD student in the Scientific Computing and Imaging Institute at the University of Utah. For my research, I build web-based visualization tools for scientists to explore their data.

I have been building a tool this year in collaboration with the Harmon lab (<http://lukejharmon.github.io>) at the University of Idaho for evolutionary biologists to explore mechanisms driving character shifts and identification of convergence in their phylogenetic data.

We are currently looking for other labs that might be interested in the tool -- I am reaching out because your lab's research [XXXX] seems to be a good fit. If you are interested in potentially incorporating interactive visualization tools into your workflow I'd love to find a time to chat to learn about your data and to show you what we are designing for the Harmon Lab. I've included a screenshot of the tool below to give you a sense of the interface.



“many voices, datasets, contexts,”

**Reached out to other labs
outside of the core lab we
were collaborating with
*(this did not work out)**

TRANSPARENT:
“Transparent reporting should be self-critical and include errors, failures, analytical dead ends – the joys and mistakes,”
(Meyer & Dykes, 2019)

Working at Harmon Lab	
05/13/2019	Initial Meeting with Luke Harmon, Initial Meeting with Jack Sullivan, Initial Meeting with Hector, Notes on Analysis, Sketchbook
05/14/2019	Initial Meeting with Austin Patton, Initial meeting with Megan, Initial Meeting with Larry
05/15/2019	Sketchbook
05/17/2019	Examples of Work from Lab, Initial Interview with Eva Top, Gave Presentation to Group
05/18/2019	Domain Visual Examples
05/20/2019	Initial Meeting with Thibault, Initial Meeting with Amanda
05/21/2019	Planning Creativity Workshop, Started uncertainty example collection, Visit Maren's Lab at WSU
05/22/2019	Meet with Roni to Discuss Multinet, Sketch Brainstorm Ideas
05/24/2019, Workshop Notes
05/27/2019	Summarized Needs from Workshop
05/28/2019	Sketchbook
05/29/2019	Skype with James and Luke, Meeting with Luke, Sketchbook
05/30/2019	Draft Initial Requirements Document, Sketchbook
05/31/2019	Sketching Tool Features, Sketchbook
06/01/2019	Sketchbook
06/03/2019	Familiarizing Myself with the Domain Analysis workflow
06/06/2019	Status check on progress with Roni and Curt
06/07/2019	Familiarizing myself with Arbor Workflows
06/10/2019	Multinet Tech Meeting, Sketchbook
06/12/2019	Second Meeting with Austin
06/13/2019	Taking Notes on Ancestral State Reconstruction, Sketchbook
06/14/2019	Second Meeting with Hector
06/18/2019	Created Illustrator Mockups, Meeting with Alex
06/21/2019	Update on Path Design

started to record audio

**TRANSPARENT:
Created an audit trail
containing annotated
records of the study
process.**

**Includes:
Notes,
Email correspondence,
Sketchbook scans,
Photos of collaborator sketches,
Links to papers,
Notes from audio**

