

A Phylogeny From the Fish Up

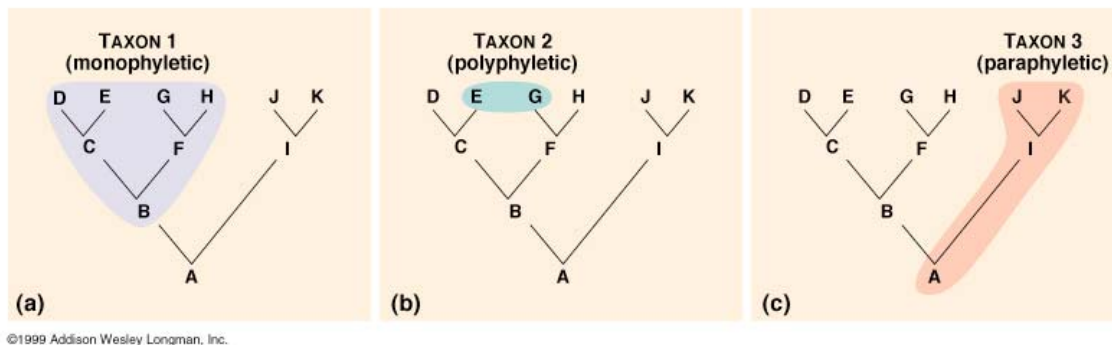
I. Definitions

Parsimony: Parsimony is the principle that the best answer to a problem is the simplest. This principle can be applied to reconstructing a phylogenetic tree (i.e., hypotheses of how organisms are related to one another) by choosing the tree with the fewest evolutionary steps or transformations.

Monophyly: A monophyletic group is one in which all members are descended from one common ancestor (**a** in the figure below).

Polyphyly: A polyphyletic group is one in which its members are descended from multiple most common recent ancestors (**b** in the figure below).

Paraphyly: A paraphyletic group is one in which not all descendants of a common ancestor are included (**c** in the figure below).



Pleisiomorphy: A primitive or ancestral character.

Sympleisiomorphy: A shared primitive character.

Apomorphy: A derived character.

Synapomorphy: A shared derived character.

Autapomorphy: A unique derived character.

II. Building a Phylogeny from the Fish Up

Break in to groups of two (three is okay if there is an odd number of people). Each group will be supplied a jar of assorted fish specimens, each specimen will have a letter (A-G) tied around the caudal peduncle.

Examine the specimens carefully for the following character states (in some cases, since we will not be dissecting any of the specimens, you may need to use what you have learned in lecture to determine the state of characters):

1. Fins present as paired appendages (0=absent, 1=present)
 2. Placement of pelvic fins (0=abdominal, 1=thoracic or jugular)
 3. Pectoral-fin fusion (0=absent, 1=present)
 4. Anal-fin spine(s) (0=absent, 1=present)
 5. Dorsal fin containing spines (0=absent, 1=present)
 6. Scales (0=none, 1=placoid, 2=cycloid or cteonid)
 7. Head with spines (0=absent, 1=present)
 8. Insertion of pectorals (0=ventral, 1=high on body)
 9. Jaws (0=absent, 1=present)
 10. Gill opening (0=one per side, 1=two or more per side)
 11. Caudal tail (0=heterocercal, 1=homocercal)
2. Prepare a table showing the distribution of the character states (i.e. a character matrix). Arrange your taxa (A-G) along the vertical axis, and your characters along the horizontal axis (see whiteboard for an example).
3. Sum the rows, and then order them from lowest (primitive) to highest (derived) score. *Please note* that this is a step of convenience. Summation of character states does not indicate a taxon's placement in a phylogeny. If we asked you to search all of "tree space" for the most parsimonious tree, we would be here for days.
4. Construct the most parsimonious cladogram (i.e., one that requires the FEWEST changes) using the scores you assigned to each taxon. Your tree does not need to be perfectly symmetrical or asymmetrical (see whiteboard), it just needs to be the most parsimonious cladogram.
5. Label the taxa and character state changes (see whiteboard for example) on your final tree.
6. Answer the following questions, regarding the phylogeny you created, in your lab notebook:
- a. Name an autapomorphy for taxon E. Is this character parsimony informative (does the information it provides change the shape of the tree)? Why or why not?
 - b. Name a synapomorphy for the sister group to taxon D.
 - c. Can you use the presence or absence of jaws to determine if taxa B and G maintain a sister relationship? Why or why not?
 - d. According to a current classification, the rajiid (taxon E) and the squalid (taxon A) are both placed in the class Chondrichthyes. Based on your phylogeny, is this a monophyletic assemblage? How do you know?
 - e. You come across an order of fishes in the literature that includes salmonids (taxon F), gadids (C), and scorpaenids (G) to the exclusion of serranids (B). Based on your phylogeny, would you deem this group mono-, para-, or

polyphyletic? Briefly explain, or illustrate, how you determined this.

