

1 Reading assignment 8. Section 4.5

Find the file TA-PrStrPr.t2v3.rar with another section of the material.
Read Section 4.5 and answer the following:

1. What do values in a PAM matrix represent?

The probability that one amino acid is mutated into another different. The amount of evolutionary change in a protein sequence i.e. the frequency with which two amino acids have replaced each other during evolution. Its formula is the following:

$$\frac{f_{ij}}{f_i \times f_j}$$

2. In a PAM matrix, suppose $f_{\text{ALA}}=0.3$, $f_{\text{SER}}=0.2$, $f_{\text{ALA-SER}}=0.3$. Calculate the entry ALA-SER of the PAM matrix.

$$\text{PAM}' = \frac{f_{\text{ALA-SER}}}{f_{\text{ALA}} \times f_{\text{SER}}} = \frac{0.3}{0.3 \times 0.2} = 5$$

3. How are BLOSUM matrices derived?

Using local alignments of very conserved regions in homologous proteins. Is derived from alignments such that all sequences sharing more than N% identity with any other sequence in the alignment are averaged and represented as a single sequence.

4. Consider the matrices PAM100, PAM250, BLOSUM62, BLOSUM90. Which matrices would be best for comparing closely related sequences?

PAM100 (low ratio) and BLOSUM 90 (high ratio).

5. What is the purpose of programs such as BLAST?

To search a data base for protein sharing a significant sequence similarity to a query protein evaluating whether the score obtained in the real alignment is likely to belong to its random distribution to compute the reported probability values.