Trevo: An Exploratory Tool for Tree and Trait Data

Pre-Paper Talk

THE UNIVERSITY OF UTAH

SCI
www.sci.utah.edu

Photo by Edi Libedinsky on Unsplash

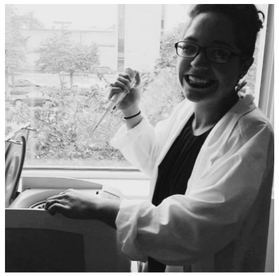
Background and Motivation

(these black and white title pages are to be explicit about the content sections. They won't be in the final presentation)

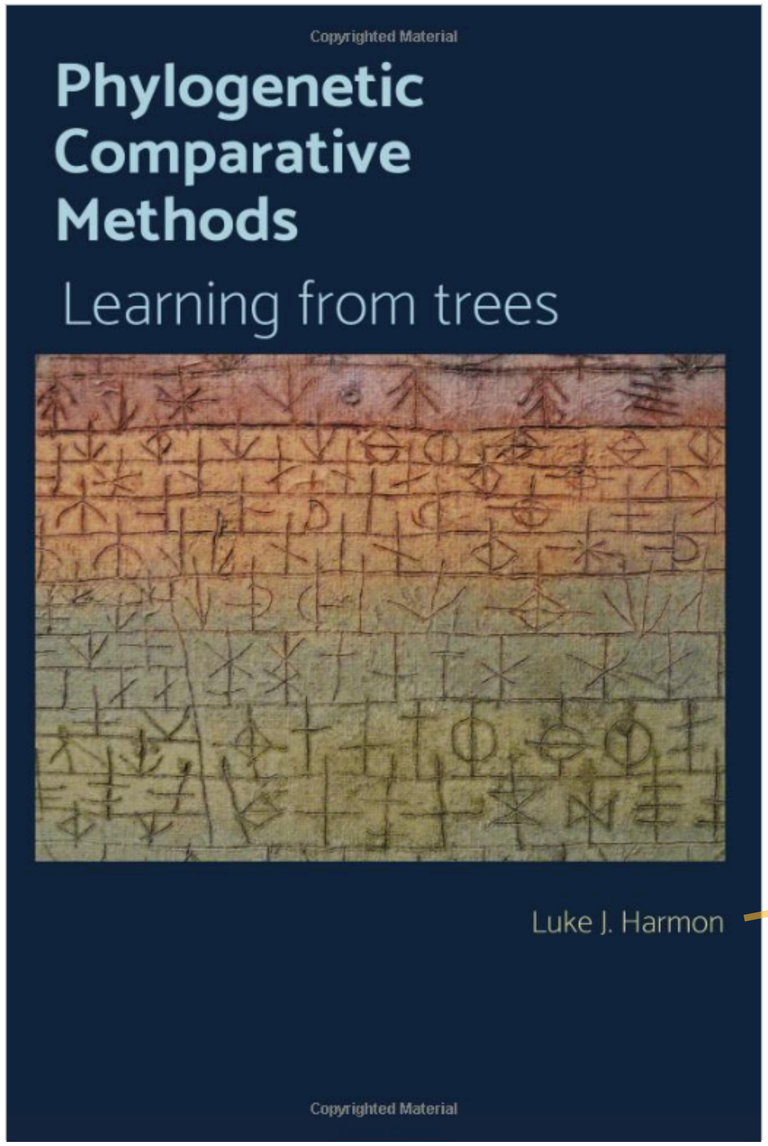
***Evolution**: the process of change in a given characteristic spanning generations*