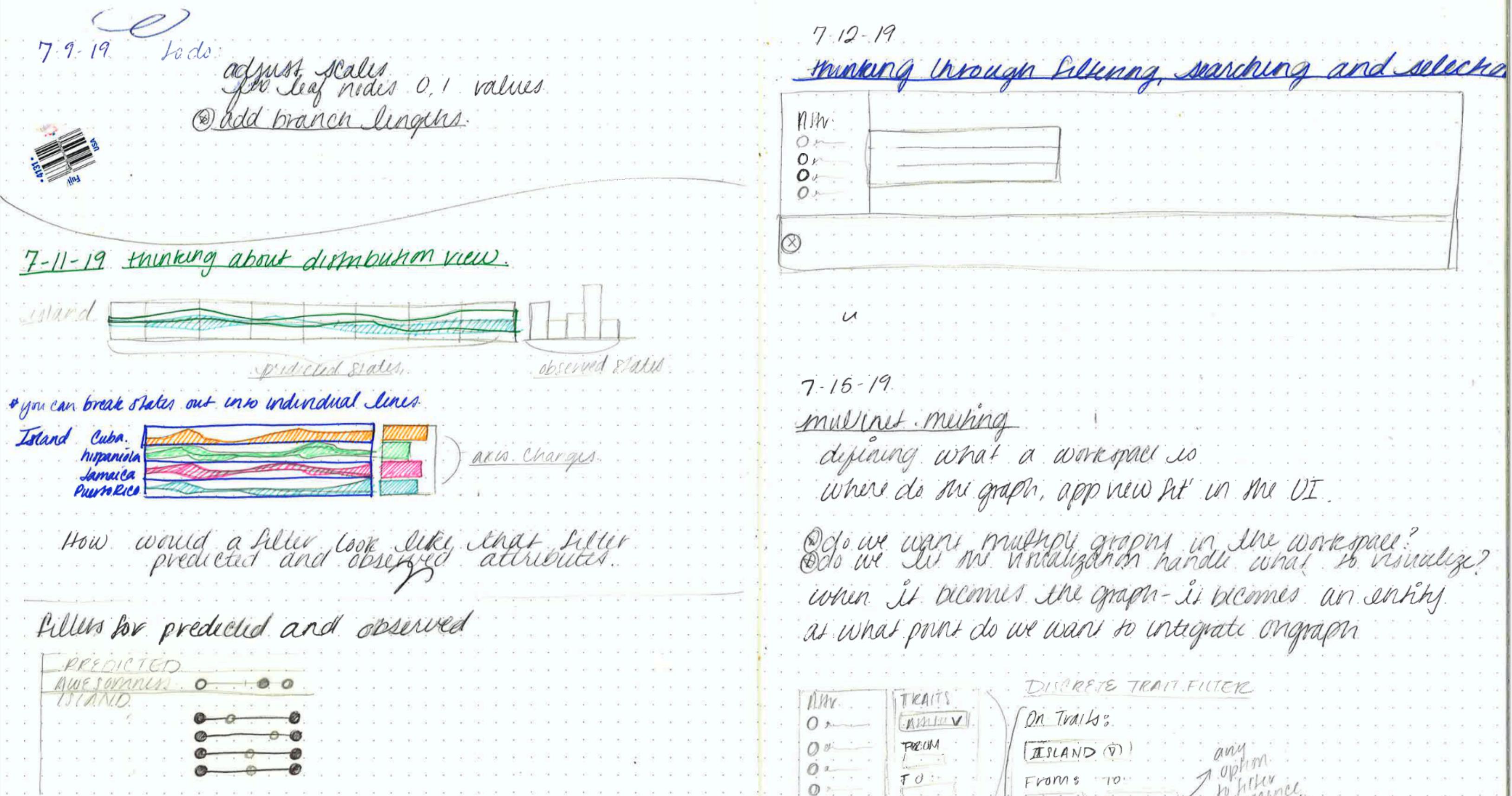
Evolutionary Biology Design Study Audit Trail

