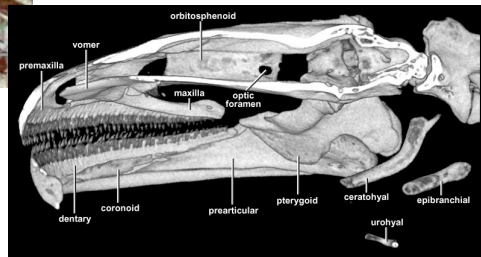


Parallel Coordinate Plots
towards
Phylogenetic Trees

Rebecca Shapley
SIMS 247, Spring 2004

How to make a tree

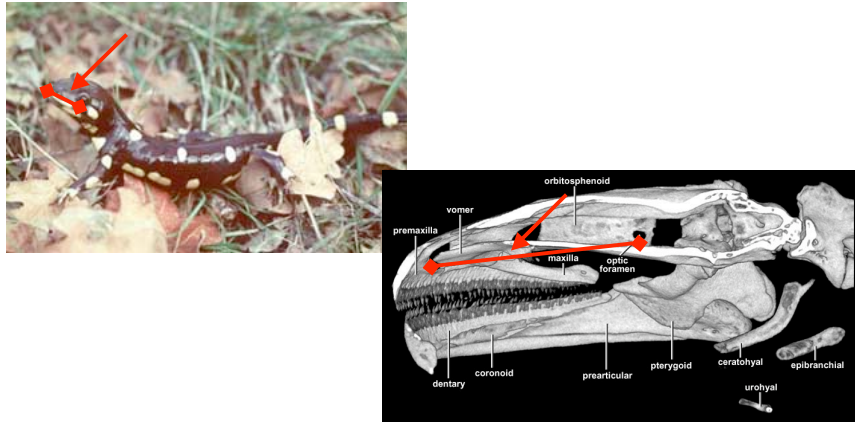
Start with organisms.



To make an evolutionary tree, biologists take organisms, identify attributes that have different values due to evolutionary events, (characters and character states) and make a huge matrix encoding which state each organism has, then using algorithms to convert the matrix into a tree. My interest is in determining if parallel coordinates visualization can assist in the early part of the process, for identifying attributes/characters and values/states.

How to make a tree


Identify characters.



To make an evolutionary tree, biologists take organisms, identify attributes that have different values due to evolutionary events, (characters and character states) and make a huge matrix encoding which state each organism has, then using algorithms to convert the matrix into a tree. My interest is in determining if parallel coordinates visualization can assist in the early part of the process, for identifying attributes/characters and values/states.

How to make a tree

Code which specimens have which characters.



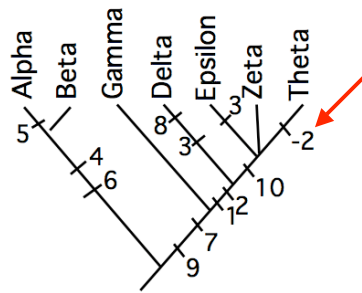
	Characters									
	1	2	3	4	5	6	7	8	9	10
Outgroup	0	0	0	0	0	0	0	0	0	0
Alpha	0	0	0	1	1	1	0	0	0	0
Beta	0	0	0	1	0	1	0	0	0	0
Gamma	0	0	0	0	0	0	1	0	1	0
Delta	1	1	1	0	0	0	1	1	1	0
Epsilon	1	1	1	0	0	0	1	0	1	1
Zeta	1	1	0	0	0	0	1	0	1	1
Theta	1	0	0	0	0	0	1	0	1	1

To make an evolutionary tree, biologists take organisms, identify attributes that have different values due to evolutionary events, (characters and character states) and make a huge matrix encoding which state each organism has, then using algorithms to convert the matrix into a tree. My interest is in determining if parallel coordinates visualization can assist in the early part of the process, for identifying attributes/characters and values/states.