***Evolutionary Biology**: explain the evolutionary process that gives the diversity of the living world.*

Worked with evolutionary biologists at the University of Idaho

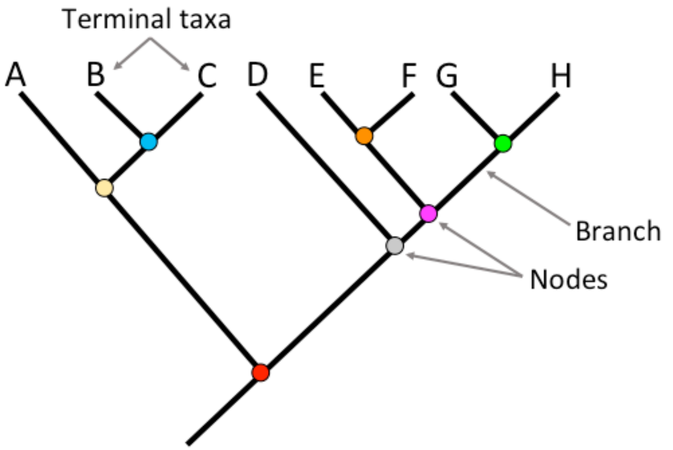




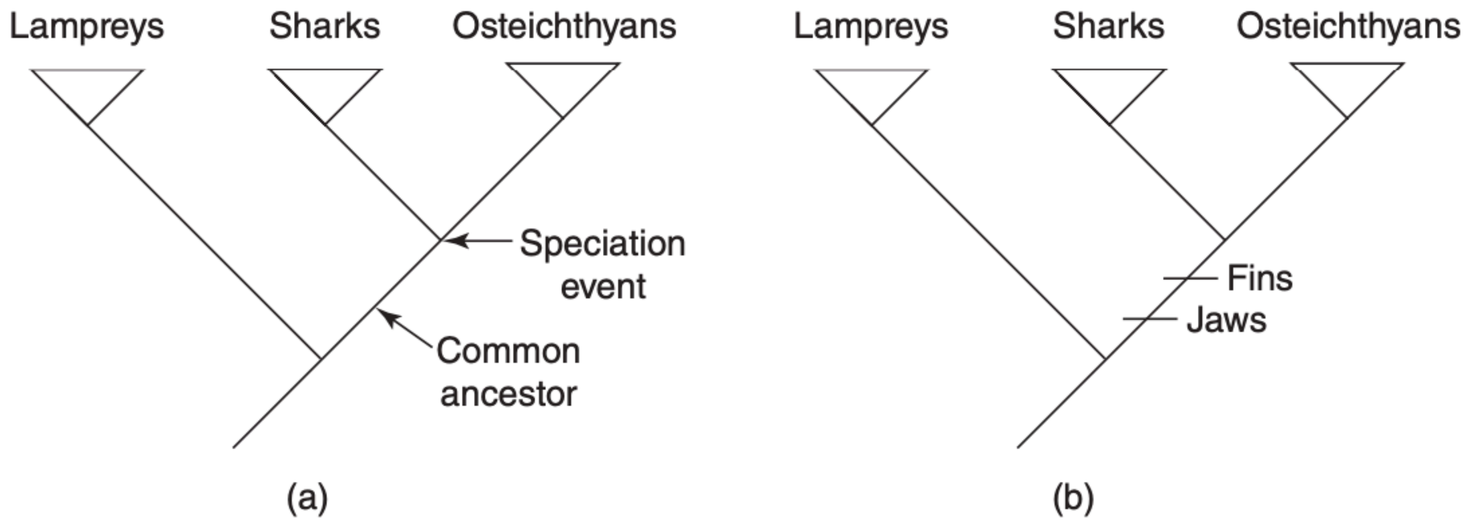
Wrote the book on comparative methods

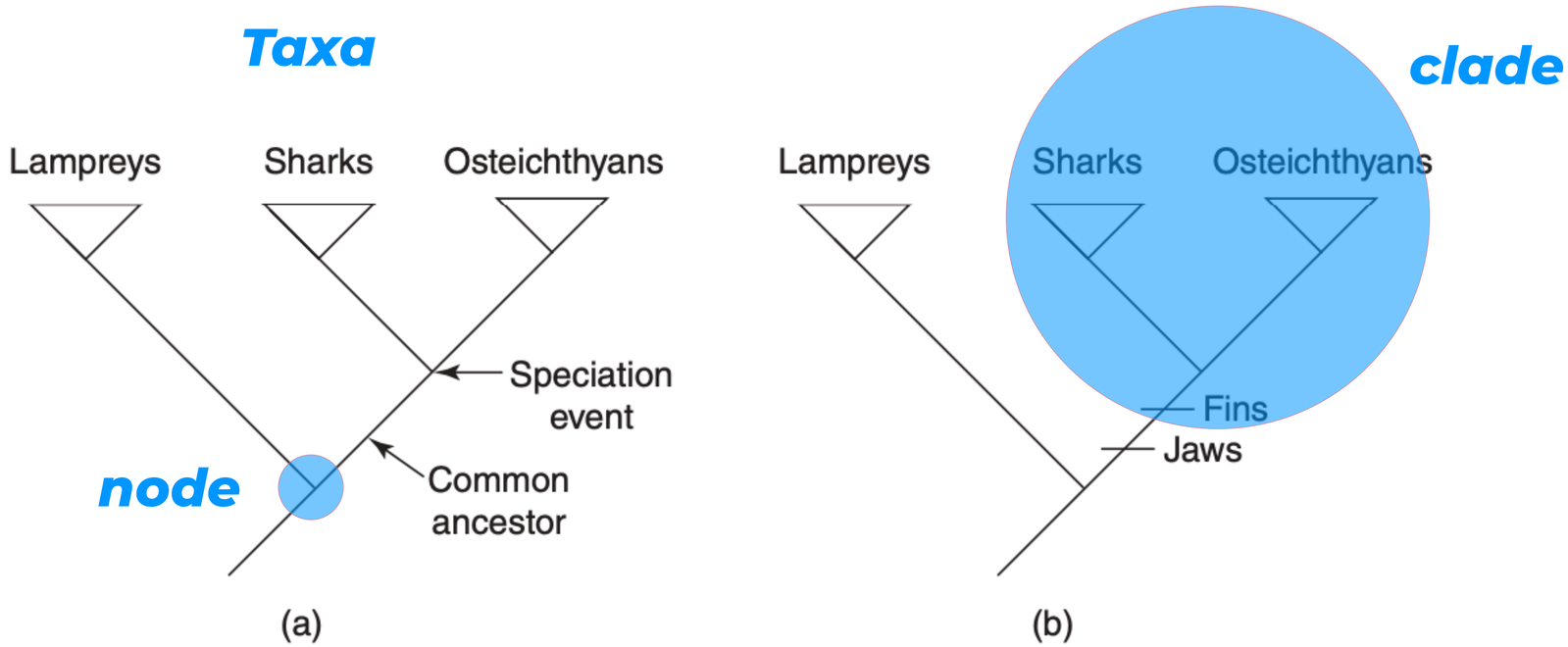
*They work extensively with **phylogenetic trees***



They work extensively with **phylogenetic trees**

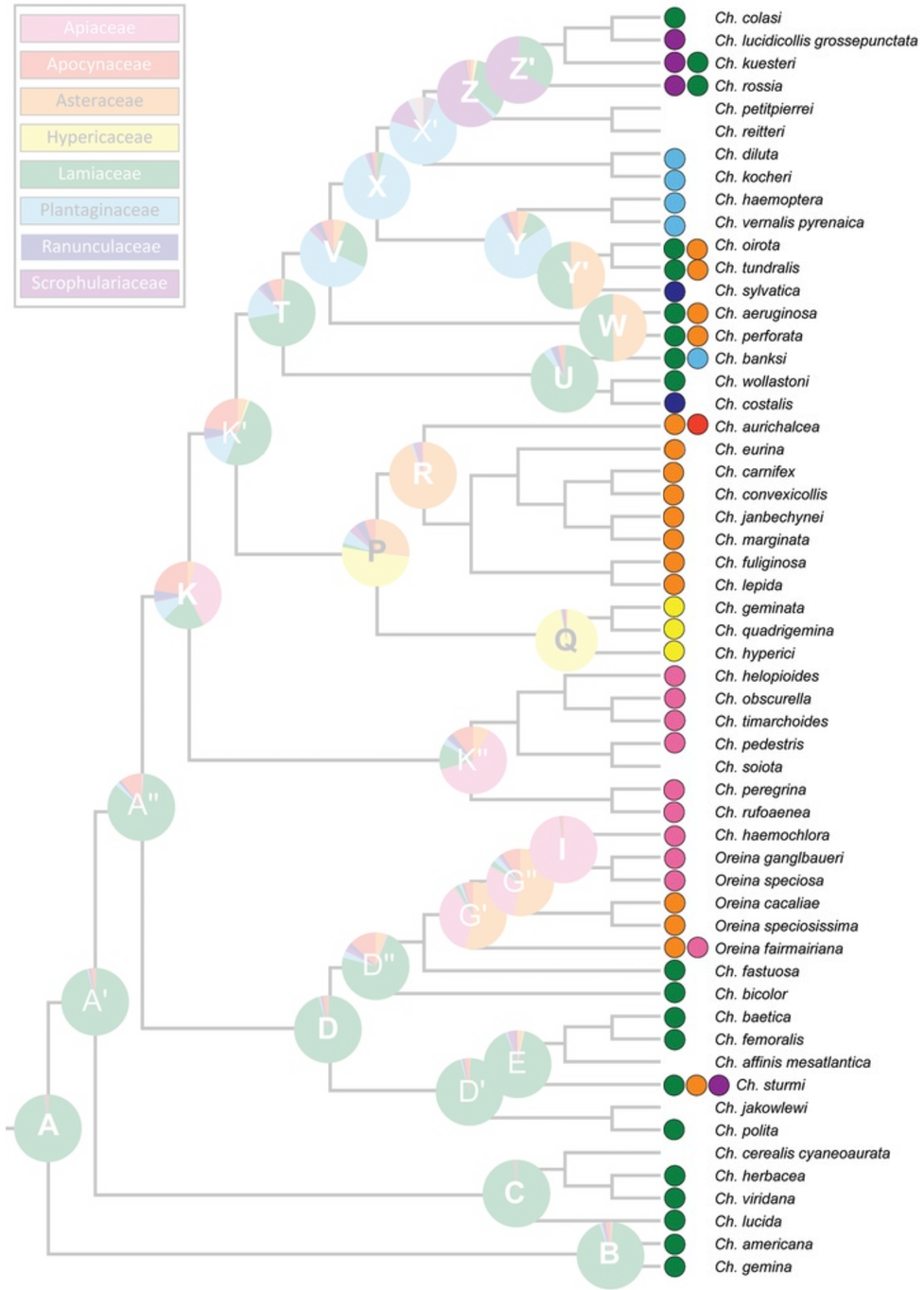
Depict hypothesized evolutionary relationships




Taxa
Lampreys
Sharks
Osteichthyans
Speciation event
node
Common ancestor
(a)
clade
Lampreys
Sharks
Osteichthyans
Fins
Jaws
(b)

To understand the evolutionary process, they reconstruct histories through ancestral state reconstruction.

