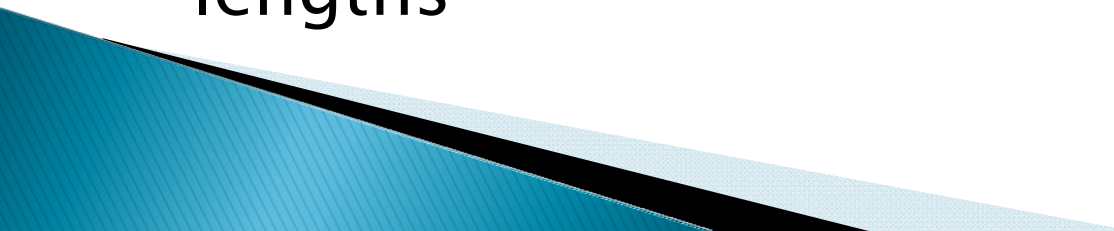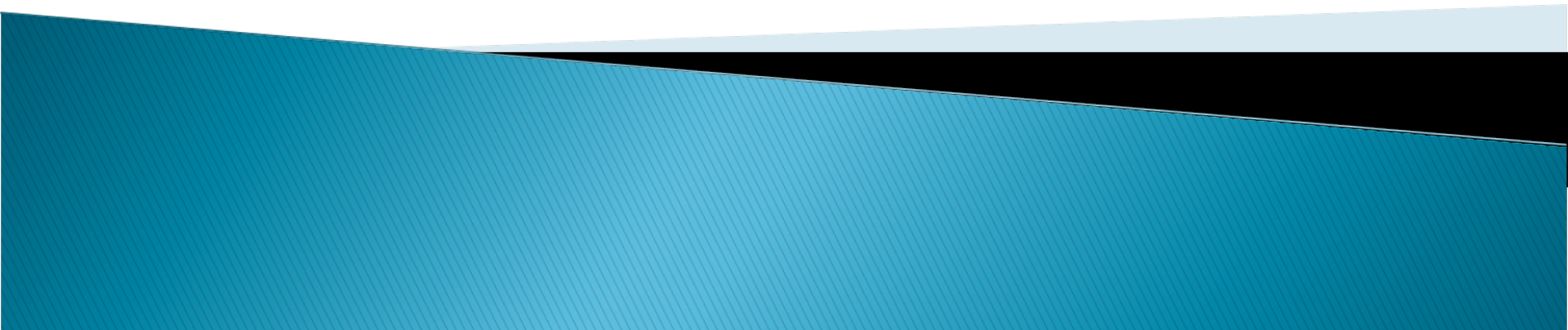# Neighbor Joining Method

Karthik Pasupathy R
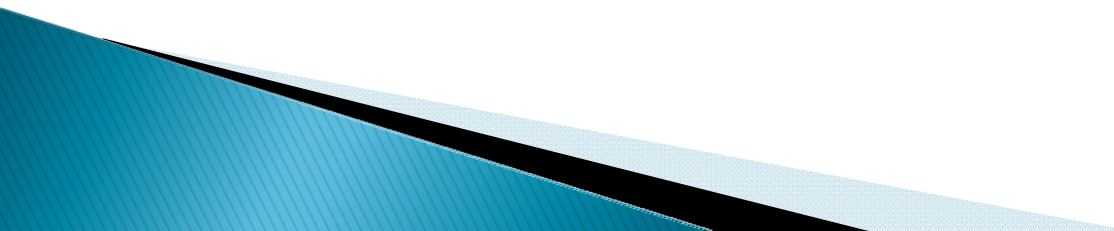
Roll.no: 1801110002

What is NJ method?

- Nj method is a clustering approach for the reconstruction of phylogenetic trees

- It is based on the ST method (Sattath and Tversky) method of tree construction

- The input is the 'n' number of taxa
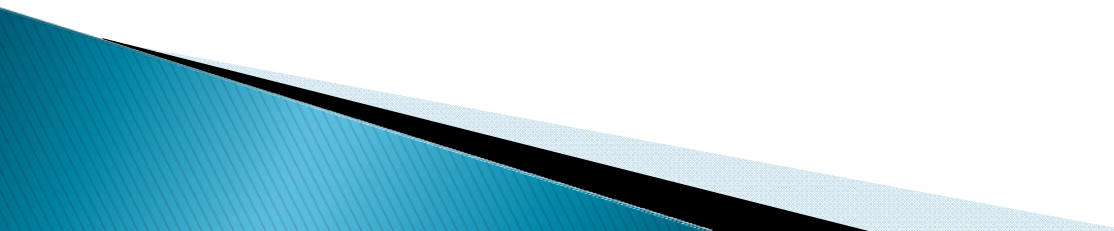
- The output is an unrooted tree with branched lengths

Other algorithms

- Distance Wagner method

- Modified Farris method (FM)

- Sattath and Tversky method (ST)

- Fitch method

ST method vs NJ method

- In ST method , the objective of the algorithm is to construct the topology of a phylogenetic tree.

