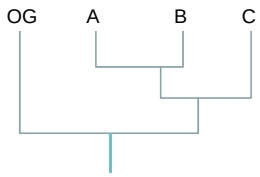
Intro to Cladistic Analysis

I. Overview

A. Scope

1. Ingroup vs. outgroup

e.g., angiosperms vs. non-angiosperms
e.g., primates vs. non-primates



I. Overview

B. Phylogenetic evidence
1. Characters & Character states

a. Morphological Characters

e.g., Phyllotaxy: 0 = alternate; 1 = opposite; 2 = whorled.

e.g., Carpel Presence: 0 = absent; 1 = present.

I. Overview

B. Phylogenetic Evidence
1. Characters & Character states

b. Molecular Characters
(e.g., DNA sequence characters)

Protein coding genes (e.g., cox1 or rbcL)
Structural RNA genes (e.g., 18S or 26S rDNA)

I. Overview

B. Phylogenetic Evidence
1. Characters & Character states

b. Molecular Characters
(e.g., DNA sequence characters)

Protein coding genes (e.g., cox1 or rbcL)
Structural RNA genes (e.g., 18S or 26S rDNA)

e.g., Position 36 in rbcL gene: 0 = A; 1 = G; 2 = C; 3 = T.

or

e.g., Position 36 in rbcL gene: A; G; C; T.

I. Overview

B. Phylogenetic Evidence
1. Characters & Character states

b. Molecular Characters

(e.g., amino acid sequence characters)

e.g., Position 5 in rubisco: 0 = glutamine; 1 = proline; etc.

I. Overview

B. Phylogenetic Evidence

c. Primary Homology Assessment (during character coding and scoring)

e.g.,
PETAL COLOR:
0 = red; 1 = blue

Position 4 in rbcL
A ; G; C; or T

I. Overview

B. Phylogenetic Evidence

c. Primary Homology Assessment (during character coding and scoring)

e.g.,
PETAL COLOR:
0 = red; 1 = blue

Position 4 in rbcL
A ; G; C; or T

d. Data Matrix

	VERTEBRAE	LEGS	ENDOTHERMY	FUR	MAMMARY GLANDS	OPPOSABLE THUMB	TAIL
Snail	0	0	0	0	0	-	-
Fish	1	0	0	0	0	-	1
Lizard	1	1	0	0	0	0	1
Bird	1	1	1	0	0	0	1
Cow	1	1	1	1	1	0	1
Monkey	1	1	1	1	1	1	1
Gorilla	1	1	1	1	1	1	0
Human	1	1	1	1	1	1	0

*Note: a "-" denotes that the character is inapplicable—i.e., not relevant—for that species.

I. Overview

C. Recognizing synapomorphies to resolve cladogram

1. Uses the OG to root the cladogram (cladogram is rooted Between the OG and the IG)

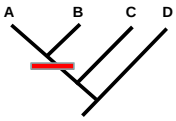
2. Parsimony

Leaf shape

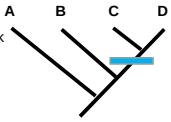
IG Sp. A cordate
IG Sp. B cordate
IG Sp. C elliptic
IG Sp. D elliptic

Problem: These are equally parsimonious

1 step on this rooted network



1 step on this rooted network

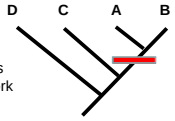


Leaf shape

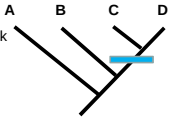
IG Sp. A cordate
IG Sp. B cordate
IG Sp. C elliptic
IG Sp. D elliptic

Problem: These are equally parsimonious

1 step on this rooted network

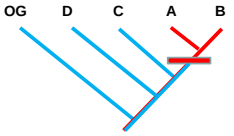


1 step on this rooted network

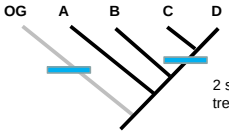


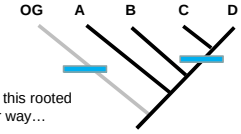
Leaf shape

IG Sp. A cordate
IG Sp. B cordate
IG Sp. C elliptic
IG Sp. D elliptic
OG Sp. elliptic



1 step on this rooted network





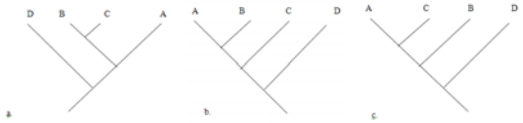
2 steps on this rooted tree, either way…