III. A Computer Generated Phylogeny of the Dragonfishes

When your TA determines that Part II has been successfully completed, he/she will take everyone to the computer lab and review the instructions for Part III. Please take with you the additional handout and appendix.

In this part of the lab, you'll explore how phylogenies are hypothesized using modern phylogeny software. In particular, we'll analyze a morphological dataset taken from W. L. Fink's (1985) examination of the most derived stomiiform family, Stomiidae, or dragonfishes. As part of this, you'll also make use of this phylogeny to comment on the validity of an older classification of these genera.

How difficult is it to find the "right" tree?

You may have wondered when producing a phylogeny in Part I of the lab, when would a systematist make it easier or him- or herself and abandon the scratch paper and start using a computer? Nowadays, it's universal to see even the smallest datasets taken from even only a few taxa analyzed with phylogeny software. Finding the best tree (i.e., one with the fewest steps when using the parsimony method) is one of the more intense operations in all of biology. Why? **The number of trees that exist increases exponentially with each taxon we consider.** Here's a table describing this increase (from Felsenstein, 2004):

Number of species	Number of trees
1	1
2	1
3	3
4	15
5	105
6	945
7	10,395
8	135,135
9	2,027,025
10	34,459,425
15	213,458,046,676,875
20	8,200,794,532,637,891,559,537
50	2.75292×10^{76}

So, for a dataset that considers 50 species, we have to consider a number of solutions (i.e., trees) that approaches the number of electrons in the universe!!

Fortunately, we'll use an algorithm that trims the possible number of trees considerably. This algorithm, the branch-and-bound method, allows the computer to look only for better solutions. For example, if the computer finds a tree with 45 steps, it won't consider any sets of trees that incur more steps. We'll implement this approach using the PENNY program in **PHYLIP**, a free phylogeny program created and made available by **Joe Felsenstein** of UW's Genome Sciences department.

About the dataset

The data you'll analyze in PHYLIP are taken from a study of dragonfish relationships published by W. L. Fink in 1985. The family contains over 200 species in 27 genera. We'll consider only 14 of these genera—analyzing data from all 27 genera would take many, many, many hours, even with the world's fastest computers. The data are largely descriptions of the osteology of these fishes, and in particular, the bones of the neurocranium. **Osteological data** are the most ubiquitous form of morphological data used in phylogenetic studies of vertebrates. Below you'll find several example drawings taken from Fink's study that describe the osteology of these genera.

Just as you produce a matrix for Part I of the lab, these data are arranged in tabular form with each character numbered. "0" is the primitive state (represented in each character by the outgroup, a primitive stomiiform), "1" is derived, and "?" means the author was unsure or the character was unclear. Each character is described in detail in the appendix under "Character argumentation."

Classification of the Stomiidae

Before Fink's (1985) study, the Stomiidae, as they are now recognized, were divided among six families. The following is a pre-Fink classification of the genera you'll analyze:

Suborder Stomiatoidei

Family **Chauliodontidae**

Genus *Chauliodus*

Family **Stomiidae**

Genus *Stomias*

Family **Astronesthidae**

Genus *Astronesthes*

Genus *Heterophotus*

Genus *Rhadinesthes*

Family **Melanostomiatidae**

Genus *Melanostomias*

Genus *Echiostoma*

Genus *Photonectes*

Genus *Tactostoma*

Genus *Eustomias*

Family **Idiacanthidae**

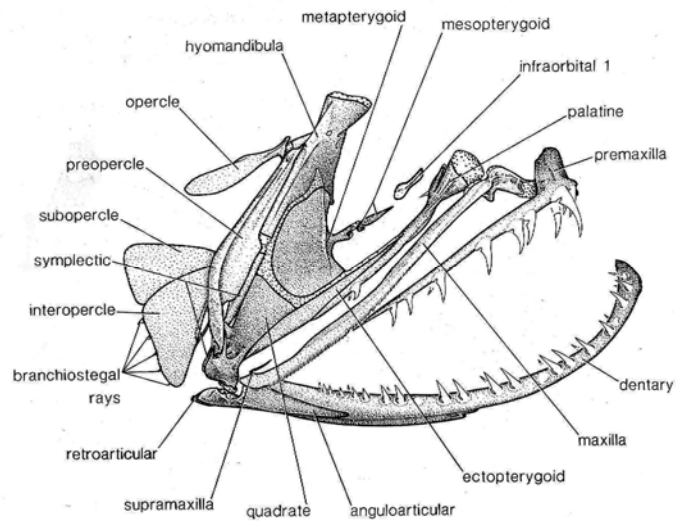
Genus *Idiacanthus*

Family **Malacosteidae**

Genus *Malacosteus*

Genus *Photostomias*

Genus *Aristostomias*



The jaws and suspensorium of *Stomias lampropeltis* (from Fink, 1985).

Using PHYLIP to create a tree

PHYLIP is a simple, command-line operated package. It's really no frills. There are several other phylogeny packages available, both free and for purchase. Essentially, they all work the same, with identical mathematical algorithms.

Open the PHYLIP folder on the Desktop. Open the text file named “311_Stomiidae.txt” The “15” and “68” refer to the number of taxa and number of characters, respectfully. PHYLIP has a limit to the number of characters in a taxon name—that’s why the names are incomplete. Close the text file.

To load the data and specify and output file:

1. Open “PENNY.exe” This is the program that with search for the best tree using parsimony as the criterion and using the branch-and-bound algorithm.
2. PENNY will tell you, “can’t find input file ‘infile’ Please enter a new file name” Enter “311_Stomiidae.txt”
3. PENNY will then tell you, “the file ‘outfile’ you wanted to use as output file already exists. Do you want to replace it, Append to it, write to a new File, or Quite?” Type “f” to create a new file.
4. PENNY will then tell you to enter a new file name. Enter “311_Out” Hit enter.
5. The PENNY screen now appears.

The PENNY settings screen looks pretty complicated, but for the most part, the default settings are exactly what we want. These settings will root the tree with the first taxon in the list, our “outgroup”, and this will conveniently make our “outgroup” the outgroup. PENNY will also print the tree to the outfile and a tree file, too.

Running PENNY:

1. Enter “Y” to tell PENNY that the settings are correct.
2. PENNY will tell you, “the file ‘outtree’ that you wanted to use as output tree file already exists. Do you want to Replace it, Append to it, write to a new File, or Quite?” Enter “F” to create a new file.
3. PENNY will then tell you to enter a new file name. Enter “311_Outtree” Hit enter. PENNY will close and a computation screen will flash.
4. Congratulations!! You’ve created a computer phylogeny!!

Now open the output file “311_Out”. You may have to select a text-reading program to access the file. **How many steps (i.e., changes) does the tree involve?** Close this file.

Creating a cladogram:

1. Now open “DRAWGRAM.exe”. This program will produce a tree in the form of a cladogram. DRAWGRAM will tell you that the font file can’t be found. Enter “font1” and hit enter.
2. The DRAWGRAM screen will appear with a list of settings. Enter “s” to change the tree style. DRAWGRAM will then ask you to choose from a list. Enter “c” for cladogram and then hit enter.
3. Enter “H” to so that the tree grows vertically rather than horizontally.
4. Now enter “Y” to accept the change and draw the cladogram. A screen with a printed tree will now appear. Your TA will now hand you a copy of this tree.