- NJ method not only provides topology but also provides final tree with branched lengths

Algorithm

- Construct a unresolved tree with star topology

- Construct a Distance matrix and calculate the value of Sa , Sb , Sc.....Sn

- Group neighbors based on distance

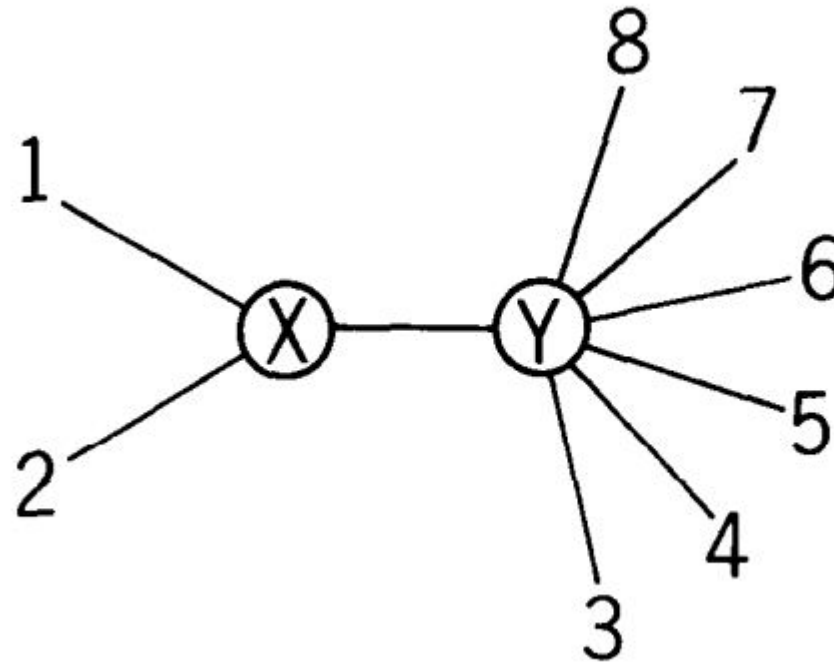- Repeat the steps until you get the final tree