II. Cladogram Construction

A. Choose best cladogram

1. Start with all possible ingroup topologies

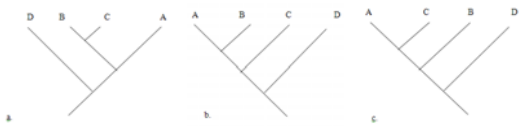
	char. 1	char. 2	char. 3	char. 4
Species A	0	0	1	1
Species B	1	1	0	0
Species C	1	0	1	1
Species D (outgroup)	1	1	1	0



II. Cladogram Construction

2. Use parsimony to map characters

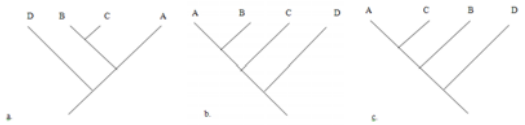
	char. 1	char. 2	char. 3	char. 4
Species A	0	0	1	1
Species B	1	1	0	0
Species C	1	0	1	1
Species D (outgroup)	1	1	1	0



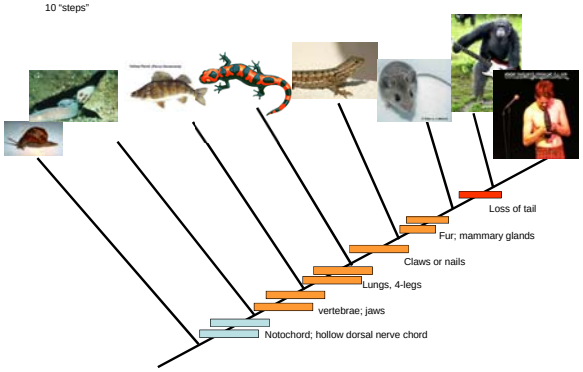
II. Cladogram Construction

3. Choose cladogram with fewest steps

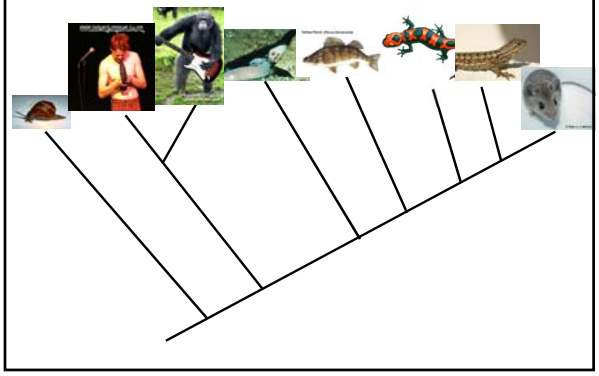
	char. 1	char. 2	char. 3	char. 4
Species A	0	0	1	1
Species B	1	1	0	0
Species C	1	0	1	1
Species D (outgroup)	1	1	1	0



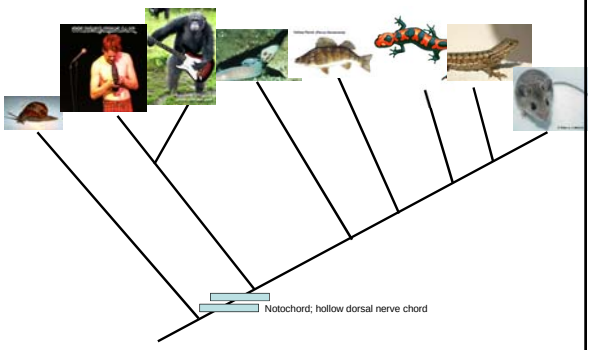
10 "steps"



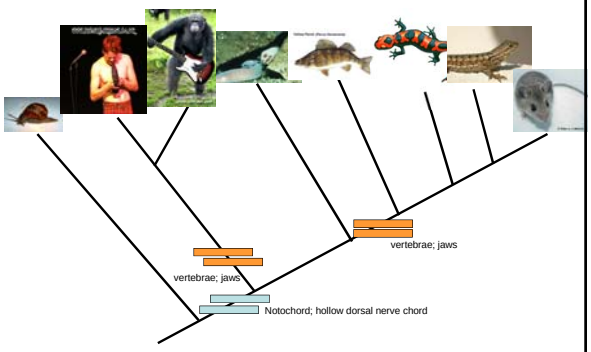
How to evaluate an alternative cladogram? Parsimony (number of steps)