Diagrams from Lipscomb, Diana. [Basics of Cladistic Analysis](#). George Washington University. Washington, D.C. 1998

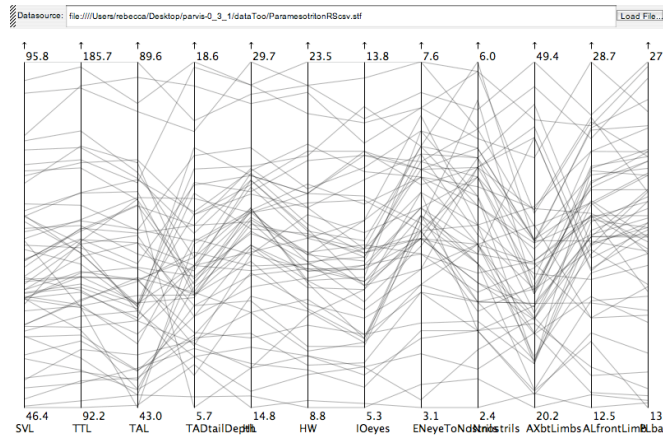
How to make a tree

Apply algorithm to
reconcile groupings
and propose a tree.



To make an evolutionary tree, biologists take organisms, identify attributes that have different values due to evolutionary events, (characters and character states) and make a huge matrix encoding which state each organism has, then using algorithms to convert the matrix into a tree. My interest is in determining if parallel coordinates visualization can assist in the early part of the process, for identifying attributes/characters and values/states.

Data: 60 Salamanders

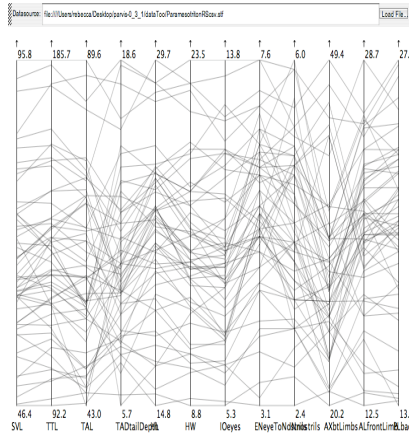


SVL: snout-vent length
 TTL: total length
 TAL: tail length
 TAD: maximum tail depth
 HL: head length
 HW: head width
 IO: distance between from corner of eyes
 EN: distance from front corner of eye to nostril
 IN: distance between nostrils
 AX: distance between limbs on right side
 AL: length of front limb
 PL: length of back limb

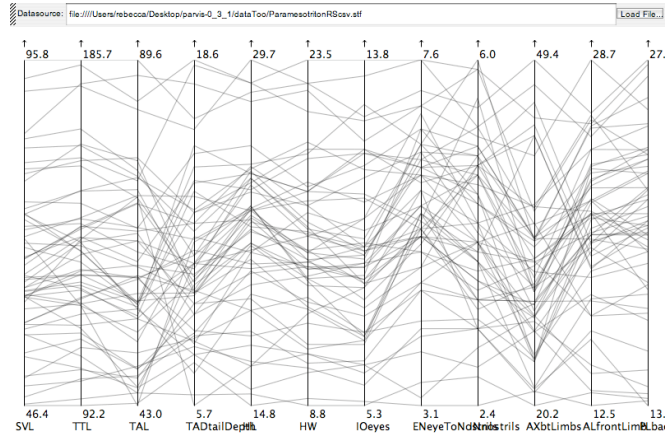
-From Sean Rovito, IB, UC Berkeley

Phylogenetic Features

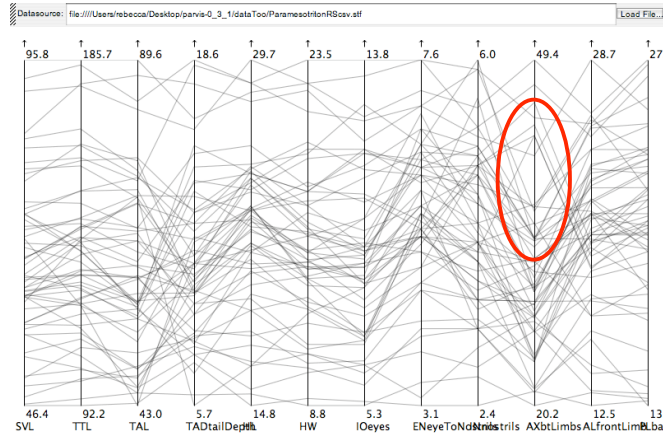
- Gaps
- Phylogenetically informative characters
- OTU signature
- Non-independence: combine for new one
- Non-independence: need to scale



