

# Classification and Phylogeny



The diversity of life is great. To communicate about it, there must be a scheme for organization.

There are many species that would be difficult to organize without a scheme based on consistent criteria.



**Classification** - an organized scheme for grouping organisms - a tool for communication -

**Taxonomy** - the practice of naming and classifying organisms.

Linnaeus developed the classification scheme we use today

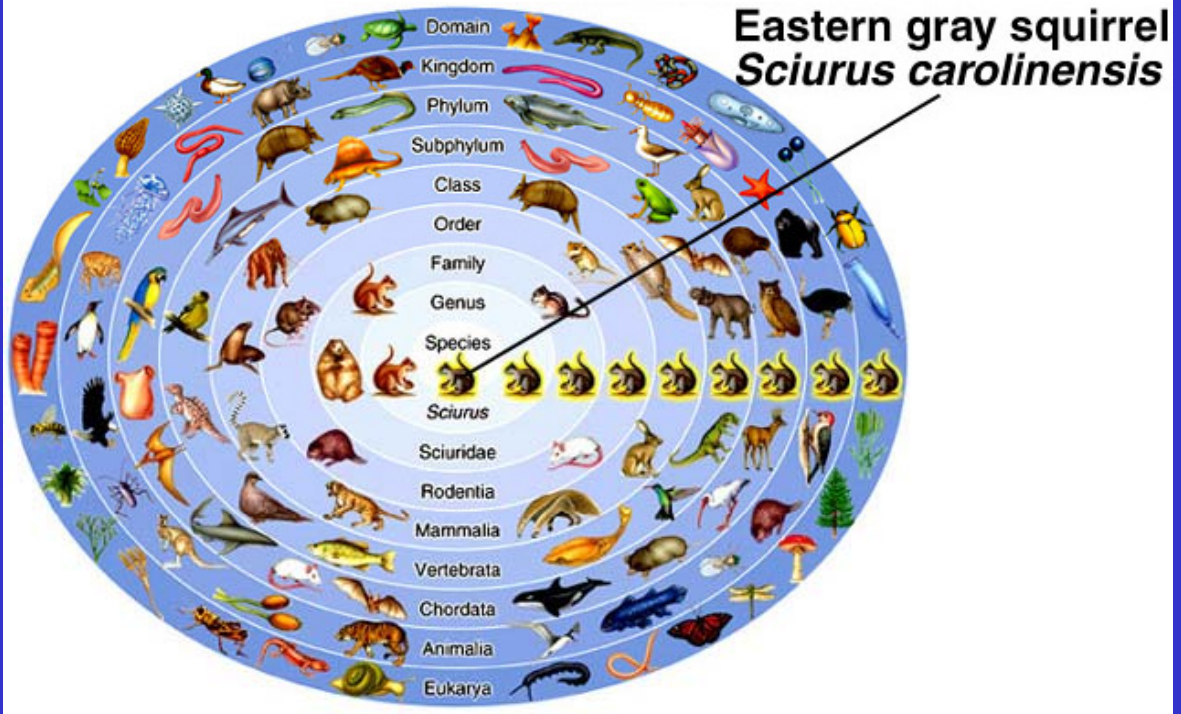
Species have **binomial** names (*Genus species*) and are arranged in a hierarchical classification - Linnaeus saw a genus as a general kind of organism and a species as a more specific kind of organism. Linnaeus grouped organisms by their degree of similarity (as he saw it). “God creates, Linnaeus organizes”

**Hierarchical** - a series of successive and inclusive rankings

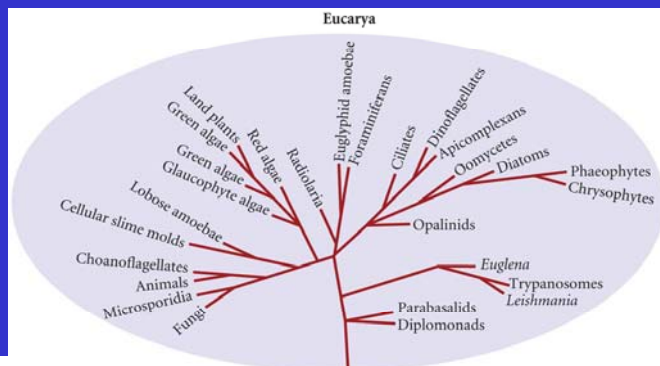
The major categories used in our current hierarchical classification are **Domain, Kingdom, Phylum, Class, Order, Family, Genus, Species** - many intermediate subcategories are necessary for organizing the diversity within groups with many species

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

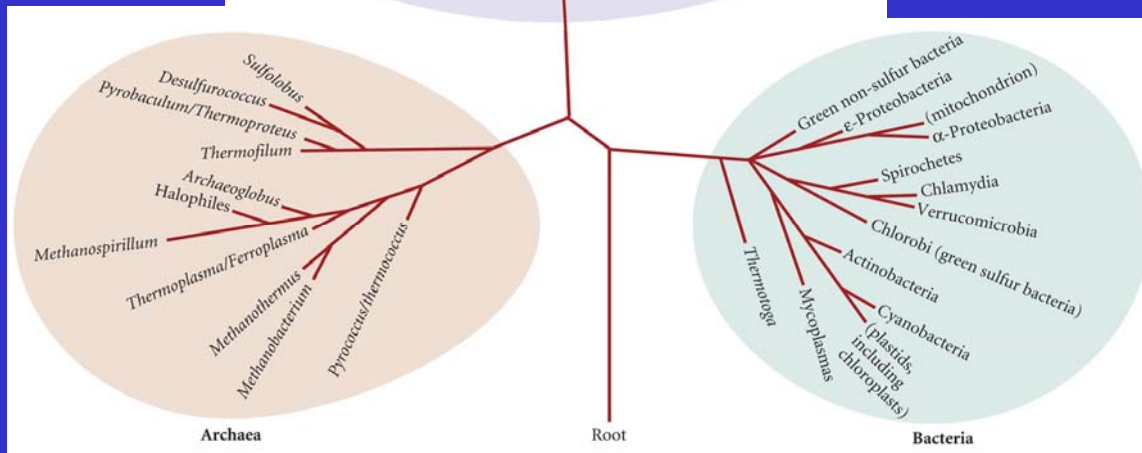
# Hierarchical System of Organism Classification



Evolution through anagenesis and cladogenesis has produced the diversity of living things today, which can be represented as a tree

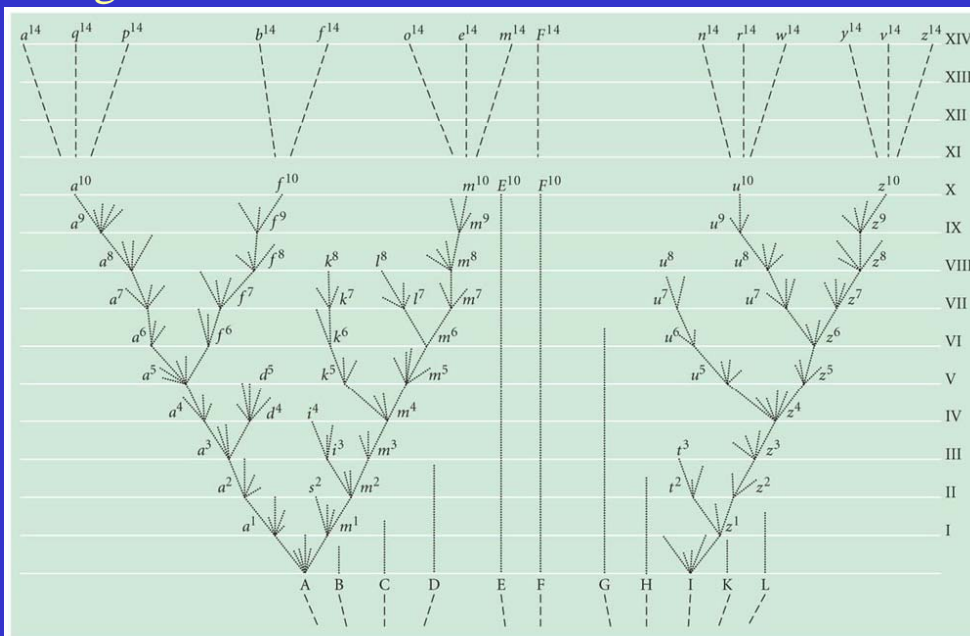


The currently used classification has 3 Domains - two prokaryotic and one eukaryotic.



A tree that represents the pattern of cladogenesis (but not necessarily the amount of anagenesis) is called a **cladogram**.

A **phylogeny** is a representation of the hypothetical evolutionary history of a group of organisms that may be a cladogram, but may also include information about the amount of evolution in each lineage.



Darwin pictured the process of evolution with variable rates of anagenesis and cladogenesis and many extinctions

The process of developing an estimate of the evolutionary history of a group of organisms is called phylogenetic inference.

The unit of data used in phylogenetic inference is the **character**. Characters are any physical characteristics of a group of organisms. They can be anatomical (morphological), behavioral, physiological, biochemical, molecular.

Characters may be useful for phylogenetic inference if they come in alternate **character states** - different shapes of the same bone, enzymes that differ in their amino acid sequence, genes that differ in their nucleotide sequence, etc.



The states of a character may represent evolutionary changes. For any two states of a character one may have originated earlier in the evolutionary history of the group and one may have originated later.

Character states that originated earlier in the evolutionary history of a group can be called **primitive** or **ancestral** or **plesiomorphic**.

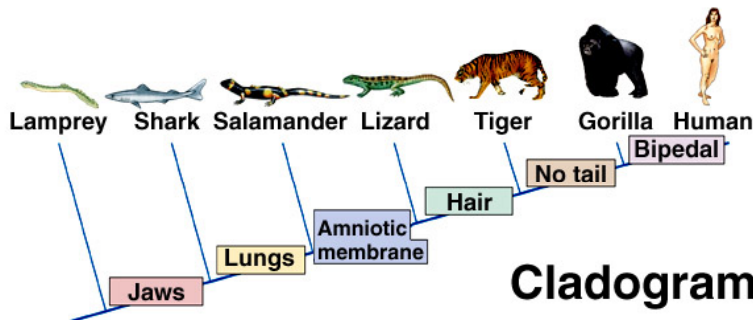
Character states that originated later are called **advanced** or **derived** or **apomorphic**.

Derived states of a character provide information about the degree of relatedness of two or more species. A derived state that is shared by two or more species is called a **synapomorphy**.

Synapomorphies allow species to be grouped and organized into a phylogeny.

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

Traits: Organism	Jaws	Lungs	Amniotic membrane	Hair	No tail	Bipedal
Lamprey	0	0	0	0	0	0
Shark	1	0	0	0	0	0
Salamander	1	1	0	0	0	0
Lizard	1	1	1	0	0	0
Tiger	1	1	1	1	0	0
Gorilla	1	1	1	1	1	0
Human	1	1	1	1	1	1



In this example

0 = ancestral state

1 = derived state

Lungs are a synapomorphy that allow all vertebrates with four legs to be grouped together.

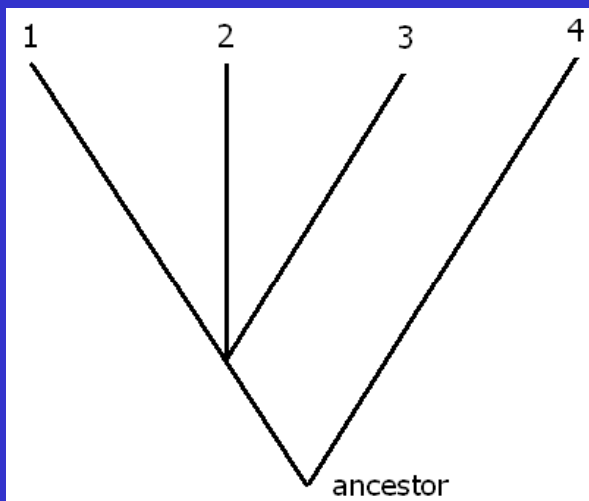
A synapomorphy can identify a **monophyletic group** - all the descendants of the single common ancestor.

The common ancestor possessed the derived state of the character.

**Phylogenetic inference** - using character states to develop an estimate of evolutionary relationship

A character state matrix

Character	a	b	c	d	e	f	g	h	i	j
Species										
1	0	0	1	1	1	0	0	0	0	0
2	0	0	1	0	0	1	1	1	1	0
3	0	0	1	0	0	1	1	0	0	1
4 (outgroup)	1	1	0	0	0	0	0	0	0	0
ancestor	0	0	0	0	0	0	0	0	0	0



The task is determine the relatedness of a group of species in the “ingroup.”

An outgroup - a group that is related to the ingroup, but clearly not a member of the ingroup, provides orientation.

The common ancestor of the whole group is hypothesized to have had the ancestral state of all characteristics.

# A simple example - a constant rate of evolution with no evolutionary reversals or multiple origins of the derived state

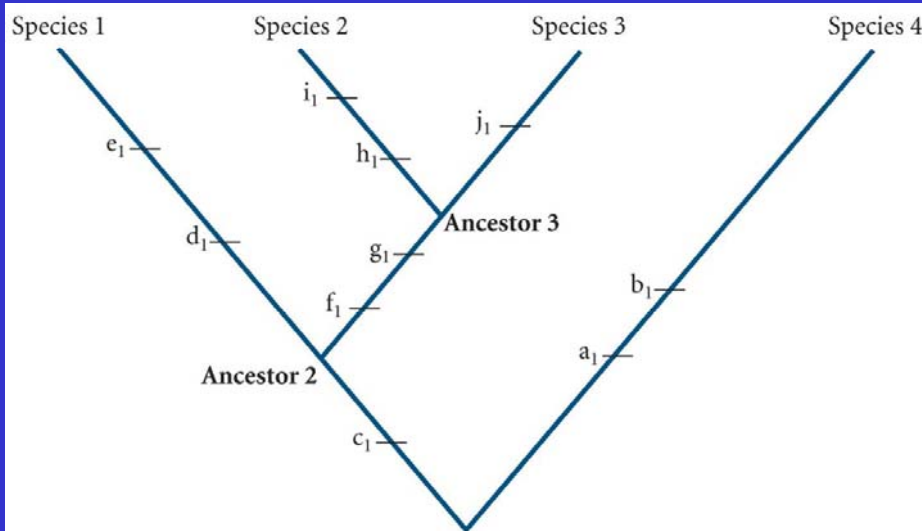
Character	a	b	c	d	e	f	g	h	i	j
Species										
1	0	0	1	1	1	0	0	0	0	0
2	0	0	1	0	0	1	1	1	1	0
3	0	0	1	0	0	1	1	0	0	1
4 (outgroup)	1	1	0	0	0	0	0	0	0	0
ancestor	0	0	0	0	0	0	0	0	0	0

Shared character states

	1	2	3	4
1	-	4	5	5
2		-	7	3
3			-	4

Shared derived character states

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



A derived state that has evolved only once is called **homologous** in all the species that share it.

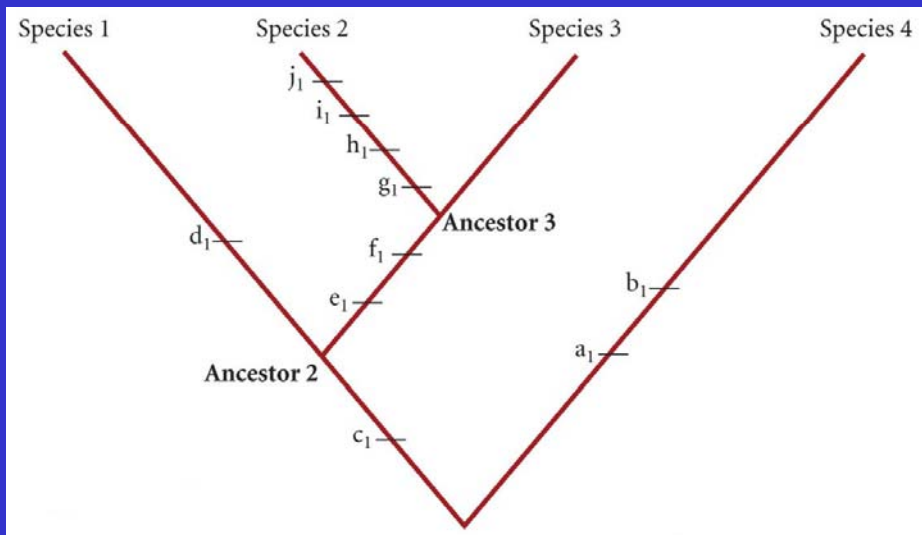
## An example with variable rates of evolution among branches.

Character	a	b	c	d	e	f	g	h	i	j
Species										
1	0	0	1	1	0	0	0	0	0	0
2	0	0	1	0	1	1	1	1	1	1
3	0	0	1	0	1	1	0	0	0	0
4 (outgroup)	1	1	0	0	0	0	0	0	0	0
ancestor	0	0	0	0	0	0	0	0	0	0

Shared character states

	1	2	3	4
1	-	3	7	6
2		-	6	1
3			-	5

Overall sp. 1  
& sp. 3 are  
most similar



Shared derived character states

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

Synapomorphies  
and not overall  
similarity reflect  
evolutionary  
history

An example with evolutionary reversals and multiple origins of the derived state of some characters.

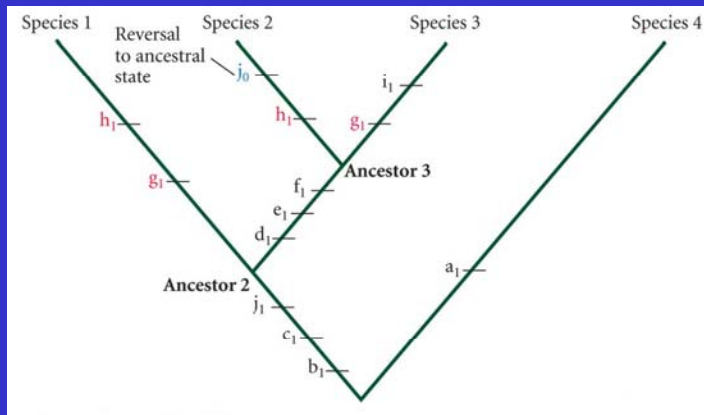
Character	a	b	c	d	e	f	g	h	i	j
Species										
1	0	1	1	0	0	0	1	1	0	1
2	0	1	1	1	1	1	0	1	0	0
3	0	1	1	1	1	1	1	0	1	1
4 (outgroup)	1	0	0	0	0	0	0	0	0	0
ancestor	0	0	0	0	0	0	0	0	0	0

Shared character states

	1	2	3	4
1	-	5	5	4
2		-	6	3
3			-	1

Shared derived character states

	1	2	3	4
1	-	3	4	0
2		-	5	0
3			-	0



Character states that arise more than once are called **homoplasious**.

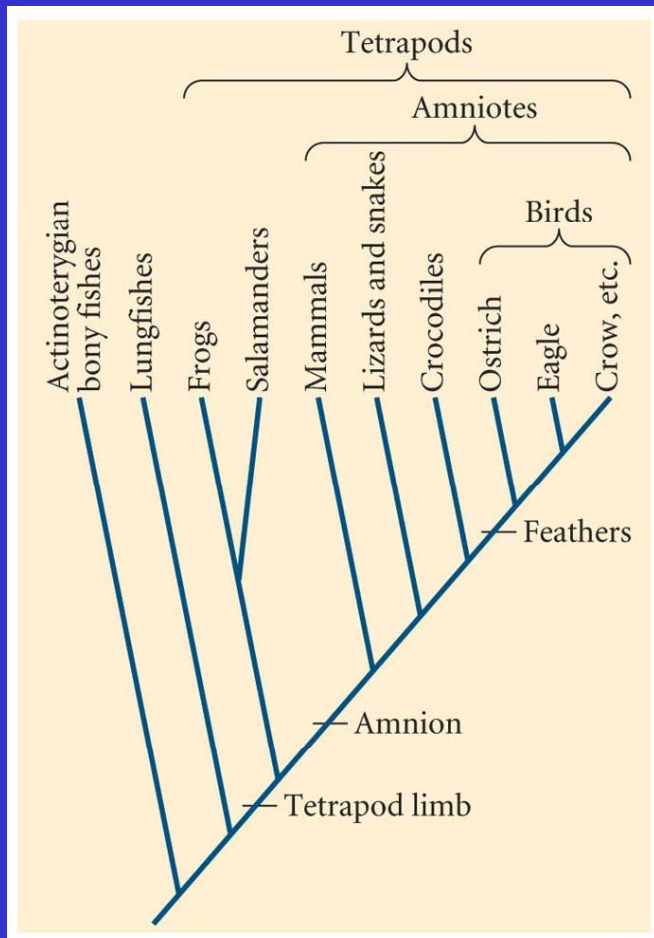
The derived state of characters g and h evolved twice - they are examples of **convergent evolution**.

The ancestral state of character j is present in species 2. This is an **evolutionary reversal**.

How can the ancestral and derived states of a character be identified?

Outgroup comparison is the most commonly used method. The state of a character found in species outside the group being analyzed is the ancestral state.

For example - among tetrapods (vertebrates with four limbs – amphibians, reptiles, birds, and mammals) there are those with amniotic development (the embryo forms and develops within an amniotic sac) and those without. The tetrapods themselves provide no information about the direction of evolution because the amnion could have been gained or lost during evolution. The outgroup, all other vertebrates (the fishes), does not have amniotic development. Therefore, it is most likely that amniotic development is the derived state in tetrapods. The derived state (having amniotic development) can then be used as a potential synapomorphy and used in constructing a phylogeny.

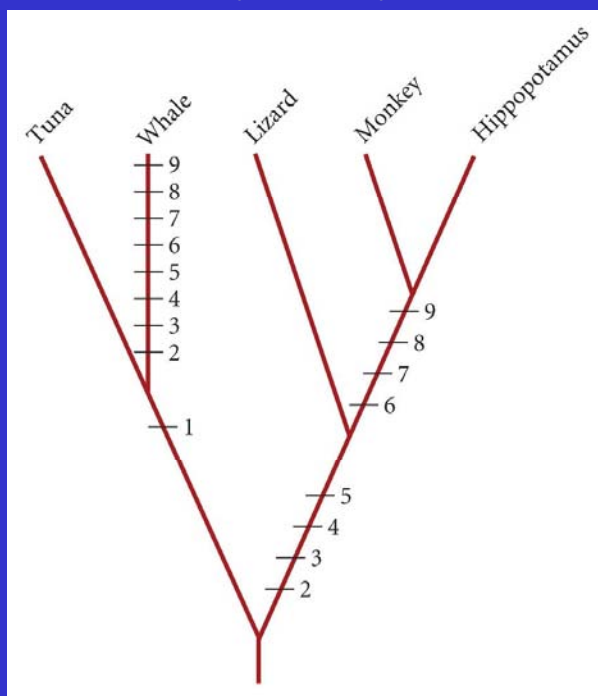


Synapomorphies are the basis for modern classifications. Organisms that share the derived state of a characteristic that evolved only once can be classified as belonging to the same group.

Ancestral states do not provide information about evolutionary relatedness and are usually not used for classifying organisms.

Convergent evolution and evolutionary reversals provide misleading information about the relationships of species.

How can one decide among competing hypothesis of evolutionary history?

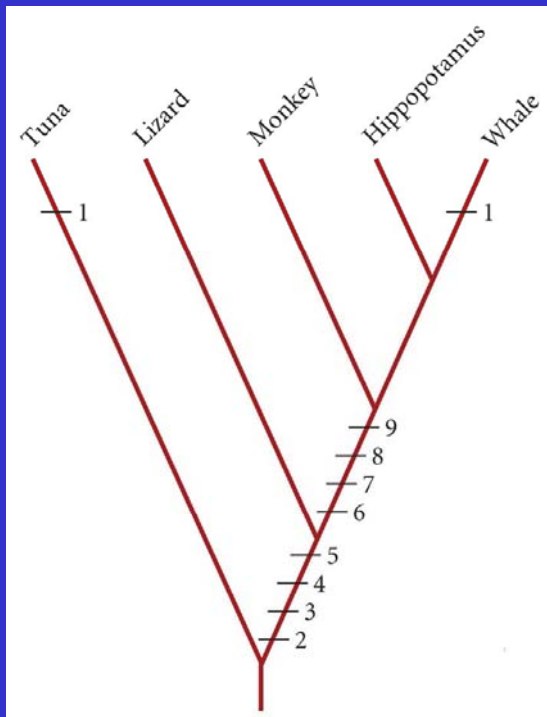


Derived character states:

- 1: the presence of dorsal fin
- 2: the presence of a pectoral girdle
- 3: the presence of limbs
- 4: the presence of lungs
- 5: the presence of cervical vertebrae
- 6: the presence of a single aortic arch
- 7: the presence of a dentary jawbone
- 8: the presence of mammary glands
- 9: the presence of a 4-chambered heart

Simpler hypotheses are preferred - the principle of parsimony.

Here dorsal fins evolved twice but all other derived character states evolved once.



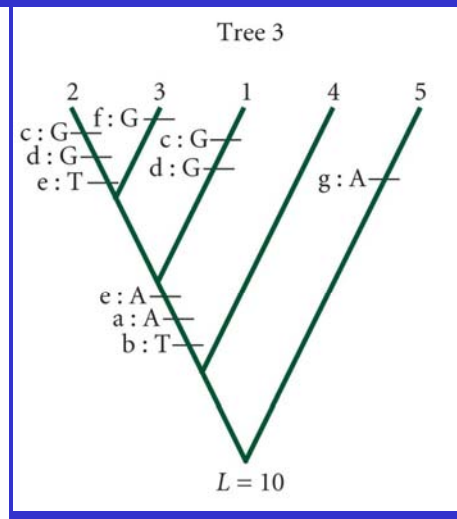
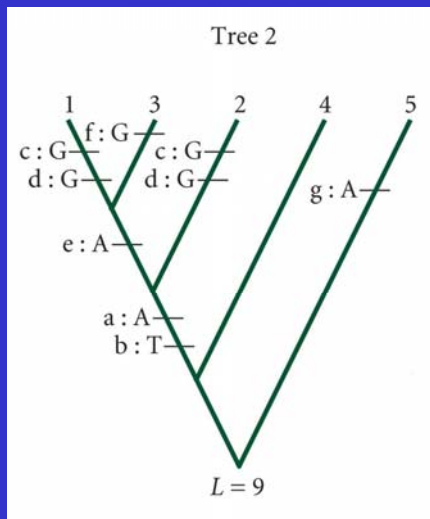
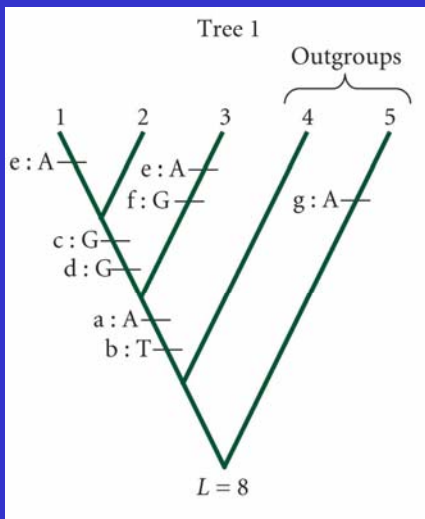
Derived character states:

- 1: the presence of dorsal fin
- 2: the presence of a pectoral girdle
- 3: the presence of limbs
- 4: the presence of lungs
- 5: the presence of cervical vertebrae
- 6: the presence of a single aortic arch
- 7: the presence of a dentary jawbone
- 8: the presence of mammary glands
- 9: the presence of a 4-chambered heart

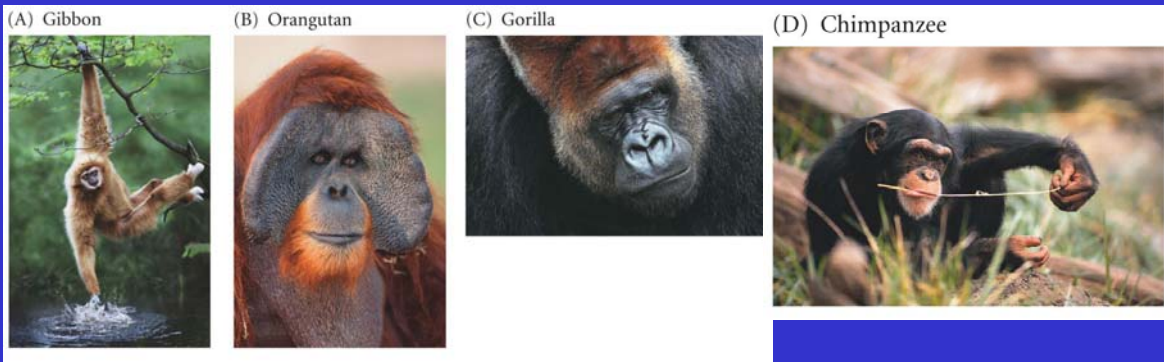
**The method of maximum parsimony** - the estimate of evolutionary relationship that requires the fewest character state changes is the simplest hypothesis.

Species	Character						
	a	b	c	d	e	f	g
1	A	T	G	G	A	C	T
2	A	T	G	G	T	C	T
3	A	T	T	C	A	G	T
4	C	G	T	C	T	C	T
5	C	G	T	C	T	C	A

Among species 1, 2, and 3 - three different hypotheses are possible. Each can be evaluated for the number of character state changes.

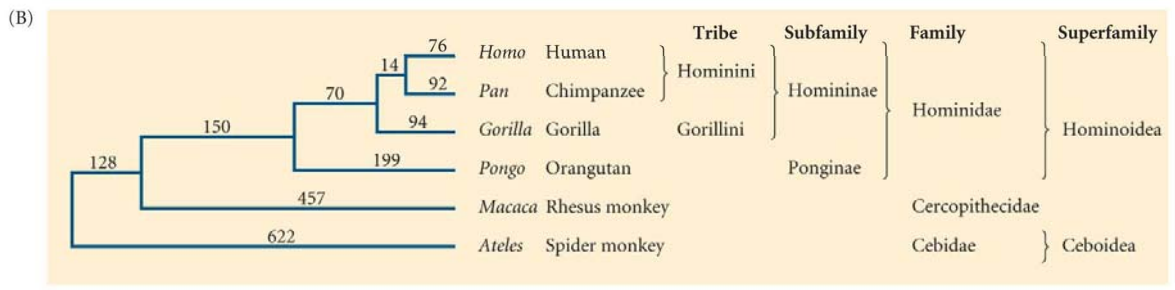


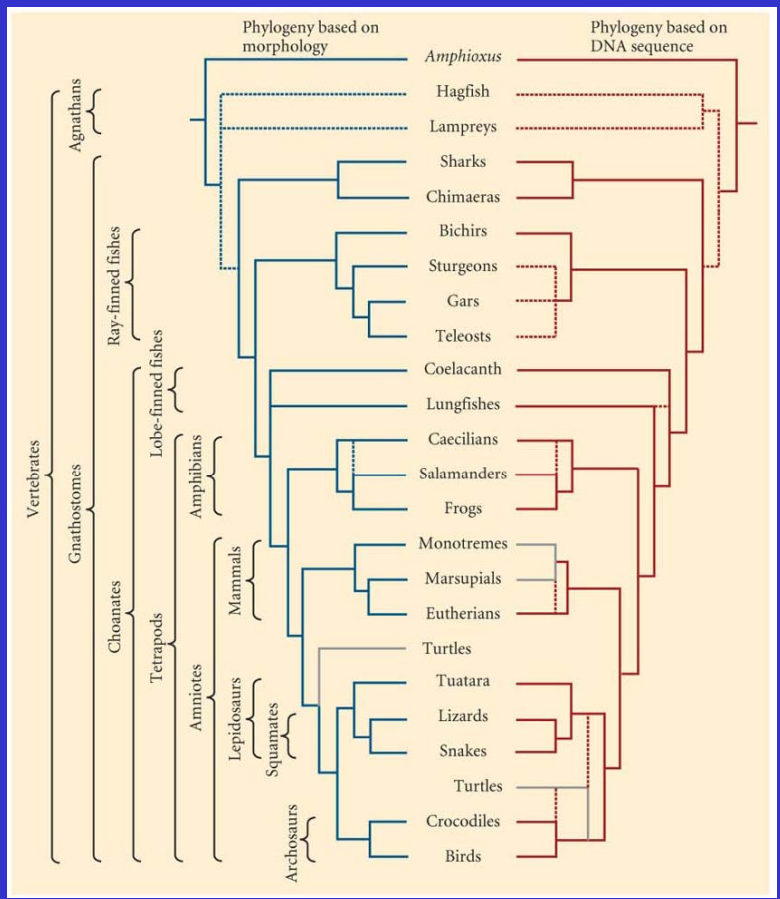
# Read: An example of phylogenetic analysis - p 29-31



(A) DNA sequence alignment for Gibbon, Orangutan, Gorilla, Chimpanzee, and other primates. The sequence is: *Homo* \* \* \* \* C A C A A T A . . . T G A G C . . . G A A G A G A T G . . . G T G A A A G . . . A \* \* \* \* \*  
*Pan* \* \* \* \* A A A G C . . . G C G G T G A A A G . . . A \* \* \* \* \*  
*Gorilla* T A A T A A T T T T G G T T G A A A T A T A  
*Pongo* T A A T A A T T T T C G T T C A A A T A T A  
*Macaca* T A A T C G T T T T C A T T G A A T A T A  
*Ateles* T A A T A C T T T T A T T G A C T A T A

3903                      3913                      5361                      5365                      6367                      6375                      8224                      8230                      8468                      8474

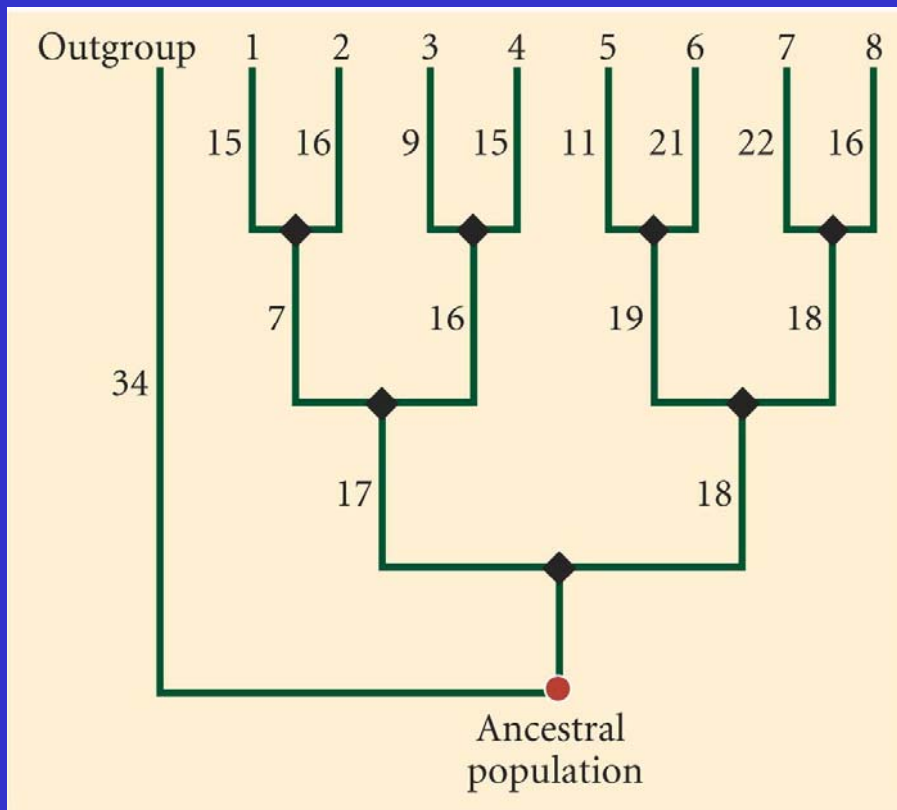




Hypotheses of evolutionary relationship can be modified or further validated by gathering additional data.

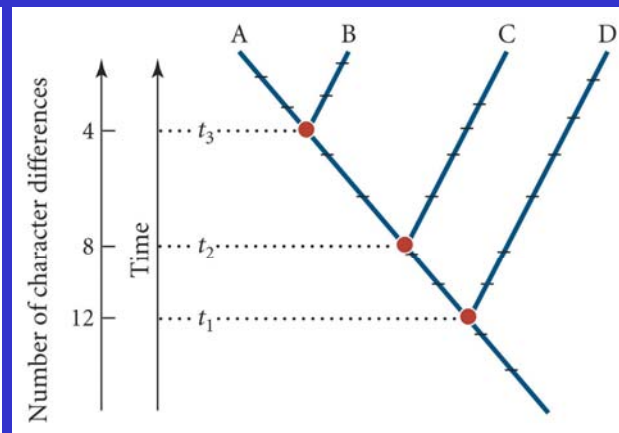
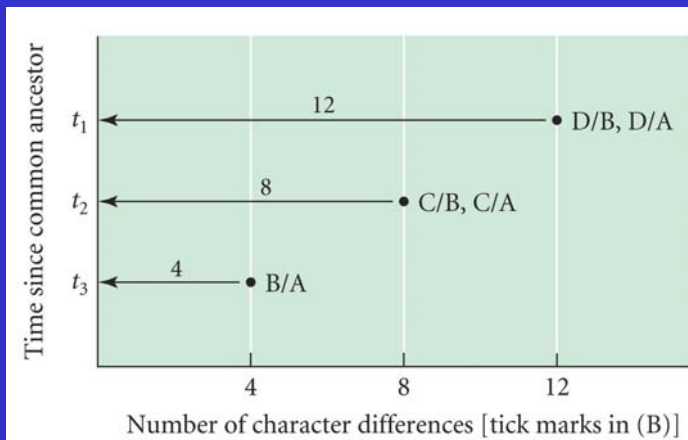
The robustness of a phylogenetic hypothesis can be evaluated by comparing hypotheses created with independent data sets.

Read about the Hillis et al. (1992) experimental creation of evolution in phage T7 and phylogenetic reconstruction.



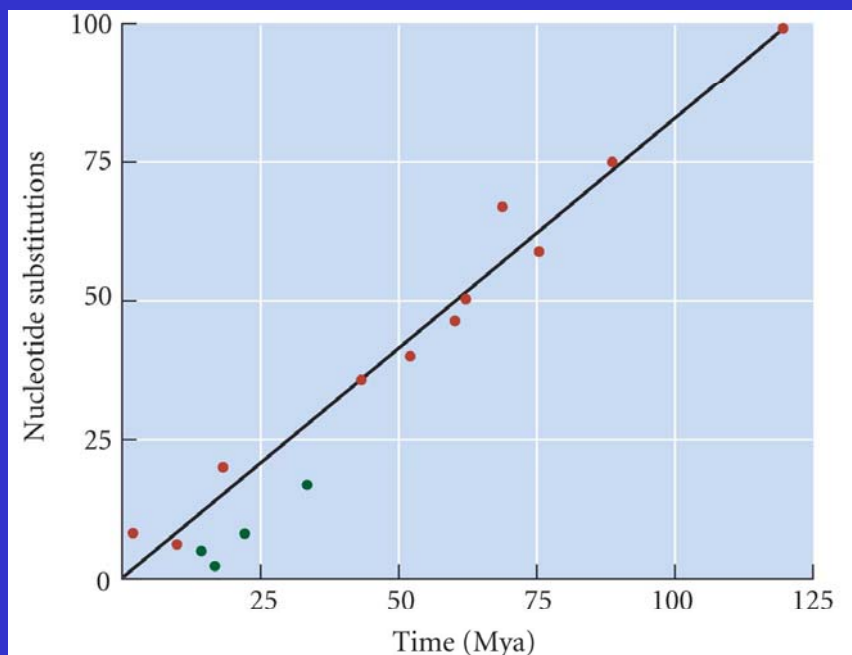
**The molecular clock** - early in the history of the study of molecular evolution it was hypothesized that molecular evolution may be different than morphological evolution. Molecular evolution may proceed at a constant rate.

If the molecular clock hypothesis is true, the time since divergence of two species could be estimated by the number of DNA base pair differences between species.



The rate of molecular evolution can be estimated by referring to the fossil record. The fossil record gives an approximate date of divergence of genetic lineages. The rate of molecular change can then be estimated.

Not all groups appear to evolve at the same rate.

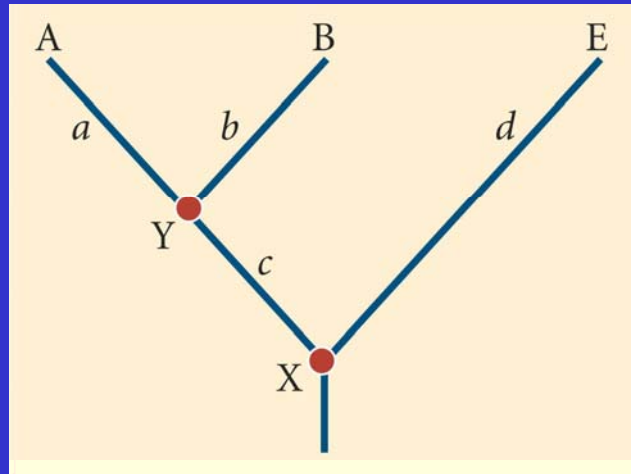


### The relative rate test:

If there is a constant rate of molecular evolution then the amount of molecular change (genetic distance) should be equal for any pair of descendants of a common ancestor.

A and B should each be equally different from E:

$$a + c + d = b + c + d$$

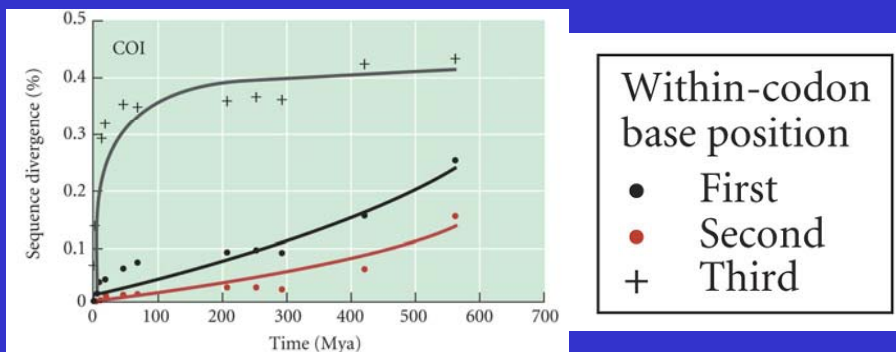


Studies of closely related species suggest that the rate of molecular evolution is often nearly constant.