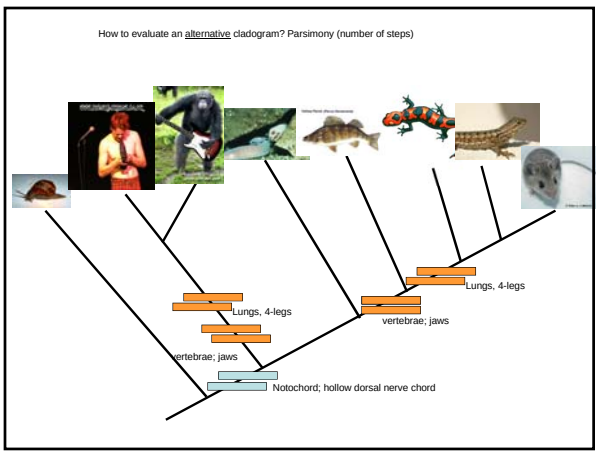
How to evaluate an alternative cladogram? Parsimony (number of steps)



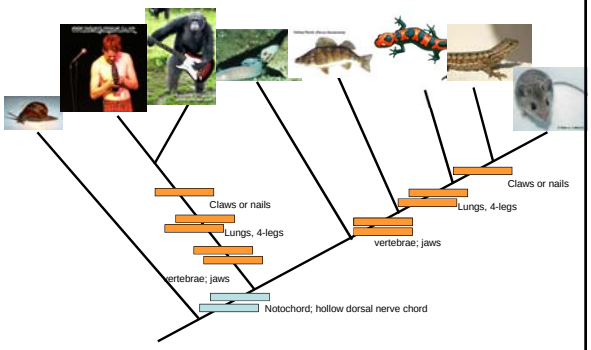
How to evaluate an alternative cladogram? Parsimony (number of steps)



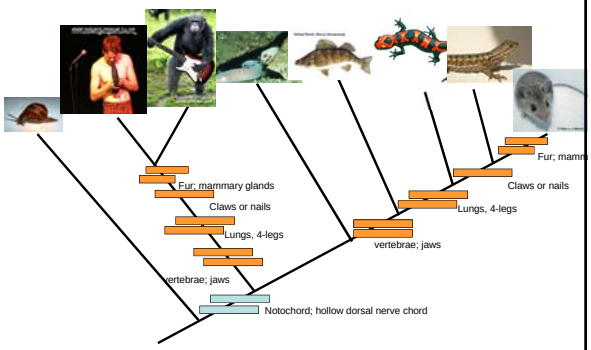
How to evaluate an alternative cladogram? Parsimony (number of steps)



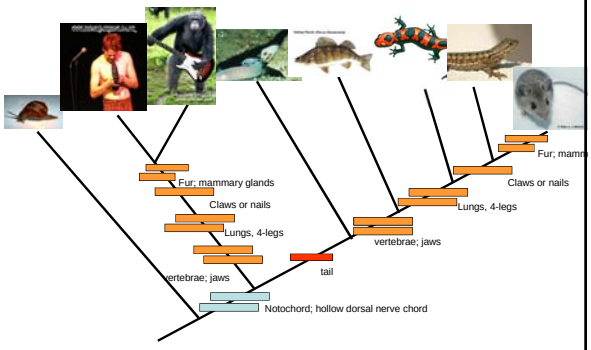
How to evaluate an alternative cladogram? Parsimony (number of steps)



How to evaluate an alternative cladogram? Parsimony (number of steps)

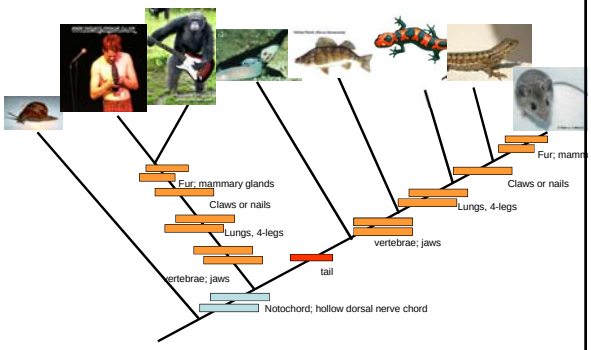


How to evaluate an alternative cladogram? Parsimony (number of steps)



How to evaluate an alternative cladogram? Parsimony (number of steps)