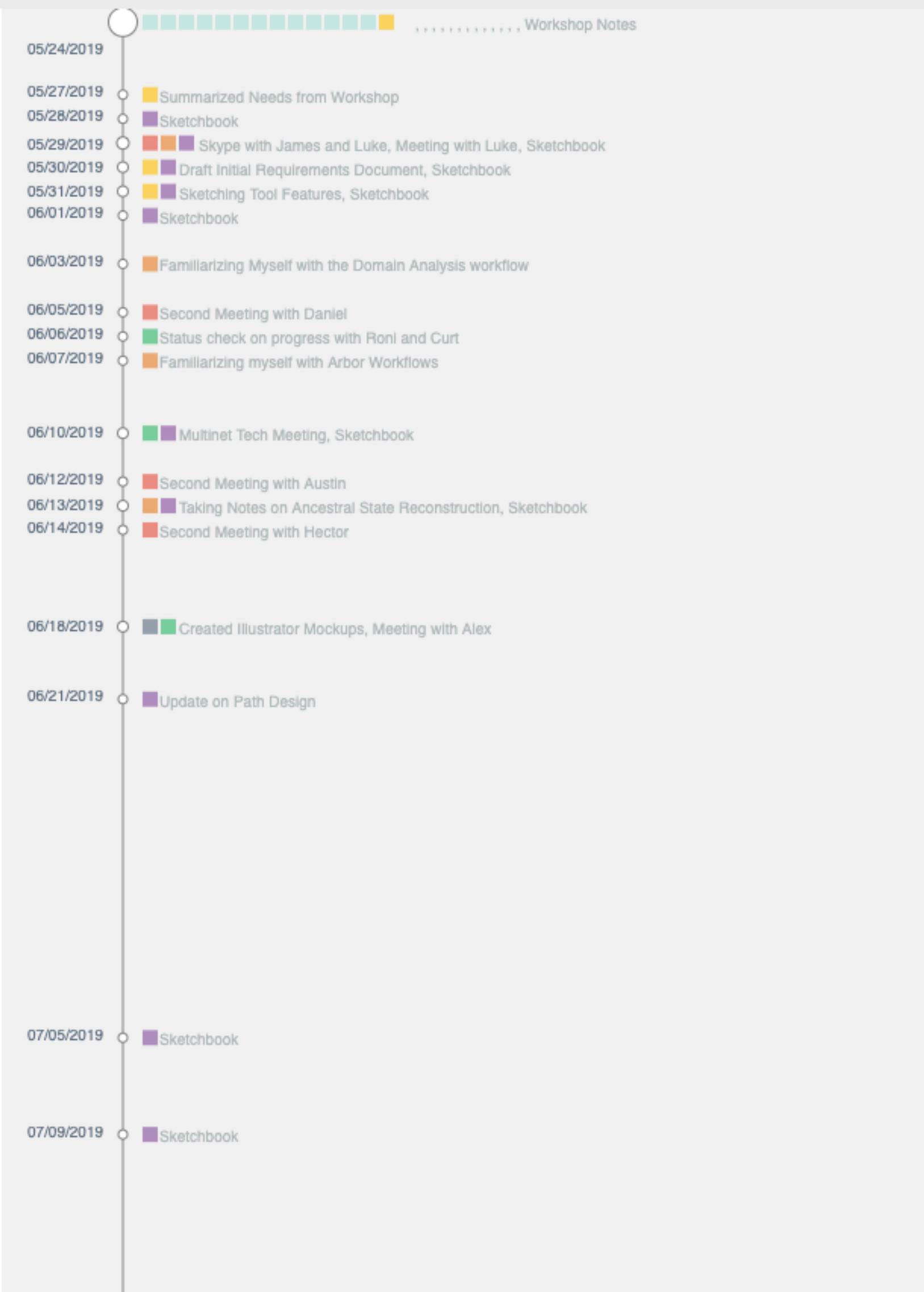
[TIMELINE](#)
[REFERENCES](#)
[SKETCHBOOK](#)

06/21/2019	Update on Path Design
07/05/2019	Second Meeting with Daniel, Sketchbook
07/09/2019	Sketchbook
07/16/2019	Meeting with Alex, Sketchbook
07/17/2019	Third Meeting with Austin
07/22/2019	Update of Sidebar Tree
07/23/2019	Meeting with Luke
07/25/2019	Development of Summary View, Update on sort functionality
07/29/2019	Sketchbook
07/30/2019	Sketchbook
07/31/2019	Sketchbook
08/07/2019	Assesement of Development and Requirements

Sketchbook Tue Jul 09 2019 00:00:00 GMT-0600 (Mountain Daylight Time)