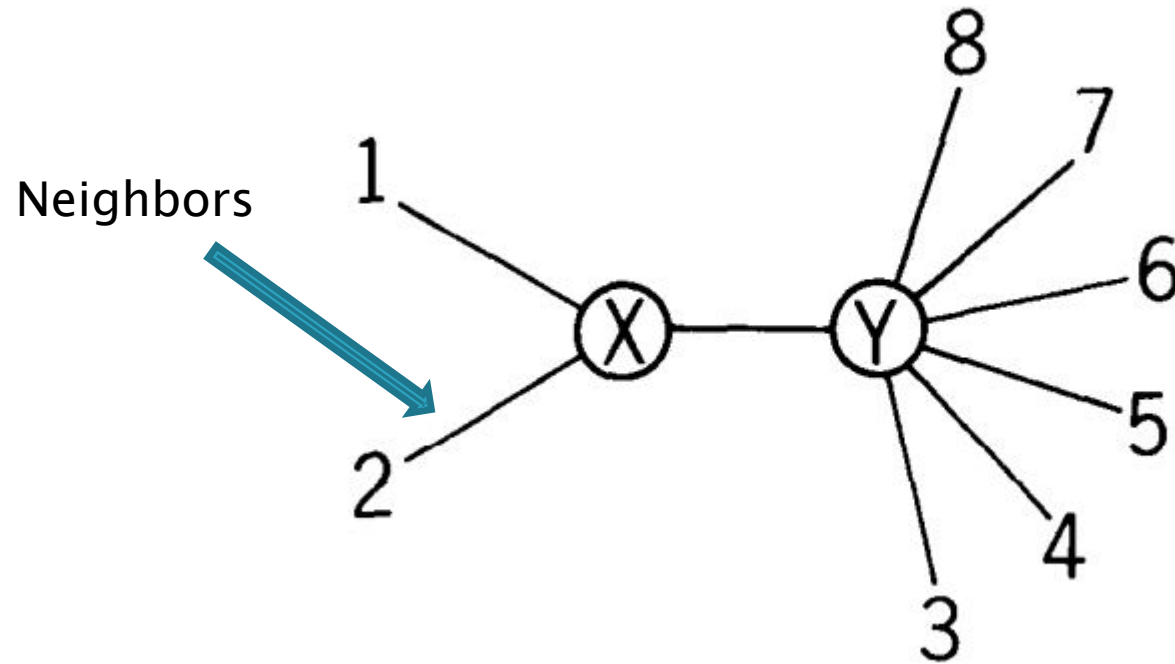# Algorithm

▸ construct an unresolved tree with star topology
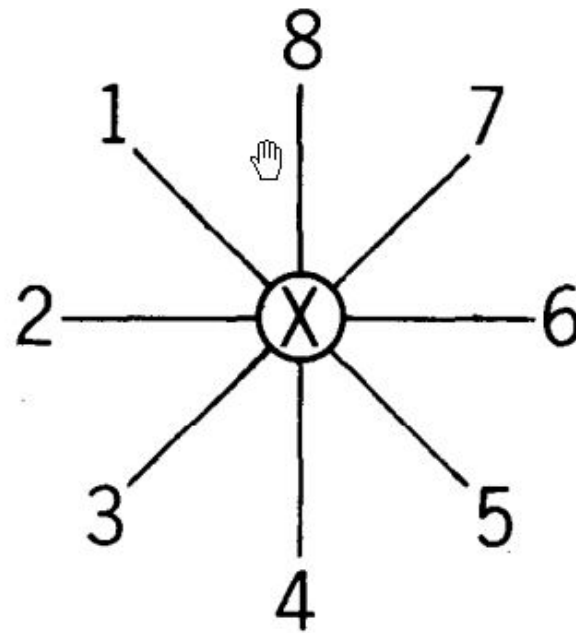
Algorithm

- Join the closest neighbors (OTUs with similar characters)

Algorithm

- Join the closest neighbors (OTUs with similar characters)

Distance matrix

	A	B	C	D	E
B	5				
C	4	7			
D	7	10	7		
E	6	9	6	5	
F	8	11	8	9	8

	U_1	C	D	E
C	3			
D	6	7		
E	5	6	5	
F	7	8	9	8

	U_1	C	U_2
C	3		
U_2	3	4	
F	7	8	6

	U_2	U_3
U_3	2	
F	6	6

	U_4
F	5

Step 1

S calculations

$S_x =$ (sum all D_x)/$(N-2)$, where N is the # of OTUs in the set.

$S_A = (5+4+7+6+8)/4 = 7.5$
$S_B = (5+7+10+9+11)/4 = 10.5$
$S_C = (4+7+7+6+8)/4 = 8$
$S_D = (7+10+7+5+9)/4 = 9.5$
$S_E = (6+9+6+5+8)/4 = 8.5$
$S_F = (8+11+8+9+8)/4 = 11$

$S_{U_1} = (3+6+5+7)/3 = 7$
$S_C = (3+7+6=8)/3 = 8$
$S_D = (6+7+5+9)/3 = 9$
$S_E = (5+6+5+8)/3 = 8$
$S_F = (7+8+9+8)/3 = 10.6$

$S_{U_1} = (3+3+7)/2 = 6.5$
$S_C = (3+4+8)/2 = 7.5$
$S_{U_2} = (3+4+6)/2 = 6.5$
$S_F = (7+8+6)/2 = 10.5$

$S_{U_2} = (2+6)/1 = 8$
$S_{U_3} = (2+6)/1 = 8$
$S_F = (6+6)/1 = 12$

Because $N - 2 = 0$, we cannot do this calculation.

Step 2

Calculate pair with smallest (M), where $M_{ij} = D_{ij} - S_i - S_j$.

Smallest are
$M_{AB} = 5 - 7.5 - 10.5 = -13$
$M_{DE} = 5 - 9.5 - 8.5 = -13$
Choose one of these (AB here).

Smallest is
$M_{CU_1} = 3 - 7 - 8 = -12$
$M_{DE} = 5 - 9 - 8 = -12$
Choose one of these (DE here).

Smallest is
$M_{CU_1} = 3 - 6.5 - 7.5 = -11$

Smallest is
$M_{U_2F} = 6 - 8 - 12 = -14$
$M_{U_3F} = 6 - 8 - 12 = -14$
$M_{U_2U_3} = 2 - 8 - 8 = -14$
Choose one of these ($M_{U_2U_3}$ here).