Data: 60 Salamanders

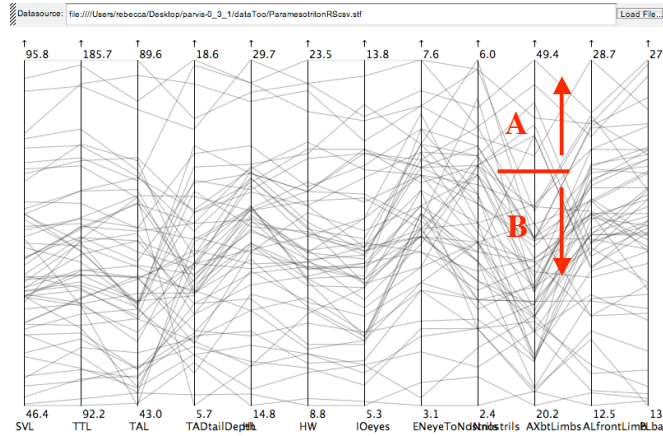


Data: 60 Salamanders



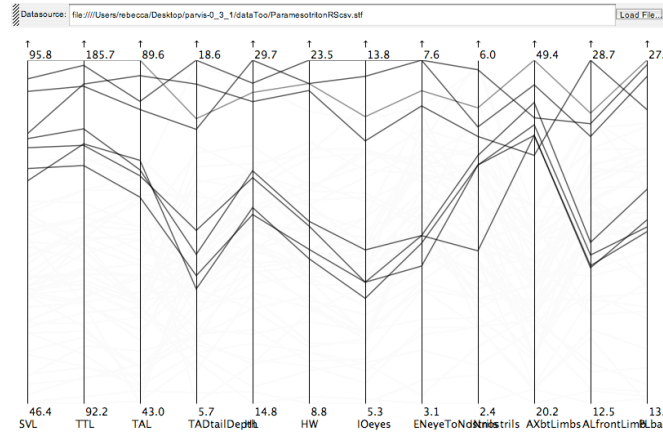
Notice the Gap.

Data: 60 Salamanders



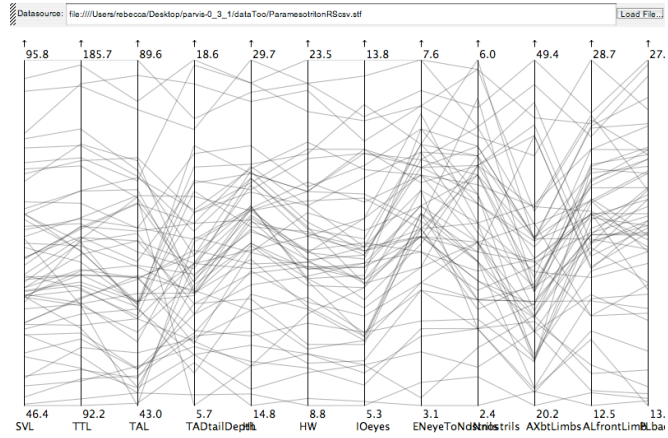
Notice the Gap.

Data: 60 Salamanders

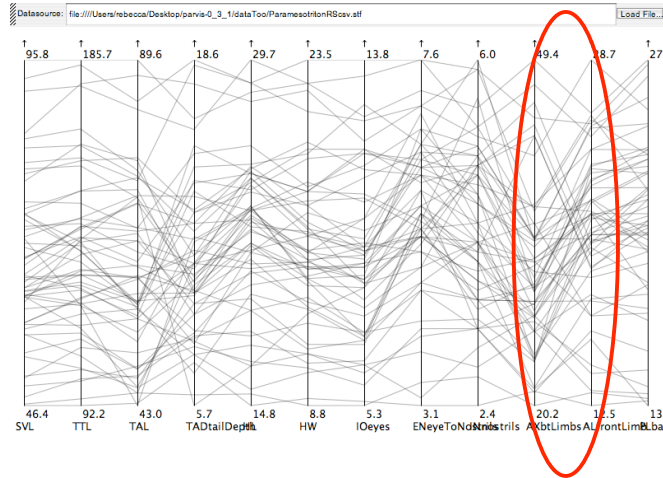


Brushing for “A”, can see the a likely OTU and that A is probably a parsimoniously informative character.