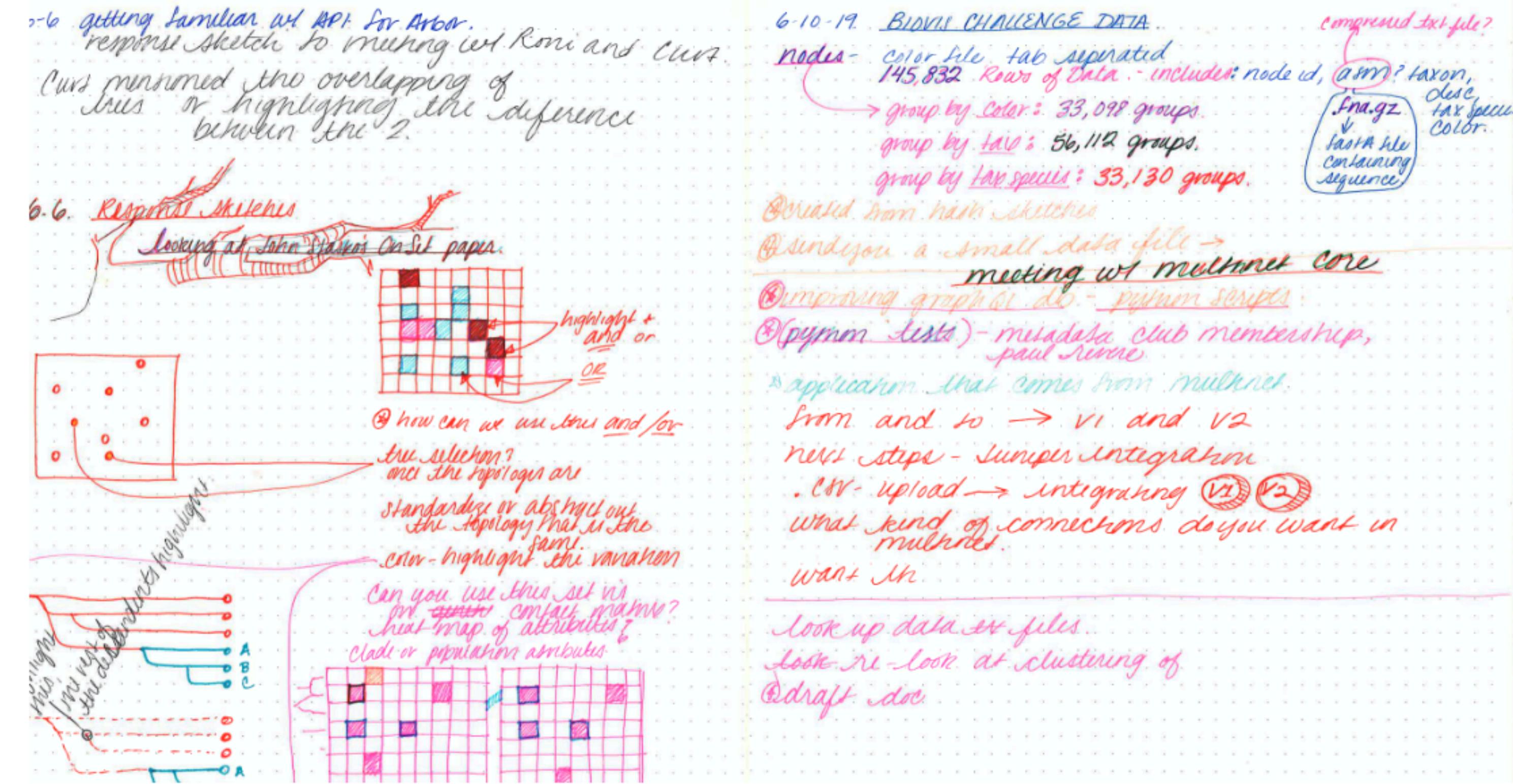
Tags:
sketch, brainstorm,





Sketchbook Mon Jun 10 2019 00:00:00 GMT-0600 (Mountain Daylight Time)

Tags:
sketch, brainstorm,



Audit trail website

<http://vdl.sci.utah.edu/evo-bio-audit/>

Lessons Learned:

Lessons learned:

- 1. Explicit, systematic reflection was demonstrably productive.***

Reflection-in-Action: Systematic and regular

***Reflected before and after meetings
Helped propel the process forward***

Reflection-in-Action: Was reflexive

Identify limitations that could be acted upon

"I have not been recording these interviews as I am in the first week and I do not want to be intrusive. I wanted to keep this meeting open ended - the ultimate goal was getting a baseline understanding of what they do..."

"I will ask Austin if he minds if I record the interview on my phone. I believe the recording will help me capture more than I can get from my note taking, and maybe more importantly, be more engaged in the interview process."

