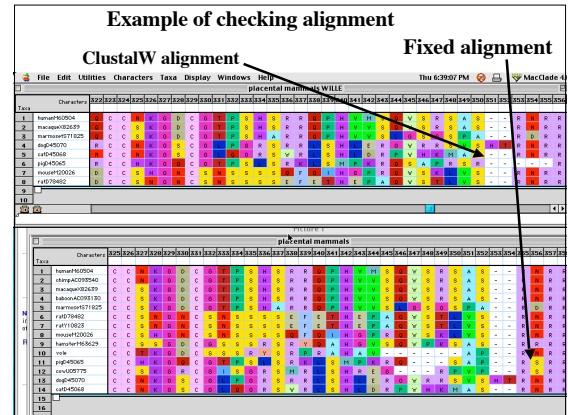
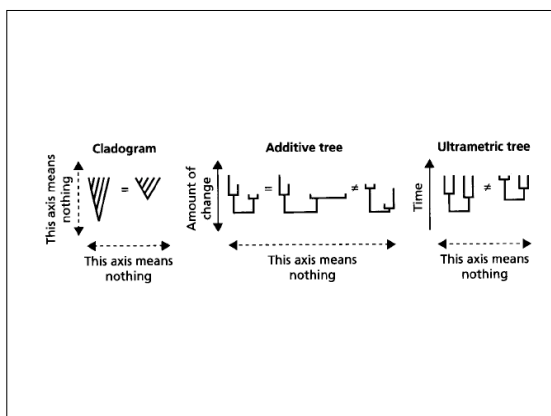
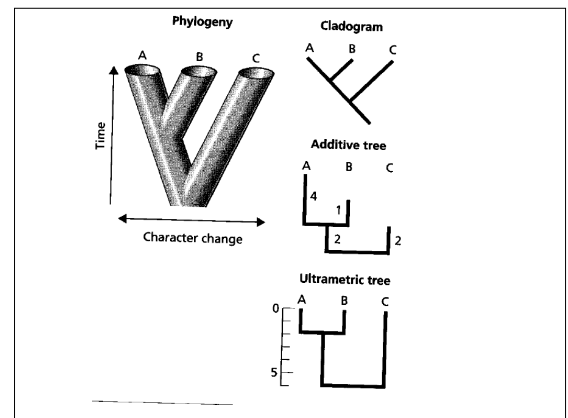
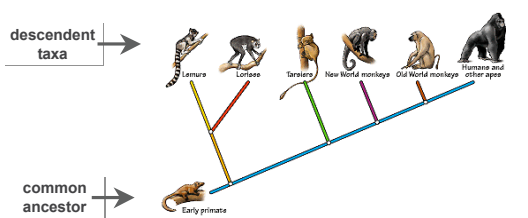


PHYLOGENETIC TREES



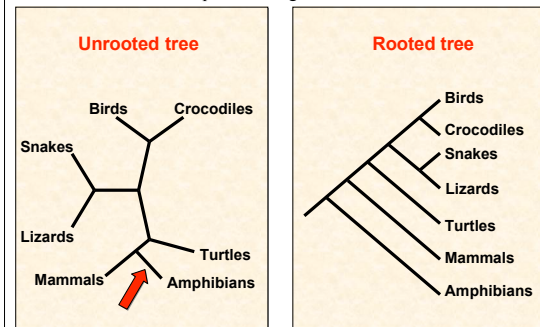
What is a phylogeny?

Phylogeny is the history of a group of taxa described as an evolutionary tree with a common ancestor at the base and descendent taxa as branch tips

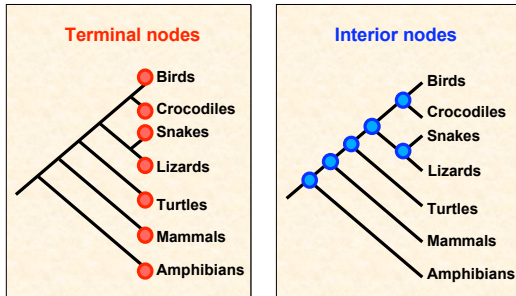


Terms to describe phylogenetic trees

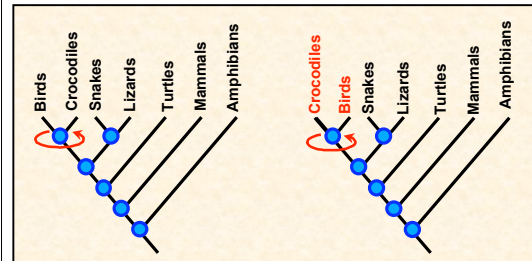
Remember ClustalW midpoint rooting.



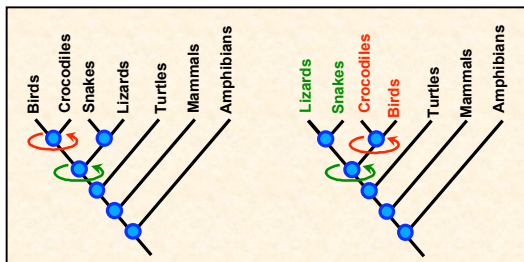
Terms to describe phylogenetic trees



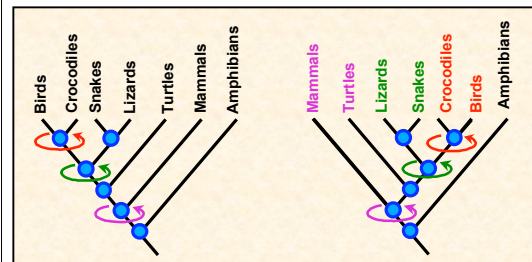
Taxa can rotate around interior nodes without changing relationships



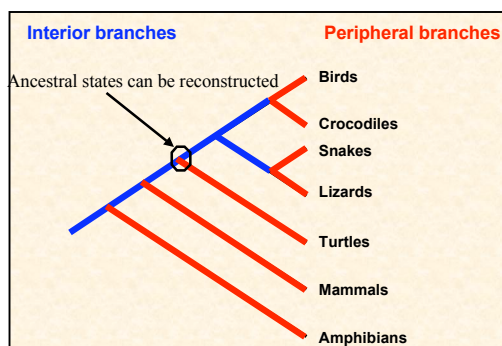
Taxa can rotate around interior nodes without changing relationships



Taxa can rotate around interior nodes without changing relationships



Terms to describe phylogenetic trees



It is incorrect to say "high or low homology" or "percent homology."

Homology is a **hypothesis** that similarity of a **trait** in two or more species indicates descent from a common ancestor.

a trait could be a morphological, behavioral, or molecular character

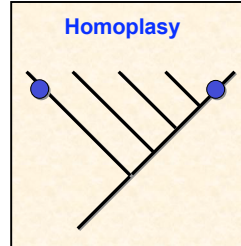
Homologous features are derived from common ancestors.

It seems reasonable to assume that things that look similar are related.

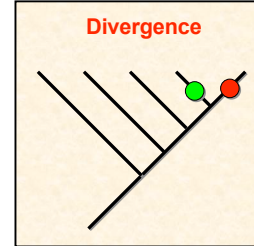
But, appearances can be misleading.

Two common reasons for deceptive appearances

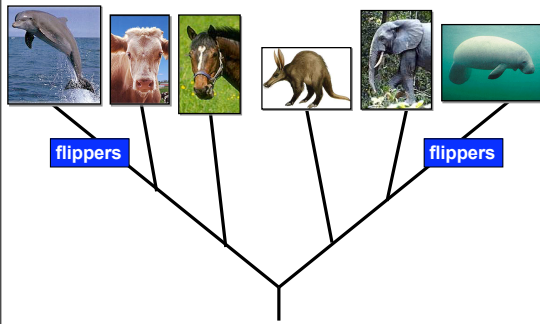
Homoplasy



Divergence

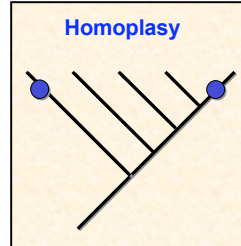


Example of Homoplasy

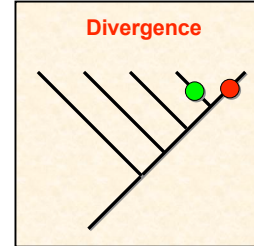


Two common reasons for deceptive appearances

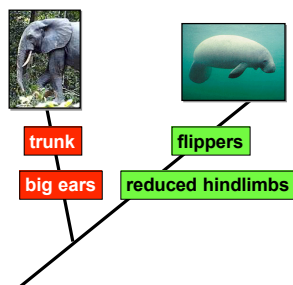
Homoplasy



Divergence



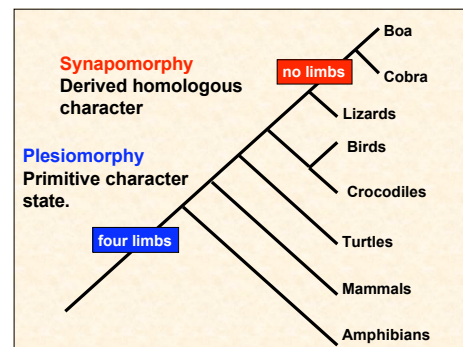
Example of Divergence



Terms to describe character states

Synapomorphy
Derived homologous character

Plesiomorphy
Primitive character state.



Terms to describe character states

Plesiomorphy, synapomorphy, and homoplasy are relative terms and may change depending on the taxonomic problem.

How are phylogenies constructed?

Three commonly used methods



Parsimony

Distance

Maximum Likelihood

Parsimony Methods

Few assumptions:

1. Simpler hypotheses are preferred to complicated ones.
2. Only synapomorphies are informative about relationships.

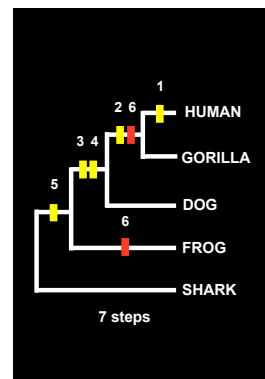
The basic idea of parsimony is to minimize the number of changes occurring on a phylogeny. Check all possible tree topologies and choose the shortest.



Morphology

	1	2	3	4	5	6
HUMAN	1	1	1	1	1	1
GORILLA	0	1	1	1	1	1
DOG	0	0	1	1	1	0
FROG	0	0	0	0	1	1
SHARK	0	0	0	0	0	0

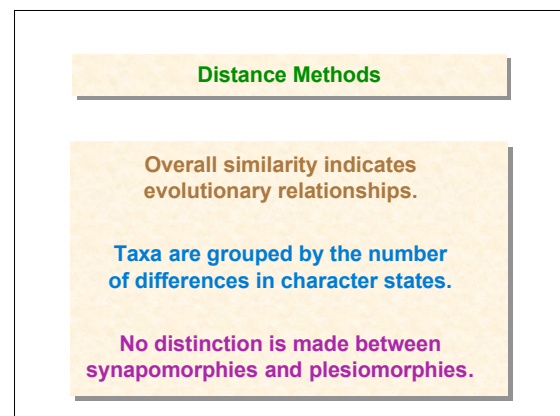
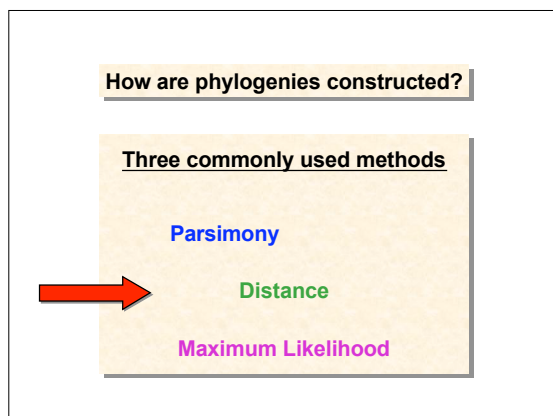
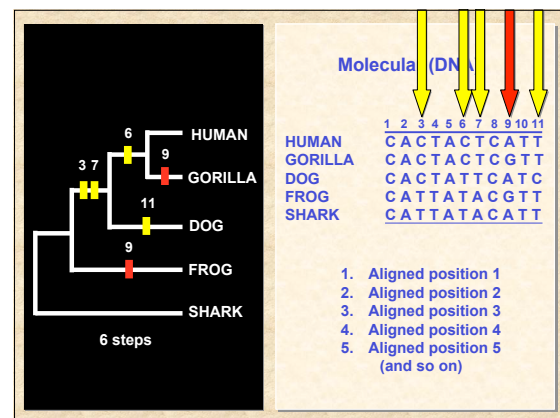
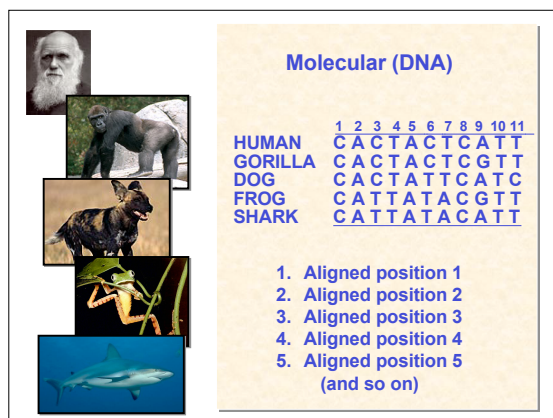
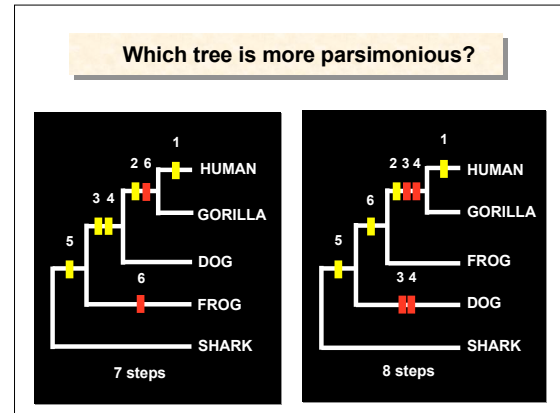
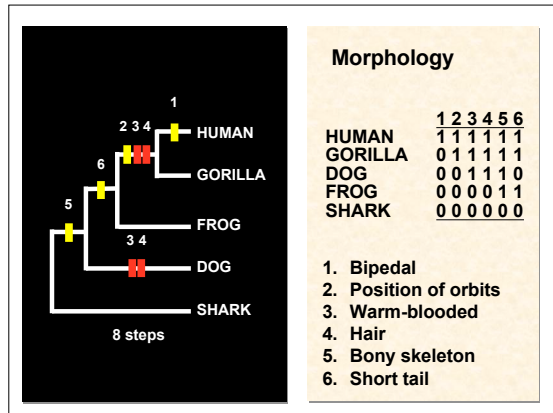
1. Bipedal
2. Position of orbits
3. Warm-blooded
4. Hair
5. Bony skeleton
6. Short tail



Morphology

	1	2	3	4	5	6
HUMAN	1	1	1	1	1	1
GORILLA	0	1	1	1	1	1
DOG	0	0	1	1	1	0
FROG	0	0	0	0	1	1
SHARK	0	0	0	0	0	0

1. Bipedal
2. Position of orbits
3. Warm-blooded
4. Hair
5. Bony skeleton
6. Short tail

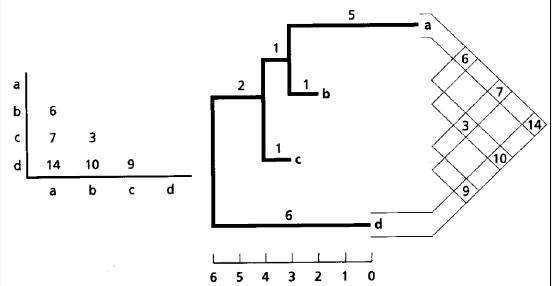


Distance methods start by calculating the distances between pairs of species.

	Morphology	Molecular (DNA)
HUMAN	1 1 1 1 1 1	CACTACTCATT
GORILLA	0 1 1 1 1 1	CACTACTCGTT
DOG	0 0 1 1 1 0	CACTATTTCATC
FROG	0 0 0 0 1 1	CATTATACGTT
SHARK	0 0 0 0 0 0	CATTATACATT

	Human	Gorilla	Dog	Frog
Human				
Gorilla	2			
Dog	5	5		
Frog	8	6	7	
Shark	9	9	6	3

Distance matrix



How are phylogenies constructed?

Three commonly used methods

Parsimony

Distance

Maximum Likelihood

Maximum Likelihood Methods

ML methods find the tree most likely to produce the data observed given a specific model of how character states change.

ML methods are almost always used with only molecular data.

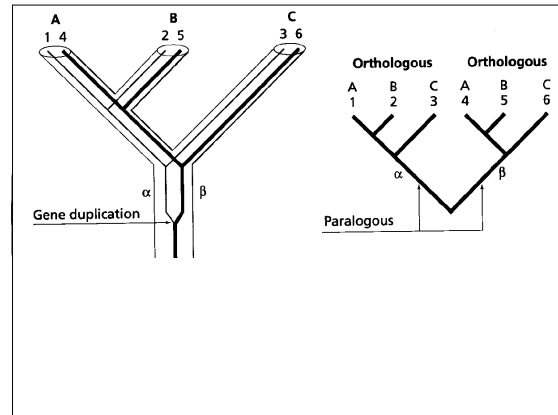
Which method should I use?

- Depends on the data being analyzed.
- Try all methods to determine if results are robust to model assumptions.
- My preference is for ML.

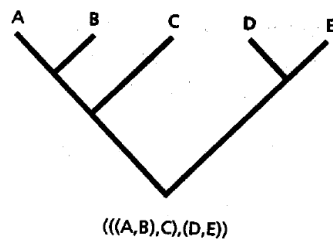
Some things we ignored

- Base composition
- Transition/transversion bias
- DNA substitution models
- Multiple substitutions
- Character weighting
- Choice of taxa

Does a tree generated from a gene sequence always give the species tree?



Representation of tree using nested parentheses



Is the branching order significant?

- Bootstrap test for significance.
- Resample data with replacement
- Reconstruct phylogeny
- Redo 100s of times
- Build consensus tree, bootstrap values are percent on times that branch was recovered.

ORIGINAL DATA	AAATCTGCTAGCATGCA
	AAATCTGCTACCATGCA
BOOTSTRAPPED DATA SETS	
AAAAACCCGGCTCAGCA	AATTGGAAAGCAAGGCA
AAAAACCCGGCTCAGCA	AATTGGAAAGCAAGGCA
AATTTTCCCCATTGCC	AAAATCCCATGGGAA
AATTTTCCCCATTGCC	AAAATCCCATGGGAA
AATTTCAAACCCCTCA	AATTGGCAAGCAAAAC
AATTTCAAACCCCTCA	AATTGGCAAGCAAAAC
AATTTGCCCCCAGAA	AAATCTTCCTGGGAA
AATTTGCCCCCAGAA	AAATCTTCCTGGGAA
AAATCCCGGTACTTGCA	AAAATTTCCAAGATTCA
AAATCCCGGTACTTGCA	AAAATTTCCAAGATTCA