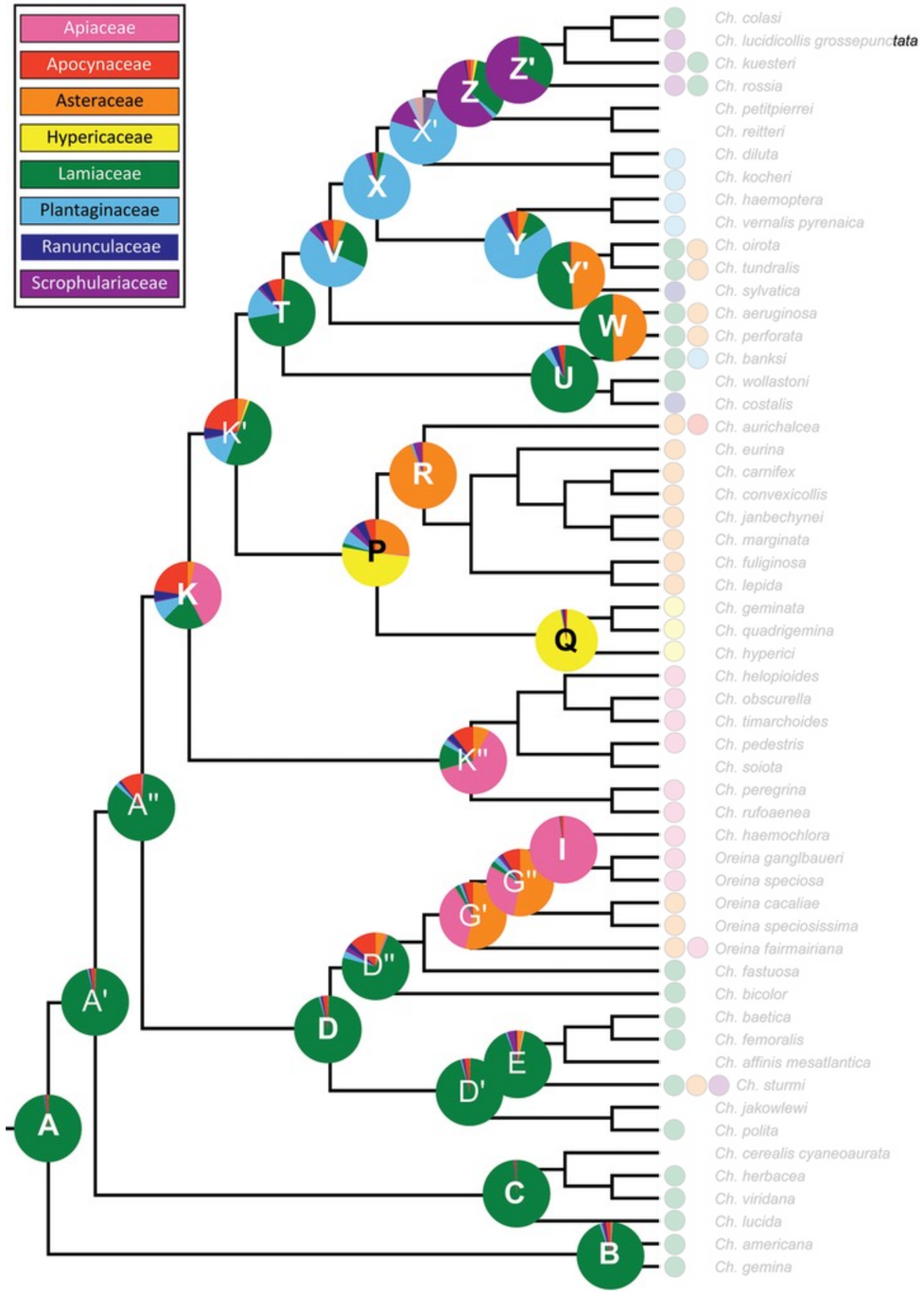
How and when did traits evolve?



Ancestral State Reconstruction

Use observed attributes from species samples

Jurado-Rivera JA, Petitpierre E. New contributions to the molecular systematics and the evolution of host-plant associations in the genus Chrysolina (Coleoptera, Chrysomelidae, Chrysomelinae). ZooKeys. 2015(547):165.



Ancestral State Reconstruction

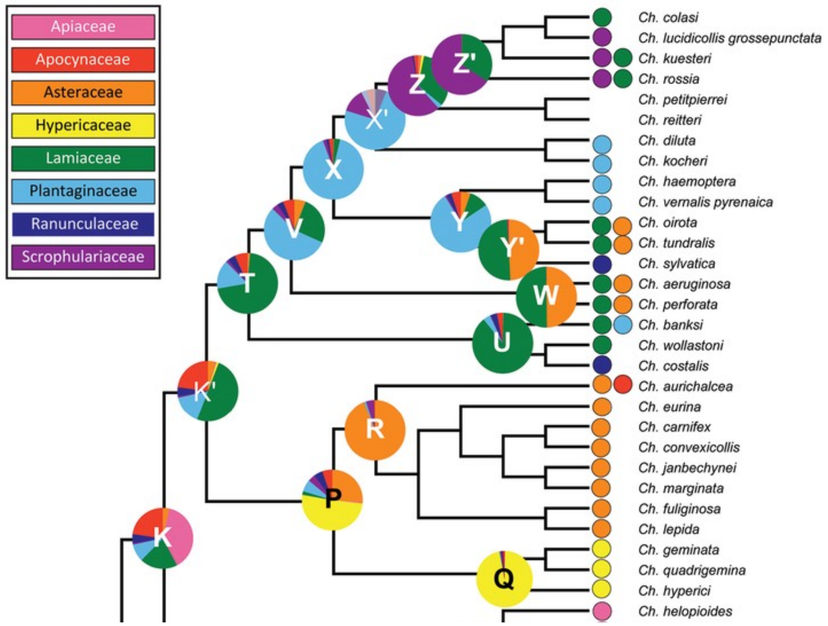
Calculate state probabilities (discrete) or estimated values (continuous) for the common ancestors up the tree.

Jurado-Rivera JA, Petitpierre E. New contributions to the molecular systematics and the evolution of host-plant associations in the genus Chrysolina (Coleoptera, Chrysomelidae, Chrysomelinae). ZooKeys. 2015(547):165.

Trait Data:

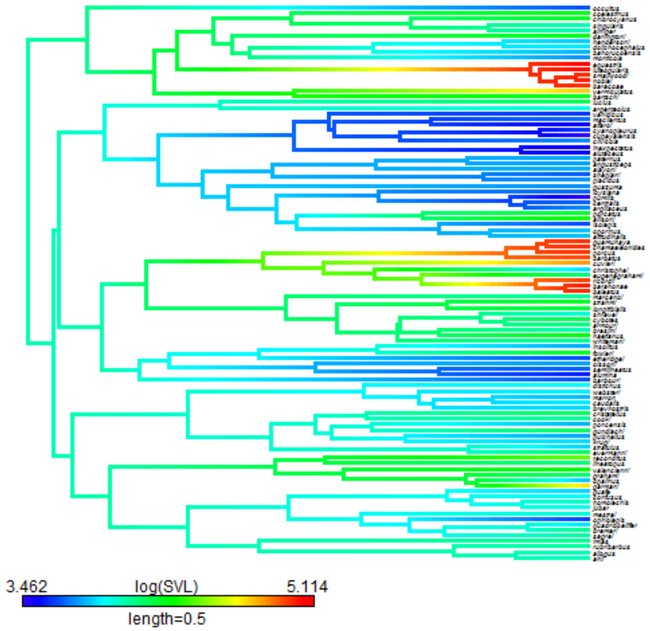
Discrete

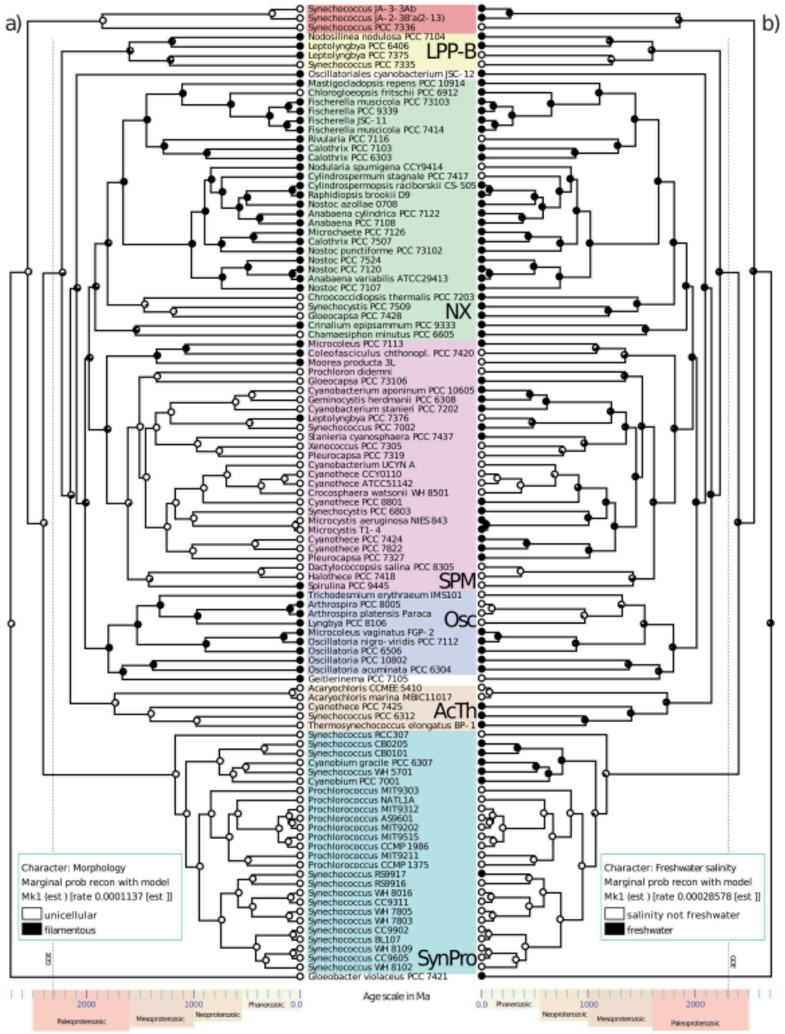
Probability that internal node is in a category



Continuous

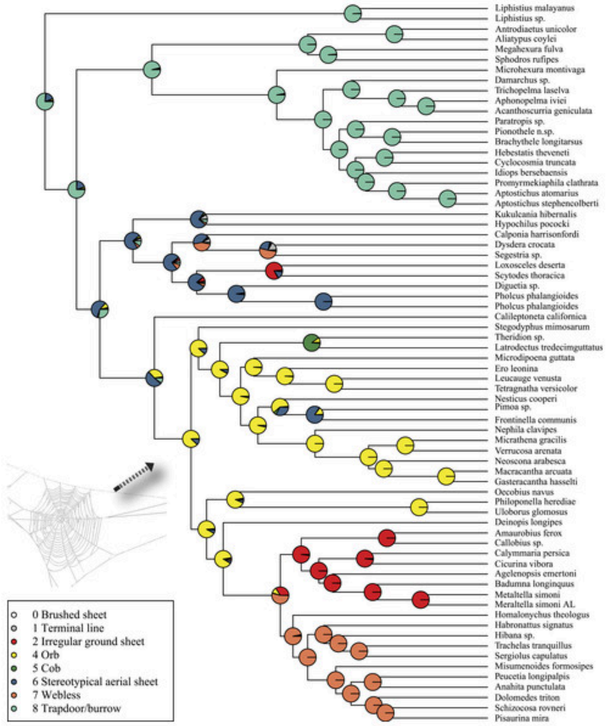
Estimated value and 95% confidence interval

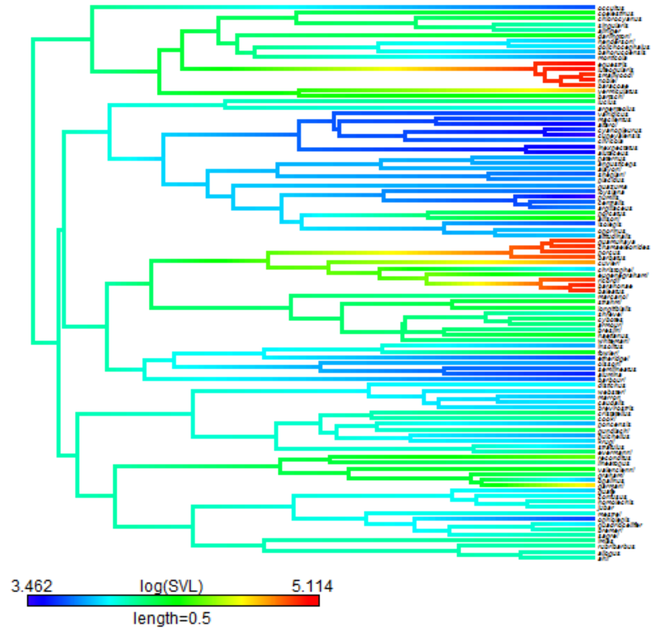




They have to view states of different attributes on separate trees.

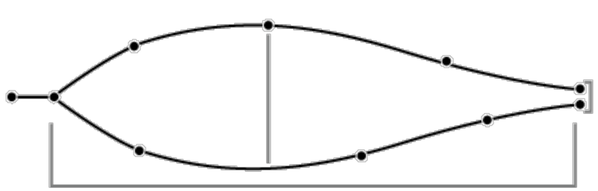
“If you have one continuous trait you can do things. If you two - ok. If you have three or four or five, there is nothing really sufficient”

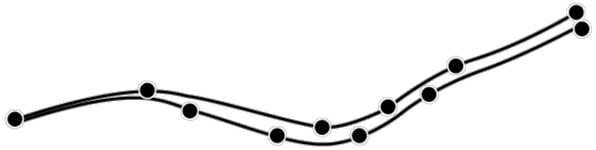


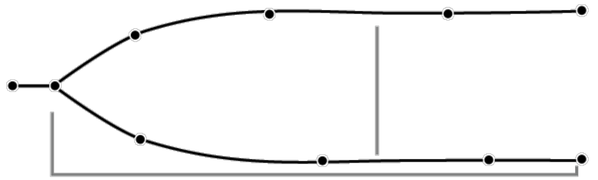


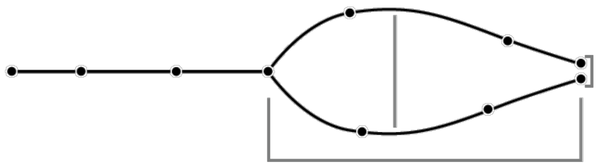
Do trait changes follow any established evolutionary hypotheses?