Recording allowed for more abundant reflection-on-action

Reflection-on-Action: Reflection via Reflective Transcription

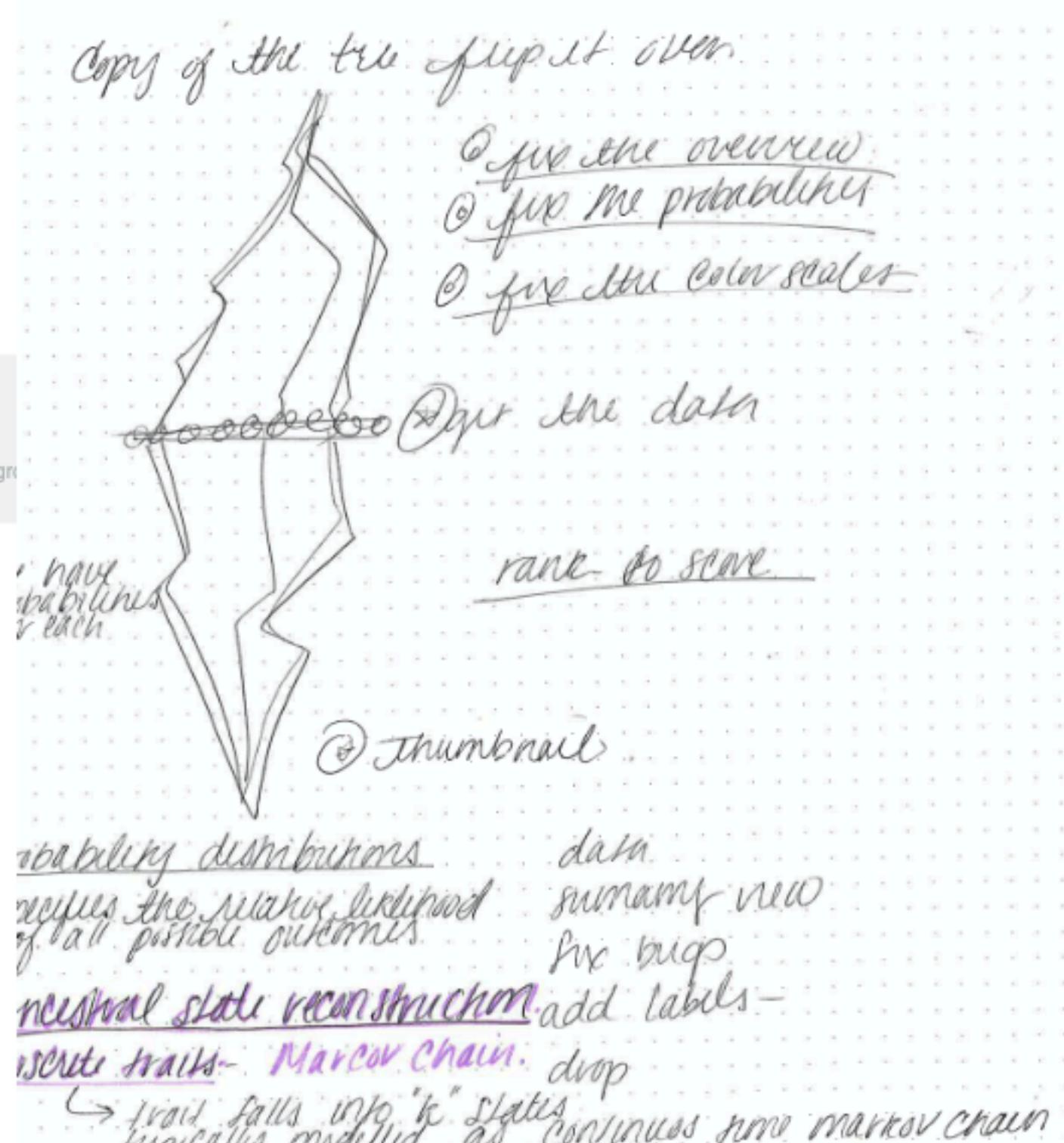
Happened the same day or next day.

***Audio facilitated reflection-Captured more than
note-taking could.***

Reflection-on-Action: Reflection via artifact Curation

Sketchbook Fri Nov 01 2019 00:00:00 GMT-0600 (Mountain Daylight Time)

Tags:
sketch, brainstorming,



- ① detailed document of design process
- ① open doors and coll. eman and nina
- ① supplementary material w/ doc.
- How to describe it as multiple levels
main highlights and structure
have areas that we can dive into.
pair that is the lens for 2 criteria
and "abundance"
- ① link to google doc links
- Color code the uncertainty? abstract information
- ① think broadly about doc you are collecting
- ① what have you captured about or discussion
- ① how do I include the ls review
- ① how do I represent this
- ① Evolutionary biology: multiple data, museum, look at other data

Curation facilitated reflection-on-action that played a vital role in design development.

Ex: revisiting of distribution view

“When does reflection occur, what triggers it, and what should we be capturing from it?”

“When does reflection occur, what triggers it, and what should we be capturing from it?”

