Homologous traits shared by organisms is the signature of evolutionary history.

Organisms possess traits because they inherited them

Organisms with similar traits are usually similar because they inherited the traits from shared ancestors

Generally, the more closely related organisms are, the more similar they are and the more homologies they share.

We can use this logic to help reconstruct evolutionary histories.
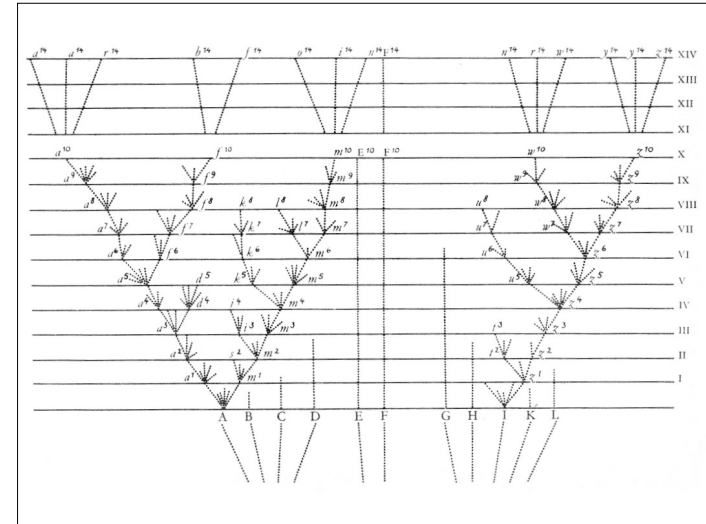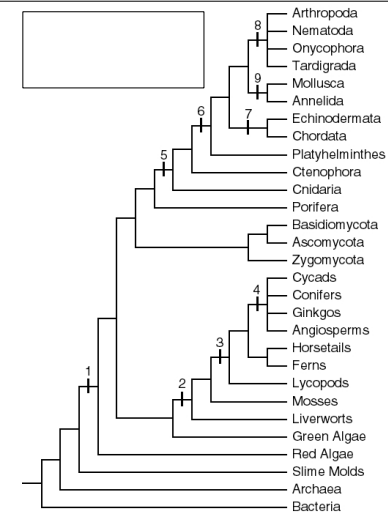

A branching tree-like diagram that estimates evolutionary relationships between taxa is called a *phylogeny* or a *phylogenetic tree*. You may also hear it called a "cladogram", or simply a "tree".

Trees have branches, tips, and nodes.
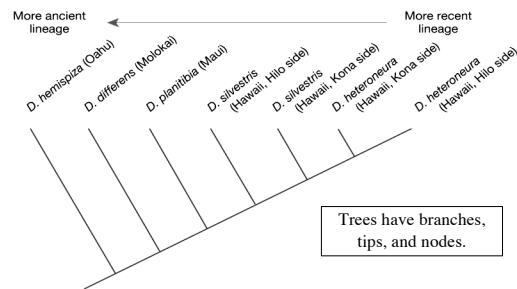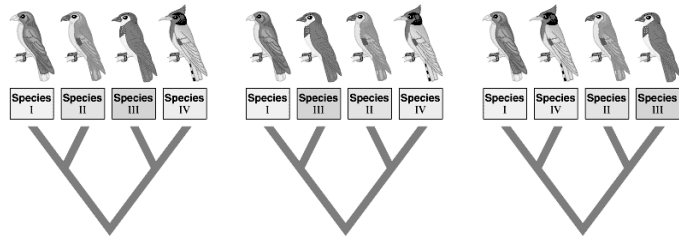
A *monophyletic group*, also called a *clade*, is defined as an ancestor and all of its descendants. True taxa (i.e. real ones) are clades.

Any phylogenetic diagram can be a true statement of history, regardless of what lineages are omitted from it, or what level of detail is shown.

Note on this particular tree, individual species are not shown, though they are implied.

1

Different phylogenies make different statements about evolutionary history

Species I Species II Species III Species IV

Species I Species III Species II Species IV

Species I Species IV Species II Species III
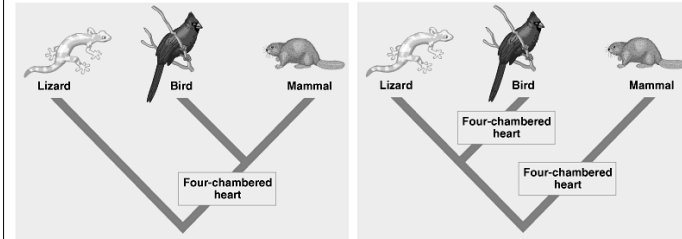
Copyright © Pearson Education, Inc., publishing as Benjamin Cummings.

Hypothesis I   Hypothesis II   Hypothesis III

Different phylogenies can also make different statements about the evolution of traits.

Lizard   Bird   Mammal

Four-chambered heart

(a) Mammal–bird clade

Lizard   Bird   Mammal

Four-chambered heart

Four-chambered heart

(b) Lizard–bird clade

Copyright © Pearson Education, Inc., publishing as Benjamin Cummings.

Box 21.2 Figure 1a

Analogy: When similarities result from convergent evolution

Common dolphin   1m

Ichthyosaur   1m

Synapsids Monotremes Marsupials Rodents Primates Whales and dolphins Elephants Pterosaurs Dinosaurs Birds Ichthyosaurs Lizards

The members of lineages between the dolphins and ichthyosaurs do not have adaptions like:
- streamlined bodies
- long jaws filled with teeth
- fins and flippers
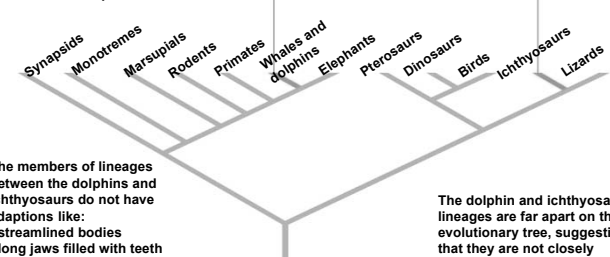
The dolphin and ichthyosaur lineages are far apart on the evolutionary tree, suggesting that they are not closely related

---

One of the most powerful tools in modern biology is phylogenetic systematics: the science of reconstructing any portion of the true tree of life, usuallly using data from extant species

A phylogeny can estimate history and can provide an objective framework for making meaningful comparisons between organisms. Why do we use primates and rodents in biomedical research? Because of phylogenetic relatedness.

Nowadays, most data used to estimate phylogeny come from genetic sequences: DNA, RNA, and/or amino acids (proteins), but many still do phylogenetic systematics using phenotypic traits.

Slide 1

# An example of a tree estimated using phenotypic data

TAXA

CHARACTERS	Lancelet (outgroup)	Lamprey	Tuna	Salamander	Turtle	Leopard
Hair	0	0	0	0	0	1
Amniotic (shelled) egg	0	0	0	0	1	1
Four walking legs	0	0	0	1	1	1
Jaws	0	0	1	1	1	1
Vertebral column (backbone)	0	1	1	1	1	1

(a) Character table

(b) Cladogram

Lancelet (outgroup)   Lamprey   Tuna   Salamander   Turtle   Leopard

Hair

Amniotic egg

Four walking legs

Jaws

Vertebral column

Copyright © Pearson Education, Inc., publishing as Benjamin Cummings.

Slide 2

What are doing when we "do" phylogenetic systematics?

- We *reconstruct* or *estimate* a phylogenetic tree from data (we don't "make" it)

- We do this by selecting a number of *taxa* of interest (species, genes, individuals in a population, viruses, etc., depending on our question)

- We then collect and *align* lot of data on *characters* (traits) that we believe to be *homologous* between our taxa (structures, behaviors, DNA or protein sequences), and create a *character matrix*

- We select an *optimality criterion* based on our beliefs on how evolution has occurred among our taxa and characters

- We analyze many possible *topologies* (possible tree arrangements), searching for the one that best fits our pre-determined optimality criterion

- We then assess the *robustness* (strength) of our best result, using statistics

- If we believe there are errors in our *phylogenetic estimate* (tree), we reassess our taxon sample, our character matrix, and our optimality criterion, and do the analyses again to see if we get a new estimate that is more consistent with other data (such as the fossil record, geography, etc.)

Slide 3

How are these organisms related?

	Lion	Turtle	Shark	Rabbit
Hair	1	0	0	1
4 legs	1	1	0	1

Slide 4
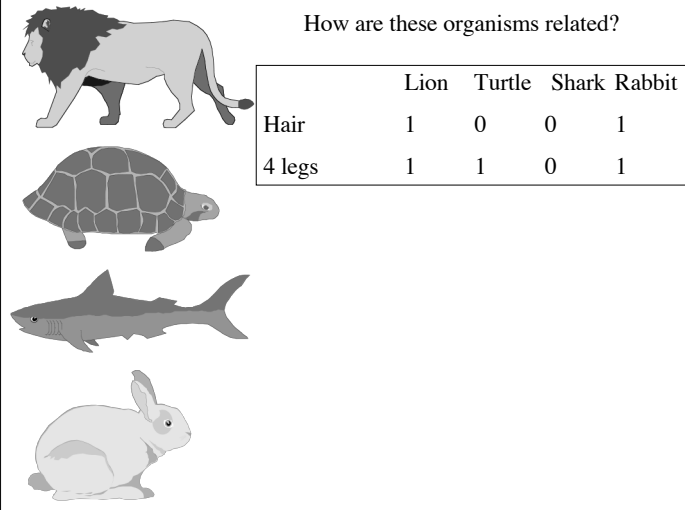
A more realistic data matrix

Character Matrix "Bembidion matrix"

Graphics Text Parameters Modules Citations

Taxon \ Character	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1 balli	1	1	1	0	–	0	0	1&2	1	0	0	0	1	1	1	0	0	0	
2 foveum	1	1	1	0	–	0	0	1&2	1	0	0&1	1	1	1	1	0	0	0	
3 argenteolum	1	3	2&3	0&1	–	0	0	2	0	0	0	2	1	1	0	0	0	0	
4 alaskense	1	3	2&3	0	–	0&1	0	2	0	0	1	0	1	1	0	1	0	0	
5 *semenovi	1	2	2	0	–	0	0	2	0	0	0	3	1	1	0	0	0	0	
6 stenoderum	1	3	3	0&2	0	0	0	1&2	0	0	0	2	1	1	0	0	0	0	
7 carinula	1	3	0&1&2	0	–	0	0	2	1	0	0&1	1	1	1	0	0	0	0	
8 velox	1	3	0&1&2	1	–	0	0	2	1	0	0	1	1	1	0	0	0	0	
9 lapponicum	1	3	0&1&2&3	0&2	0	0	0	2	1	0	0	2	1	1	0	0	0	0	
10 punctatostriatum	1	3	0&1&2	1	–	0	0	0	1	0	0	2	1	1	0	0	0	0	
11 hesperium	1	2&3	3	0&2	0	0&1	0&1&2	1	0	0	0	1	1	1	0	0	0		
12 lorquinii	1	3	3	1	–	0&1	0	0&1	0	0	0	1	1	1	0	0	0	0	
13 zephyrum	1	3	3	2&3	0	0&1	0	2	1	0	1	0	1	1	0	0	0	0	
14 l. levettei	1	3	3	3	0	1	0	2	1	0	0&1	0	1	1	0	0	0	0	

Data matrix from D.R. Maddison, 1993. Systematics of the Holarctic beetle subgenus Bracteon and related Bembidion (Coleoptera: Carabidae). Bulletin of the Museum of Comparative Zoology, 53(3):143-299.

Annotation above refers to Bembidion matrix

How are these organisms related?

	Lion	Turtle	Shark	Rabbit
Hair	1	0	0	1
4 legs	1	1	0	1

We need an optimality criterion to choose a best estimate (tree)

Parsimony: begins with the assumption that the simplest hypothesis that explains the data is probably the correct one. Assume that change is rare, and select the tree that requires the least amount of change along its branches to produce the data.

(In this example, we use simple morphological characters, but this method is also used with molecular sequence data.)

	Lion	Turtle	Shark	Rabbit
Hair	1	0	0	1
4 legs	1	1	0	1

At least 2 gains (or losses) of hair
At least 1 gain (or loss) of 4 legs
Total minimum changes: 3

	Lion	Turtle	Shark	Rabbit
Hair	1	0	0	1
4 legs	1	1	0	1

At least 2 gains (or losses) of hair
At least 1 gain (or loss) of 4 legs
Total minimum changes: 3

At least 2 gains (or losses) of hair
At least 1 gain (or loss) of 4 legs
Total minimum changes: 3

At least 1 gain (or loss) of hair
At least 1 gain (or loss) of 4 legs
Total minimum changes: 2

	Lion	Turtle	Shark	Rabbit
Hair	1	0	0	1
4 legs	1	1	0	1

This is our **most parsimonious** branching hypothesis, as it requires the least amount of change to explain the observed data.

However, in order to relate this diagram to history, we must polarize the direction of change through time. **We must root the tree.**

One way to root is to choose an outgroup: a taxon which you believe shares the most ancient common ancestor with all the others. It then becomes the sister taxon of the clade formed by all of the other taxa.

NOTE THAT ALL OF THESE ROOTED TREES ARE IDENTICAL, IN THAT THEY ALL SAY THE SAME THING ABOUT HISTORY.

What history can we infer from this phylogenetic tree?

Rabbits and Lions share a more recent common ancestor with one another than with turtles or sharks. Mammals.

Rabbits, Lions, and Turtles share a more recent common ancestor with each other than with sharks. Tetrapods.

Sharks and turtles are NOT "older" or "ancestors" of mammals.

What else can we infer from this phylogenetic tree?

The hair of rabbits is probably homologous with the hair of lions. Hair is a synapomorphy of mammals.

The 4 limbs of turtles are probably homologous with those of mammals. Having 4 limbs is a synapomorphy of tetrapods

With more taxa, we could infer that sharp carnivorous teeth in lions are probably NOT homologous with those of sharks.

5
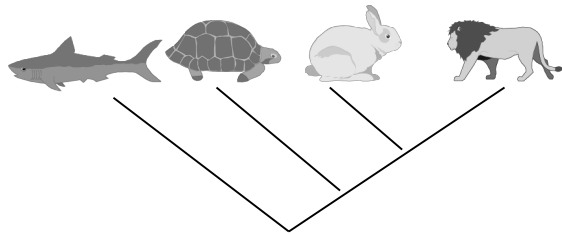
What else can we infer from this phylogenetic tree?

Patterns of ancestry. Mammals form a *monophyletic group*, also called a *clade* (an ancestor and all of its descendants)

Tetrapods in this case (include turtles) are also a clade

---

Biologists have learned a lot from phylogenetic trees.

For example: "reptiles" are not a monophyletic group (i.e., not a clade), because the common ancestor to all the animals we call "reptiles" was also the common ancestor to birds and mammals.
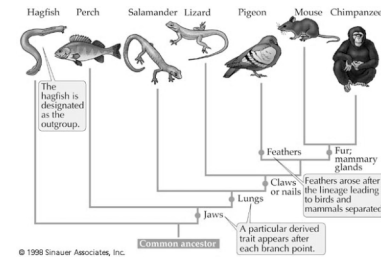
"Reptiles", therefore, is not a *natural group*, unless it includes birds and mammals. It is a *paraphyletic group*.

Modern phylogenetic systematics is shaking some phylogenetic trees
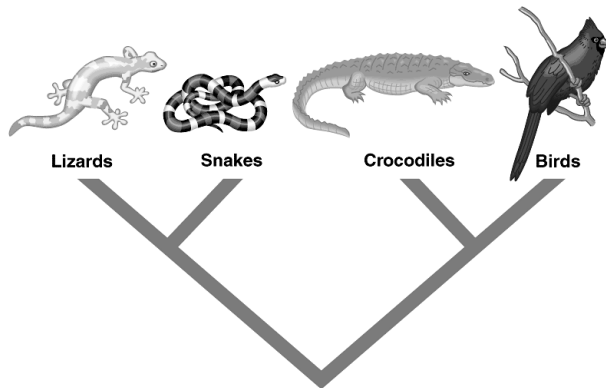
Lizards   Snakes   Crocodiles   Birds

Considerations:

Polytomies (hard polytomy; soft polytomy)

Multiple equally optimal solutions

Tree Consensus

Two kinds of parsimony

- When mapping characters and counting changes

- When selecting an optimal tree

Other optimality criteria used to choose a best estimate (tree)

Distance: Based on the assumption that closely related organisms are going to be more similar. Construct a distance matrix, and select the tree that *minimizes the differences* between taxa.
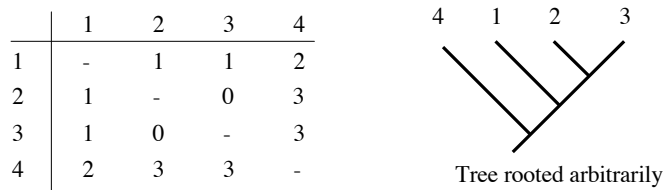
1 ATGCGTCTC
2 ATGAGTCTC
3 ATGAGTCTC
4 TTGCGTCCC

	1	2	3	4
1	-	1	1	2
2	1	-	0	3
3	1	0	-	3
4	2	3	3	-

Tree rooted arbitrarily

Other optimality criteria used to choose a best estimate (tree)

Distance: Based on the assumption that closely related organisms are going to be more similar. Construct a distance matrix, and select the tree that *minimizes the differences* (distances) between taxa.

1 ATGCGTCTC
2 ATGAGTCTC
3 ATGAGTCTC
4 TTGCGTCCC

	1	2	3	4
1	-	1	1	2
2	1	-	0	3
3	1	0	-	3
4	2	3	3	-

⇐ C->A
⇐ T->A
⇐ T->C

Tree rooted arbitrarily
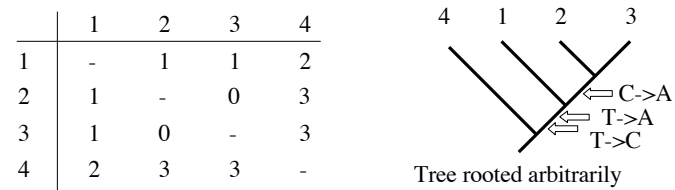
Other optimality criteria used to choose a best estimate (tree)

Maximum Likelihood (for DNA sequence data): Start with a model of nucleotide evolution, then begin examining possible trees. Ask: what is the likelihood that a given tree would have produced the actual observed sequence data under the model of evolution? The most optimal tree is the one with the highest likelihood score.

1 ATGGGTCTC
2 ATGAGTCTC
3 ATGAGTCTC
4 ATGCGTCTC

	A	T	G	C
A	-	.17	.20	.12
T	.17	-	.14	.20
G	.20	.14	-	.17
C	.12	.20	.17	-

Tree rooted arbitrarily