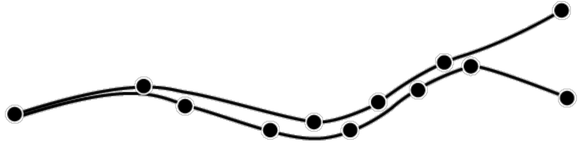
Established Evolutionary Hypotheses:

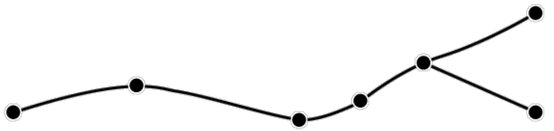




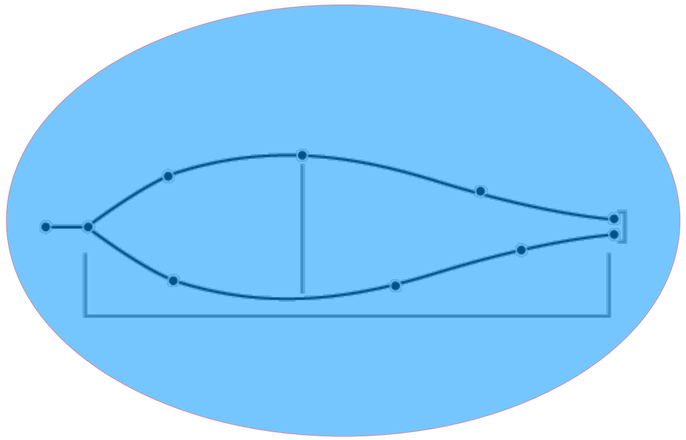


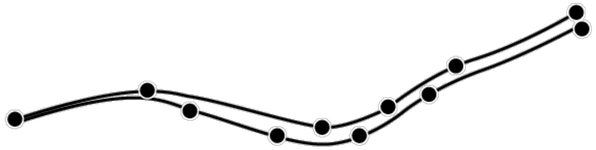


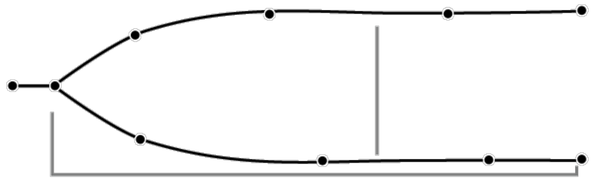


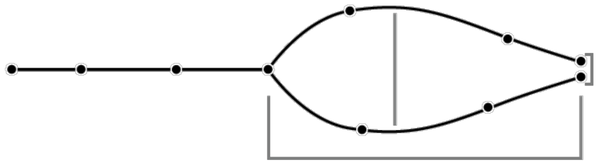


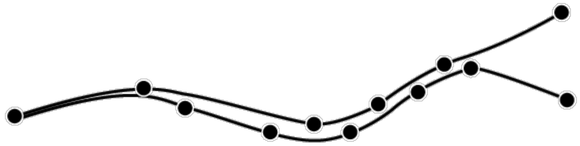
Established Evolutionary Hypotheses:

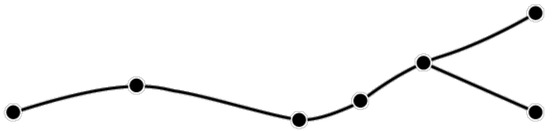






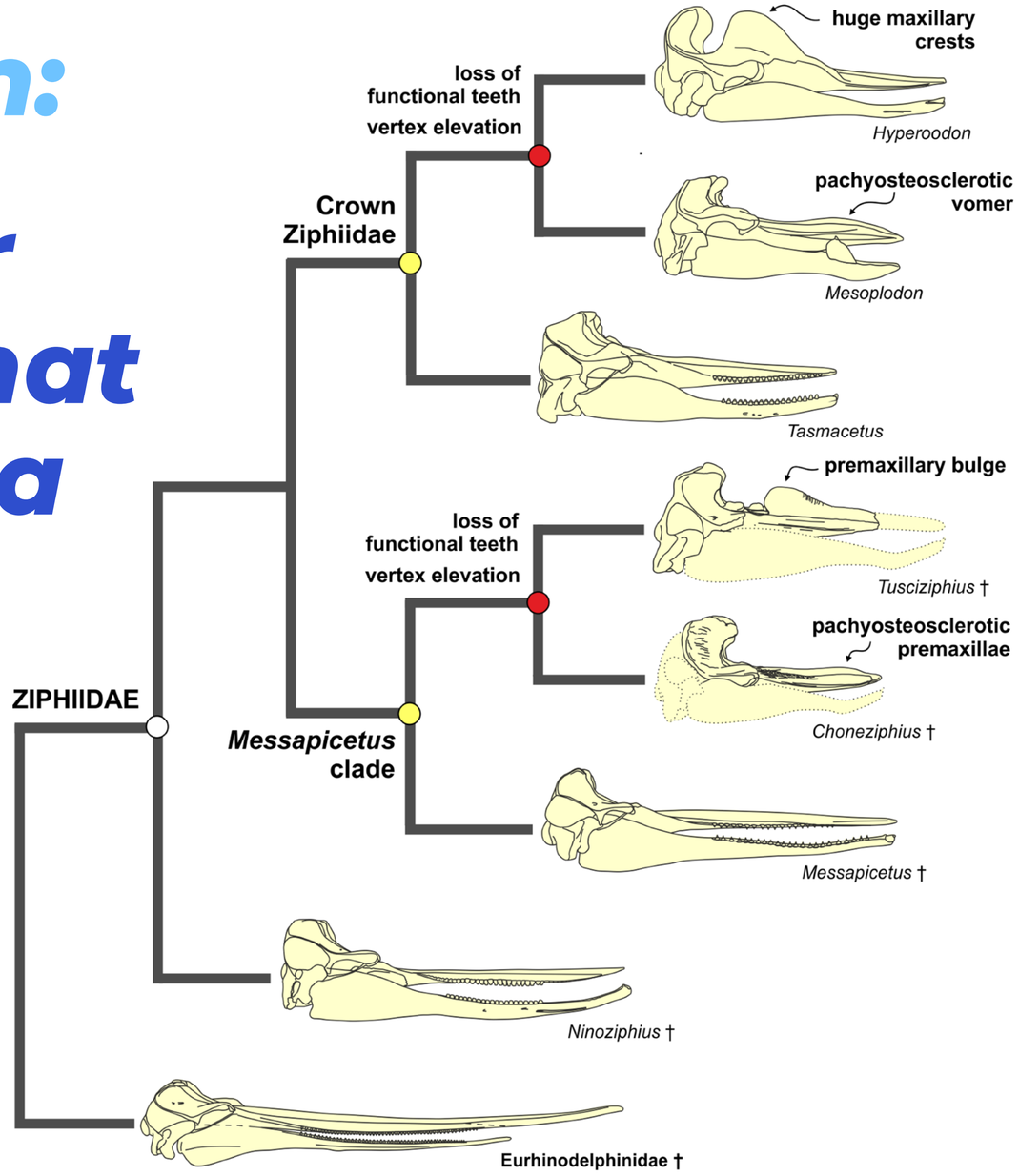


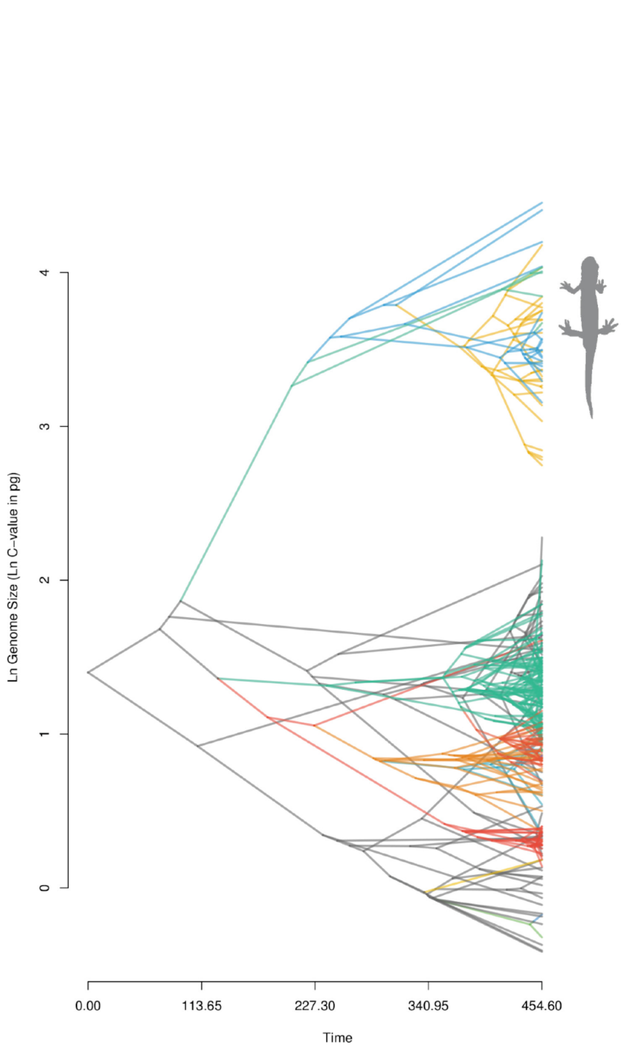


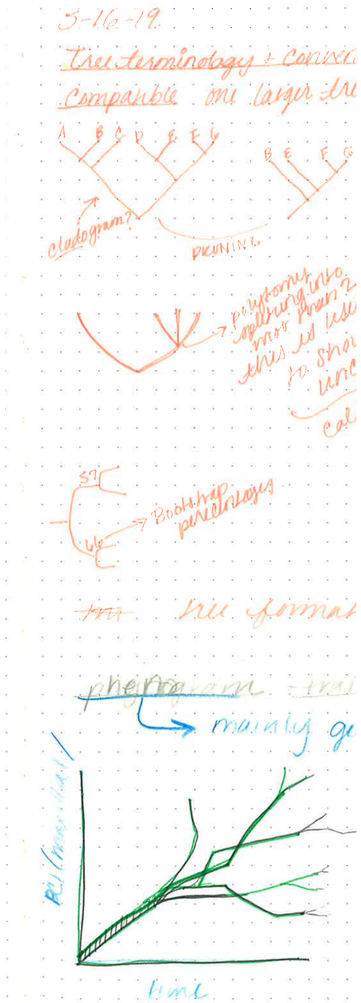


Convergent Evolution:

Emergence of similar features in species that have not come from a common ancestor







Traitgrams express the change in value across time in a continuous path in the context of the topology.

Limitations of current representations:

Lack the tools to view discrete and continuous characters together.

Limitations of current representations:

Comparisons are limited by topology.

Limitations of current representations:

No tool that allows visual testing of established evolutionary hypotheses.

Contribution

Contribution:

Visual exploratory tool that allows analysts to visualize multivariate phylogenetic data independent of the topology of the tree.

Ranking system that allows analysts to test established hypotheses regarding the relationships of species histories.

Lessons learned on design and development of an applied visualization tool informed by a reflective process through the duration of the design study.

Task Characterization

Explain the evolutionary process that gives the diversity of the living world.

we defined domain requirements that fall within this goal and operationalized tasks that support these requirements.

Task Characterization

1) Understand how and when traits evolved.

2) Identify evolutionary events and trends

3) Visual data check of distributions and outliers

Task Characterization

- **Understand how and when traits evolved**
  - See how trait value change over time
    - Plot trait values of internal nodes through time (path view)
  - Identify species that share similar value trends
    - Sort by volatility and value change scores of internal node values (path view)
    - subtree definition by trait values. (expanded sidebar)
    - Copy species names from internal node brush (summary view)

Task Characterization

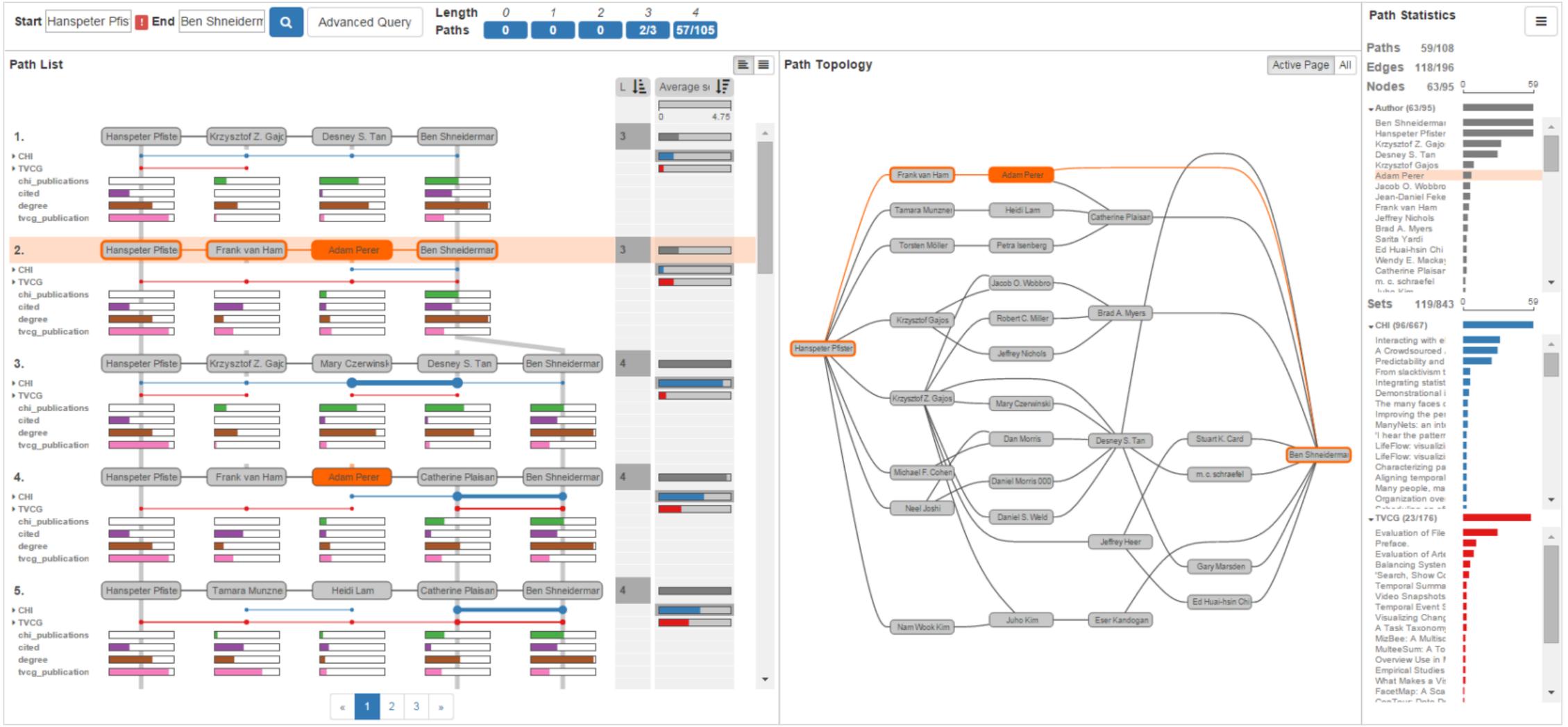
- Detect exceptional evolutionary events and trends
  - **Identify Macro-Evolutionary Modes (Test historical hypotheses about evolution)**
    - Identify the mode by the trait relationship between species (rank view)
    - Ability to quantify these relationships- which species stand out as exceptionally similar or different? (rank view)
  - **Identify situations of re-evolution**
    - View state changes in discrete traits through time of a given species (path view)
  - **Identify positive and negative trends**
    - View changes of value in continuous traits (path view)
  - **Identify which traits follow the topology of the tree, and which do not**
    - quantify ‘phylogenetic signal’ (volatility) in path view
      - Traits that don't have a phylogenetic signal - scattered varied (path view)
      - Traits that do - clades have consistent trait values (path view)