Exactly when reflection-on-action should occur remains an open question, - should be done during the process of design and development

Regular reflection-in-action provided an abundance of material to speculate and draw from

Recommendations:

- **Reflect before and after meetings with collaborators**
- **Record audio**
- **Curation of artefacts is a process that facilitates reflection. Record insight emergent from this.**

Lessons learned:

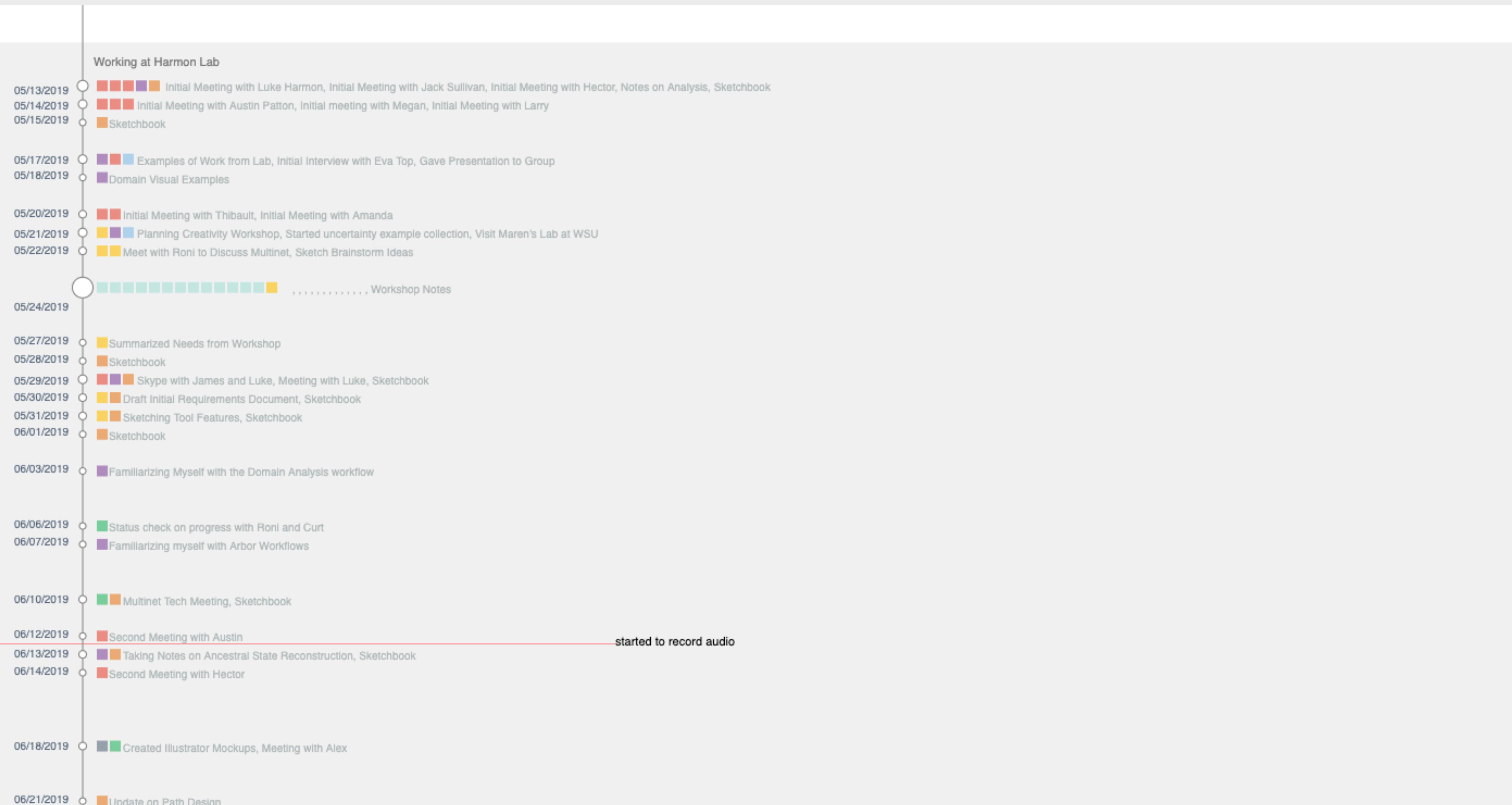
2. When artifacts are discoverable, they support transparency and reflection