Step 3

Create a node (U) that joins pair with lowest M_{ij} such that $S_{iU} = D_{ij}/2 + (S_i - S_j)/2$.

U_1 joins A and B:
$S_{AU_1} = D_{AB}/2 + (S_A - S_B)/2 = 1$
$S_{BU_1} = D_{AB}/2 + (S_B - S_A)/2 = 4$

U_2 joins D and E:
$S_{DU_2} = D_{DE}/2 + (S_D - S_E)2 = 3$
$S_{EU_2} = D_{DE}/2 + (S_E - S_D)/2 = 2$

U_3 joins C and U_1:
$S_{CU_3} = D_{CU_1}/2 + (S_C - S_{U_1})/2 = 2$
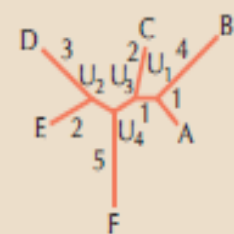$S_{U_1U_3} = D_{CU_1}/2 + (S_{U1} - S_C)/2 = 1$

U_4 joins U_2 and U_3:
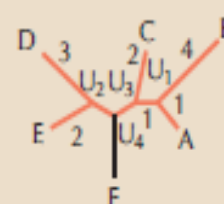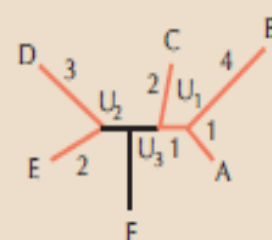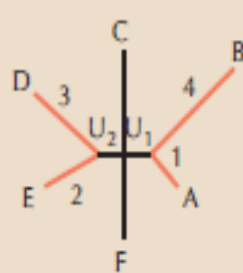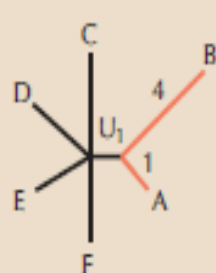$S_{U_2U_4} = D_{U_2U_3}/2 + (S_{U2} - S_{U3})/2 = 1$
$S_{U_3U_4} = D_{U_2U_3}/2 + (S_{U3} - S_{U2})/2 = 1$.

For last pair, connect U_4 and F with branch length = 5.

Step 4

Join i and j according to S above and make all other taxa in form of a star. Branches in black are of unknown length. Branches in red are of known length.

Step 5

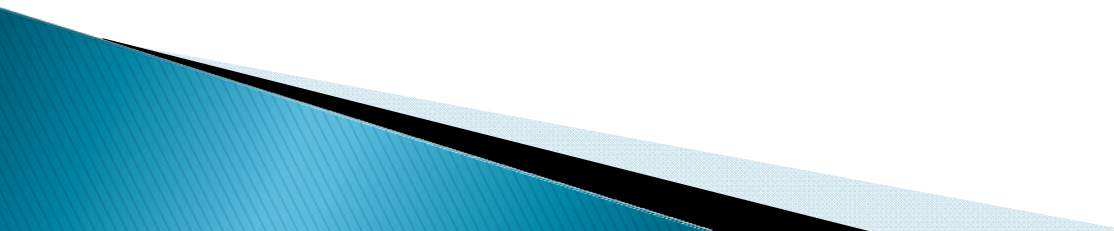Calculate new distance matrix of all other taxa to U with
$D_{xU} = D_{ix} + D_{jx} - D_{ij}$,

Comments

Note this is the same tree we started with (drawn in unrooted form here).

Advantages

- Provides sub-optimal solution that is close to the optimal tree

- Faster tree Building method

- It is a polynomial – time algorithm. So it can also be used for large data sets

- Statistically consistent

Disadvantages

- Cannot find common ancestor as the final output is an unrooted tree

- Tests only a single tree

- May not find true topology since it is a greedy algorithm

- Intermediate ancestors can be ignored (as it does not follow step by step approach)

References

- The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Tree Naruya Saitou2 and Masatoshi Nei ; Mol. Biol. Evol. 4(4):406-425.1987.

- http://mbe.oxfordjournals.org/content/4/4/406.long

- http://en.wikipedia.org/wiki/Neighbor_joining

- http://www.icp.ucl.ac.be/~opperd/private/neighbor.html

- http://www.megasoftware.net/mega3/WebHelp/phylogenetictrees_rh/rh_neighbor_joining_method.htm

THANK YOU