Studies of more distantly related species suggest that the rate of molecular evolution can vary greatly among groups. For example, within primates and within rodents the rates of evolution have been relatively constant. However, primates have evolved relatively slowly compared to rodents.

## Difficulties in phylogenetic analysis

1. Deciding what is a character and what are its states is difficult (example: teeth in mammals)
2. Homoplasy is common. This results in many equally supported phylogenetic estimates
3. The process of evolution can erase evidence of prior evolution. If a character goes through several state changes, it can be difficult to associate it with the original ancestral state. (example: multiple base pair changes at a single site (A → C → T → G))



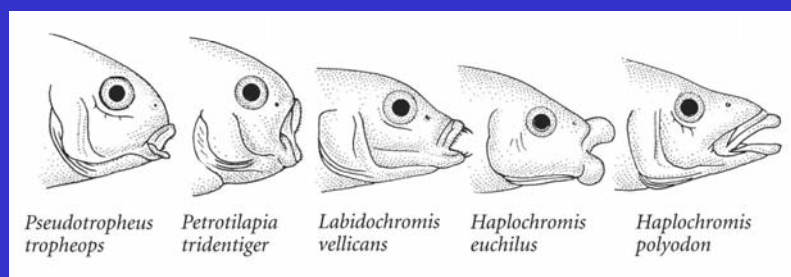
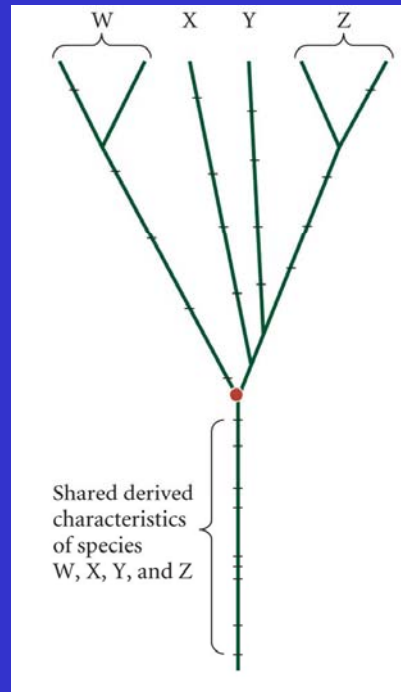
Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

**Table 15.1 The Genetic Code**

First Letter	Second Letter								Third Letter
	U	C	A	G					
U	UUU	Phenylalanine	UCU	Serine	UAU	Tyrosine	UGU	Cysteine	U
	UUC		UCC		UAC		UGC		C
	UUA	Leucine	UCA		UAA	Stop	UGA	Stop	A
	UUG		UCG		UAG	Stop	UGG	Tryptophan	G
C	CUU	Leucine	CCU	Proline	CAU	Histidine	CGU	Arginine	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	Glutamine	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	Isoleucine	ACU	Threonine	AAU	Asparagine	AGU	Serine	U
	AUC		ACC		AAC		AGC		C
	AUA	ACA	AAA		Lysine	AGA	Arginine	A	
	AUG	ACG	AAG			AGG		G	
G	GUU	Valine	GCU	Alanine	GAU	Aspartate	GGU	Glycine	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	Glutamate	GGA		A
	GUG		GCG		GAG		GGG		G

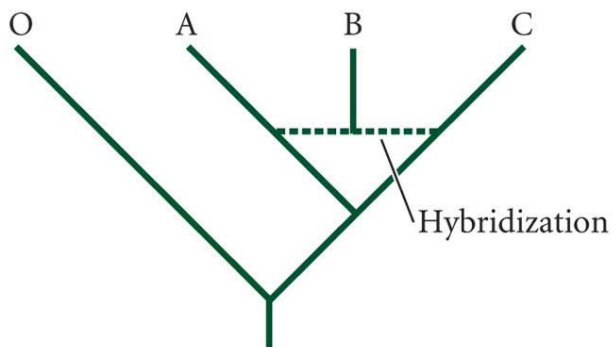
4. Some lineages evolve so rapidly that no distinctive synapomorphies evolve to link them.

Adaptive radiations - sudden bursts of evolution, usually in response to introduction of a species into a new environment, result in such rapid evolution that relationships are difficult to determine.

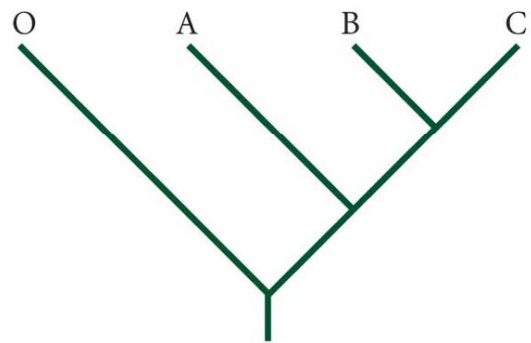


Hybridization can result in the formation of new species. This can result in reticulate, rather than bifurcating, evolution.

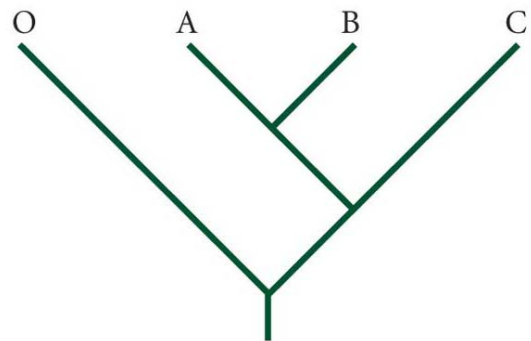
(C) Reticulate phylogeny



(A) Phylogeny based on sequence of gene 1



(B) Phylogeny based on sequence of gene 2



Horizontal gene transfer can provide false information about relationships among species.

