**Course: Advanced Bioinformatics**

**Module title: Neighbor Join**

**Module no. : 48**

This module gives evidence on neighbor joins.

The neighbor-joining method is very similar to the Fitch-Margoliash method. The sequences that should be joined are chosen to give the best least-squares estimates of the branch lengths that most closely reflect the actual distances between the sequences.

The neighbor-joining method begins by creating a star topology in which no neighbors are joined:

**B**

**A**

**C**

**D**

**E**

The tree is modified by joining pairs of sequences. The pair to be joined is chosen by calculating the sum of the branch lengths for the corresponding tree. The sum of the branch lengths is calculated as follows:



where i,j represent all sequences except m and n, and i < j.

For example, consider the tree when A and B are joined:

**B**

**A**

**C**

**D**

**E**



The pair that results in the smallest branch length is then chosen to be the pair that is joined. Based on this choice, the Fitch-Margoliash algorithm is used to compute the actual branch lengths.

After the pair has been joined, a new distance table is created with the recently joined sequences now entered as a composite. The neighbor-joining algorithm chooses the next pair of sequences to join, and the F-M algorithm computes the branch lengths.

The process continues until the correctly branched tree and distances have been identified.