Task Characterization

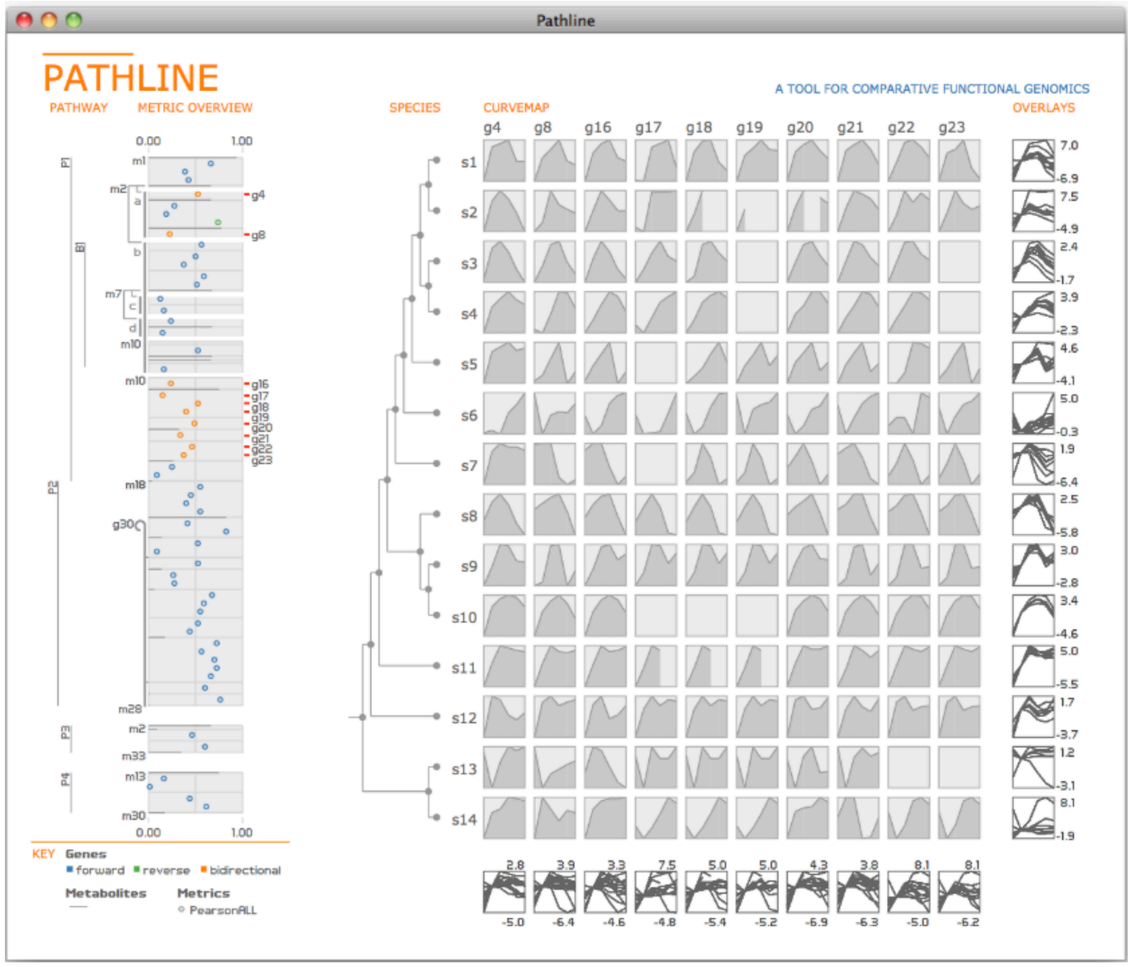
- Visual check of distribution of data and outliers
  - **Identify evolutionary outliers -**
    - Make outliers of trait distributions explicit (summary view/ ranked volatility in path)
    - Plot trait distributions in binned time segments (summary view).

Related Work

Related Work



Related Work



Related Work

Design highlights that support the tasks.


DATA: CENTRARCHID
SUMMARY VIEW
COLLAPSE ATTRIBUTES
SWITCH TO DISCRETE BARS
SHOWN ATTRIBUTES
SHOWN SUBTREES
LOG SCALE
COLOR BY TRAIT
PHENOGRAM
CLADE VIEW
LgibA
Volatility
Min Val
Max Val
AM_mass 0.00025 0.093 0.11
Close_Li 0.0080 2.6 2.8
Gape_w 0.87 8.8 11
LP_mass 0.000085 0.0041 0.013
Lo 0.87 12 14
Open_Li 0.0045 2.9 3.0
SL 9.1 89 93
UJ_Pro 0.035 2.1 2.5
max_TL 6.8e+3 2.7e+2 3.9e+2
LmicA
Volatility
Min Val
Max Val
AM_mass 0.0010 0.093 0.11
Close_Li 0.0057 2.6 2.8
Gape_w 0.77 8.8 11
LP_mass 0.000024 0.0041 0.013
Lo 0.76 12 14
Open_Li 0.032 2.9 3.1
SL 14 89 96

DATA: ANOLIS | PAIR VIEW | SHOWN ATTRIBUTES | SHOWN SUBTREES | TRAIT: BODY_WIDTH

CONVERGENCE (ANCIENT)

