

The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees

Naruya Saitou Masatoshi Nei

4 de dezembro de 2016

Agenda

- 1 Introdução
- 2 Métodos
- 3 Eficiência para recuperação de topologias
- 4 Resultados

Objetivo

Construção de árvores filogenéticas através do princípio de construção da árvore de evolução de topologia mínima ou máxima parsimônica^a.

^aAto ou costume de economizar, de poupar

Métodos comparados

- Distância de Wagner (DW)
- Método de Li
- Modificação do Método de Farris (MF)
- Proximidade^a de vizinhanças pelo método de Sattath and Tversky (ST)

^aneighborliness

Neighbor-Joining

- Semelhante ao método ST, o qual o objetivo é a construção de topologia de árvores
- diferencial: considerar os comprimentos dos ramos da árvore final
- Unidades Taxonômicas Operacionais (OTUs¹ [=neighbors])
- Para cálculo:
 - $N/2$ para N par e $(N-1)/2$ para N ímpar
 - L_{ij} - tamanho do ramo entre i e j
 - $D_{12} = L_{1X} + L_{2X}$ - Distância entre nós folhas

$$S_0 = \sum_{i=1}^N L_{iX} = \frac{1}{N-1} \sum_{i < j} D_{ij} \quad (1)$$

OTUs e topologias

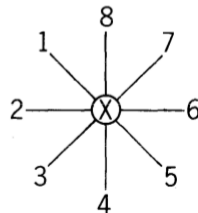
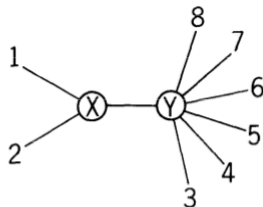
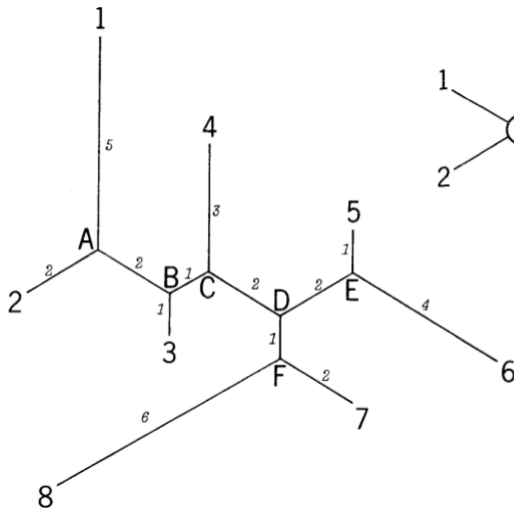
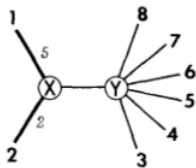


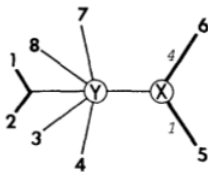
Table 1
Distance Matrix for the Tree in Figure 1

OTU	OTU						
	1	2	3	4	5	6	7
2 ..	7						
3 ..	8	5					
4 ..	11	8	5				
5 ..	13	10	7	8			
6 ..	16	13	10	11	5		
7 ..	13	10	7	8	6	9	
8 ..	17	14	11	12	10	13	8

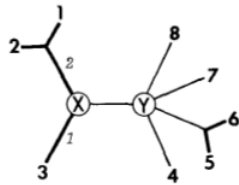
Construção da topologia



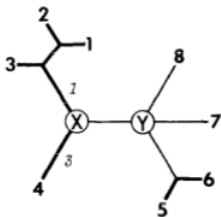
(a)



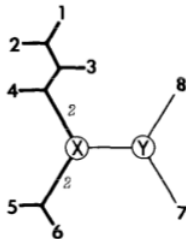
(b)



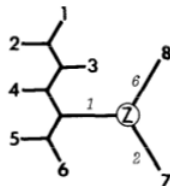
(c)



(d)



(e)



(f)

Critério de Evolução Mínima de Topologia

Assume-se como critério a seguintes condições:

- $D_{12} + D_{ij} < D_{1i} + D_{2j}$

Assim, considerando este critério, a equação 1 poderá conter a menor evolução de topologias.

Modelos de árvores

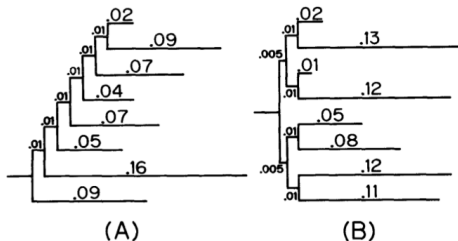
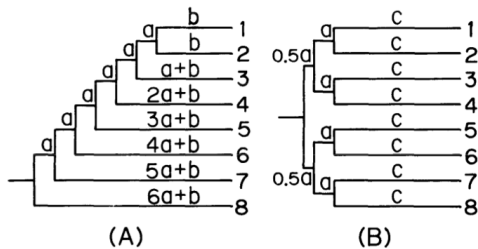


Tabela comparativas de nucleotídeos

Table 3

P_c and d_T (in parentheses) for Six Tree-making Methods for the Case of $a = 0.01$, $b = 0.04$, and $c = 0.07$

METHOD	MODEL TREE A ^a			MODEL TREE B ^a		
	300	600	900	300	600	900
UPGMA:						
p^b	14 (3.18)	36 (1.72)	58 (0.98)	14 (4.54)	36 (2.74)	51 (1.68)
d^c	15 (3.18)	34 (1.74)	56 (1.04)	13 (4.56)	35 (2.70)	52 (1.60)
MF:						
p	39 (1.76)	73 (0.58)	95 (0.10)	24 (2.86)	51 (1.30)	67 (0.76)
d	38 (1.92)	72 (0.62)	95 (0.10)	19 (2.94)	48 (1.42)	64 (0.86)
DW:						
p	42 (1.70)	75 (0.54)	96 (0.08)	26 (2.36)	55 (1.12)	79 (0.48)
d	37 (1.74)	74 (0.58)	95 (0.10)	28 (2.36)	58 (1.06)	79 (0.46)
LI:						
p	41 (1.58)	71 (0.70)	94 (0.12)	40 (2.04)	70 (0.78)	90 (0.22)
d	36 (1.84)	66 (0.82)	89 (0.24)	39 (2.10)	70 (0.78)	90 (0.26)
ST:						
p	48 (1.26)	75 (0.54)	97 (0.06)	45 (1.66)	75 (0.62)	91 (0.22)
d	44 (1.48)	70 (0.62)	96 (0.08)	43 (1.62)	74 (0.64)	91 (0.22)
NJ:						
p	48 (1.36)	76 (0.54)	97 (0.06)	46 (1.64)	76 (0.60)	91 (0.20)
d	41 (1.60)	70 (0.62)	96 (0.08)	45 (1.62)	75 (0.60)	91 (0.20)

^a As shown in fig. 6.

^b Trees reconstructed from data on the proportion of different nucleotides between the sequences compared.

^c Trees reconstructed from the Jukes-Cantor distance.

Após as comparações

Dependendo dos valores dos parâmetros p e d

$$MF \ll UPGMA \approx DW \ll Li < ST < NJ \quad (2)$$

$p \rightarrow$ proporção da diferença entre os nucleotídeos entre duas sequências

$d \rightarrow$ distância de Jukes-Cantor