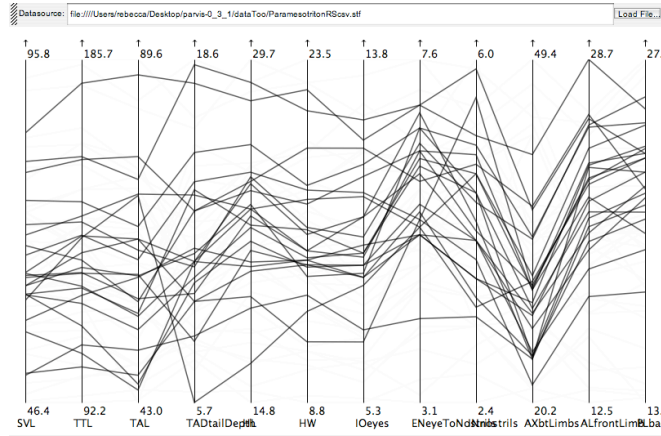
Data: 60 Salamanders



Data: 60 Salamanders

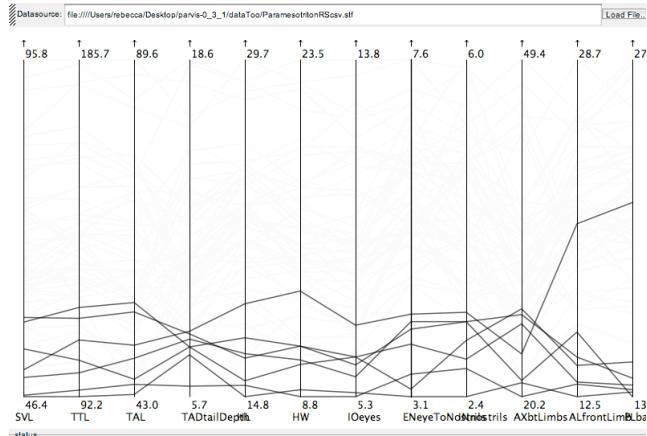


Data: 60 Salamanders



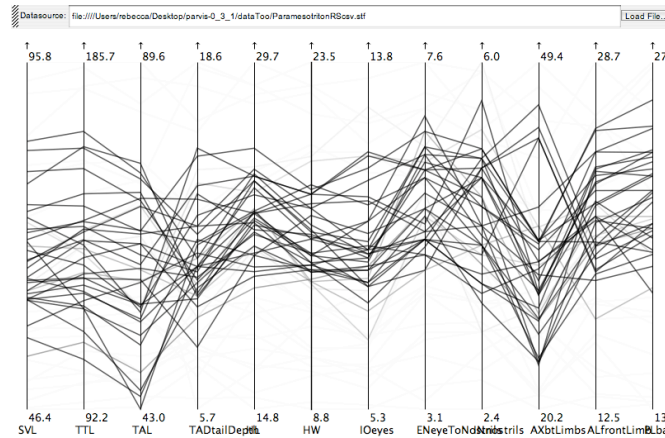
Just the records with a positive slope between AX and AL are selected.

Data: 60 Salamanders



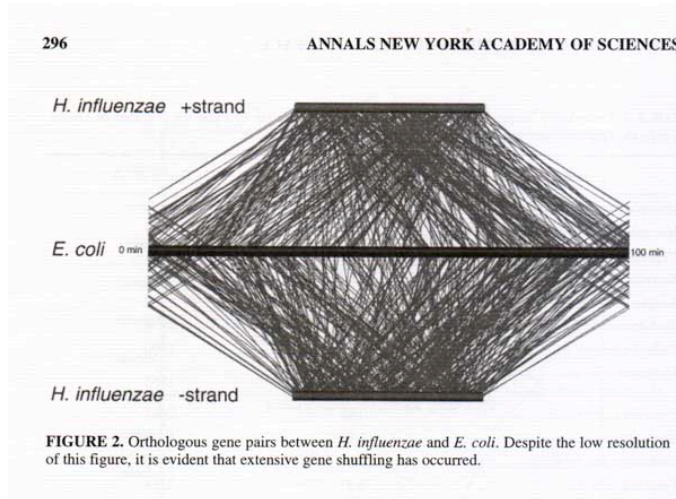
Notice - the lower values on most of the axes tend to be shared by the same lines...

