

Name \_\_\_\_\_ Period \_\_\_\_ Date \_\_\_\_\_

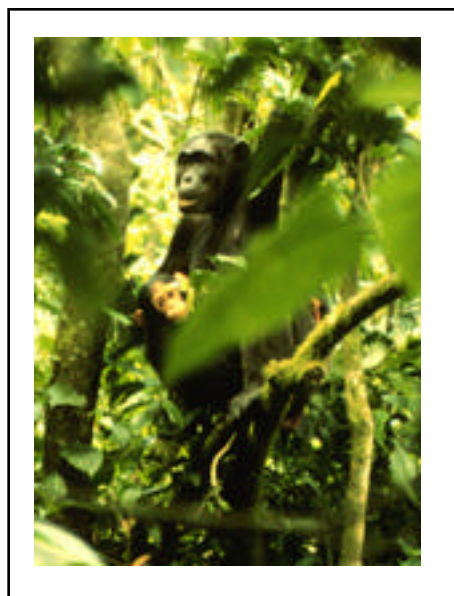
# Molecular Biology and Primate Phylogenetics

by John Banister-Marx, Wright Center for Science Education, Tufts University, Medford, MA  
In collaboration with Dr. Martin Nickels, Illinois State University Normal, Illinois

**INTRODUCTION:** In general, the function of living things depends upon utilizing energy and materials to support processes that take place within the organism. These processes are directed by a molecule called **DNA**, the hereditary material found in virtually every cell. First discovered prior to the 1900's, it's structure would not be deciphered until the 1950's. Since then a revolution affecting medicine, agriculture, and other fields has been sparked by using this "master molecule" and the powerful information held within the sequence of its **nitrogen bases**. These building blocks of DNA, read three at a time as a triplet, provide information about the precise sequencing of amino acids for the assembly of a protein.

As early as the 1960's scientists began to relate molecular differences between living organisms to geological time. Today, molecular differences among creatures can be quantified using one of four techniques: whole-protein immunology, counts of amino acid differences, nucleic acid or DNA hybridization, and counts of nucleotide difference. . The work of molecular biologists in recent years has been increasingly important in further clarifying our understanding of not only evolutionary histories (**phylogenies**) for all life on earth, but also the possible rates of mutation and speciation.

Among the first proteins to have its amino acid sequence deciphered was hemoglobin, and it remains today one of the most investigated of all proteins. Found in nearly all living things, hemoglobin is an oxygen-transport protein that gives blood its familiar red color. In vertebrates, hemoglobins are usually the primary protein of red blood cells, making them relatively easy to isolate and purify in large quantities. It is the virtually universal presence of this molecule that makes it so valuable in understanding evolutionary relationships among creatures. (Image source: <http://www.indiana.edu/~primate/primates.html>)



Though it is sometimes hard to imagine, molecules do evolve, and the change that occurs does accumulate through time. For example, the amount of difference in the sequence of amino acids of a given protein found in a pair of organisms being compared is a function of how much time has passed, and been made available for accumulating mutations, since any two creatures diverged from a common ancestor.

**PURPOSE:** To examine the amino acid sequence differences for a specific protein (beta hemoglobin) in several primate species, and from this infer ancestral relationships.

**MATERIALS:** (each student) \_\_\_\_ 4 page student handout \_\_\_\_ pencil

## **Data sources:**

Species A-C: Morris Goodman (1992). Reconstructing Human Evolution From Proteins. In: Steve Jones et al (1992), The Cambridge Encyclopedia of Human Evolution. Cambridge University Press. pp 307-312

Species D-H: Genbank, National Center for Biotechnology Information. <http://www.ncbi.nlm.nih.gov>

## PROCEDURE:

### PART 1: Comparing Amino Acid Sequences of Beta Hemoglobin

Listed below in Table #1 is a comparison of the 146 amino acid beta chain of the hemoglobin molecule in 7 selected primate species, plus one non-primate species H. Notice that the amino acid position numbers may not be continuous. This is because those that are the same for ALL species have been left off the chart to save space. Begin by selecting the species A and B. Compare the letters at each position and count the number of differences in the letters at any position down the chain. Enter the total number of differences you count in the appropriate box in Table #1. Count the number of amino acid *differences* between each of the possible pairs of organisms and then enter it in the appropriate box in the Table #2. Do not enter any data yet for creature H. You will enter in column H of Table #2 later

**Table #1:** Selected amino acid sequences for beta hemoglobin protein

Amino acid position	Species							
	A	B	C	D	E	F	G	H
1	V	V	V	V	V	V	T	V
2	H	H	H	H	H	H	L	Q
4	T	T	T	T	T	T	S	S
5	P	P	P	P	P	G	A	G
6EE			E	E	E	D	EE	
8	K	K	K	K	K	K	D	K
9	S	S	S	S	D	A	A	A
10	A	A	A	A	A	A	H	A
12	T	T	T	T	T	A	T	L
13	A	A	A	A	T	A	SA	
21	D	D	D	D	D	E	E	E
22	E	E	E	E	E	D	K	E
33	V	V	V	V	L	V	V	V
50	T	T	T	T	S	T	S	N
52	D	D	D	D	D	D	SG	
56	G	G	G	G	G	N	S	G
69	G	G	G	G	G	G	S	H
73	D	D	D	D	D	D	E	E
75	L	L	L	L	L	L	L	V
80	A	A	A	A	D	T	H	H
87	T	T	T	K	Q	Q	Q	A
104	R	R	L	R	K	R	K	R
111	V	V	V	V	V	V	S	V
112	C	C	C	C	C	C	A	V
113	V	V	V	V	V	V	E	V
114	L	L	L	L	L	L	S	L
115	A	A	A	A	A	A	E	A
116	H	H	H	H	H	H	L	R
120	K	K	K	K	K	K	H	K
121	E	E	E	E	E	E	D	D
122	F	F	F	F	F	F	K	F
123	T	T	T	T	T	T	S	T
125	P	P	P	Q	Q	Q	A	E
126	V	V	V	V	V	L	V	L
130	Y	Y	Y	Y	Y	Y	F	Y

### KEY:

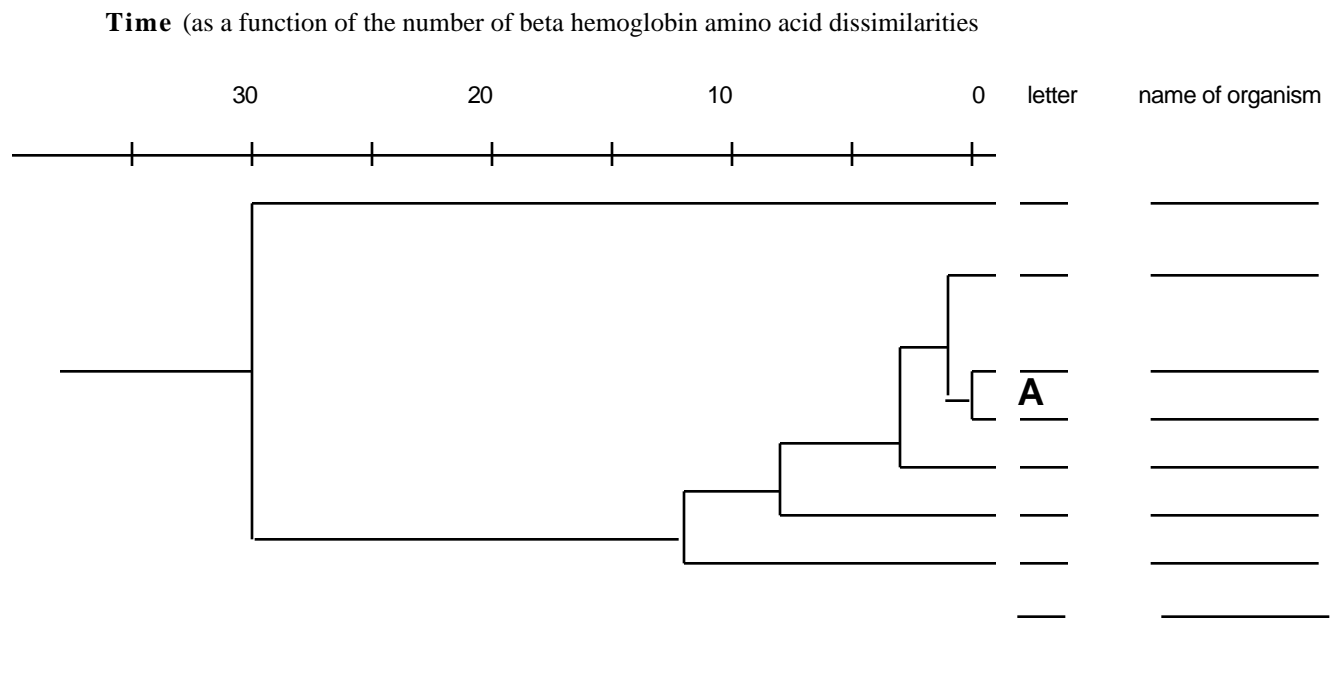
A = alanine N = asparagine  
R = arginine D = aspartic acid  
C = cysteine Q = glutamine  
G = glycine E = glutamic acid  
L = leucine H = histidine  
K = lysine F = phenylalanine  
P = proline S = serine  
Y = tyrosine T = threonine  
V = valine

**Table #2:** A comparison of the number of amino acid differences between 7 primate species. Enter your data below by counting the total number of differences in each pair of organisms indicated in Table #1

	A	B	C	D	E	F	G	H
A								
B								
C								
D								
E								
F								
G								
H								

**PART 2:** The data that you have gathered and entered in Table #2 represents the amount of time (in amino acid substitutions) since one of the pair of organisms has diverged from a common ancestor with the other organism. A fewer number of amino acid differences between any two creatures implies that the two organisms share a relatively recent common ancestor, whereas a larger number of amino acid differences implies the two organisms share a relatively distant common ancestor. Scientists can take this information to construct a phylogenetic tree that shows the branching patterns from shared common ancestry. Using the data in Table #2, enter the appropriate letter (A, B, C, D, E, F, or G) in each of the appropriate spaces to the right side of the phylogenetic tree in Graph #1 below, under the word “letter.” Notice the horizontal time axis on top of the graph. Then answer the discussion questions.

Graph #1: Evolutionary tree for seven primate species based on a comparison of amino acid *dissimilarities* of beta hemoglobin.



**Time** (in millions of years ago) to be scaled as a part of question 9

### **DISCUSSION:**

1. Which two species have the greatest similarity in the amino acid sequence for this protein? \_\_\_\_ & \_\_\_\_.
2. According to the phylogenetic tree which 2 organisms share the most recent common ancestor? \_\_\_\_ & \_\_\_\_.
3. After reviewing the basic anatomy of primates from your text and/or references, studying the amino acid difference in Table #1, and viewing the tree above, enter the *names* of the 7 primates into their appropriate positions at the right under “name of organism.” Chose from human, chimpanzee, gorilla, gibbon, rhesus monkey (Old World monkey form), squirrel monkey (New World monkey form) and lemur. (A is for human).

DISCUSSION: (continued)

4. Why are the organisms A-F all approximately the same number of amino acid differences positions to organism G? How can they possibly be equidistant from the lemur when it *appears* that a squirrel monkey looks more like a lemur than does a human? (The data *is* actually what we would predict).
5. Based on the pattern you see in Table #1, predict the number of differences for all species (A-G) with a new species H, a horse. Enter your predictions on the last row in Table #1 and draw a line of descent on the tree.
6. At a molecular level, if humans and chimpanzees are about 99% identical, a horses and zebras are about 96% identical, what does this suggest about which **pair** of organisms is more closely related? Which pair shares the most recent common ancestor? Explain your reasoning.
7. How does your answer in #5 compare with the actual data of approximately 19 amino acid differences with each of the organisms A-F? (compare the actuals by counting the differences between species H and all other creatures in Table #1 and entering a small numerical note in the appropriate box in Table #2) Should this one piece of anomalous data call into question all of the previous data comparing the primates? Why or why not?
8. How would checking to DNA triplet codon for all of the amino acid sites be a logical first step in determining the source of this apparently aberrant data when comparing this group of primates to a horse? (Hint: Look at your triplet codon/anticodon/amino acid tables).
9. By examining evidence from fossils, and molecular and anatomical comparisons, evolutionary biologists have inferred that lemurs and humans share a common ancestor at about 43 million years ago, and that humans and chimpanzees share a common ancestor at about 6 m.y.a. Use this information to completely scale a second x axis at the bottom of the graph to correspond with the upper axis for amino acid mutations. What appears to be the rate of mutation for beta hemoglobin? 1 amino acid mutation per \_\_\_\_\_ million years.
10. Do chimpanzees and squirrel monkeys share a common ancestor 4 million years ago? Why or why not?