

Molecular Biology and Primate Phylogenetics

Introduction

Studying the nucleotide sequences of DNA, and/or the amino acid sequences of proteins, gives scientists one of a growing number of ways to analyze relationships and infer ancestry for life on earth. The work of molecular biologists has been important in further clarifying and refining our understanding of not only family trees (phylogenies) for all life on earth, but also the possible **rate** of mutation, selection, and speciation.

Among the first proteins to yield its amino acid sequence was hemoglobin, and it remains today one of the most investigated of all proteins. The basic unit of hemoglobin consists of an iron-containing porphyrin (heme) that can reversibly bind oxygen attached to a globin polypeptide chain that is usually no less than 140 amino acids long. In vertebrates, hemoglobins are usually the primary protein of red blood cells, making them relatively easy to isolate and purify in large quantities. (Strickberger, M. W. *Evolution*. Jones and Barlett Publishers, 1990)

Purpose

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

Procedure Part A:

Listed below is a comparison of the 146 amino acid **beta** chain of the hemoglobin molecule in 7 selected primate species. Notice that the amino acid position numbers may not be continuous. This is because those that are the same for ALL seven species have been left off the chart to save space. Count the number of amino acid **differences** between each of the possible pairs of organisms using the data below and then enter it in Table #1. Draw a diagonal line from the upper left to lower right of the grid and then cross out all of the squares in the upper right side of the table. No need to duplicate your work. You will enter in row H later.

Table 1: A comparison of the number of amino acid differences between seven primate species.

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

Amino acid position

Human
 Chimp
 Gorilla
 Gibbon
 Rhesus Monkey
 Spumet Monkey
 Lemur
 SPECIES
 A B C D E F G

1 VAL VAL VAL VAL VAL VAL THR
 2 HIS HIS HIS HIS HIS HIS LEU
 4 THR THR THR THR THR THR SER
 5 PRO PRO PRO PRO PRO GLY ALA
 6 GLU GLU GLU GLU GLU ASP GLU
 8 LYS LYS LYS LYS LYS LYS ASP
 9 SER SER SER SER ASN ALA ALA
 10 ALA ALA ALA ALA ALA ALA HIS
 12 THR THR THR THR THR ALA THR
 13 ALA ALA ALA ALA THR ALA SER
 21 ASP ASP ASP ASP ASP GLU GLU
 22 GLU GLU GLU GLU LU ASP LYS
 33 VAL VAL VAL VAL LEU VAL VAL
 50 THR THR THR THR SER THR SER
 52 ASP ASP ASP ASP ASP ASP SER
 56 GLY GLY GLY GLY GLY ASN SER
 69 GLY GLY GLY GLY GLY GLY SER
 73 ASP ASP ASP ASP ASP ASP GLU
 76 ALA ALA ALA ALA ASN THR HIS
 80 ASN ASN ASN ASP ASN ASN ASN
 87 THR THR THR LYS GLN GLN GLN
 104 ARG ARG LEU ARG LYS ARG LYS
 111 VAL VAL VAL VAL VAL VAL SER
 112 CYS CYS CYS CYS CYS CYS ALA
 113 VAL VAL VAL VAL VAL VAL GLU
 114 LEU LEU LEU LEU LEU LEU SER
 115 ALA ALA ALA ALA ALA ALA GLU
 116 HIS HIS HIS HIS HIS HIS LEU
 120 LYS LYS LYS LYS LYS LYS HIS
 121 GLU GLU GLU GLU GLU GLU ASP
 122 PHE PHE PHE PHE PHE PHE LYS
 123 THR THR THR THR THR THR SER
 125 PRO PRO PRO GLN GLN GLN ALA
 126 VAL VAL VAL VAL VAL LEU VAL
 130 TYR TYR TYR TYR TYR TYR PHE

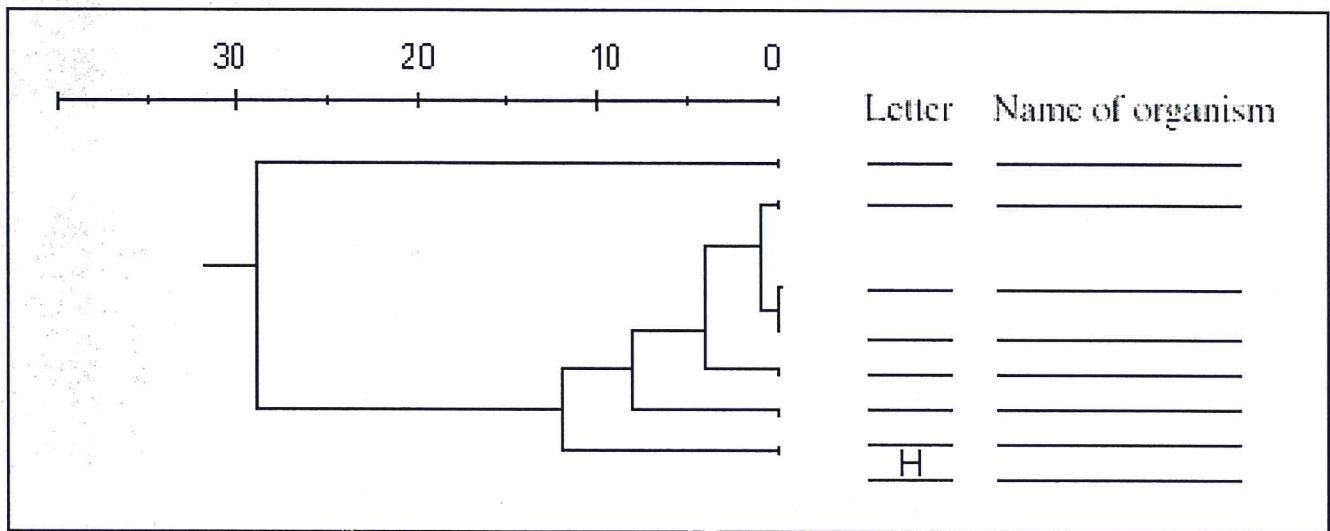
Key

- ALA = alanine
- ASN = asparagine
- ARG = arginine
- ASP = aspartic acid
- CYS = cysteine
- GLN = glutamine
- GLY = glycine
- GLU = glutamic acid
- LYS = lysine
- PHE = phenylalanine
- PRO = proline
- SER = serine
- TYR = tyrosine
- THR = threonine
- VAL = valine

Procedure Part B

The data that you have gathered and entered in Table #1 represents the amount of time (in amino acid substitutions) since one organism has diverged from a common ancestor with another 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 large 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 of descent. Using the data from Table #1, 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 under "letter." Notice the horizontal time axis on the top of the graph. Then answer the discussion questions below.

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 part of question 9

Discussion

- Which two species have the greatest similarity in the amino acid sequence for this protein? ____ and ____
- According to the phylogenetic tree which two organisms share the most recent common ancestor? ____ and ____ .
- After reviewing the basic anatomy of primates from your text and/or references, studying the amino acid differences 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 human). *Given on previous page*
- 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 spider monkey looks more like a lemur than does a human? (This data is actually what we would predict)
- 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, and horses and zebras are about 96% identical, what does this suggest about which pair of organisms is more closely related? Which pair shares a more recent common ancestor? Explain your reasoning.
7. How does your answer in #5 compare with the actual data of approximately 28 amino acid differences with each of the organisms A-F? 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 the DNA triplet codon for all of the amino acid sites be a logical first step in determining the source of this aberrant data? (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?