Data: 60 Salamanders



As do the middle values... and the upper values. Lots of horizontal movement means a need to scale one value by other values to remove the redundant information about the absolute size of the organism, and focus instead on proportions.

Parallel Coordinates in Biology



Bellgard, Matthew I.; Itoh, Takeshi; Watanabe, Hidemi; Imanishi, Tadashi; Gojobori, Takashi. Dynamic evolution of genomes and the concept of genome space. Caporale, L. H., Ed. ANNALS OF THE NEW YORK ACADEMY OF SCIENCES; Molecular strategies in biological evolution. Vol:870, page:293-300, 1999.

Parallel Coordinates in Biology

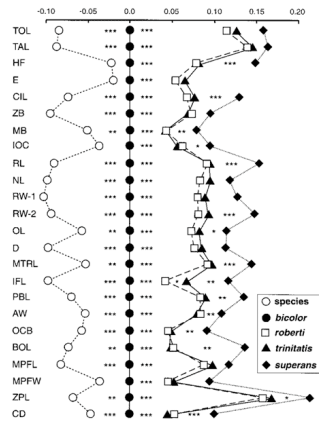
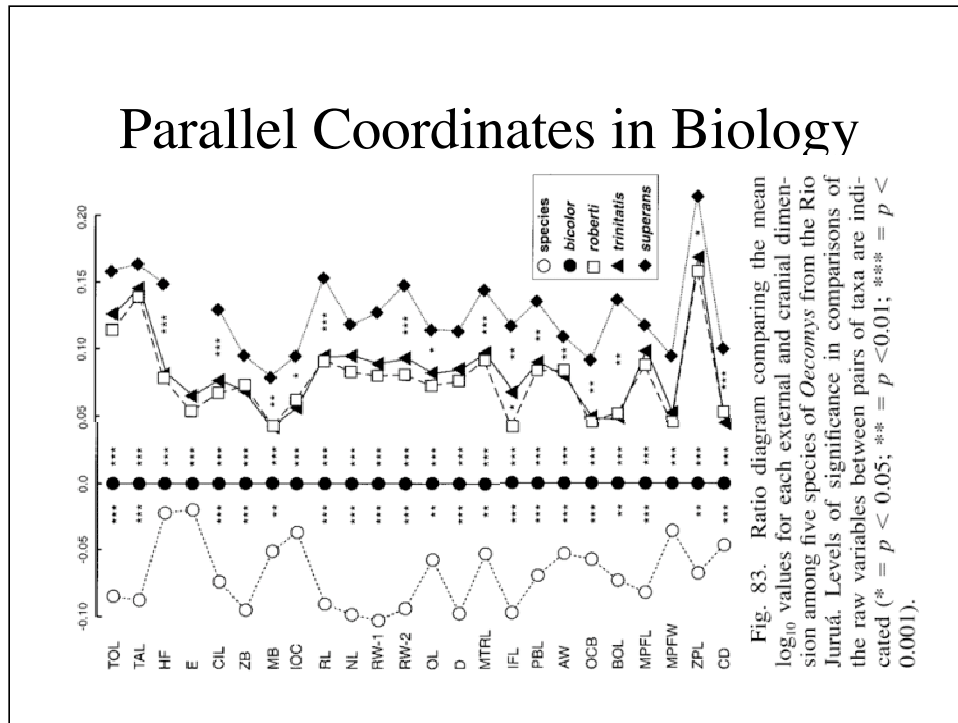


Fig. 83. Ratio diagram comparing the mean \log_{10} values for each external and cranial dimension among five species of *Occomys* from the Rio Juruá. Levels of significance in comparisons of the raw variables between pairs of taxa are indicated (* = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$).

Patton, James L.; Da Silva, Maria Nazareth F.; and Jay R. Malcom. Mammals of the Rio Juruá and the evolutionary and ecological diversification of Amazonia. BULLETIN OF THE AMERICAN MUSEUM OF NATURAL HISTORY. Vol:244, page1-306, 2000.

Parallel Coordinates in Biology



The measurements across the bottom are standard body and skull measurements for small mammals.

Patton, James L.; Da Silva, Maria Nazareth F.; and Jay R. Malcom. Mammals of the Rio Juruá and the evolutionary and ecological diversification of Amazonia. BULLETIN OF THE AMERICAN MUSEUM OF NATURAL HISTORY. Vol:244, page1-306, 2000.