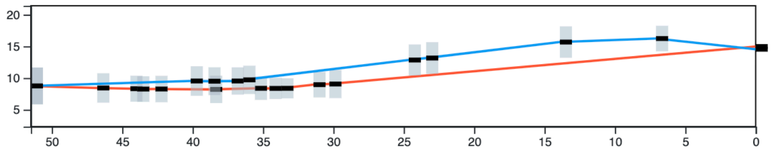
Distance -1.0 -0.5 0.0 0.5 1.0

Delta -1.0 -0.5 0.0 0.5 1.0

Closeness -1.0 -0.5 0.0 0.5 1.0

COLOR BY TRAIT | VIEW PHYLOGENY | CLADE VIEW

petersii + princeps

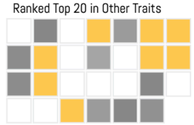


	Distance	Delta	Closeness
Score	1.00	0.562	0.978
Value	103	7.86	0430

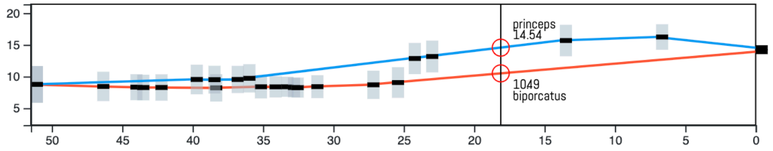
Distinctiveness Score 2.54

Rank 1

Ranked Top 20 in Other Traits



biporcatus + princeps

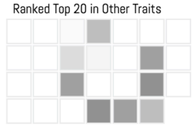


	Distance	Delta	Closeness
Score	1.00	0.514	0.968
Value	103	7.19	0.630

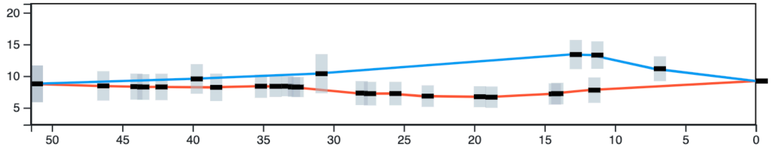
Distinctiveness Score 248

Rank 2

Ranked Top 20 in Other Traits



poecilopus + roquet

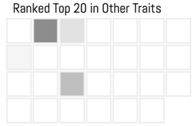


	Distance	Delta	Closeness
Score	1.00	0445	0.999
Value	103	6.23	0.0100

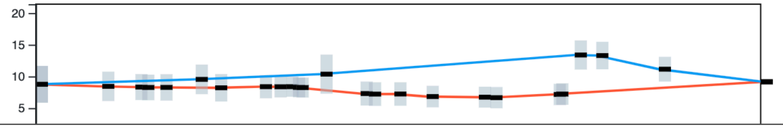
Distinctiveness Score 244

Rank 3

Ranked Top 20 in Other Traits



oxylophus + roquet

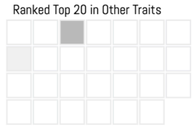


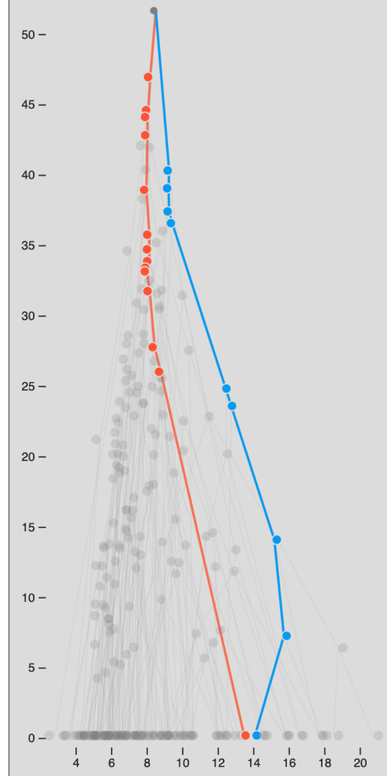
	Distance	Delta	Closeness
Score	1.00	0445	0.996
Value	103	6.23	0.0700

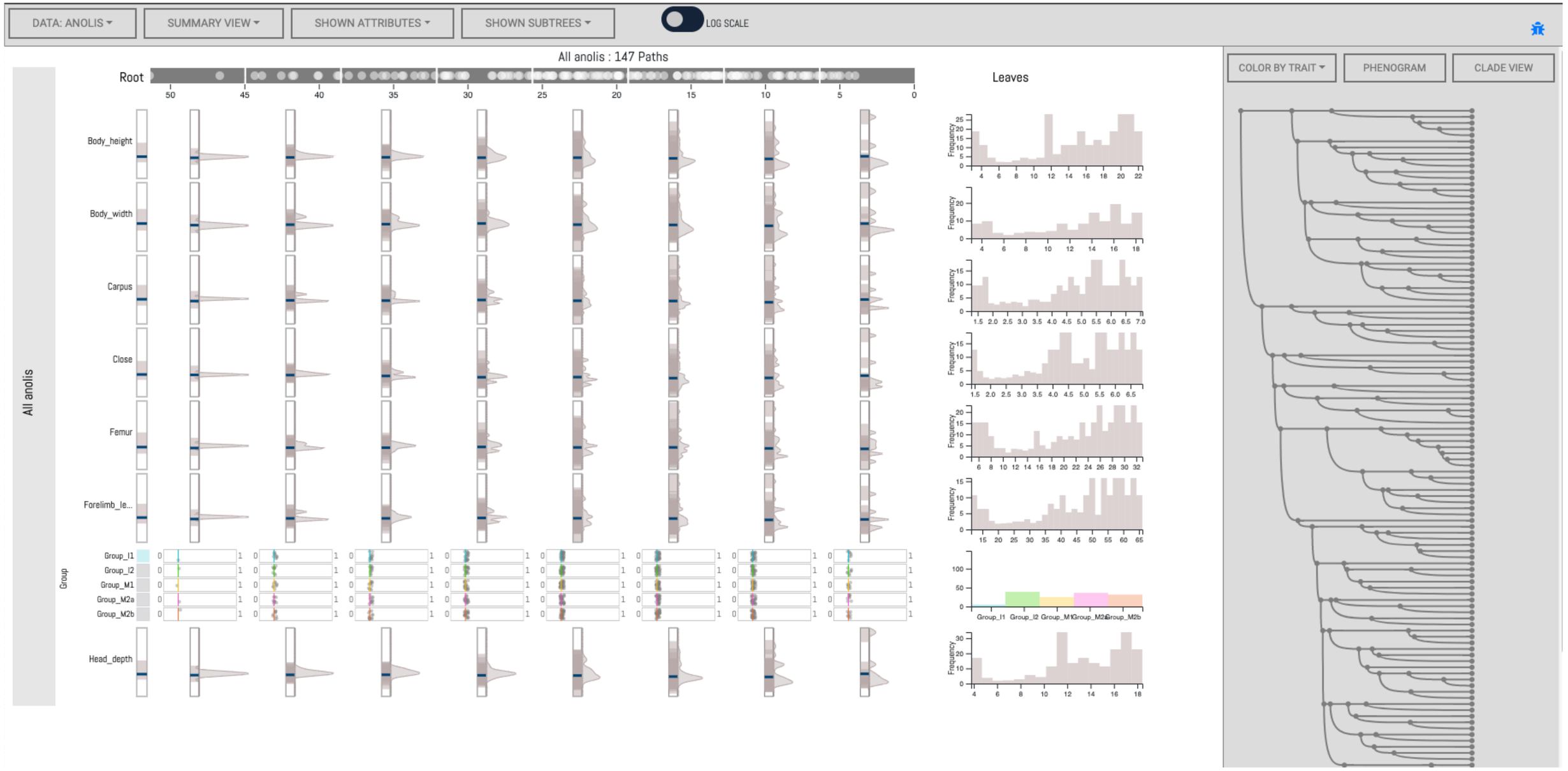
Distinctiveness Score 244

Rank 4

Ranked Top 20 in Other Traits






DATA: ANOLIS
SUMMARY VIEW
SHOWN ATTRIBUTES
SHOWN SUBTREES
LOG SCALE
All anolis : 147 Paths
Root
Leaves
All anolis
Body_height
Body_width
Carpus
Close
Femur
Forelimb_le...
Head_depth
Group
Group_I1
Group_I2
Group_M1
Group_M2a
Group_M2b
Frequency
COLOR BY TRAIT
PHENOGRAM
CLADE VIEW

The study process.

Immersed myself in the lab.

Familiarized myself with R Studio and tools they use for analysis.

Conducted interviews around the department.

Ran a CVO workshop.

Reflection is a large part of the development process

Process: Reflection through the study

Diverging from established design study methodology, we reflected through the process of the design study

Process: Reflection through the study

Shon - reflective conversation

Process: Reflection through the study

Uta - design as a speculative process

Lessons learned:

Reflection allowed reframing of ideas early in the process

Reviewing notes early in the process - things that seemed unimportant initially had new found value

Lessons learned:

Failed views could act as a catalyst for new insight

Lessons learned:

Looking at the plotted timeline of notes, the design development happened mainly in areas where reflection took place.

Abundance and Transparency

Evolutionary Biology Design Study Audit Trail

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

06/05/2019 Second Meeting with Daniel
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

07/05/2019 Sketchbook

07/09/2019 Sketchbook

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/

Evaluation

Challenge of evaluating design studies

The Challenge of Information Visualization Evaluation - plaisant

Evolutionary biologists tested the tool on their own computer

Case studies

Future work: