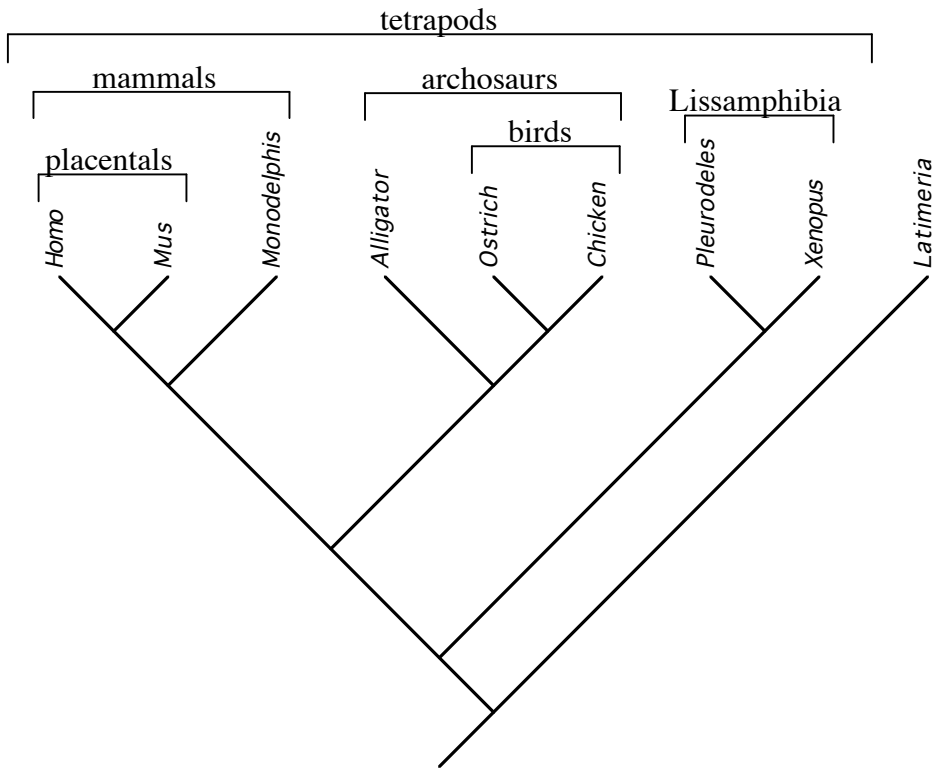


NAMES _____

Molecular Clocks

I. Discovery of the molecular clock

Relationships among the major vertebrate groups are well established, and shown by the following tree:



One way to estimate when divergences among these taxa occurred is to use data from the fossil record:

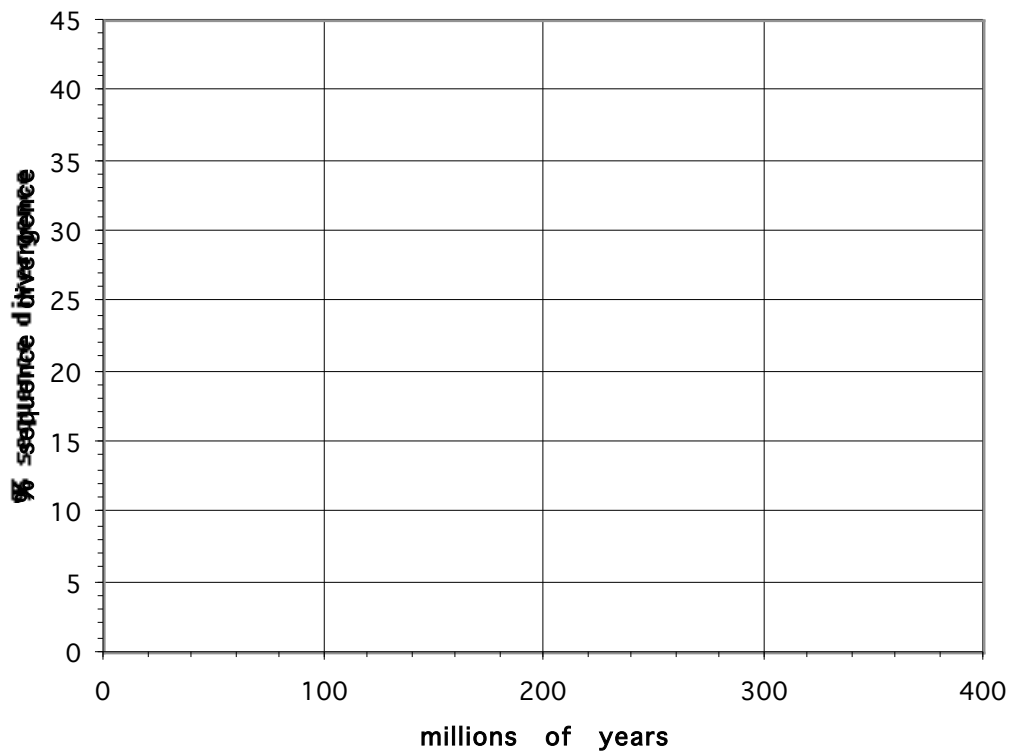
oldest tetrapod fossil	ca. 365 million years
stem amphibians (i.e. fossils that are more closely related to amphibians than to any other living group)	ca. 338 million years
oldest evidence of bird-mammal split	ca. 310 million years
oldest placental mammal fossil	between 143 and 178 million years