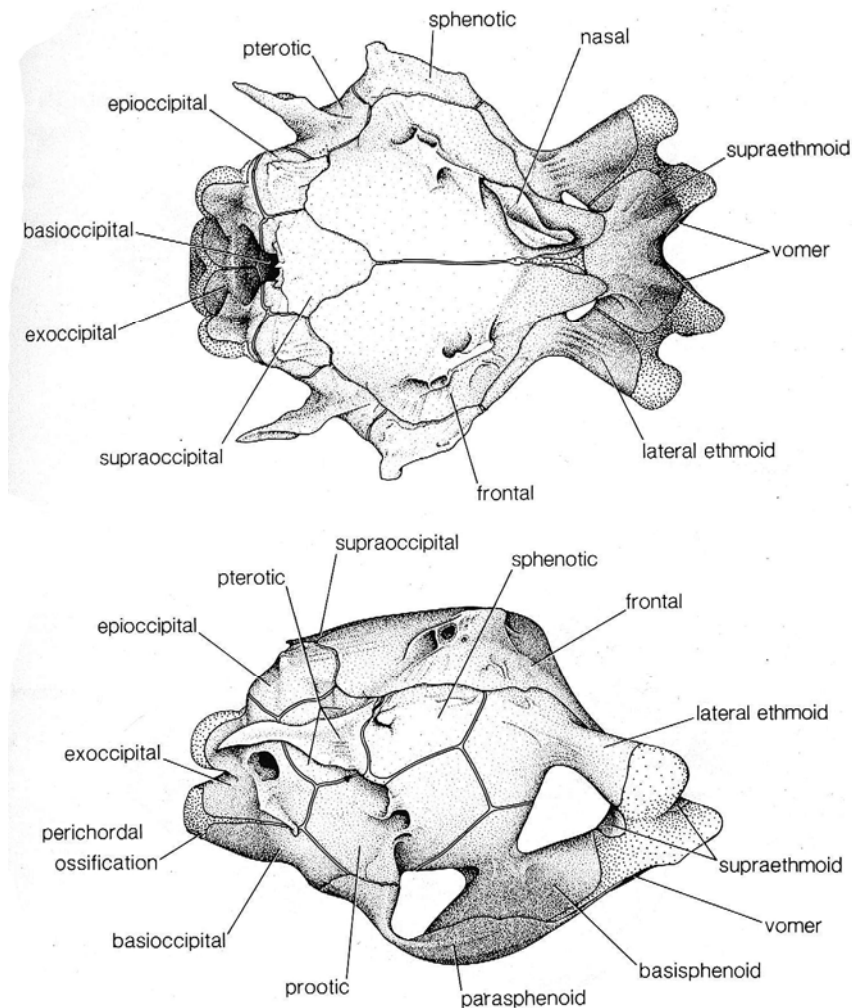
Reconciling the old classification of these genera with your phylogeny

Look carefully at your tree. Does the phylogeny support the older classification of these genera? Is each family included in the classification above **monophyletic**, that is, do members of that family share a most recent common ancestor (i.e., single node)? Or are any of these families **paraphyletic**, that is, does a family contain genera from another family? Devise a classification that reflects the phylogeny you produce using PENNY.

Literature Cited

Felsenstein J. 2004. *Inferring Phylogenies*. Sunderland, MA: Sinauer.

Fink WL. 1985. Phylogenetic interrelationships of the stomiid fishes (Teleostei: Stomiiformes). *Miscellaneous Publications Museum of Zoology University of Michigan* 171:i-vii, 1-127.



The neurocranium of *Photostomias guernei* (from Fink, 1985).

IV. Phylogeny for the Faint of Heart: a Tutorial

This is essentially homework. It is not to be done during lab time unless you find yourself with nothing else to do, but you probably won't. Please read Sandra Baldauf's (2003), *Phylogeny for the Faint of Heart: a Tutorial* and, in your lab notebook, answer the following questions:

1. What is an OTU? Name three OTUs from the tree you created in Part II.
2. What is the "root" of a phylogenetic tree? How do you root a tree? What are some difficulties that might be encountered when rooting a tree?
3. You have the following DNA sequences from taxa A-F, plus the designated outgroup (OUT):

A	G	T	C	T	A	C	G	C	A	T
B	G	G	C	T	A	C	G	C	A	T
C	A	G	G	C	T	A	C	G	C	A
D	G	G	C	T	A	C	G	C	A	T
E	G	G	C	T	A	C	G	C	A	T
F	G	T	C	T	A	C	G	C	A	T
OUT	G	A	C	T	A	C	G	C	A	T

Is this data matrix ready to be analyzed? What would have to be done in order to infer a rooted tree, for taxa A-F, from this data? Once that has been done, can you tell if any of the sites are synapomorphies (just by eye, you don't need to make a tree)? If so, which group(s) share(s) it/them?

4. What is the difference between the two classes of methods, described by the author, for inferring phylogenies?
5. What is "bootstrapping"? Why would you want to use it?
6. What is "long branch attraction"? How can you mitigate this problem in your analysis?