Evolutionary Biology Design Study Audit Trail

TIMELINE

REFERENCES

SKETCHBOOK



Making artifacts discoverable

fx homogenous,recording

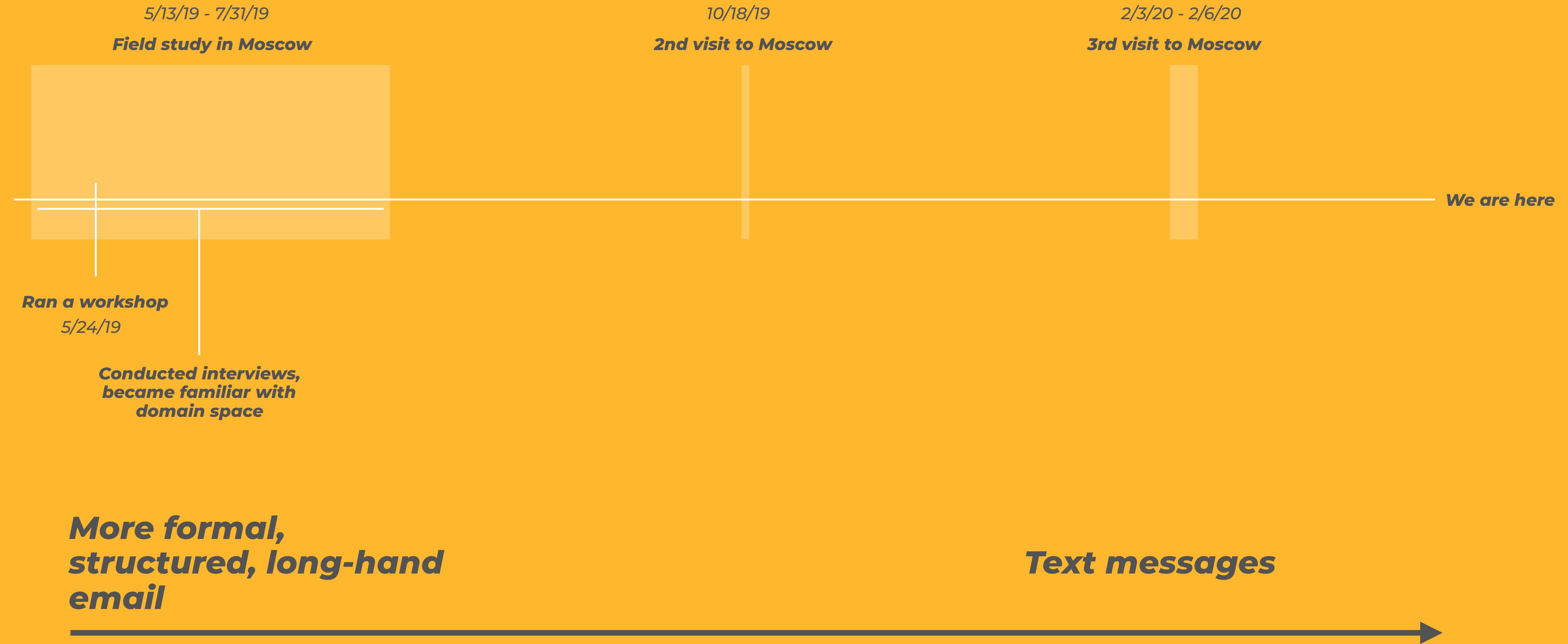
	A	B	C	D	E	F	G	H	I	convergence	1 of 5	^	v	:	X
1	Date_Range	Event	embed_link	Drive_Link	Sketch_ID	tag1	tag2	tag3	highlight						
2	4-22-2019	Meeting with Nina	https://docs.google.com	https://drive.google.com	null	interview	design studies	planning		null	null				
3	5-13-2019	Initial Meeting with Luke Harmon	https://docs.google.com	https://drive.google.com	null	interview				compare, sets, c	evolutionary events,shared traits, evolutionary integration				
4	5-13-19	Initial Meeting with Jack Sullivan	https://docs.google.com	https://drive.google.com	null	interview	chipmunks			homogenous,red	model selection,phylogeography, divergence, hybridize, bi				
5	5-13-2019	Initial Meeting with Hector	https://docs.google.com	https://drive.google.com	null	interview	domain vis			created figures,	clade, dispersal, diversification, biographic history, bio spec				
6	5-13-19	Notes on Analysis	https://docs.google.com	https://docs.google.com	null	domain	context	analysis		uncertainty,covar	maximum clade probability tree, MCC, range of branching				
7	5-13-19	Sketchbook	null	null		513	sketch			collaborator sket	tree branch range,				
8	5-14-19	Initial Meeting with Austin Patton	https://docs.google.com	https://drive.google.com	null	interview	convergence			visual analysis,re	phylogenetics, coevolution, anole lizards, divergence, morph				
9	5-14-19	Initial meeting with Megan	https://docs.google.com	https://drive.google.com	null	interview	summary stats			uncertainty,topolo	trait mediated community assembly, adapted, regional spe				
10	5-14-19	Initial Meeting with Larry	https://docs.google.com	https://drive.google.com	null	interview	meta			Epistemology	Microbial island biography				
11	5-15-2019	Sketchbook	null	null		515	sketch			domain sketches	tree representatio, phenogram, traitgram, genetic				
12	5-17-2019	Examples of Work from Lab	https://docs.google.com	https://docs.google.com	null	domain	context			visualization	dispersal, vicariance				
13	5-17-2019	Initial Interview with Eva Top	https://docs.google.com	https://drive.google.com	null	interview	genetics			Compare	Hi-C data, antibiotic resistant genes, microbes, Genotypic				
14	5-17-19	Gave Presentation to Group	https://docs.google.com	https://drive.google.com	null	presentation	expectations			null	null				
15	5-18-19	Domain Visual Examples	https://docs.google.com	https://docs.google.com	null	domain	context			null	null				
16	5-20-19	Initial Meeting with Thibault	https://docs.google.com	https://drive.google.com	null	interview	matrices			uncertainty, trajec	origins, Hi-C methodology, reconstructed plasmid-host as				