Use the fossil data to add divergence dates to the above tree for all of the nodes for which the fossil record provides an estimate. (Hint: The oldest tetrapod fossil gives a minimum estimate for the divergence between *Latimeria* and tetrapods. Why does the estimate apply to this node rather than to the node representing the earliest split within tetrapods?) **Check with me before proceeding to be sure that the nodes are properly labeled.**

The table below shows the estimated sequence divergence between representatives of some of the major groups of tetrapod vertebrates for 2745 base pairs of the recombination activating gene factor 1 (RAG-1).

	1	2	3	4	5	6	7	8
1 <i>Homo</i>								
2 <i>Mus</i>	0.14							
3 <i>Monodelphis</i>	0.23	0.27						
4 <i>Alligator</i>	0.29	0.31	0.31					
5 <i>Chicken</i>	0.32	0.34	0.33	0.19				
6 <i>Ostrich</i>	0.31	0.33	0.31	0.18	0.11			
7 <i>Xenopus</i>	0.38	0.40	0.39	0.36	0.37	0.37		
8 <i>Pleurodeles</i>	0.37	0.38	0.37	0.34	0.37	0.37	0.38	
9 <i>Latimeria</i>	0.39	0.42	0.42	0.37	0.39	0.38	0.43	0.42

Plot the % sequence divergence (on the vertical axis; the number above is the fraction of sites that differ, so multiply by 100 to get the %) versus the estimated age of divergence (on the horizontal axis) for the comparisons for which both numbers are known.



What is the general relationship between sequence divergence and time? This relationship has been called the molecular clock.

Is this general pattern consistent with the observed sequence divergence for taxa for which an estimated divergence date based on the fossil record is not available? Why or why not?

How can you explain the variation in % sequence divergence between pairs of taxa of the same age?

II. Using the molecular clock to estimate unknown divergences

Molecular clock estimates are frequently used to estimate divergence dates in cases in which the fossil record is not very good.

Explain how can you use molecular data from species with good fossil records to estimate the divergence times for taxa that are less well known in the fossil record. (Hint: Use your graph above.)

Taxon Pair	Estimated divergence time
ostrich-chicken	
<i>Pleurodeles-Xenopus</i>	
Human-mouse	

One way to test how good these estimates are is to compare estimates based on molecular and fossil data in cases in which both are available.

Suppose that the only date that was well established from the fossil record was the bird-mammal divergence date. What is the best estimate of % sequence divergence in the RAG-1 gene between archosaurs (= birds + crocodilians) and mammals?

Based on this number, what % sequence divergence do you estimate accumulates per million years between diverging lineages? This is called a molecular clock calibration.

Use the molecular clock calibration derived from the mammal-archosaur split to estimate divergence dates for the following taxa:

Taxon Pair	Molecular clock estimate	Fossil estimate
<i>Latimeria</i> -tetrapod		
Amphibian-amniote		
Alligator-bird		
Marsupial mammal-placental		

How do the estimated numbers and the numbers from the fossil record compare?

Why might you expect dates estimated using a molecular clock to be older than dates estimated from the fossil record?

How much confidence does this give you in molecular clock estimates of divergence dates?

III. Explanation for the molecular clock

The general pattern that % sequence divergence increases more or less linearly with the age of divergence between taxa was a surprise when first discovered in the early 1960s, and was one of the motivations for the development of the neutral theory of molecular evolution. Previously, it had been believed that the amount of sequence divergence would reflect how similar or different the function of a particular protein was rather than how long lineages had been diverging for.

Let's look more closely at why this pattern is predicted under the neutral theory.

Assume that a gene has a neutral mutation rate of 1×10^{-9} (given in units of mutations per site per generation; this is in the ball park of many actual estimates of mutation rate.)

How many sites do you expect to become fixed for a new base in a 1000 base pair region in 1 million years? (Assume that the generation time = 1 year)

If a speciation event occurred 1 million years ago, at how many bases (of 1000) do you expect a pair of sister taxa to differ?

What if the speciation event occurred 2 million years ago? 10 million years ago?

Use the considerations above to explain why the neutral theory predicts the molecular clock pattern. What assumptions do you need to make in this explanation?

Suppose that the neutral mutation rate is measured to be 5×10^{-9} in a gene encoding a digestive enzyme and as 1×10^{-9} in a gene encoding a DNA binding protein. Give two biological explanations for why the neutral mutation rate could be higher in the digestive enzyme gene.

How could you distinguish between these two explanations?

What effect would a higher neutral mutation rate have on the rate of sequence divergence? If you found that the DNA binding proteins were 3% divergent between two species of frogs, how divergent would you expect the digestive enzymes to be in these same species?