

## Analyzing Polypeptide Sequence Data

The amino acid sequence data for the  $\beta$  polypeptide chain of human, Rhesus monkey and Gibbon hemoglobin for three species is given in Table 1. The numbers give the position of each amino acid in the chain.

**Table 1: Amino acid sequence data for  $\beta$  globin for three species**

Species	Alignment of Amino Acid Sequences of $\beta$ -globin					
Human	1	VHLTPEEKSA	VTALWGKVVN	DEVGGEALGR	LLVVYPWTQR	FFESFGDLST
Monkey	1	VHLTPEEKNA	VTTLWGKVVN	DEVGGEALGR	LLLVPYPWTQR	FFESFGDLSS
Gibbon	1	VHLTPEEKSA	VTALWGKVVN	DEVGGEALGR	LLVVYPWTQR	FFESFGDLST
Human	51	PDAVMGNPKV	KAHGKKVLGA	FSDGLAHLDN	LKGTFAQLSE	LHCDKLHVDP
Monkey	51	PDAVMGNPKV	KAHGKKVLGA	FSDGLNHLDN	LKGTFAQLSE	LHCDKLHVDP
Gibbon	51	PDAVMGNPKV	KAHGKKVLGA	FSDGLAHLDN	LKGTFAQLSE	LHCDKLHVDP
Human	101	ENFRLLAGNVL	VCVLAHHFGK	EFTPPVQAAY	QKVVAGVANA	LAHKYH
Monkey	101	ENFKLLGNVL	VCVLAHHFGK	EFTPQVQAAY	QKVVAGVANA	LAHKYH
Gibbon	101	ENFRLLAGNVL	VCVLAHHFGK	EFTPQVQAAY	QKVVAGVANA	LAHKYH

1. Scan each sequence and circle any amino acids that do not match the human sequence.
  - a) State the number of amino acids that differ between the monkey and human sequences and between the gibbon and human sequences?
  - b) For each nonhuman species, state the percent of the amino acids that are identical to the human sequence.
2. a) Based on these data, propose a hypothesis for which of these two species is more closely related to humans. Justify your hypothesis.
  - b) What other evidence could be used to support your hypothesis?

A 2016 report from the ocean conservation advocacy group Oceana claimed that one in five of over 25,000 samples of seafood tested globally was mislabeled.

Imagine you intend to find evidence of such incorrect labelling so you purchase a piece of fish labeled as coho salmon (*Oncorhynchus kisutch*). To see if your fish is correctly labelled, you will compare a short DNA sequence from your sample to standard sequences from the same gene for three salmon species. The sequences are shown in Table 2.

**Table 2: DNA sequences from salmon species**

Sample	Sample	
Sample	labelled as Coho salmon ( <i>O. kisutch</i> )	5'-CGGCACCGCCCTAAGTCTCT-3'
Reference 1	Coho salmon ( <i>O. kisutch</i> )	5'-AGGCACCGCCCTAAGTCTAC-3'
Reference 2	Chum salmon ( <i>O. keta</i> )	5'-AGGCACCGCCCTGAGCCTAC-3'
Reference 3	Atlantic salmon ( <i>Salmo salar</i> )	5'-CGGCACCGCCCTAAGTCTCT-3'

3. State the number of bases that differ between the subject and each reference. Scan the sequences and circle any bases that do not match the sequence from the sample.
4. Based on these data, propose a hypothesis about the identity of the subject fish. Justify your reasoning.