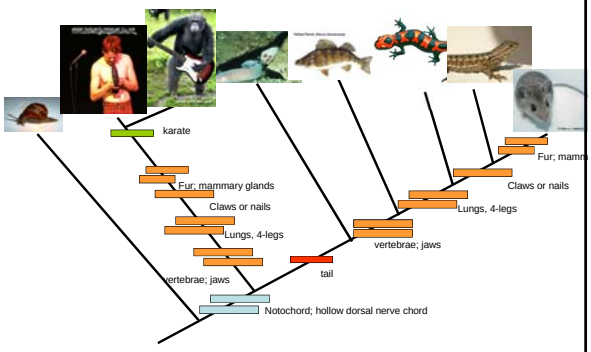
17 "steps"



How to evaluate an alternative cladogram? Parsimony (number of steps)

18 "steps"

http://www.youtube.com/watch?v=_kRb06w_XPo



II. Cladogram Construction

B. The size of the problem

1. Number of stars in the universe

Imagine Home | Ask an Astrophysicist | The Number of Stars

Ask an Astrophysicist

The Question

(Submitted January 15, 1997)

How many stars are there, named and un-named, known to exist ?

The Answer

This is a very good question! There are too many stars for scientists to actually count one-by-one, so other methods of *estimating* the total number of stars are used. We believe that there are on the order of 10^{21} stars in our Universe. If you write that number out, it looks like this: 1,000,000,000,000,000,000,000. This is a lot of stars!

Sincerely,

II. Cladogram Construction

B. The size of the problem

2. Number of possible trees

The Number of Evolutionary Trees

Joseph Felsenstein

Systematic Zoology, Vol. 27, No. 1. (Mar., 1978), pp. 27-33.

COUNTING TREES 31

TABLE 1. THE NUMBERS OF ROOTED TREES WITH n LABELLED TIPS AND WITH UNLABELLED INTERIOR NODES. THE LEFT COLUMN COUNTS ALL TREES, THE RIGHT COLUMN ONLY BIFURCATING TREES.

n	All trees	Bifurcating trees
1	1	1
2	1	1
3	4	3
4	26	15
5	236	105
6	2,752	945
7	39,208	10,395
8	660,032	135,135
9	12,818,912	2,027,025
10	282,137,824	34,459,425
11	6,939,897,856	654,729,075
12	188,666,182,784	13,749,310,575
13	5,617,349,020,544	316,234,143,225
14	181,790,703,209,728	7,905,853,580,625
15	6,353,726,042,486,112	213,458,046,676,875
16	238,513,970,965,250,048	6,190,283,353,629,375
17	9,571,020,586,418,569,216	191,898,783,962,510,625
18	408,837,905,660,430,516,224	6,332,659,870,762,850,625
19	18,522,305,410,364,569,764,416	221,643,095,476,699,771,875
20	887,094,711,304,094,583,065,206	8,200,794,532,637,891,559,375
21	44,782,218,857,751,551,087,214,592	319,830,986,772,877,770,815,625
22	2,376,613,641,928,796,906,249,519,104	13,113,070,457,687,988,603,440,625

$$1 \cdot 3 \cdot 5 \cdots (2n-3) = \prod_{k=2}^{n} (2k-3)$$

$$= \frac{(2n-3)!}{2^{n-2}(n-2)!} \quad (1)$$

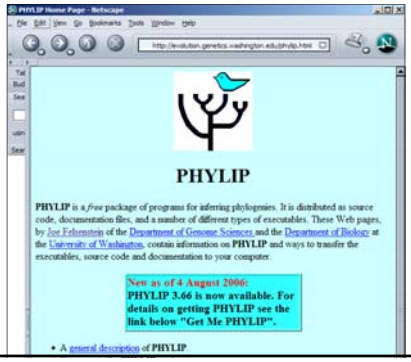
II. Cladogram Construction

C. Major Parsimony-based cladistic programs

Hennig86

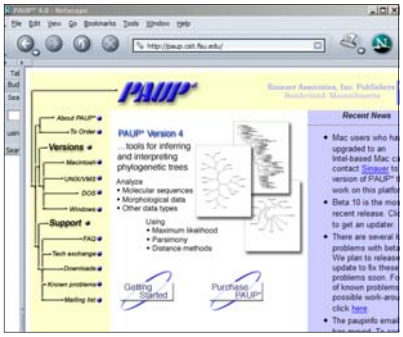
II. Cladogram Construction

C. Major Parsimony-based cladistic programs



II. Cladogram Construction

C. Major Parsimony-based cladistic programs



II. Cladogram Construction

C. Major Parsimony-based cladistic programs