Character argumentation:

1. conical bone in pocket on ventrolateral ethmoid cartilage
2. elongate groove in lateral ethmoid and pterospenoid
3. lateral wall of nasal larger than medial, forms cup
4. posterior edge of parasphenoid ends anterior to posteroventral edge of basioccipital
5. sphenotic spine extends posterior to middle of prootic and anterior edge of hyomandibula
6. post process of pterotic robust, projects posterodorsal
7. exoccipital cartilage projects anterior to post border of process
8. infraorbitals except first absent
9. ligament attaches near anterodorsal dentary looping post to anterior floor of mouth
10. palatine cart interrupted, palatine ends post in point
11. vent border of palatine arch arched dorsal in lateral view
12. Palatine teeth in two area
13. mesopterygoid reduced
14. preopercle very narrow at symplectic-hyomandibula joint
15. interhyal articulation w/ anterior margin of cart between hyomandibula & symplectic, bound to metopterygoid by a ligament
16. basihyal \geq length of anterior border of hypohyal element
17. basihyal reduced
18. depth of post urohyal $<$ depth of basibranchial series
19. blood vessel passes through hypohyal element and penetrates lateral face of element in ventral hypohyal
20. ligament between basihyal and hypohyal attaches to anterior 1/2 of hypohyal
21. margin of hyohyal element originates at 60 deg angle to dorsal margin of basibranchial 1
22. hypohyal element about 2x long anteriorly to posteriorly as dorsoventrally
23. anterior hypohyal projects anterolateral
24. gill rakers absent in adults
25. one or fewer toothpaches associated with any basibranchial ossification
26. pair of tooth plates associate with basibranch 1 fused at midline
27. basibranchial 1&2 toothpaches limited to two per basibranchial, arranged in bilateral pairs
28. basibranchial 3 about four-fifths the length of basibranchial 2 or shorter
29. fourth basibranchial square from dorsal aspect
30. cartilage-filled body of pharyngobranchial 3 distinctly concave along medial border
31. cartilage of pharyngobranchial 3 more curved
32. no branchiostegal rays articulation on ventral hypohyal near anterior border
33. branchiostegal rays deeply bifurcated dorsally
34. dorsal portion of geniohyoideus has both tendinous and muscular insertion on dentary
35. ligament from dorsolateral tip of fourth epibranchial to cleithrum near dorsal tip
36. ventral rector gill arch muscles attach to ceratobranchial
37. portion of the adductor mandibulae inserts on postorbital photophore
38. dorsal section of medial division of adductor mandibulae has origin anterior to insertion of levator arcus palatini
39. most or all of supraneurals absent
40. spinous portion of most of neural arches between anterior arch and dorsal fin org do not meet at midline
41. anterior portion of neural arch of first centrum greatly enlarged
42. caudal vertebral centra reduced in size

43. anteriormost, fully developed vertebral centrum and posteriormost symmetric centrum small
44. dorsal fin well anterior to midbody
45. hypural 6 not present in caudal skeleton
46. no posttemporal bone
47. dorsal tip of cleithrum an elongate slender spine
48. ventromedial surface of cleithrum concave near main curvature
49. edge of coracoscapular plate between proximal scapular ramus and med point of articulation of proximal radials concave
50. coracoscapular plate dorsoventrally thicker
51. fibrous remnant of mesocoracoid extends from medial margin of coracoscapular plate
52. lateral two proximal radials tightly articulated in a posterior deep concavity of coracoscapular plate
53. first proximal radial with a cartilaginous lateral border
54. proximal radial positioned posteromedial to other radials
55. all distal radials but anterior are reduced to small, oblong or round cartilage are absent
56. distal radial of first p1 fin ray, when present, remains at least partially unossified anterodorsally
57. three or fewer distal radials
58. anterior one or two fin rays reduced in length, anterior two to four rays tightly bound together (i.e., rod-ray complex)
59. anterior cartilaginous process is enlarged, extending as an elongate conical cart
60. length of anterior process roughly equal or less than length of post plate
61. posterior pelvic plate enlarged anteriorly and posteriorly
62. posterior cartilaginous process long near lateral margin

63. medial pelvic radial anteriorly elongate, extending at least as far anteriorly as articulation of 3rd & 4th fin rays
64. postorbital photophore ventral or posteroventral to eye
65. gelatinous membrane over body, scales buried within
66. hexagonal pigment in skin
67. mental barbel present, extending from urohyal
68. no skin between mandibular rami