17	5-20-19	Initial Meeting with Amanda	https://docs.google.com	https://drive.google.com	null	interview	pca			locally adapted,d	how populations vary, individual variation, evolution rate, g				
18	5-21-19	Planning Creativity Workshop	https://docs.google.com	https://drive.google.com	null	brainstorm				null	null				
19	5-21-19	Started uncertainty example collection	https://docs.google.com	https://docs.google.com	null	domain	context			null	null				
20	5-21-19	Visit Maren's Lab at WSU	https://docs.google.com	https://drive.google.com	null	presentation	interview			Transfer, Annotati	variations of branching, identify traits that have evolved, di				
21	5-22-19	Meet with Roni to Discuss Multinet	https://docs.google.com	https://drive.google.com	null	brainstorm	assesment			null	null				
22	5-22-19	Sketch Brainstorm Ideas	https://docs.google.com	https://drive.google.com	null	brainstorm				null	null				
23	5-24-19		null	null	work1	workshop	documentation	sketch		null	null				
24	5-24-19		null	null	work2	workshop	documentation	sketch		null	null				
25	5-24-19		null	null	work3	workshop	documentation	sketch		null	null				
26	5-24-19		null	null	work4	workshop	documentation	sketch		null	null				
27	5-24-19		null	null	work5	workshop	documentation	sketch		null	null				
28	5-24-19		null	null	work6	workshop	documentation	sketch		null	null				
29	5-24-19		null	null	work7	workshop	documentation	sketch		null	null				
30	5-24-19		null	null	work8	workshop	documentation	sketch		null	null				
31	5-24-19		null	null	work9	workshop	documentation	sketch		null	null				

Recommendations:

- **Make your records discoverable (by you and others).**
- **Create a system for organization and curation as soon as you begin collecting**
- **Categorization and tagging makes artefacts more discoverable and allows recurring concepts to float to the top.**

Lessons learned:

3. changes in collaboration patterns are evidence for criteria



Change in patterns of communication: Evidence of ABUNDANT and INFORMED

*More formal,
structured, long-hand
email*

Text messages



Change in patterns of communication: Evidence for ABUNDANT

***More formal,
structured, long-hand
email***

Do not know anyone

Text messages

***Building of mutual trust
over span of field study***



Change in patterns of communication: Evidence for INFORMED

***More formal,
structured, long-hand
email***

Text messages

Do not know anything

***Can hold a conversation
about mechanisms of
evolution***



Recommendations:

- **Look for evidence of immersion, such as:**
 - **Shifting modes of communication**
 - **Ways in which interaction with collaborators can be indicators/evidence of criteria.**