

Conceitos fundamentais da aula anterior:

Enraizamento: propriedades operacionais e biológicas

Termos associados a grupos:

Redefinição de grupos monofiléticos

Grupos-irmãos

Grupos-externos e internos

Termos associados aos estados de caráter:

apomorfias

plesiomorfias

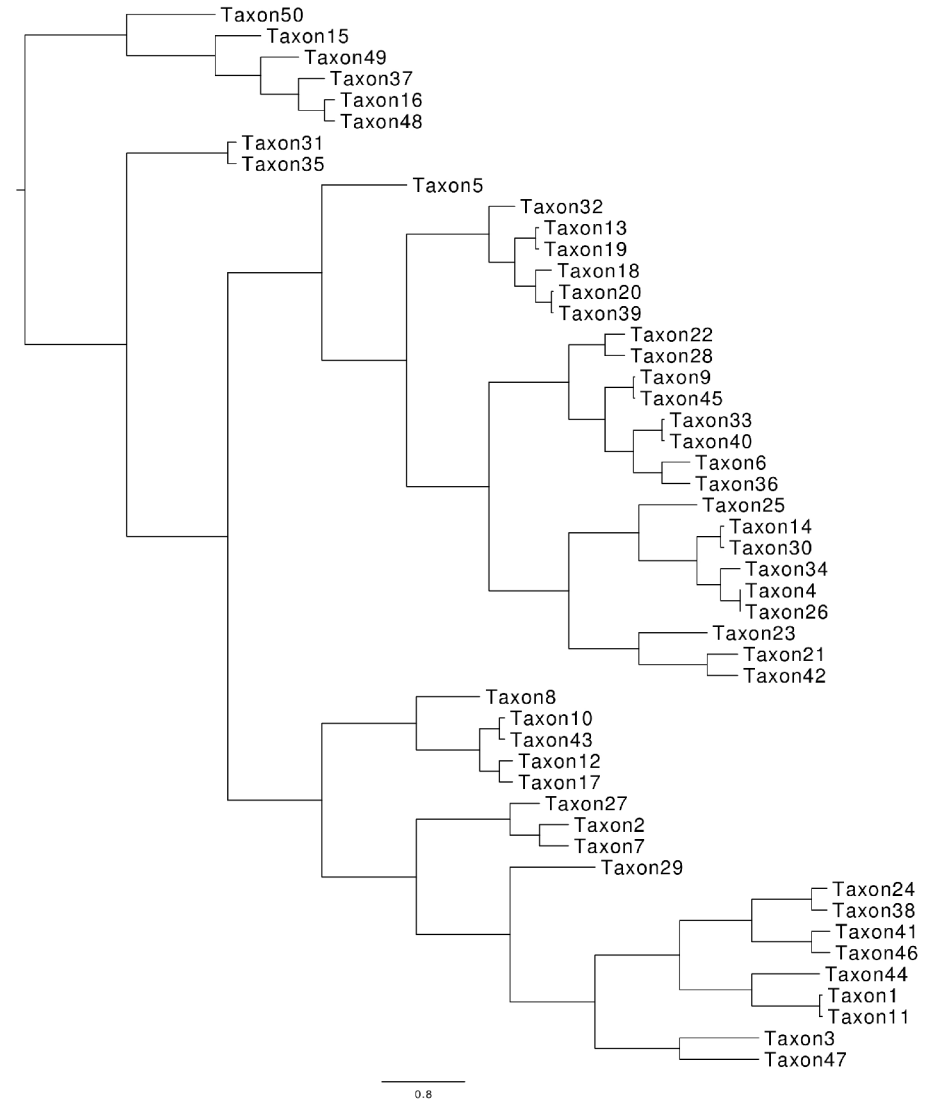
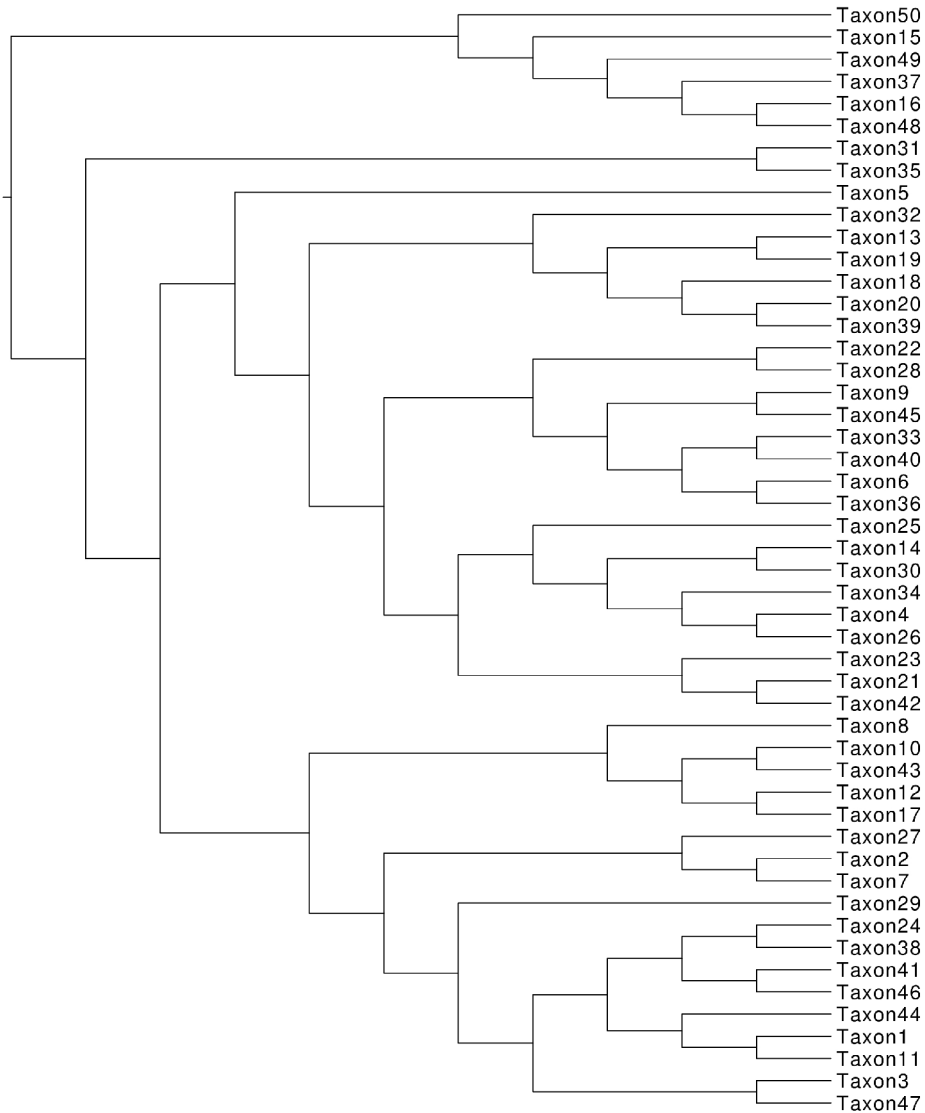
Sinapomorfias

Simplesiomorfias

Autapomorfias

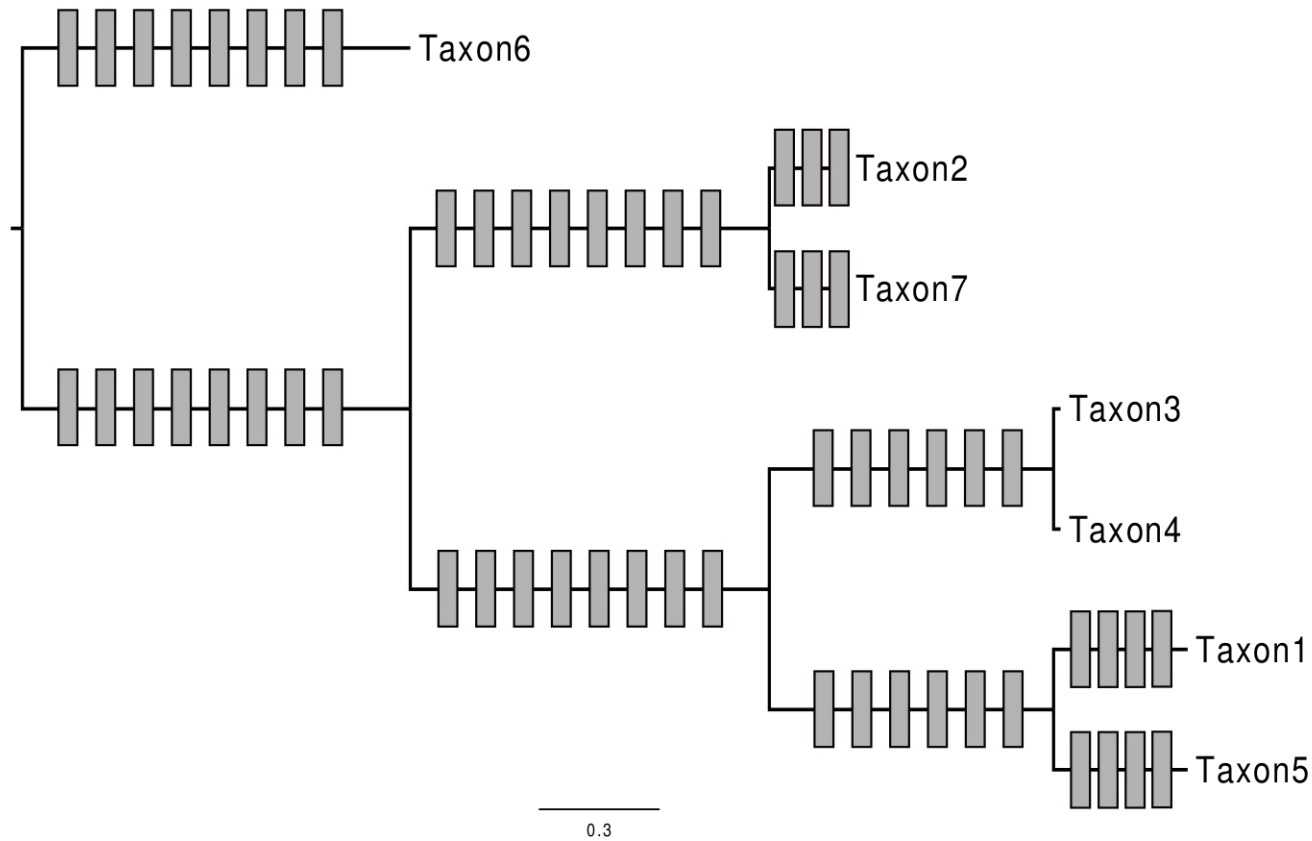
Leitura de cladograma:

Há alguma diferença quanto ao conteúdo informativo entre estes dois diagramas?



Leitura de cladograma:

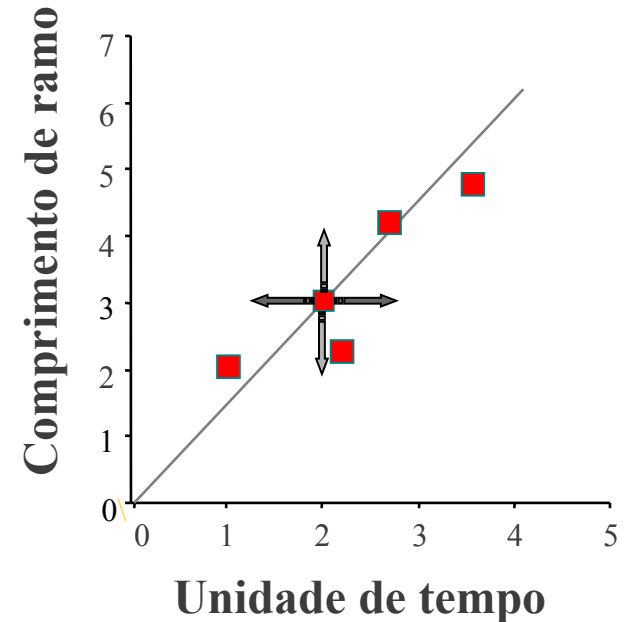
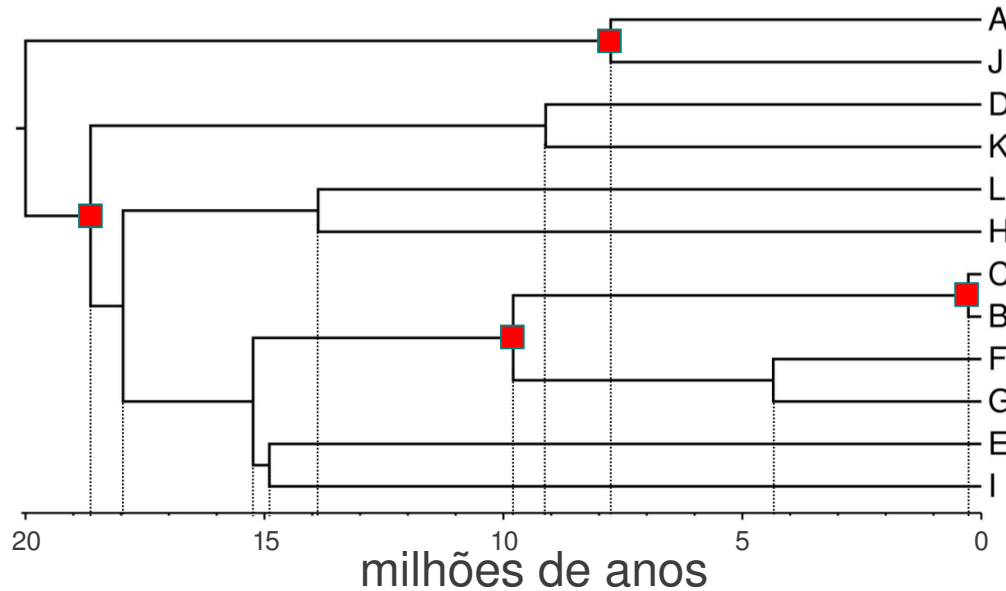
Há alguma diferença quanto ao conteúdo informativo entre estes dois diagramas?



Sim. O diagrama (cladograma) ao lado exibe os comprimentos de ramos. O comprimento dos ramos é proporcional ao número de transformações e nos fornecem uma medida de divergência relativa entre as linhagens.

Leitura de cladograma:

Estimativas de tempo absoluto

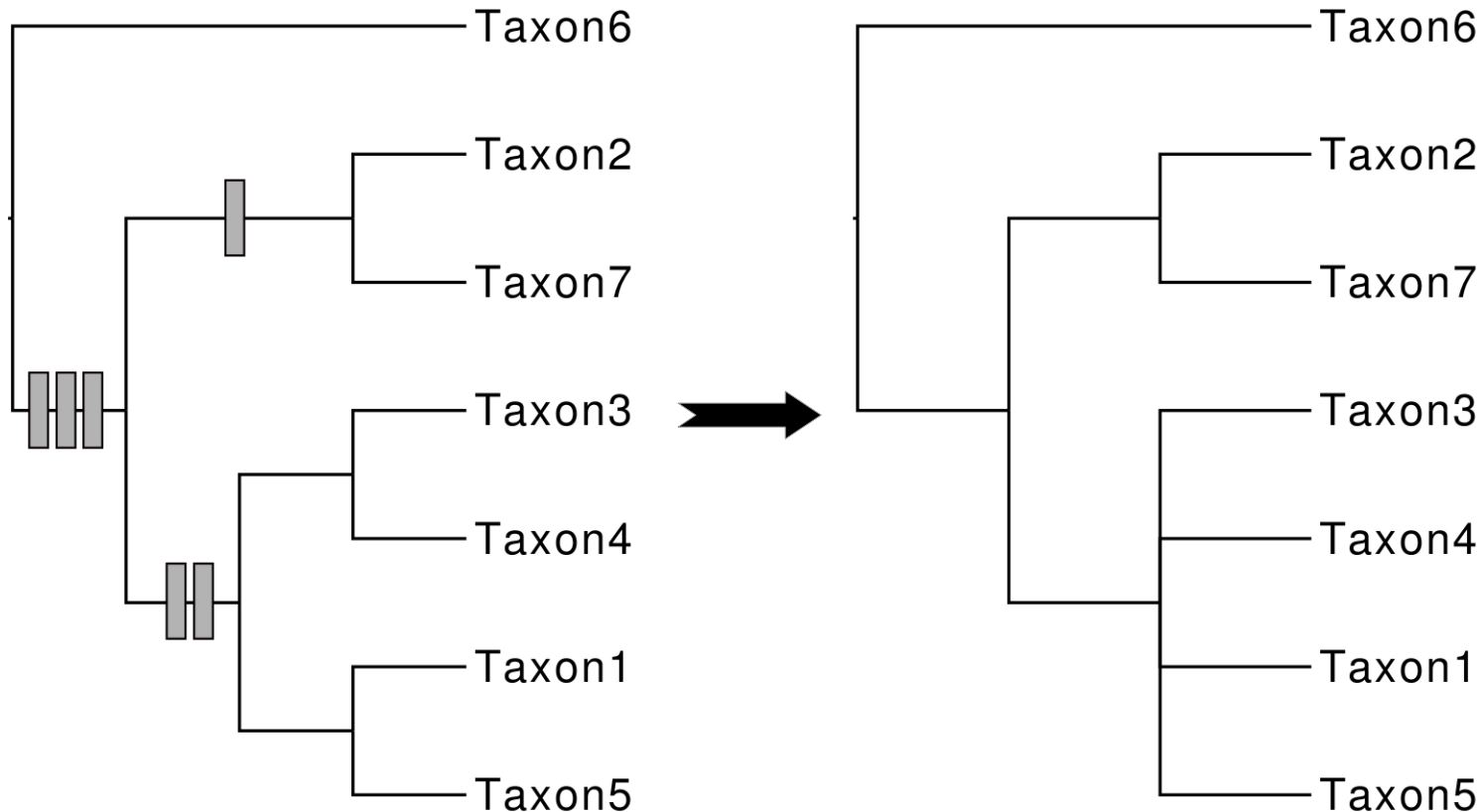


Premissas de relógios moleculares:

1. Taxas constantes de substituições (função linear) ao longo da evolução do grupo.
2. Taxa de substituição é idêntica ao longo de todas as posições e entre todas as linhagens.
3. A hipótese filogenética é construída sem nenhum erro.
4. O número de substituições ao longo de cada linhagem é reconstruída sem erro (comprimento de ramo)
5. Datas de calibração são exatas.
6. A função de regressão (comprimento do ramo/tempo) é conduzida sem erro.

Leitura de cladograma:

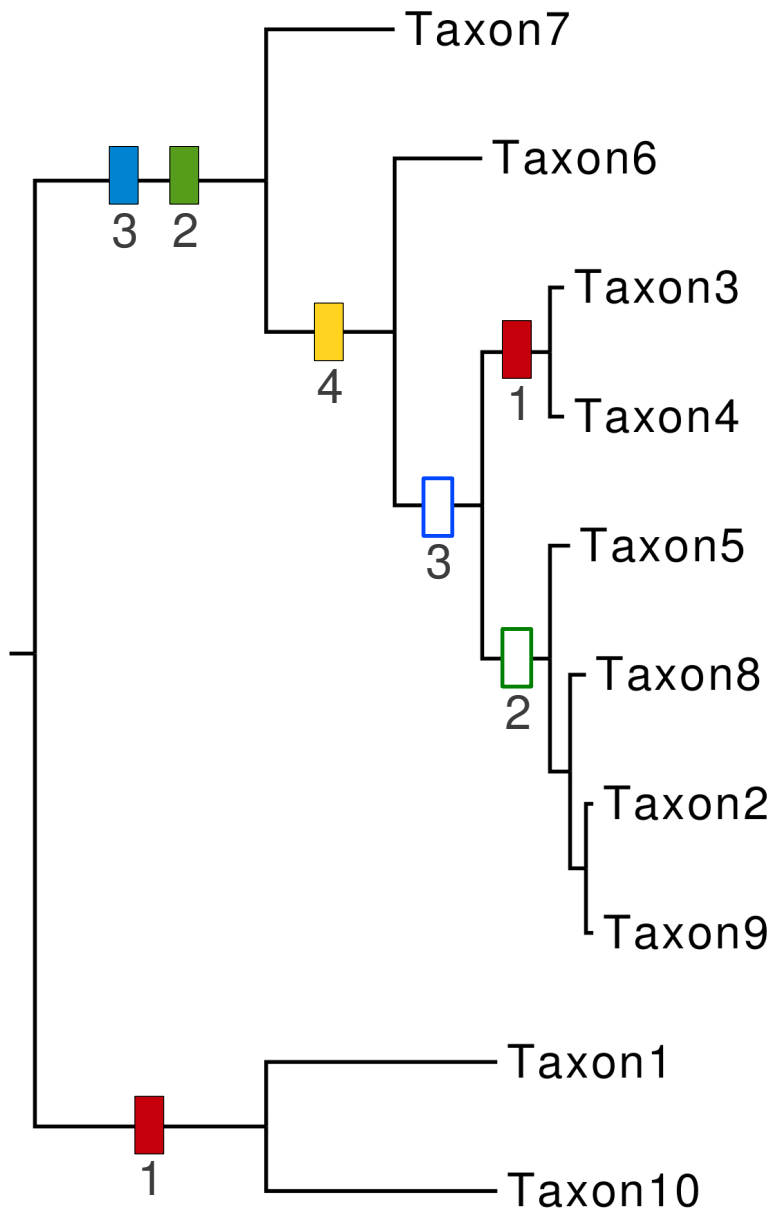
Suporte e resolução



1. Não existe suporte (evidência) para os grupos: (Taxon2, Taxon4) e (Taxon1, Taxon5).
2. Não há resolução para o nó compartilhado pelos terminais: Taxon2, Taxon4, Taxon1 e Taxon5.
3. A resolução de um nó depende da existência de pelo menos uma transformação.

Leitura de cladograma:

Suporte e resolução

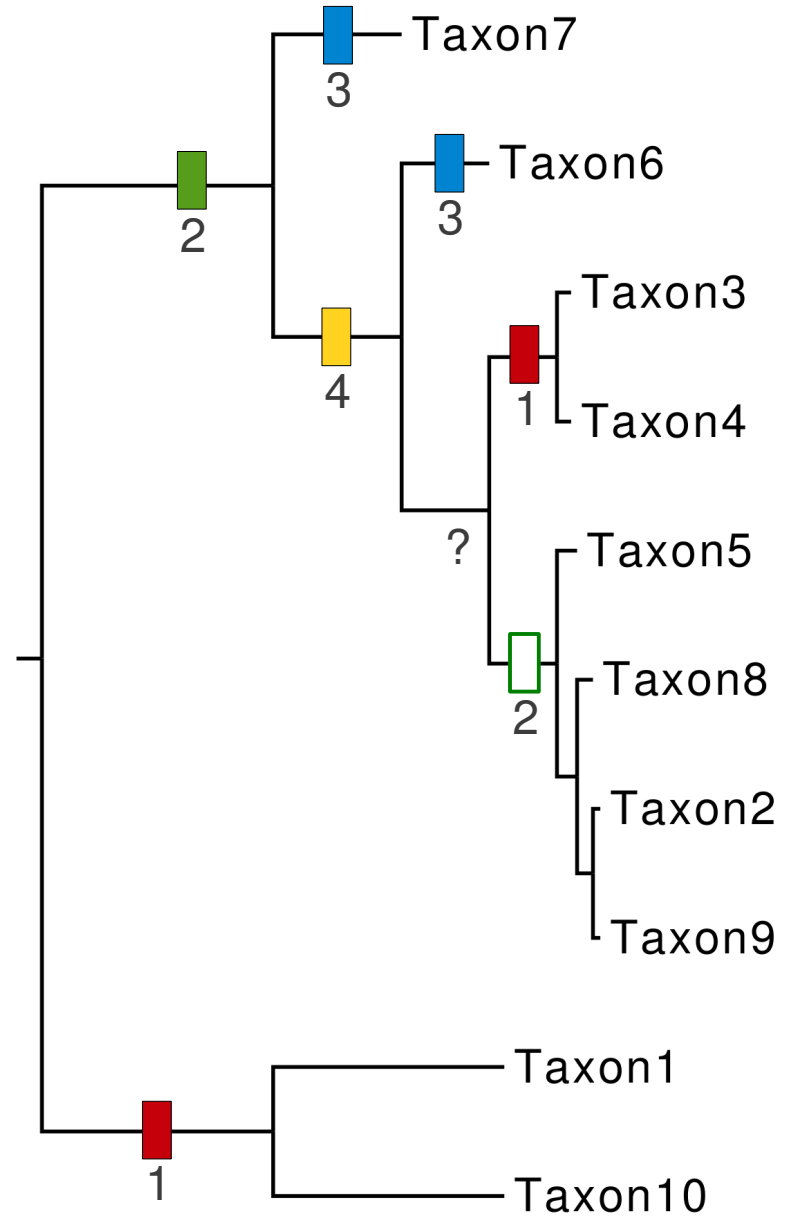
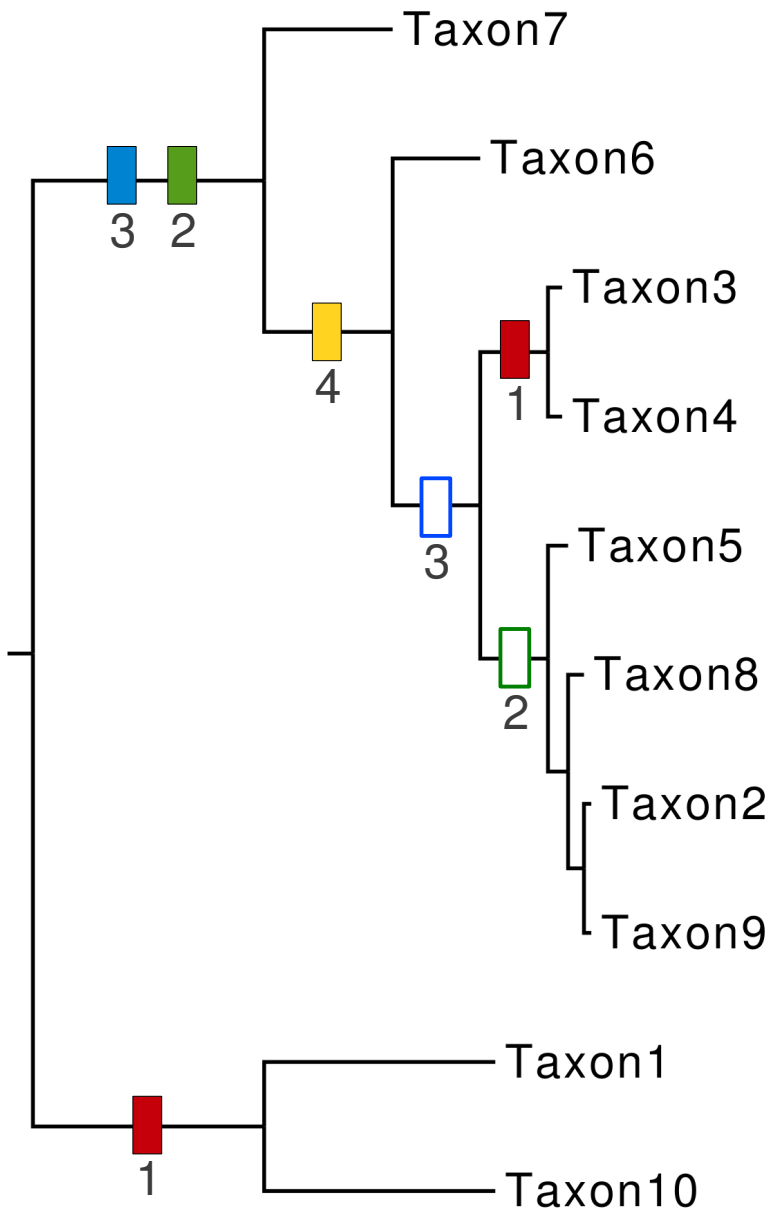


Considere os nós que estão sendo sustentados por transformações e responda:

Existe algum nó com suporte ambíguo neste cladograma?

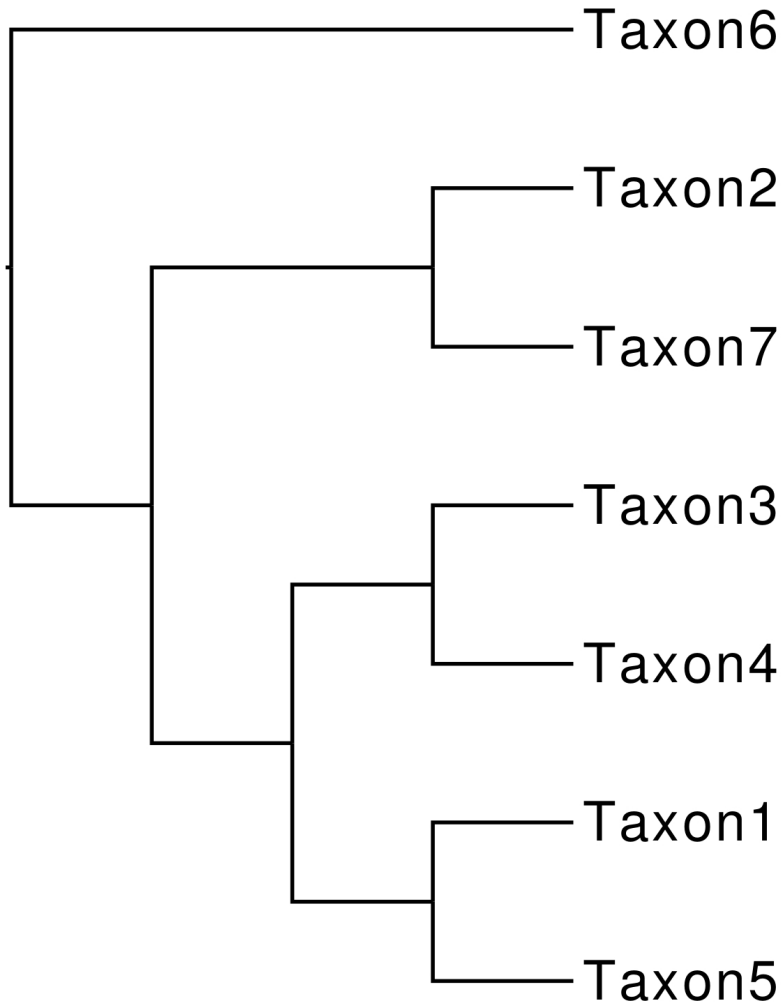
Leitura de cladograma:

Suporte e resolução



Leitura de cladograma:

Descrição das relações de parentesco
Considere:



O grupo monofilético Taxon2+Taxon7 é grupo-irmão do clado formado por Taxon3, Taxon4, Taxon1 e Taxon5. Dentro deste clado, Taxon3 e Taxon4 compartilham o mesmo nó, ao passo que Taxon1 e Taxon5 compartilham um ancestral comum exclusivo. Todos estes terminais formam o clado que é grupo-irmão de Taxon6.

Leitura de cladograma:

Descrição das relações de parentesco

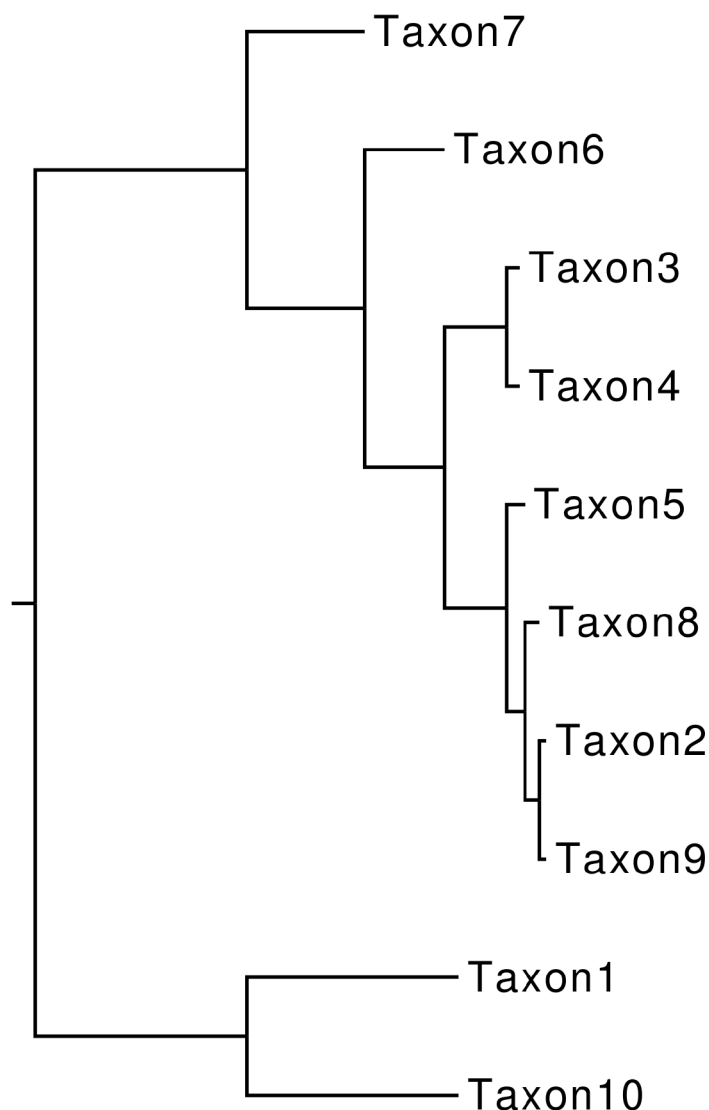
Considere a descrição abaixo e desenhe o cladograma:

O cladograma com 10 terminais (T1-T10) contém dois grupos monofiléticos: um formado por T1 e T10 e o outro formado pelos demais. Dentro deste grande clado, T3 é grupo-irmão de T4, clado este que compartilha o mesmo nó que o grupo monofilético formado por T5, T8, T2 e T9. A resolução destes quatro terminais sugere que T2 e T9 e grupo-irmão de T8 e estes três são grupo-irmão de T5. Finalmente, T6 é grupo-irmão destes 6 terminais formando um clado grupo-irmão de T7.

Leitura de cladograma:

Descrição das relações de parentesco

Considere a descrição abaixo e desenhe o cladograma:

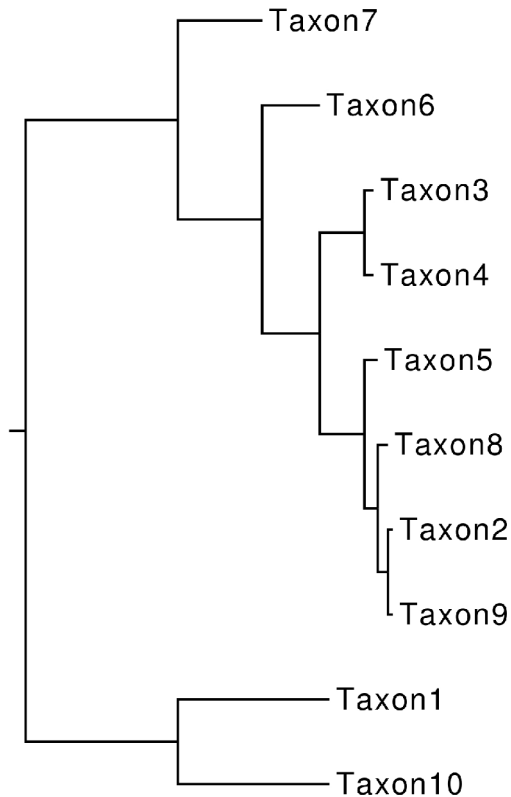


O cladograma com 10 terminais (T1-T10) contém dois grupos monofiléticos: um formado por T1 e T10 e o outro formado pelos demais. Dentro deste grande clado, T3 é grupo-irmão de T4, clado este que compartilha o mesmo nó que o grupo monofilético formado por T5, T8, T2 e T9. A resolução destes quatro terminais sugere que T2 e T9 e grupo-irmão de T8 e estes três são grupo-irmão de T5. Finalmente, T6 é grupo-irmão destes 6 terminais formando um clado grupo-irmão de T7.

Leitura de cladograma:

Descrição das relações de parentesco

Considere a descrição abaixo e desenhe o cladograma:



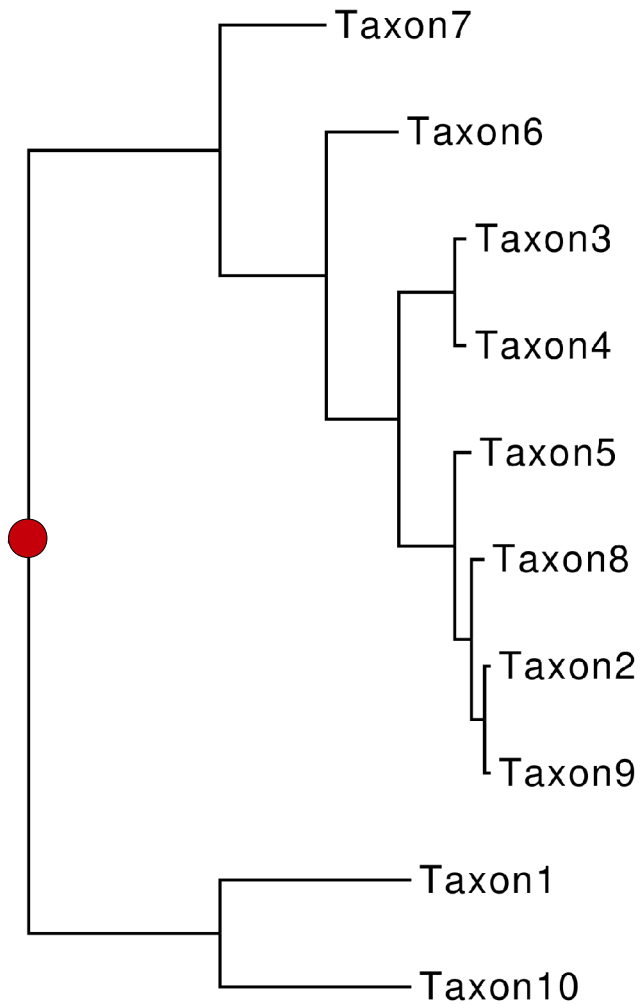
O cladograma com 10 terminais (T1-T10) contém dois grupos monofiléticos: um formado por T1 e T10 e o outro formado pelos demais. Dentro deste grande clado, T3 é grupo-irmão de T4, clado este que compartilha o mesmo nó que o grupo monofilético formado por T5, T8, T2 e T9. A resolução destes quatro terminais sugere que T2 e T9 e grupo-irmão de T8 e estes três são grupo-irmão de T5. Finalmente, T6 é grupo-irmão destes 6 terminais formando um clado grupo-irmão de T7.

Dito de outra maneira, a relação destes terminais pode ser expressa da seguinte forma:

$((T10, T1), (T7, (T6, ((T3, T4), (T5, (T8, (T2, T9))))))))$

Leitura de cladograma:

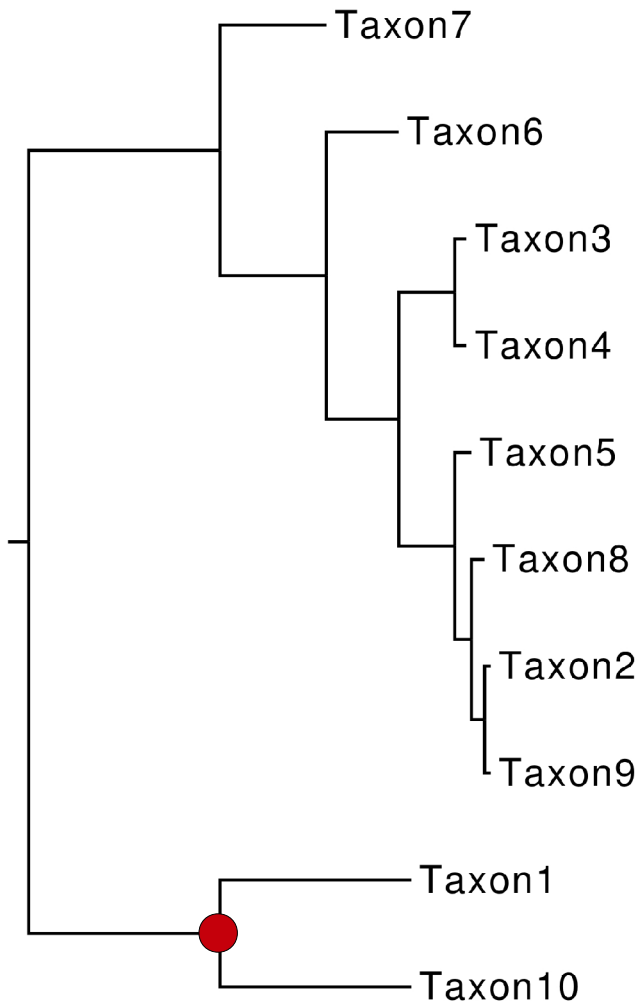
Descrição das relações de parentesco
Notação parentética



(T10,T1,T7,T6,T3,T4,T5,T8,T2,T9)

Leitura de cladograma:

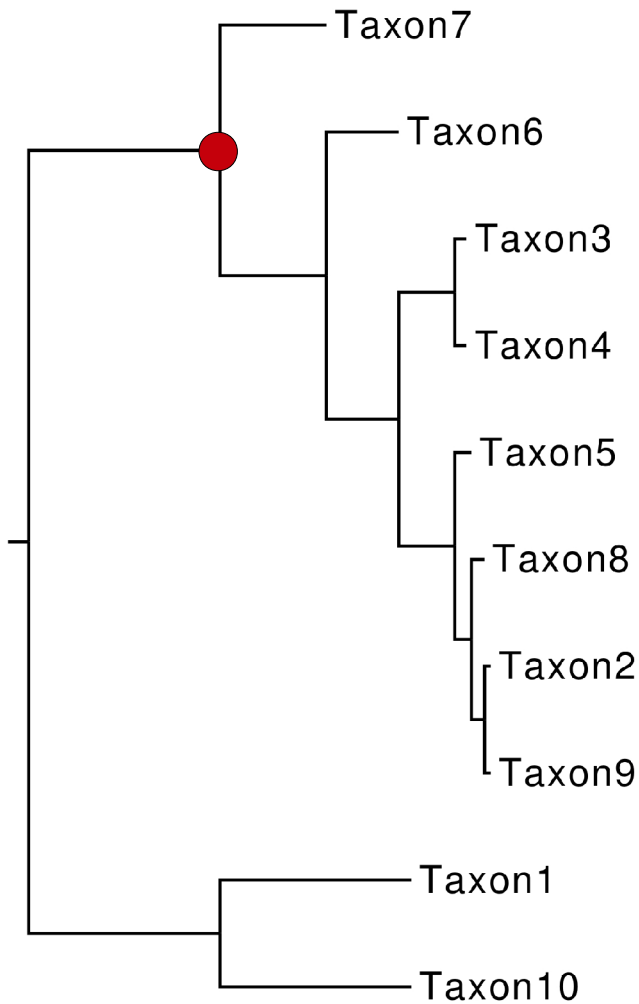
Descrição das relações de parentesco
Notação parentética



(T10,T1,T7,T6,T3,T4,T5,T8,T2,T9)
((T10,T1),T7,T6,T3,T4,T5,T8,T2,T9)

Leitura de cladograma:

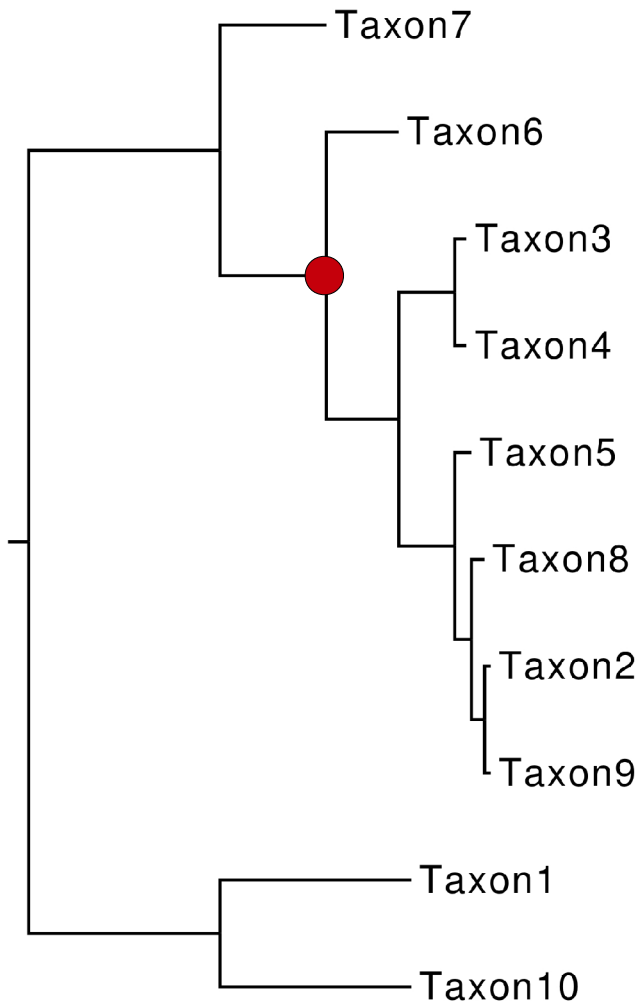
Descrição das relações de parentesco
Notação parentética



(T10,T1,T7,T6,T3,T4,T5,T8,T2,T9)
((T10,T1),T7,T6,T3,T4,T5,T8,T2,T9)
((T10,T1),(T7,T6,T3,T4,T5,T8,T2,T9))

Leitura de cladograma:

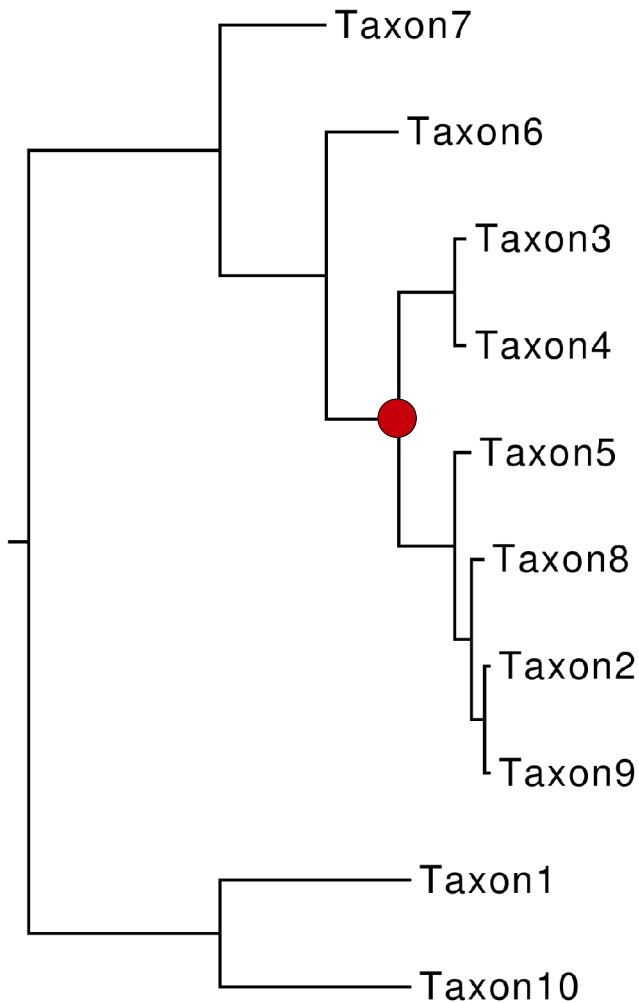
Descrição das relações de parentesco
Notação parentética



(T10,T1,T7,T6,T3,T4,T5,T8,T2,T9)
((T10,T1),T7,T6,T3,T4,T5,T8,T2,T9)
((T10,T1),(T7,T6,T3,T4,T5,T8,T2,T9))
((T10,T1),(T7,(T6,T3,T4,T5,T8,T2,T9)))

Leitura de cladograma:

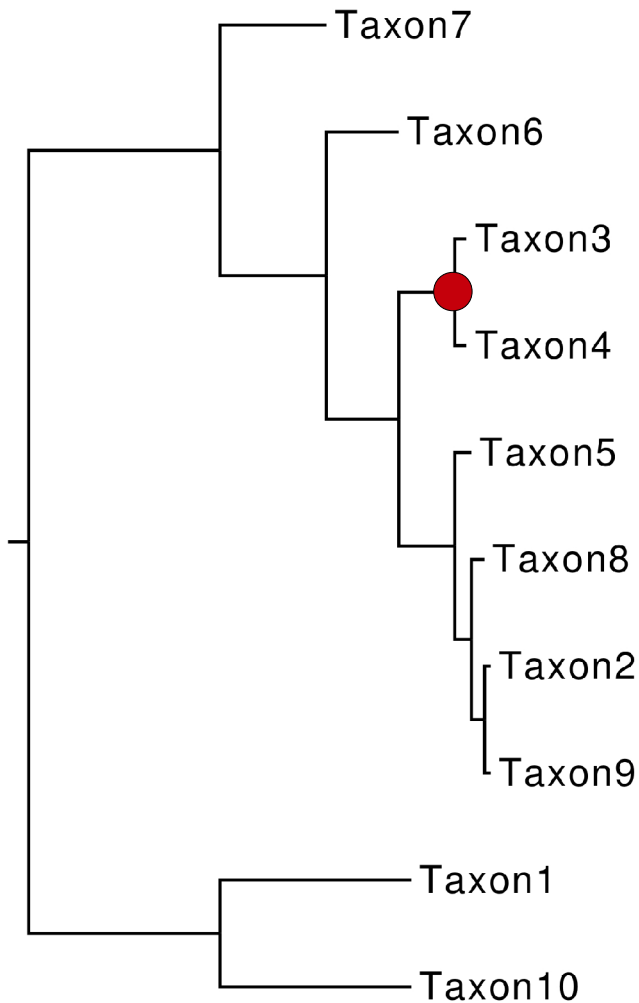
Descrição das relações de parentesco
Notação parentética



(T10,T1,T7,T6,T3,T4,T5,T8,T2,T9)
((T10,T1),T7,T6,T3,T4,T5,T8,T2,T9)
((T10,T1),(T7,T6,T3,T4,T5,T8,T2,T9))
((T10,T1),(T7,(T6,T3,T4,T5,T8,T2,T9)))
((T10,T1),(T7,(T6,(T3,T4,T5,T8,T2,T9))))

Leitura de cladograma:

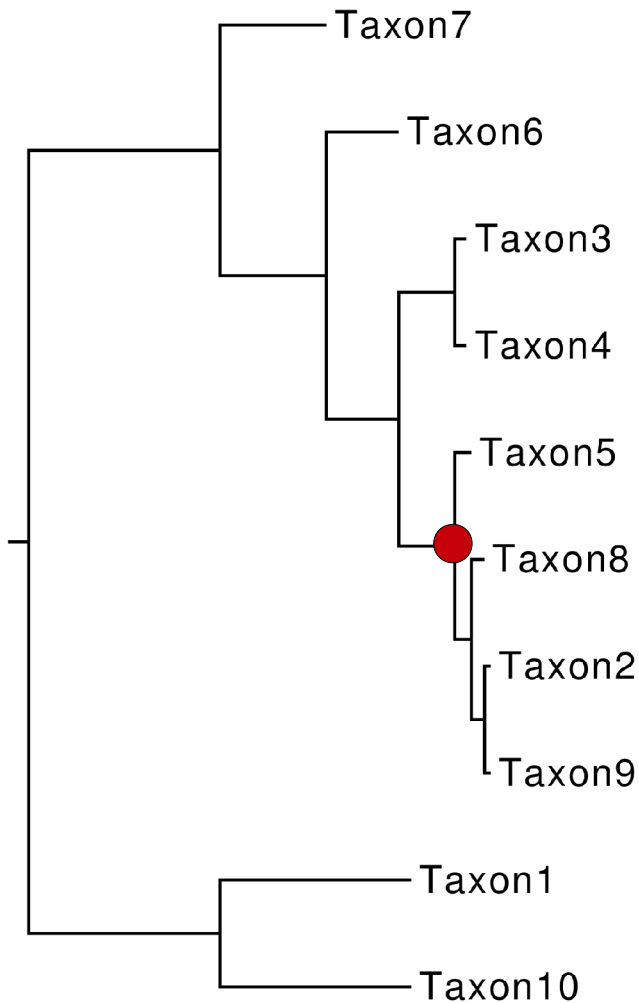
Descrição das relações de parentesco
Notação parentética



(T10,T1,T7,T6,T3,T4,T5,T8,T2,T9)
((T10,T1),T7,T6,T3,T4,T5,T8,T2,T9)
((T10,T1),(T7,T6,T3,T4,T5,T8,T2,T9))
((T10,T1),(T7,(T6,T3,T4,T5,T8,T2,T9)))
((T10,T1),(T7,(T6,(T3,T4,T5,T8,T2,T9))))
((T10,T1),(T7,(T6,((T3,T4),T5,T8,T2,T9))))

Leitura de cladograma:

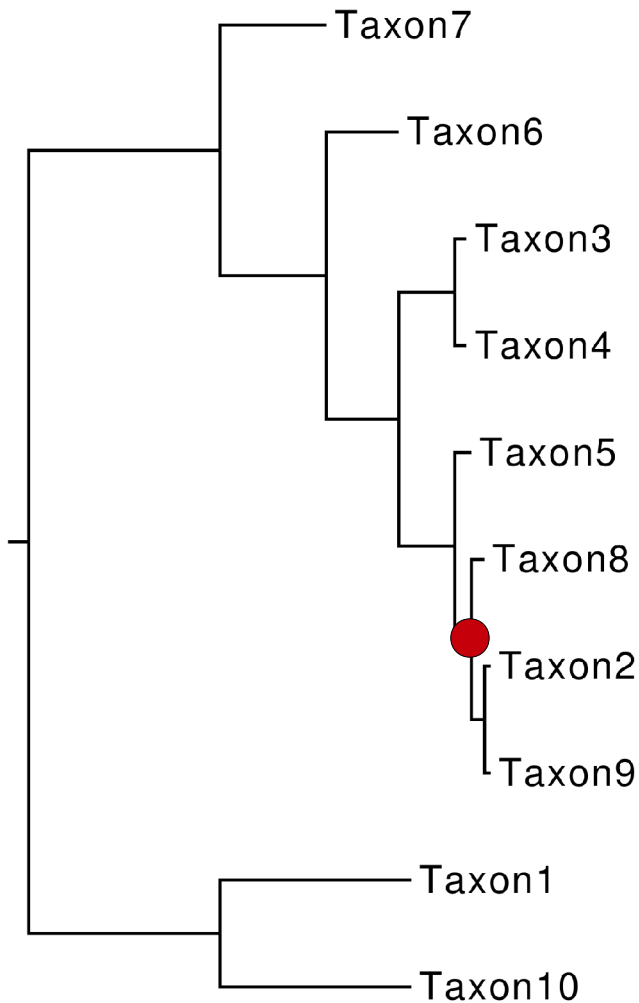
Descrição das relações de parentesco
Notação parentética



(T10,T1,T7,T6,T3,T4,T5,T8,T2,T9)
((T10,T1),T7,T6,T3,T4,T5,T8,T2,T9)
((T10,T1),(T7,T6,T3,T4,T5,T8,T2,T9))
((T10,T1),(T7,(T6,T3,T4,T5,T8,T2,T9)))
((T10,T1),(T7,(T6,(T3,T4,T5,T8,T2,T9))))
((T10,T1),(T7,(T6,((T3,T4),T5,T8,T2,T9))))
((T10,T1),(T7,(T6,((T3,T4),(T5,T8,T2,T9)))))

Leitura de cladograma:

Descrição das relações de parentesco
Notação parentética

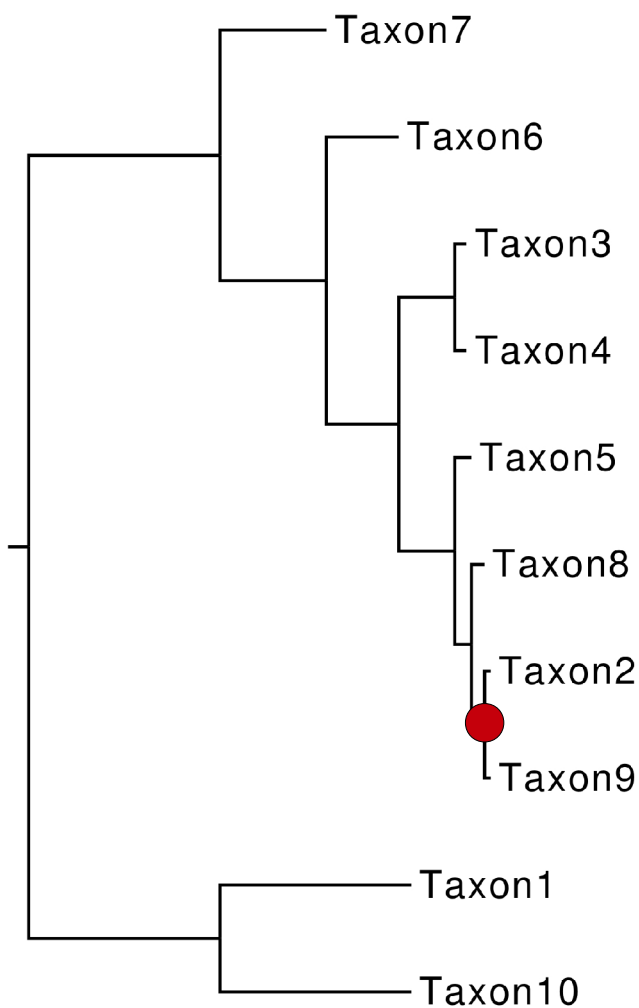


(T10,T1,T7,T6,T3,T4,T5,T8,T2,T9)
((T10,T1),T7,T6,T3,T4,T5,T8,T2,T9)
((T10,T1),(T7,T6,T3,T4,T5,T8,T2,T9))
((T10,T1),(T7,(T6,T3,T4,T5,T8,T2,T9)))
((T10,T1),(T7,(T6,(T3,T4,T5,T8,T2,T9))))
((T10,T1),(T7,(T6,((T3,T4),T5,T8,T2,T9))))
((T10,T1),(T7,(T6,((T3,T4),(T5,T8,T2,T9))))))
((T10,T1),(T7,(T6,((T3,T4),(T5,(T8,T2,T9))))))

Leitura de cladograma:

Descrição das relações de parentesco

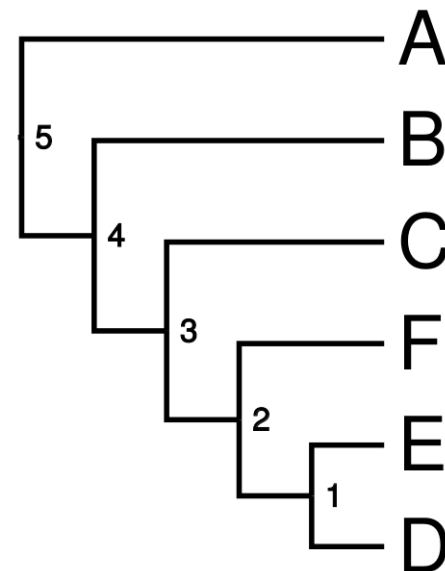
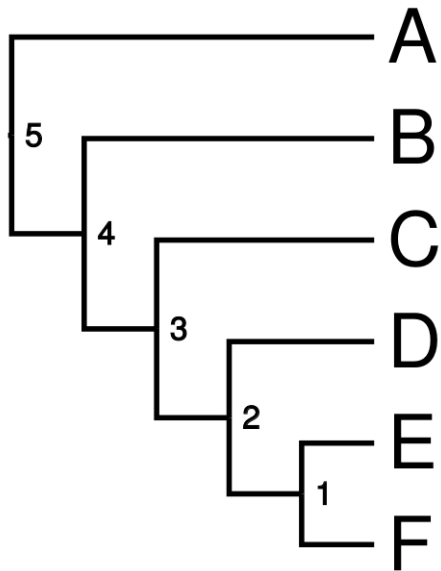
Notação parentética



(T10,T1,T7,T6,T3,T4,T5,T8,T2,T9)
((T10,T1),T7,T6,T3,T4,T5,T8,T2,T9)
((T10,T1),(T7,T6,T3,T4,T5,T8,T2,T9))
((T10,T1),(T7,(T6,T3,T4,T5,T8,T2,T9)))
((T10,T1),(T7,(T6,(T3,T4,T5,T8,T2,T9))))
((T10,T1),(T7,(T6,((T3,T4),T5,T8,T2,T9))))
((T10,T1),(T7,(T6,((T3,T4),(T5,T8,T2,T9))))))
((T10,T1),(T7,(T6,((T3,T4),(T5,(T8,T2,T9))))))
((T10,T1),(T7,(T6,((T3,T4),(T5,(T8,(T2,T9))))))))

Consenso

Considere os seguintes cladogramas fundamentais:



Qual é o conteúdo informativo comum destas topologias?

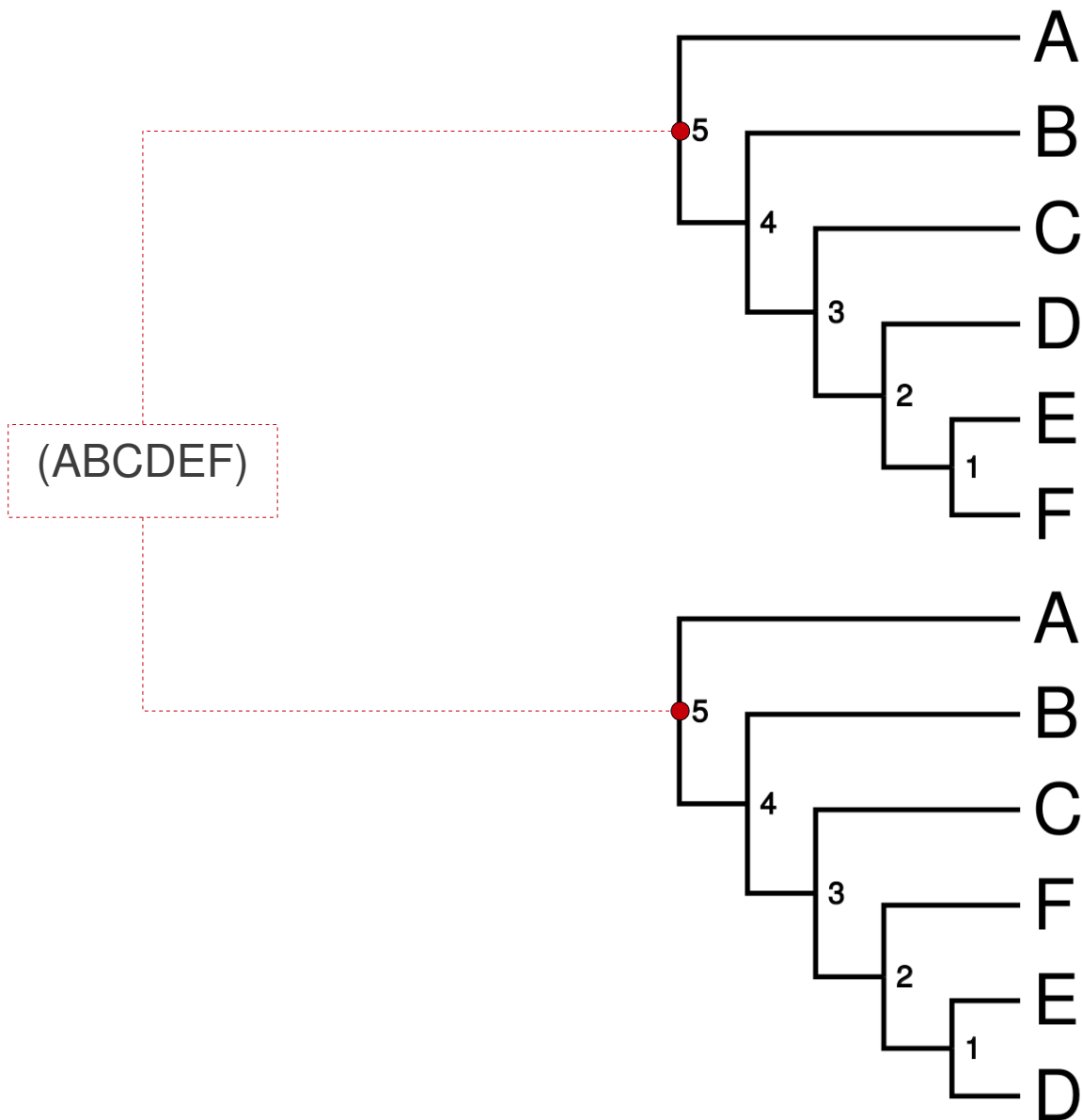
Quais componentes são comuns entre estas duas topologias?

Consenso

Considere os seguintes cladogramas fundamentais:

Componentes
comuns:

(ABCDEF)



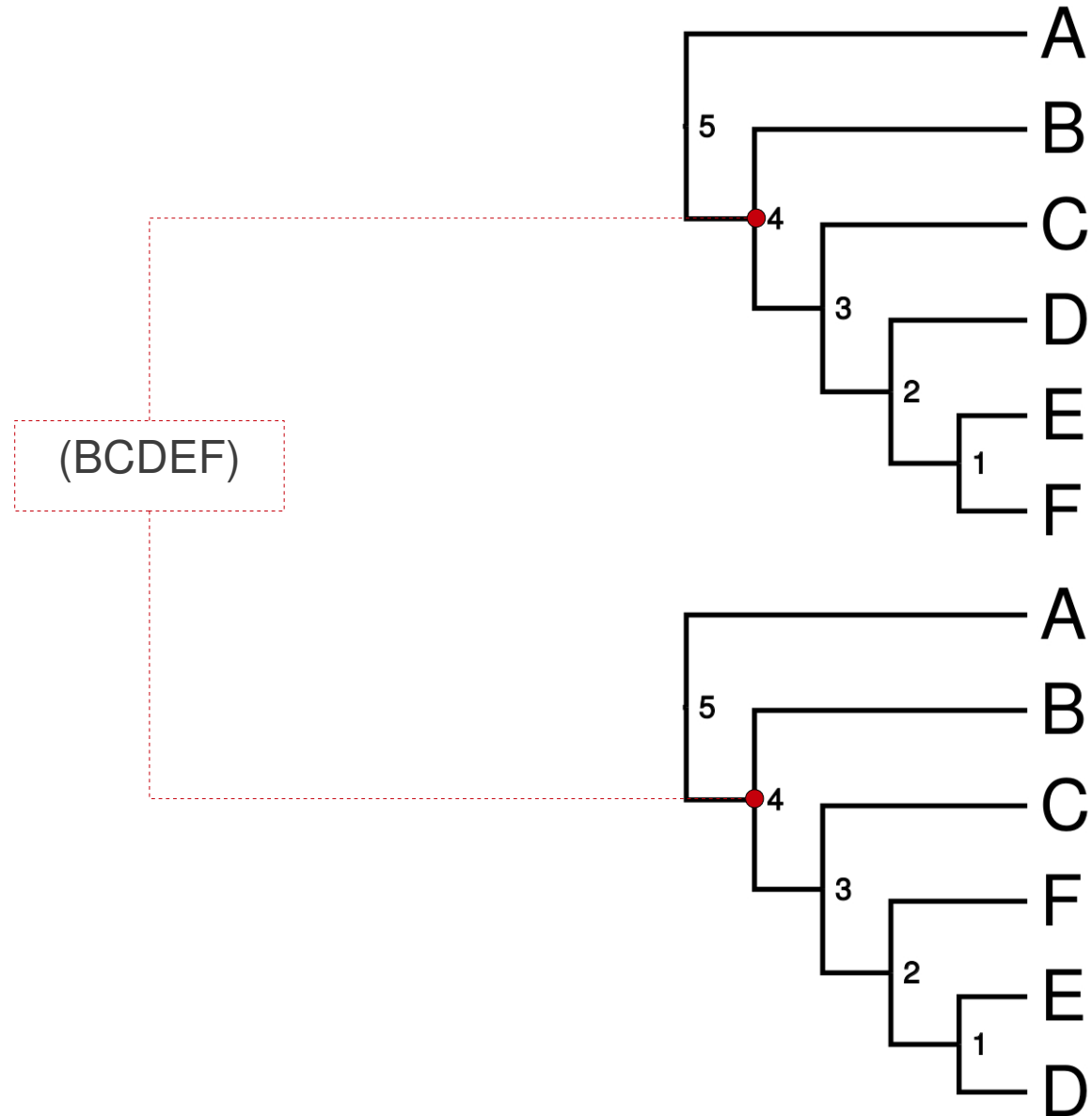
Consenso

Considere os seguintes cladogramas fundamentais:

Componentes
comuns:

(ABCDEF)
(BCDEF)

(BCDEF)



Consenso

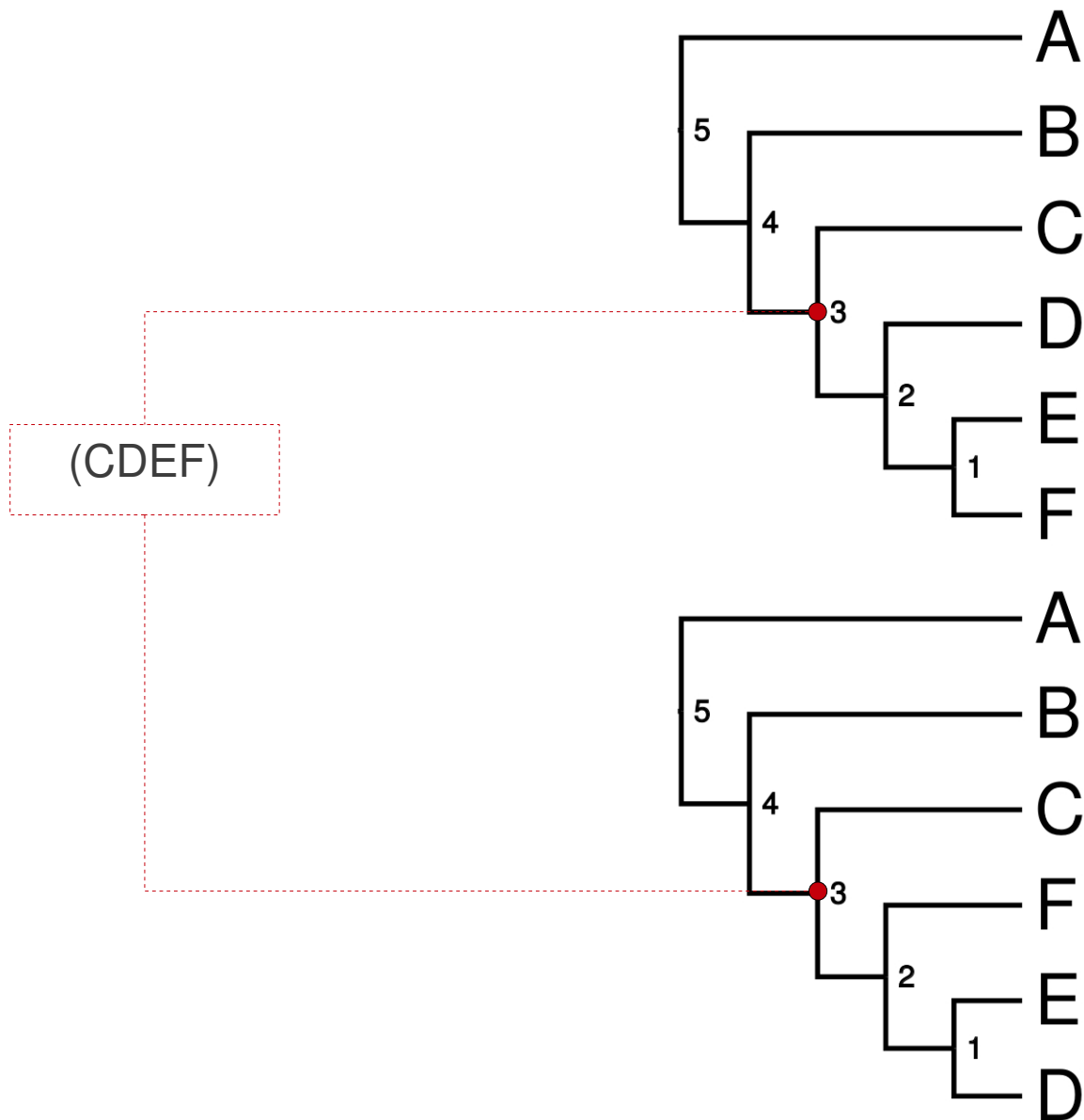
Considere os seguintes cladogramas fundamentais:

Componentes
comuns:

(ABCDEF)

(BCDEF)

(CDEF)

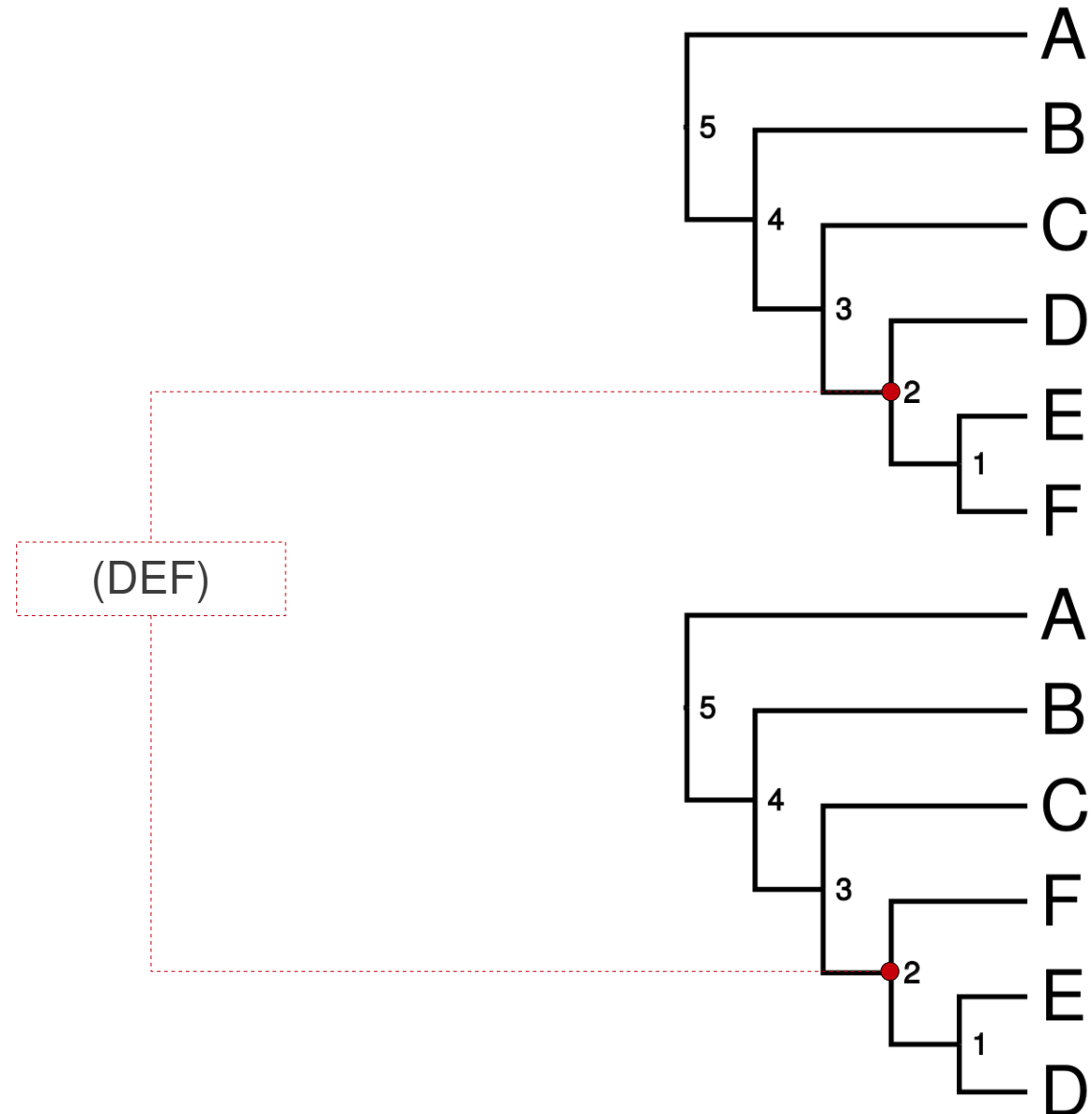


Consenso

Considere os seguintes cladogramas fundamentais:

Componentes
comuns:

- (ABCDEF)
- (BCDEF)
- (CDEF)
- (DEF)

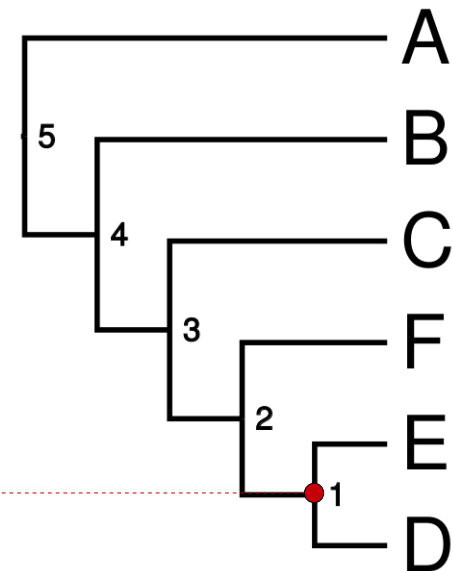
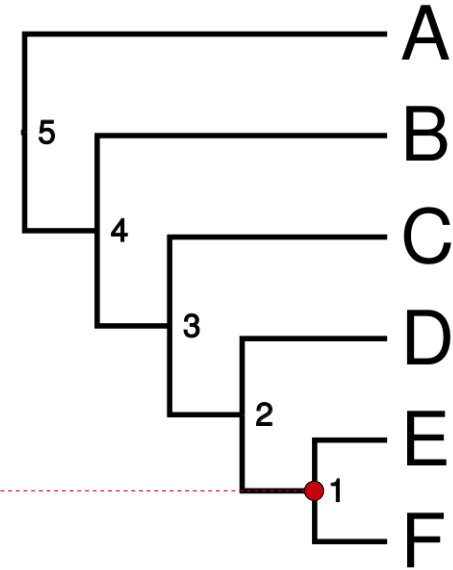
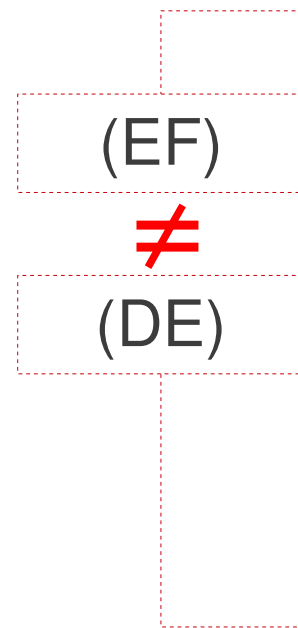


Consenso

Considere os seguintes cladogramas fundamentais:

Componentes
comuns:

(ABCDEF)
(BCDEF)
(CDEF)
(DEF)



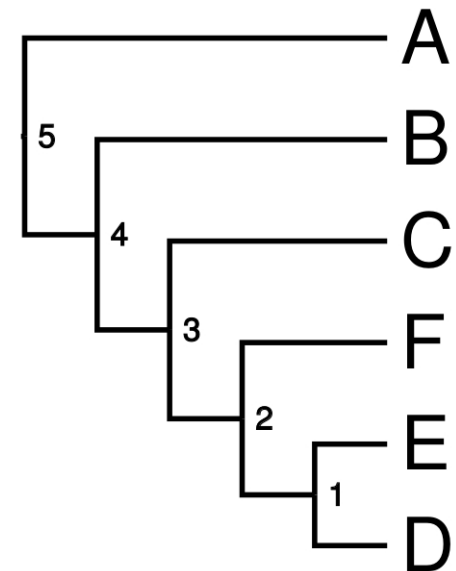
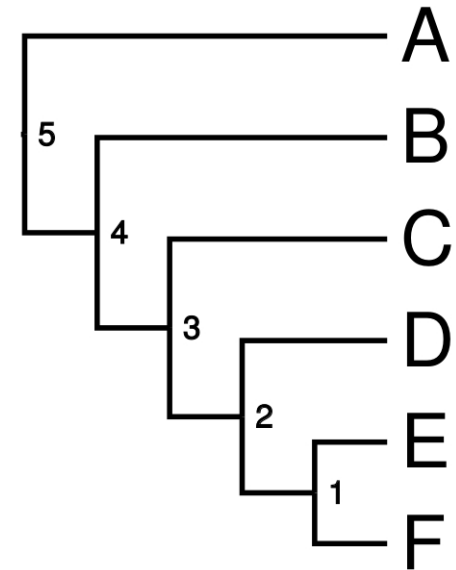
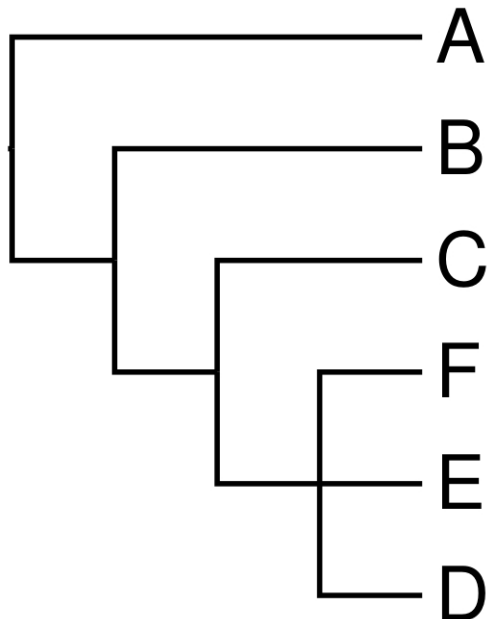
Consenso

Considere os seguintes cladogramas fundamentais:

Componentes
comuns:

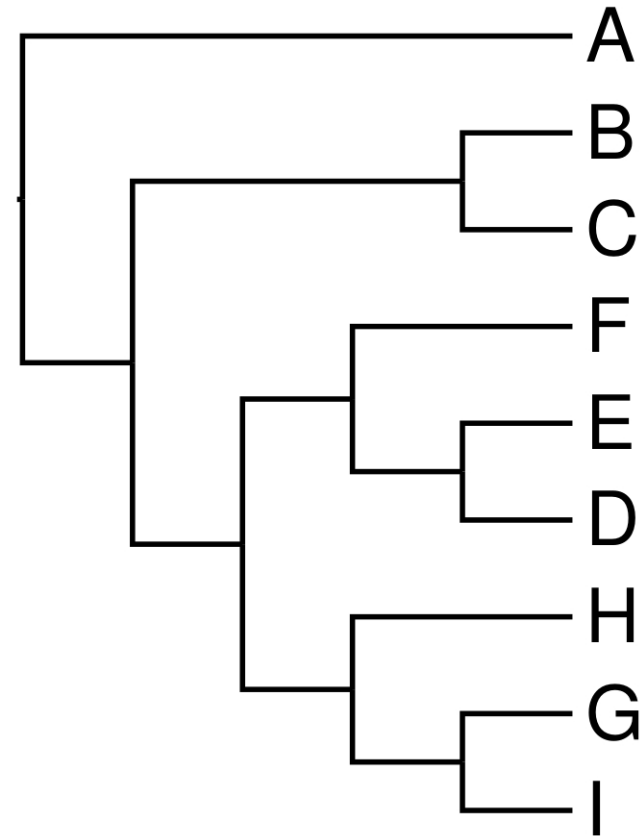
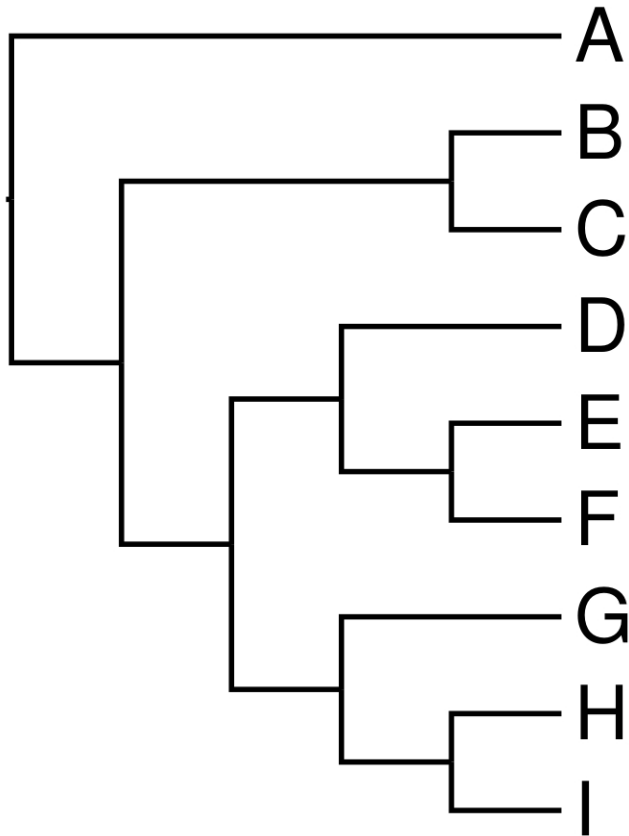
(ABCDEF)
(BCDEF)
(CDEF)
(DEF)

Consenso estrito:



Consenso

Exemplo 1

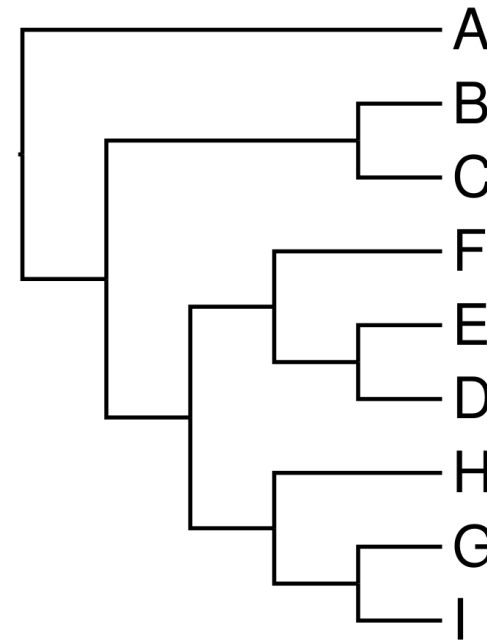
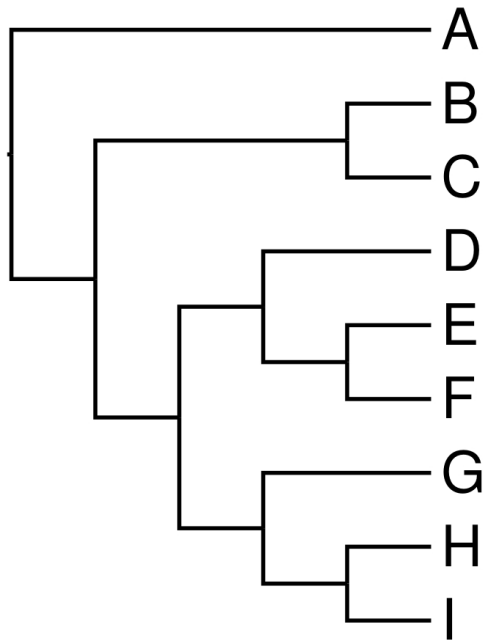


Quais componentes são comuns entre estas duas topologias?

Qual é a topologia de consenso?

Consenso

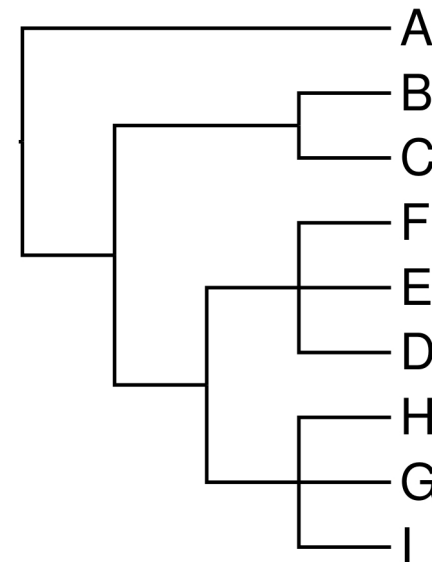
Exemplo



Quais componentes são comuns entre estas duas topologias?

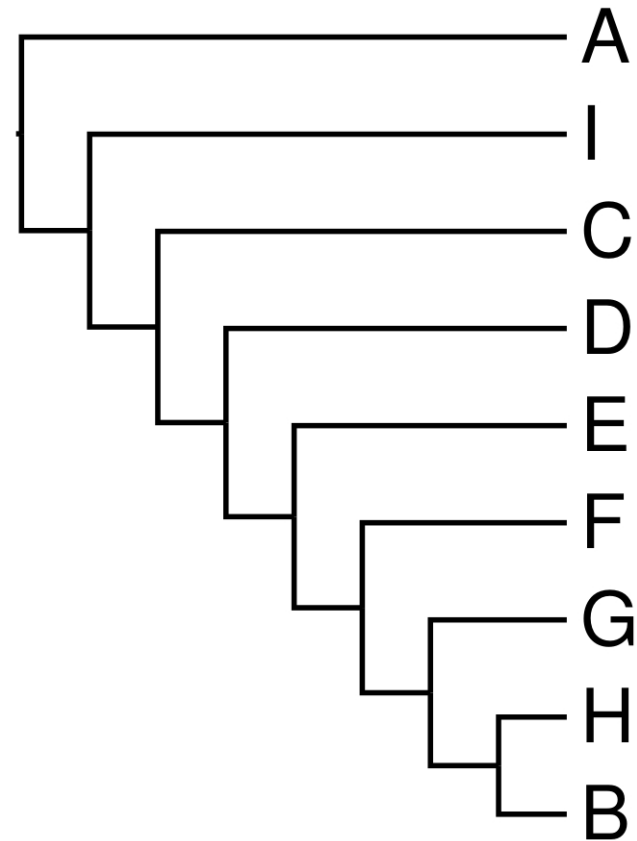
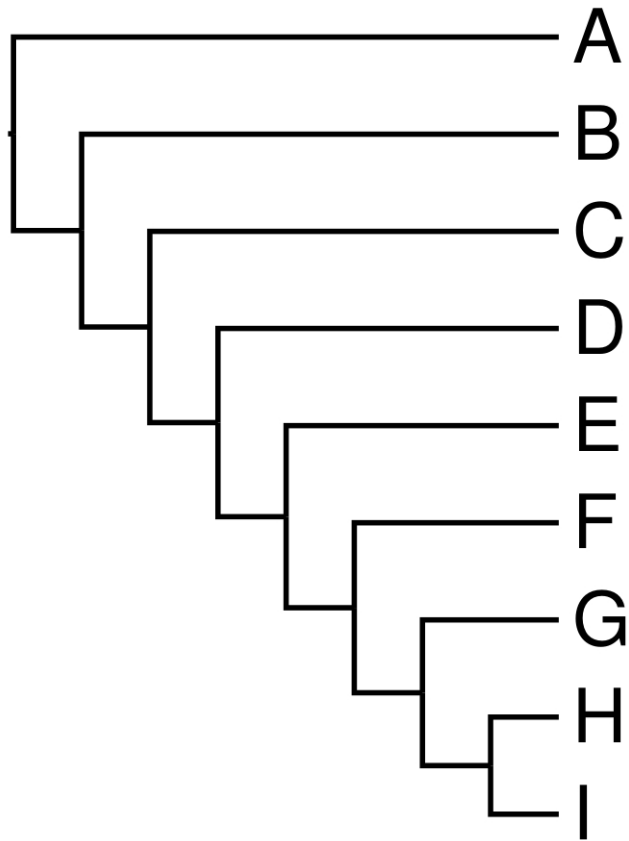
- 1: (ABCDEFGHI)
- 2: (BCDEFGHI)
- 3: (DEFGHI)
- 4: (DEF)
- 5: (GHI)
- 6: (BC)

Qual é a topologia de consenso?



Consenso

Exemplo 2



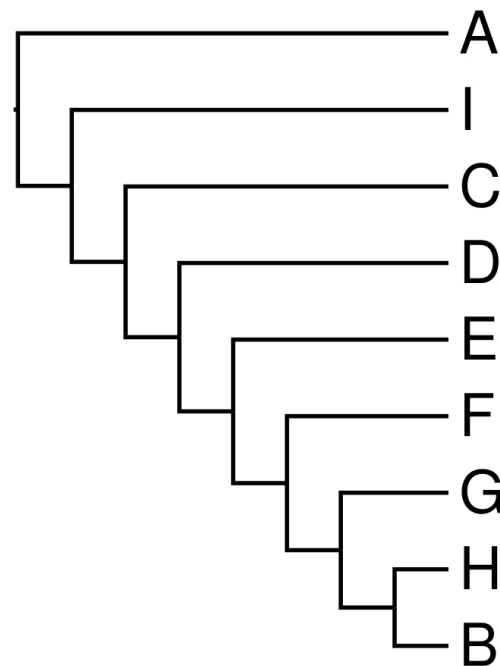
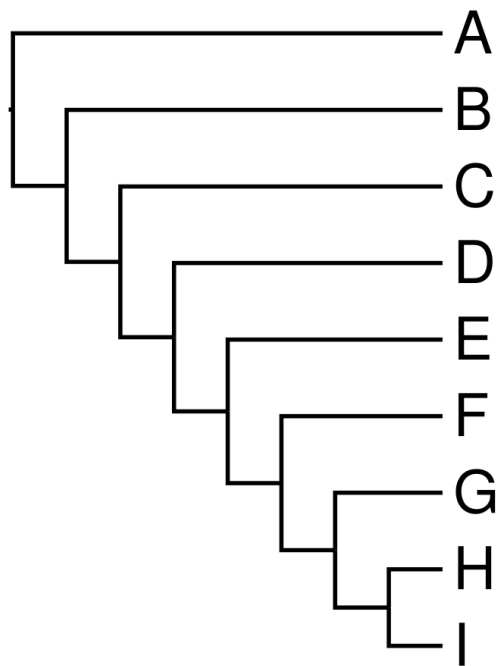
Quais componentes são comuns entre estas duas topologias?

Qual é a topologia de consenso?

Existe algum terminal que está forçando a topologia de consenso?

Consenso

Exemplo 2

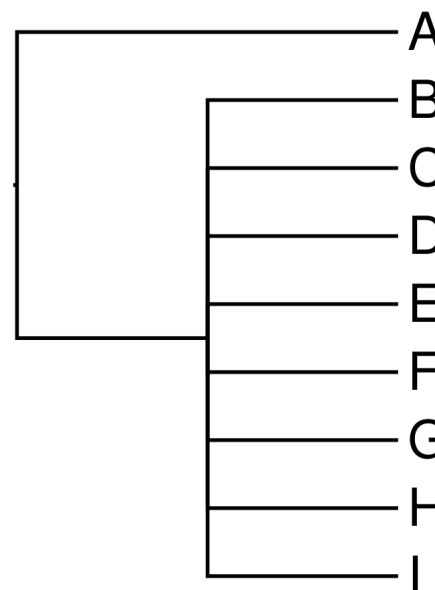


Quais componentes são comuns entre estas duas topologias?

1: (ABCDEFGHI)

2: (BCDEFGHI)

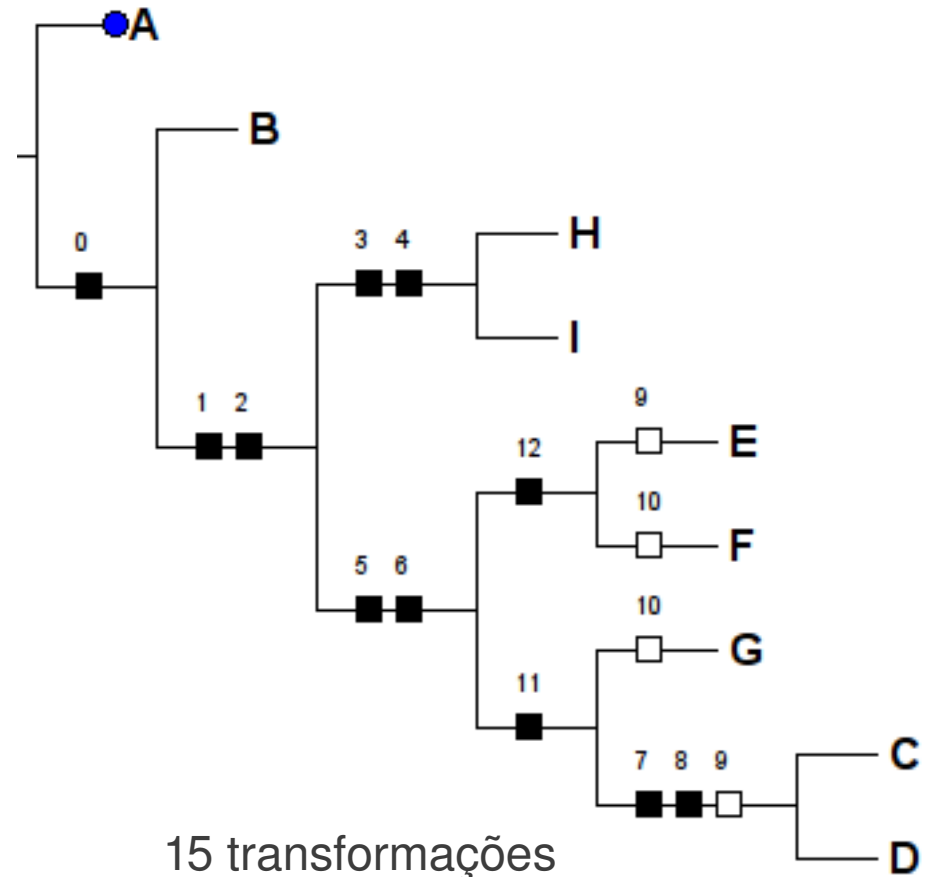
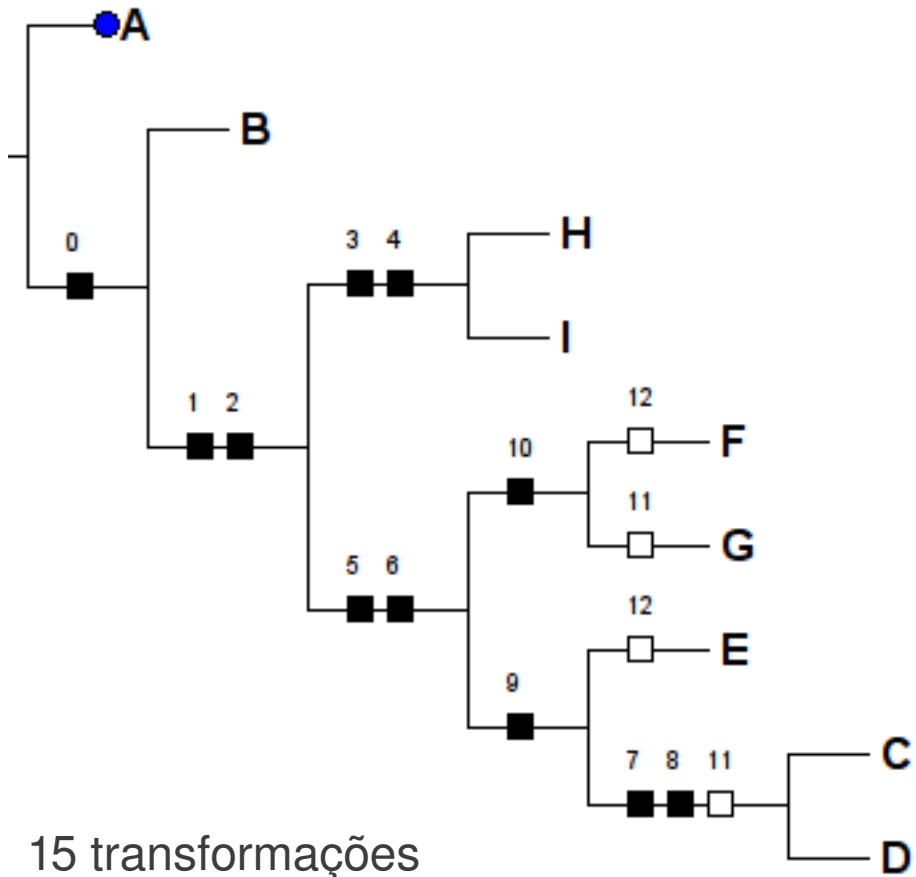
Qual é a topologia de consenso?



Consenso

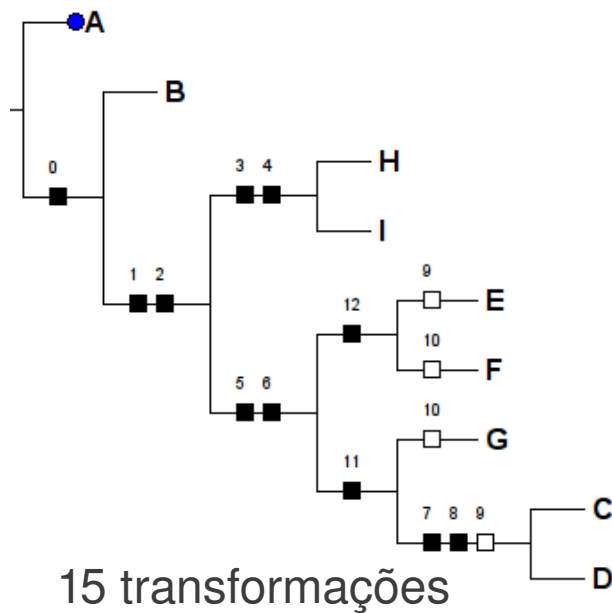
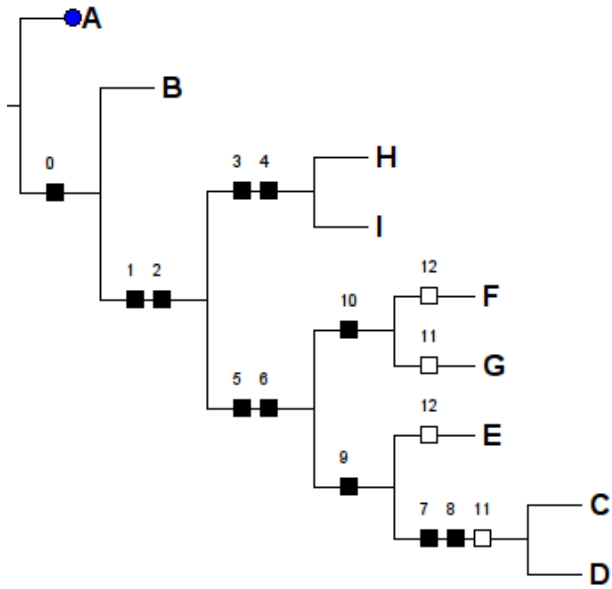
Exemplo 3

```
xread
13 9
A 0000000010000
B 1000000010000
C 1110011101010
D 1110011101010
E 1110011011001
F 1110011010101
G 1110011010110
H 1111100010000
I 1111100010000;
```

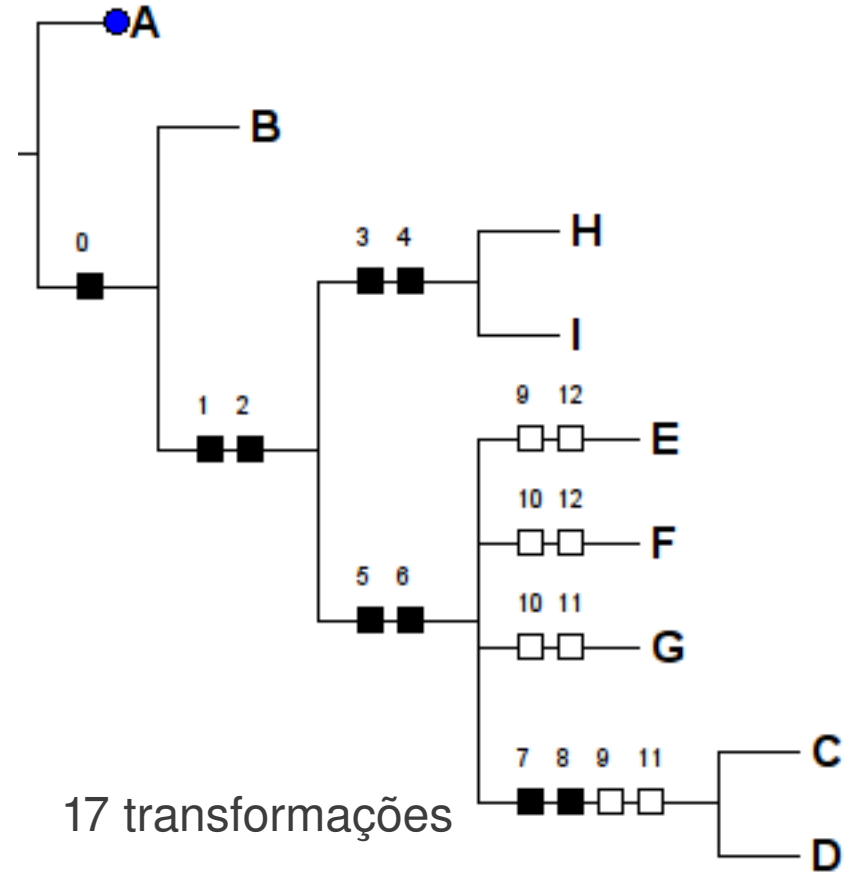


Consenso

Exemplo 3



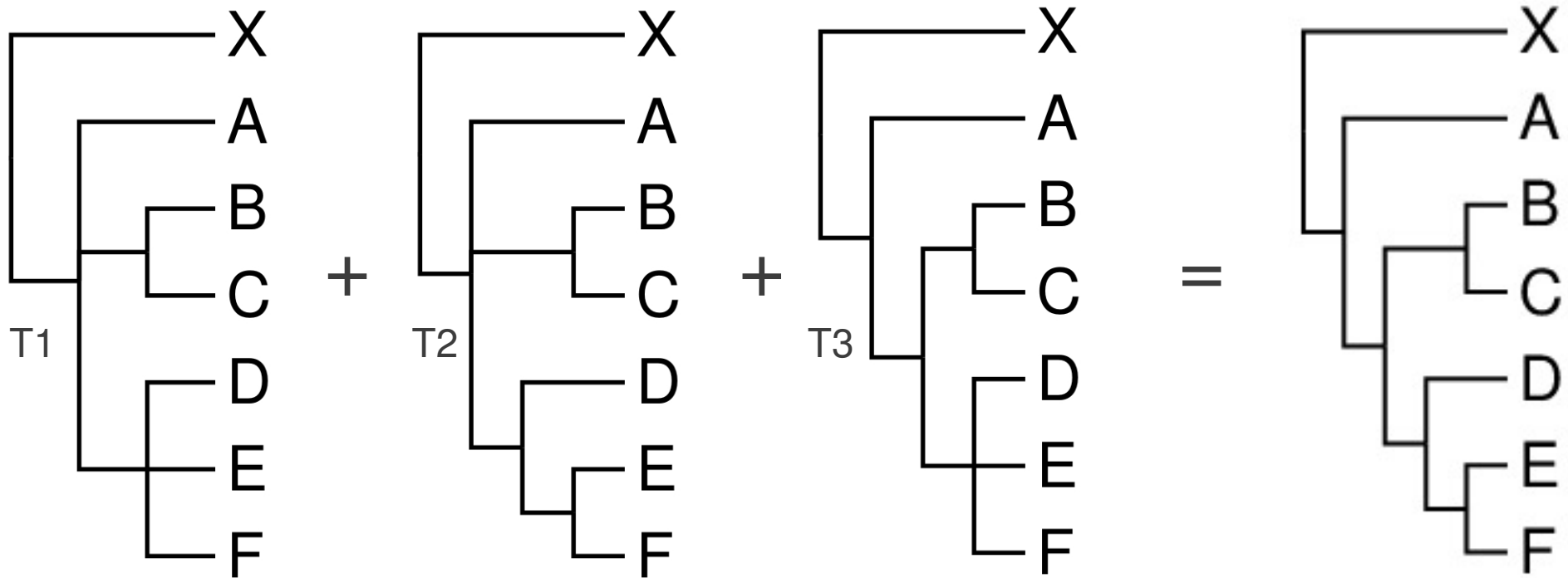
15 transformações



17 transformações

Consenso

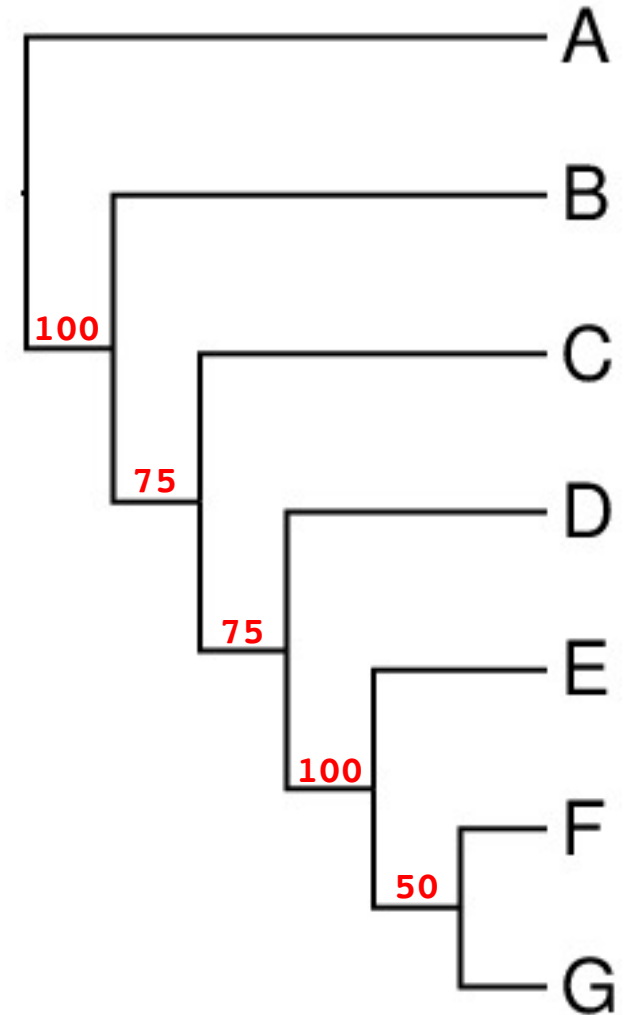
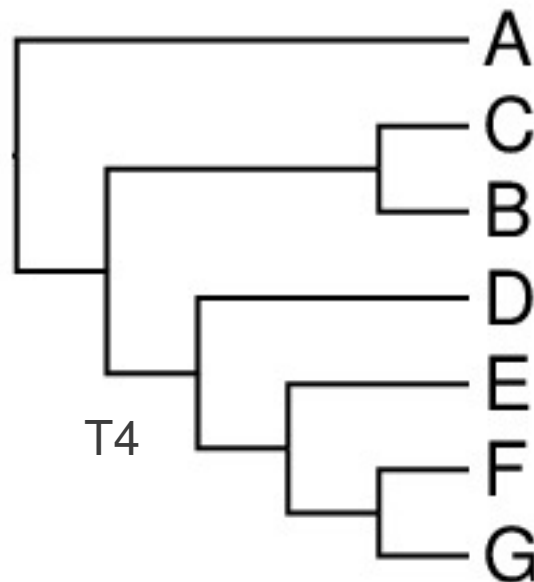
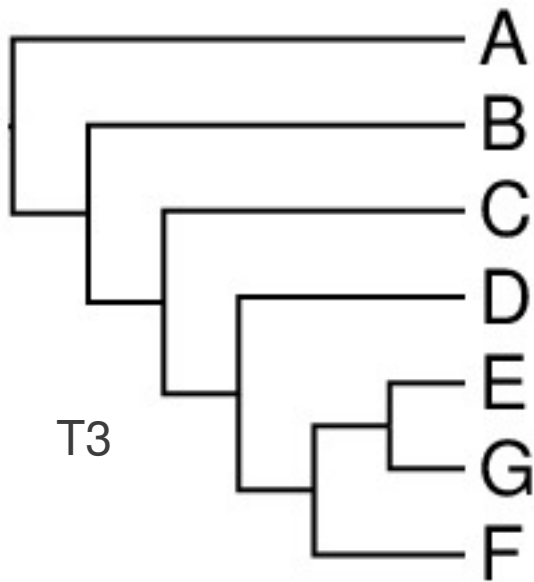
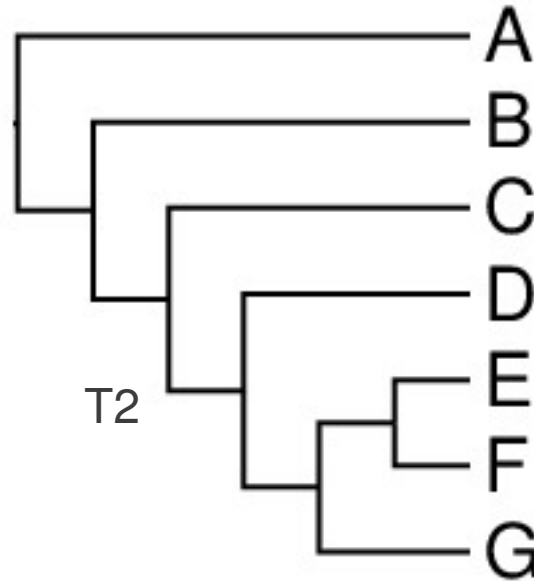
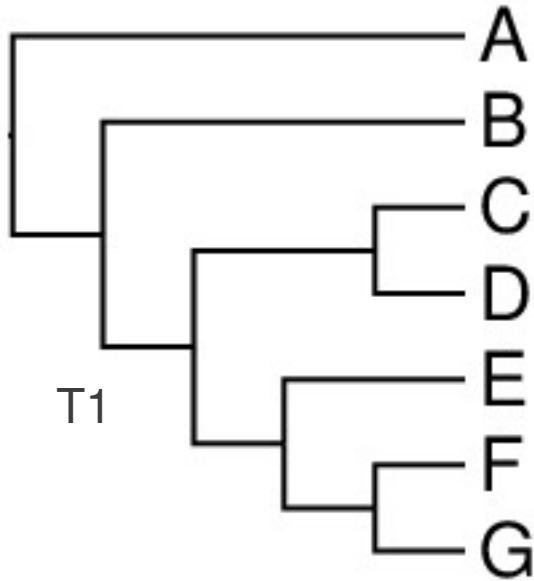
Semi-estrito (“*compatible components*”)



1. Note que o componente resolvido da T2 envolvendo (DEF) não contradiz o mesmo componente nas topologias T1 e T3.
2. Note que o componente resolvido da T3 envolvendo (BCDEF) não contradiz o mesmo componente nas topologias T1 e T2.
3. Portanto, a topologia de consenso pode ser mais resolvida do que as topologias fundamentais.

Consenso

Maioria (“majority rule”)



1. Neste consenso são considerados os componentes com maior frequência.
2. Note que a topologia de consenso não existe no conjunto de árvores fundamentais.

O desenvolvimento da Cladística

Dados moleculares e modelos probabilísticos

Aristóteles – 384-322 A.C.



Darwin
1809-1882



1859

Período essencialista

Mundo dinâmico

Resistência e Nova Síntese

Sistemática Evolutiva

1936 - 1947

1960's

Fenética

1970's

Cladística

1990's

Probabilismo

Carolus Linnaeus
1707-1778



Buffon
1707-1788



Lamarck
1744 -1829



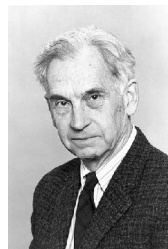
St-Hilair
1772 -1844



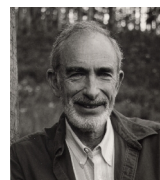
Cuvier
1769 -1832



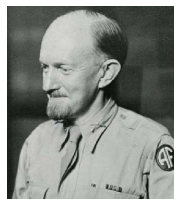
Ernest Mayr
1904 - 2005



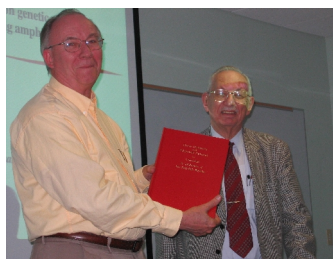
Paul Erlich



G.G. Simpson
1902 - 1984

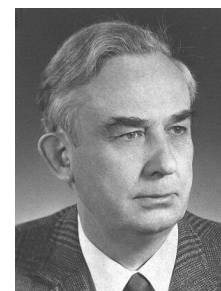


Theodosius Dobzhansky
1900 -1975



James Rohlf

R. Sokal
1926 -



Willi Hennig
1913 - 1976



Steve Farris



Joe Felsenstein



David Hillis

Motivação para o uso de Verossimilhança:

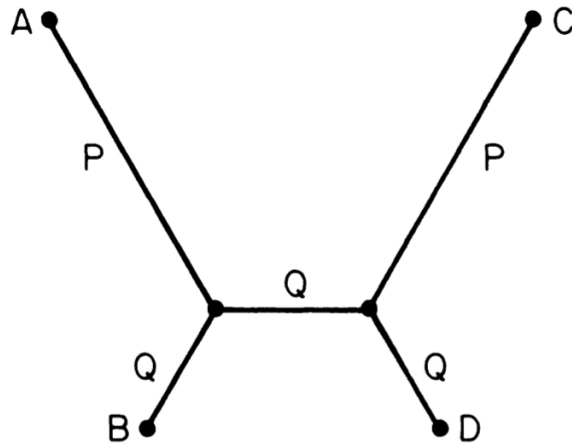
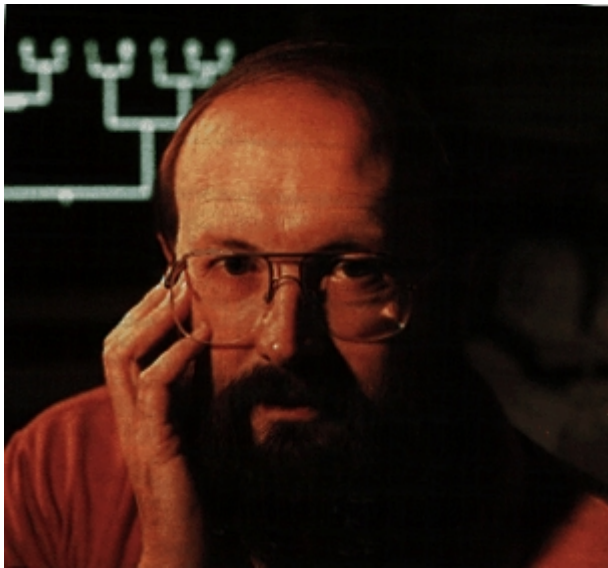


FIG. 3.—True unknown phylogeny (with root omitted) used to find cases in which unrooted Wagner tree parsimony methods will be inconsistent.



Joe Felsenstein
University of
Washington

CASES IN WHICH PARSIMONY OR COMPATIBILITY METHODS WILL BE POSITIVELY MISLEADING¹

JOSEPH FELSENSTEIN

Abstract

Felsenstein, J. (Department of Genetics, University of Washington, Seattle, WA 98195) 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.* 27:401-410.—For some simple three- and four-species cases involving a character with two states, it is determined under what conditions several methods of phylogenetic inference will fail to converge to the true phylogeny as more and more data are accumulated. The methods are the Camin-Sokal parsimony method, the compatibility method, and Farris's unrooted Wagner tree parsimony method. In all cases the conditions for this failure (which is the failure to be statistically consistent) are essentially that parallel changes exceed informative, nonparallel changes. It is possible for these methods to be inconsistent even when change is improbable a priori, provided that evolutionary rates in different lineages are sufficiently unequal. It is by extension of this approach that we may provide a sound methodology for evaluating methods of phylogenetic inference. [Numerical cladistics; phylogenetic inference; maximum likelihood estimation; parsimony; compatibility.]

Parsimony or minimum evolution methods were first introduced into phylogenetic inference by Camin and Sokal (1965). This class of methods for inferring an evolutionary tree from discrete-character data involves making a reconstruction of the changes in a given set of characters on a given tree, counting the smallest number of times that a given kind of event need have happened, and using this as the measure of the adequacy of the evolutionary tree. (Alternatively, one can compute the weighted sum of the numbers of times several different kinds of events have occurred.) One attempts to find that evolutionary tree which requires the fewest of these evolutionary events to explain the observed data. Camin and Sokal treated the case of irreversible changes along a character state tree, minimizing the number of changes

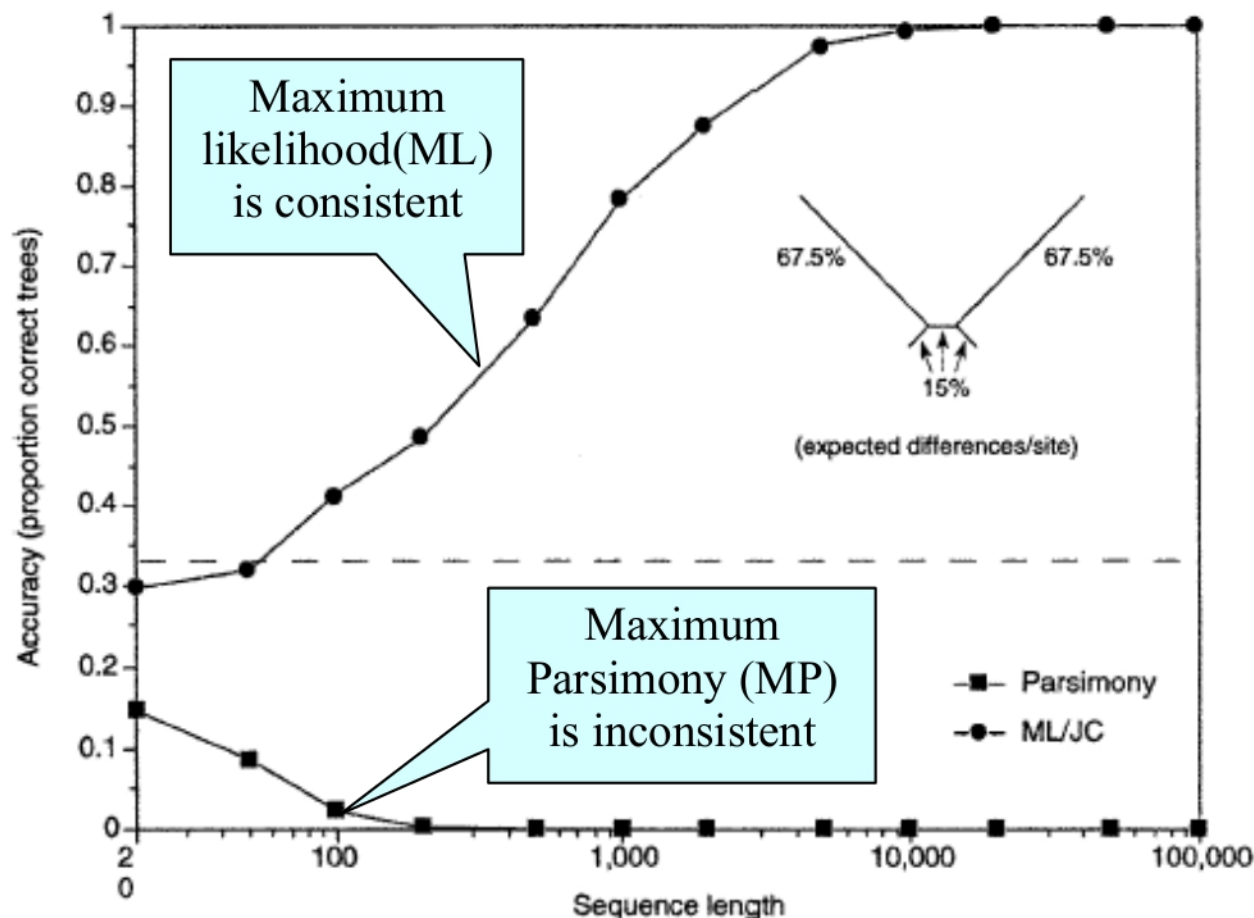
of character states required. A number of other parsimony methods have since appeared in the systematic literature (Kluge and Farris, 1969; Farris, 1969, 1970, 1972, 1977; Farris, Kluge, and Eckhardt, 1970) and parsimony methods have also found widespread use in studies of molecular evolution (Fitch and Margoliash, 1967, 1970; Dayhoff and Eck, 1968; see also Fitch, 1973). Cavalli-Sforza and Edwards (1967; Edwards and Cavalli-Sforza, 1964) earlier formulated a minimum evolution method for continuous-character data.

An alternative methodology for phylogenetic inference is the compatibility method, introduced by Le Quesne (1969, 1972). He suggested that phylogenetic inference be based on finding the largest possible set of characters which could simultaneously have all states be uniquely derived on the same tree. The estimate of the phylogeny is then taken to be that tree. While Le Quesne's specific suggestions as to how this might be done have been criticized by Farris (1969), his general approach, which is based on Camin and Sokal's (1965) concept of the compatibility of two characters, has been made rigorous and extended in a series of papers by G. F. Estabrook, C. S. Johnson, Jr., and F. R. McMorris (Estabrook,

¹ This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Department of Energy, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately-owned rights.

Motivação para o uso de Verossimilhança:

"Felsenstein Zone" Trees



Accuracy is proportion correct out of 1000 simulation replicates

Dashed line is 1/3, the probability of getting the tree correct by pure random chance

Fig. 3, p. 531, from Swofford et al. 2001. Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. *Systematic Biology* 50: 525-539.

O que é verossimilhança?



Sir Ronald Aylmer Fisher
(17 February 1890 – 29 July
1962)

$$L(\theta|x) = P(x|\theta)$$

A verossimilhança (L) de um conjunto de parâmetros (θ), dado alguma observação (x) é igual a probabilidade daquela observação ter ocorrido dados os valores daqueles parâmetros.

Parâmetros estatísticos podem ser vistos como as características numéricas de um modelo.

Um **modelo** estatístico é a formalização matemática da relação entre variáveis que correspondem a observações potenciais que inclui a descrição das incertezas sobre estas observações devido à variabilidade natural, erros ou informação incompleta.

O que é verossimilhança?



Sir Ronald Aylmer Fisher
(17/02/1890 – 29/07/1962)

$$L(\theta|x) = P(x|\theta)$$

Fisher (1912) empregou a função de verossimilhança $f(\theta|x)$ com a ideia de que o(s) valor(es) de θ que maximizam a probabilidade dos dados observados (x) seria um bom estimador de θ .

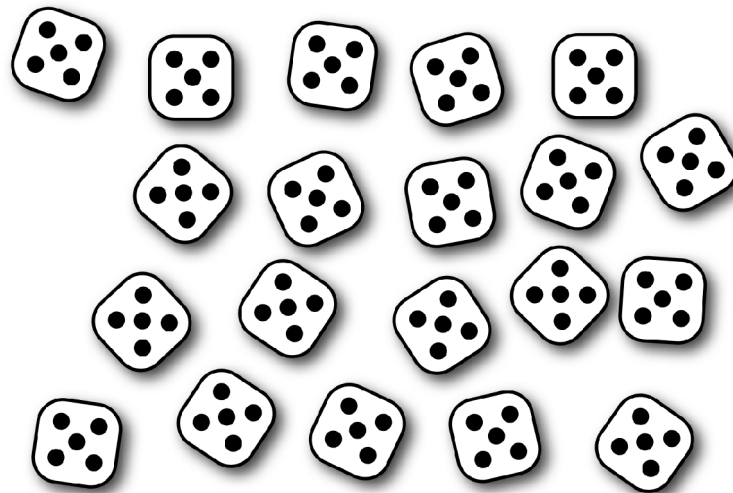
O que é verossimilhança?

$$L(\theta|x) = P(x|\theta)$$

Considere:

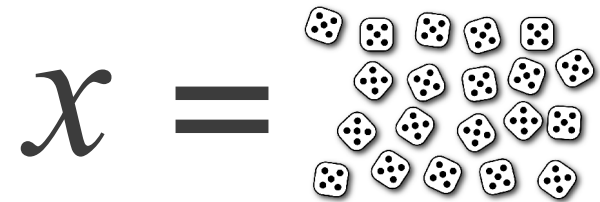
Você jogou 20 dados sobre uma mesa e todos apresentaram 5!
Portanto:

$x =$



O que é verossimilhança?

$$L(\theta|x) = P(x|\theta)$$



Considere que você deseja avaliar a L de dois modelos (θ_1 e θ_2):

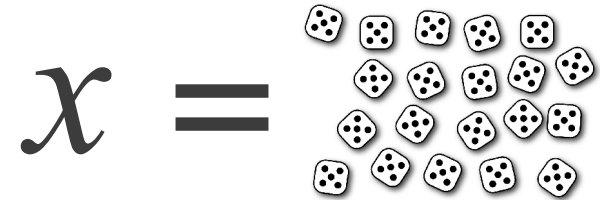
Modelo 1 (θ_1) – dado honesto – assume que cada face do dado possui um número de 1 a 6 cuja a probabilidade conjunta (i.e., $P_{(A \text{ e } B)} = P_{(A)} \times P_{(B)}$) de obter 20 dados apresentando 5 é igual a $(1/6)^{20}$.

Modelo 2 (θ_2) – dado viciado (Feito pela **Petrobras**) – assume que todas as faces do dado tem o número 5. Portanto, a probabilidade conjunta (i.e., $P_{(A \text{ e } B)} = P_{(A)} \times P_{(B)}$) de obter 20 dados apresentando 5 é igual a $(1)^{20}$.

Um **modelo** estatístico é a formalização matemática da relação entre variáveis que correspondem a observações potenciais que inclui a descrição das incertezas sobre estas observações devido à variabilidade natural, erros ou informação incompleta.

O que é verossimilhança?

$$L(\theta|x) = P(x|\theta)$$



Considere que você deseja avaliar a L de dois modelos (θ_1 e θ_2):

Modelo 1 (θ_1) – dado honesto – assume que cada face do dado possui um número de 1 a 6 cuja a probabilidade conjunta (i.e., $P_{(A \text{ e } B)} = P_{(A)} \times P_{(B)}$) de obter 20 dados apresentando 5 é igual a $(1/6)^{20}$.

$$L(\theta_1|x) = P_{(x|\theta_1)} = (1/6)^{20} = 1/3.656.158.440.062.976$$

Modelo 2 (θ_2) – dado viciado – assume que todas as faces do dado tem o número 5. Portanto, a probabilidade conjunta (i.e., $P_{(A \text{ e } B)} = P_{(A)} \times P_{(B)}$) de obter 20 dados apresentando 5 é igual a $(1)^{20}$.

$$L(\theta_2|x) = P_{(x|\theta_2)} = (1)^{20} = 1$$

Verossimilhança vs. Probabilidade

$$x = \text{[16 dice icons]} = \text{observação} = O$$

Modelo 1 (θ_1) – dado honesto = Hipótese 1 = H_1

Modelo 2 (θ_2) – dado viciado = Hipótese 2 = H_2

$$L(\theta|x) = P(x|\theta) \quad x = \text{[16 dice icons]}$$

$$P(O|H) \neq P(H|O)$$

Verossimilhança vs. Probabilidade

$$P(O|H) \neq P(H|O)$$

Considere:

O = Você está ouvindo um barulho no forro de sua casa.

H = há *gremlins* jogando boliche no seu forro.



$$L_{(H|O)} = P_{(O|H)} = \text{Alta}$$

$$P_{(H|O)} = \text{Baixa}$$

A verossimilhança desta hipótese é muito alta, pois se há *gremlins* jogando boliche no seu forro, a probabilidade de haver barulho é alta.

No entanto, certamente você não pensa que o barulho é evidência (torna-se provável) que haja *gremlins* jogando boliche no seu forro.

Voltemos a ideia original:



Sir Ronald Aylmer Fisher
(17/02/1890 – 29/07/1962)

$$L(\theta|x) = P(x|\theta)$$

Fisher (1912) empregou a função de verossimilhança $f(\theta|x)$ com a ideia de que o(s) valor(es) de θ que maximizam a probabilidade dos dados observados (x) seria um bom estimador de θ .

Estimativa de máxima verossimilhança:

OTU **CCACGTACGTACGTACGTACGTACGTACGTAC**



$$\theta = \alpha t = v$$

onde v é a "distância evolutiva entre essas sequências.

HTU **TTACGTACGTACGTACGTACGTACGTACGTAC**

Voltemos a ideia original:



Sir Ronald Aylmer Fisher
(17/02/1890 – 29/07/1962)

$$L(\theta|x) = P(x|\theta)$$

Qual o valor de θ que maximiza a probabilidade dos dados observados?

Estimativa de máxima verossimilhança:

OTU **CCACGTACGTACGTACGTACGTACGTACGTACGTAC**



$$\theta = \alpha t = v$$

onde v é a "distância evolutiva entre essas sequências."

HTU **TTACGTACGTACGTACGTACGTACGTACGTACGTAC**

Voltemos a ideia original:

$$L(\theta|x) = P(x|\theta)$$

Qual o valor de θ que maximiza a probabilidade dos dados observados?

Estimativa de máxima verossimilhança:

OTU **CCACGTACGTACGTACGTACGTACGTACGTACGTAC**



$$\theta = \alpha t = v$$

HTU **TTACGTACGTACGTACGTACGTACGTACGTACGTAC**

probabilidade de não mudança de estado*

$$P(ii) = \left[\left(\frac{1}{4} \right) * \left(\frac{1}{4} + \frac{3}{4} e^{-4v} \right) \right]$$

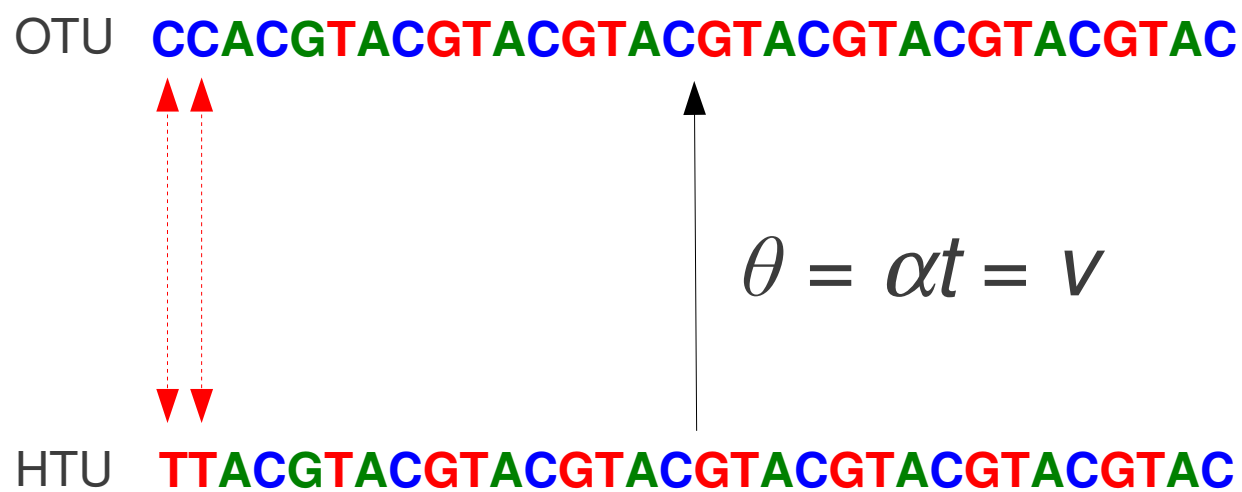
probabilidade de mudança de estado*

$$P(ij) = \left[\left(\frac{1}{4} \right) * \left(\frac{1}{4} - \frac{1}{4} e^{-4v} \right) \right]$$

* baseado no modelo de JC69 (distribuição de Poisson) e frequências iguais de pares de base.

Voltemos a ideia original:

Estimativa de **máxima** verossimilhança:



probabilidade de não
mudança de estado*

$$P(ii) = \left[\left(\frac{1}{4} \right) * \left(\frac{1}{4} + \frac{3}{4} e^{-4v} \right) \right]$$

probabilidade de
mudança de estado*

$$P(ij) = \left[\left(\frac{1}{4} \right) * \left(\frac{1}{4} - \frac{1}{4} e^{-4v} \right) \right]$$

$$L = \left[\left(\frac{1}{4} \right) * \left(\frac{1}{4} - \frac{1}{4} e^{-4v} \right) \right]^2 * \left[\left(\frac{1}{4} \right) * \left(\frac{1}{4} + \frac{3}{4} e^{-4v} \right) \right]^{30}$$

* baseado no modelo de JC69 (distribuição de Poisson) e frequências iguais de pares de base.

Voltemos a ideia original:



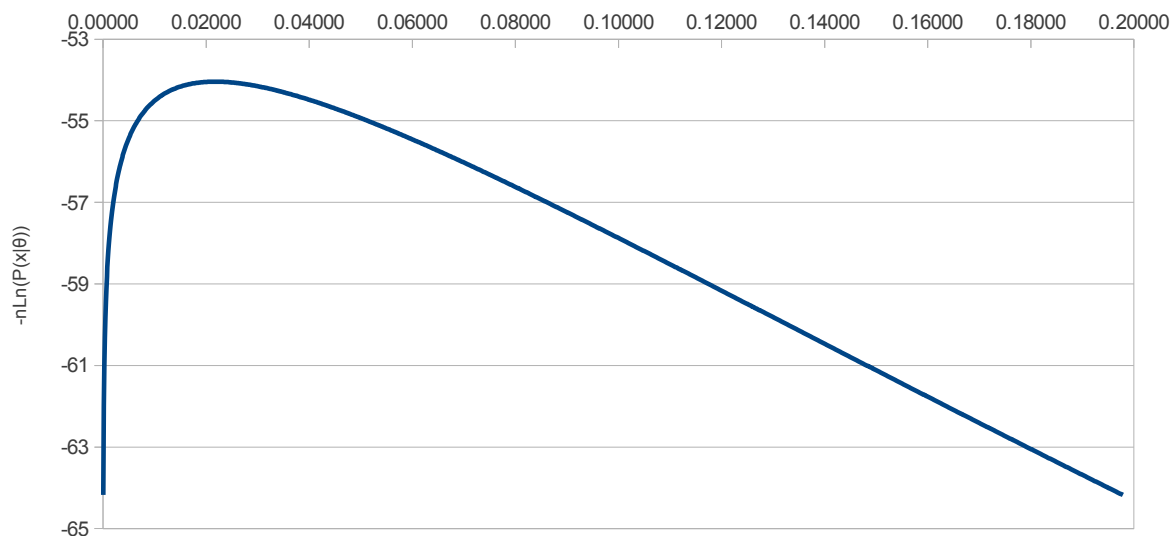
Sir Ronald Aylmer Fisher
(17/02/1890 – 29/07/1962)

$$L(\theta|x) = P(x|\theta)$$

Qual o valor de θ que maximiza a probabilidade dos dados observados?

$$L = \left[\left(\frac{1}{4} \right) * \left(\frac{1}{4} - \frac{1}{4} e^{-4v} \right) \right]^2 * \left[\left(\frac{1}{4} \right) * \left(\frac{1}{4} + \frac{3}{4} e^{-4v} \right) \right]^{30}$$

Estimativa de Máxima Verossimilhança (v)



Estimativa de distância evolutiva (v)

Valor estimado para v: 0.021752

Lógica da inferência filogenética

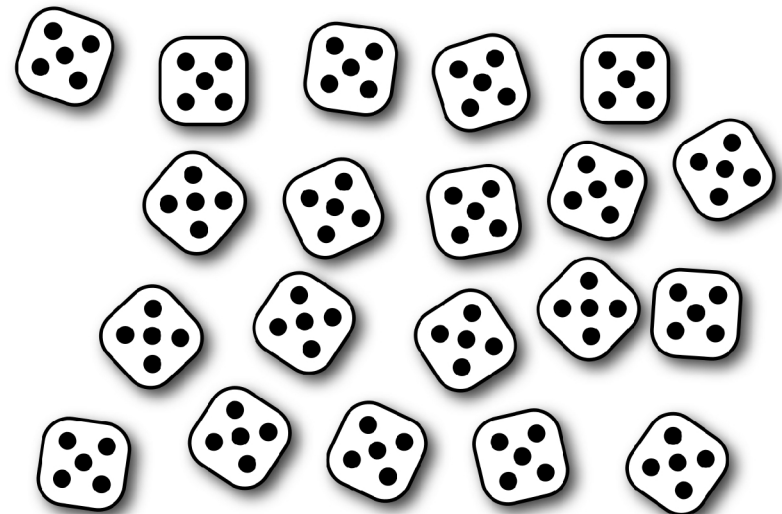
Avaliação e critério de seleção: soluções possíveis

Critério de seleção: **verossimilhança**
(*Likelihood*)

A probabilidade das observações calculadas utilizando **um modelo** nos diz o quão surpresos nós estaríamos com os dados observados

O modelo escolhido é aquele que menos surpreende!

Considere que eu jogue 20 dados sobre uma mesa e obtenha o seguinte resultados:



Obtenção de dados
[observação]



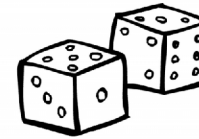
Codificação em evidências



Avaliação de evidências diante dos cenários possíveis
(critério de seleção)

Cálculo de probabilidades:

Se dois eventos não possuem soluções em comum, eles são chamados de **disjuntos** e suas probabilidades obedecem à seguinte regra:



Regra 1: Se dois eventos A e B são disjuntos, então a probabilidade de qualquer um destes eventos e a soma das probabilidades dos dois eventos: $P_{(A \text{ ou } B)} = P_{(a)} + P_{(b)}$

Regra 2: A probabilidade de um evento **A** não ocorrer é $P_{(A^c)} = 1 - P_{(A)}$



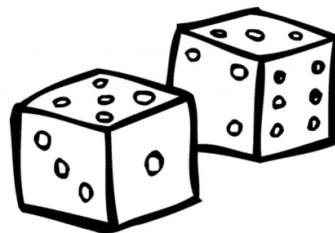
Copyright © Ron Leishman

$$P_{(\text{1})} = A/S = 1/6$$

$$P_{(\text{1})^c} = 1 - P_{(A)} = 1 - 1/6$$

Probabilidades e modelos

Considere que você jogue dois dados:



A probabilidade de obter “olhos de serpente” é expressa pela seguinte regra:











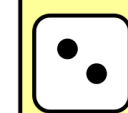
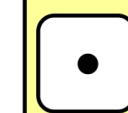
Regra 3: Se dois eventos A e B são independentes, então a probabilidade dos dois eventos ocorrerem um após o outro (i.e., intersecção de eventos, ou **probabilidade conjunta**) é o produto das probabilidades de cada evento: $P_{(A \text{ e } B)} = P_{(A)} * P_{(B)}$.

Combinando as regras:

Qual a probabilidade de se obter a soma 7 ao jogar dois dados?

E

$1/6 * 1/6$

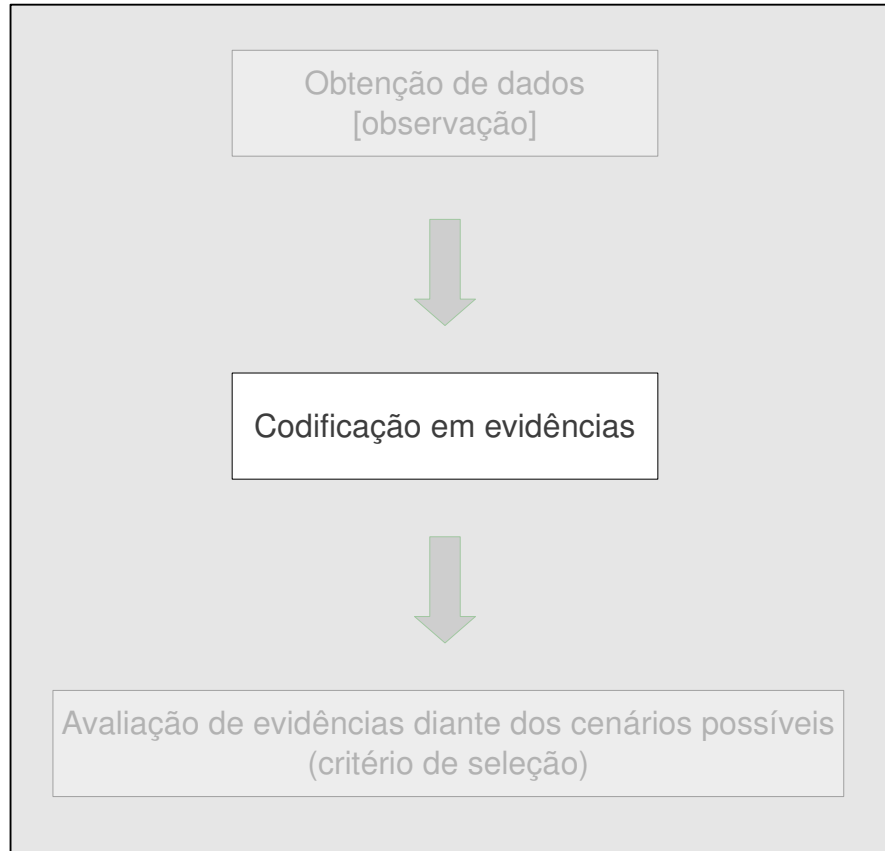
					
					

$(1/36)+(1/36)+(1/36)+(1/36)+(1/36)+(1/36) = 1/6$

OU OU OU OU OU

Lógica da inferência filogenética

↓ ↓ ↓
 sp.X CTGGCTACGT
 sp.A TGGAGTAAGT
 sp.B CCTAGCAAGT
 sp.C CCTGATTGCA



Parcimônia:

EVIDÊNCIAS: transformação de estados de caracteres

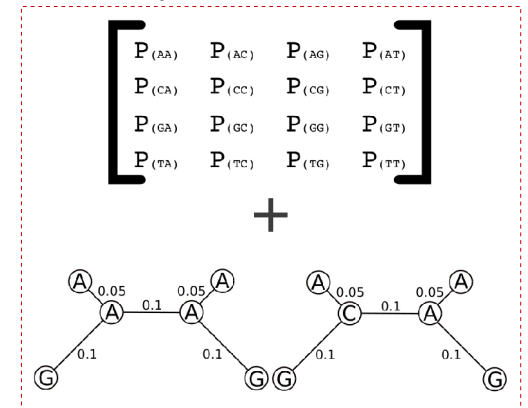
- 1a. posição: C ↔ T
- 3a. posição: G ↔ T
- 10a. posição: T ↔ A

CODIFICAÇÃO: matriz de dados

	c ₁	c ₂	c ₃	c ₄	c ₅	c ₆	c ₇	c ₈	c ₉	c ₁₀
sp.X	1	3	2	2	1	3	0	1	2	3
sp.A	3	2	2	0	2	3	0	0	2	3
sp.B	1	1	3	0	2	1	0	0	2	3
sp.C	1	1	3	2	0	3	3	2	1	0

Probabilística (ML):

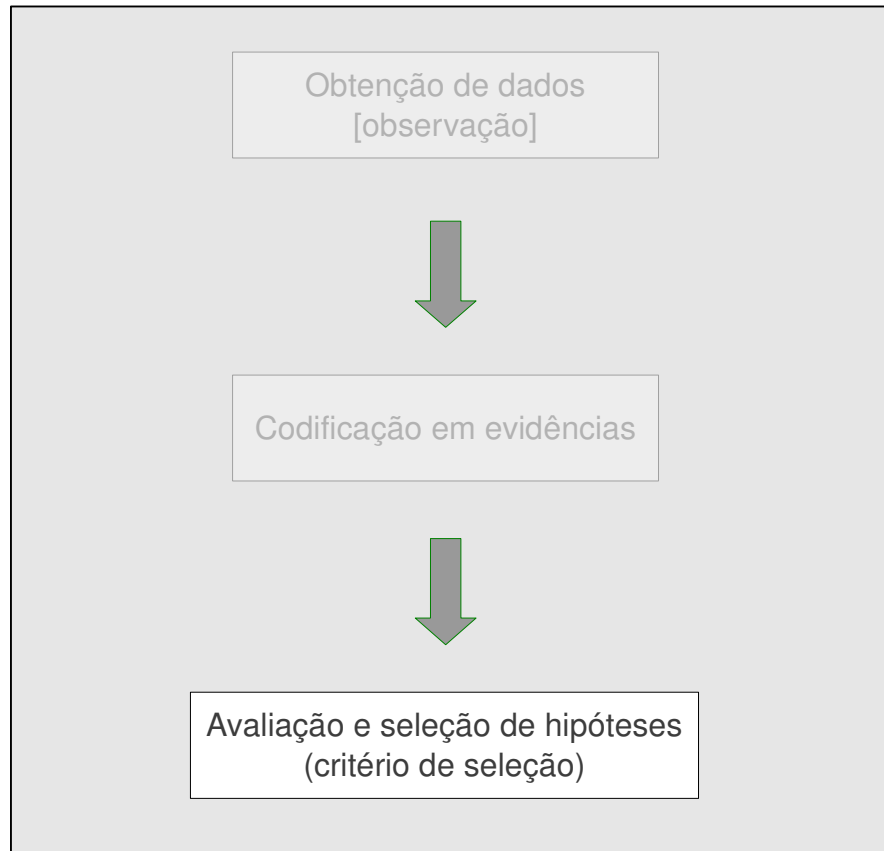
EVIDÊNCIAS: modelo de transformações + topologia que melhor explicam seus dados.



sp.X CTGGCTACGT
 sp.A TGGAGTAAGT
 sp.B CCTAGCAAGT
 sp.C CCTGATTGCA

Lógica da inferência filogenética

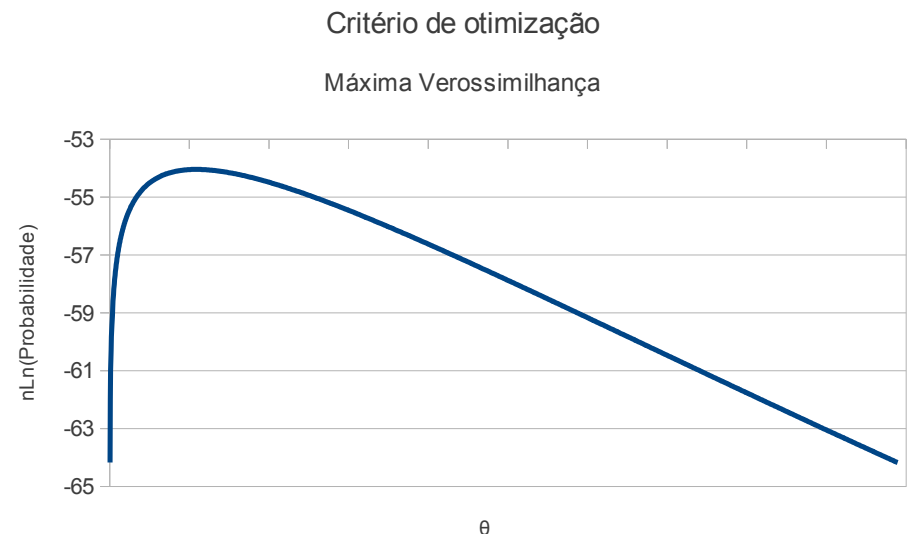
Avaliação e critério de seleção: verossimilhança máxima (ML)



$$L(\theta|x) = P(x|\theta)$$

$$L(\theta | \begin{matrix} \text{sp. X} & \text{CTGGCTACGT} \\ \text{sp. A} & \text{TGGAGTAAGT} \\ \text{sp. B} & \text{CCTAGCAAGT} \\ \text{sp. C} & \text{CCTGATTGCA} \end{matrix}) = P(\begin{matrix} \text{sp. X} & \text{CTGGCTACGT} \\ \text{sp. A} & \text{TGGAGTAAGT} \\ \text{sp. B} & \text{CCTAGCAAGT} \\ \text{sp. C} & \text{CCTGATTGCA} \end{matrix} | \theta)$$

Qual o valor de θ que maximiza a probabilidade dos dados observados?



Lógica da inferência filogenética

Avaliação e critério de seleção: verossimilhança máxima (ML)

$$L(\theta|x) = P(x|\theta)$$

$$L(\theta | \begin{matrix} \text{sp. X} & \text{CTGGCTACGT} \\ \text{sp. A} & \text{TGGAGTAAGT} \\ \text{sp. B} & \text{CCTAGCAAGT} \\ \text{sp. C} & \text{CCTGATTGCA} \end{matrix}) = P(\begin{matrix} \text{sp. X} & \text{CTGGCTACGT} \\ \text{sp. A} & \text{TGGAGTAAGT} \\ \text{sp. B} & \text{CCTAGCAAGT} \\ \text{sp. C} & \text{CCTGATTGCA} \end{matrix} | \theta)$$

Maximização do parâmetro θ :

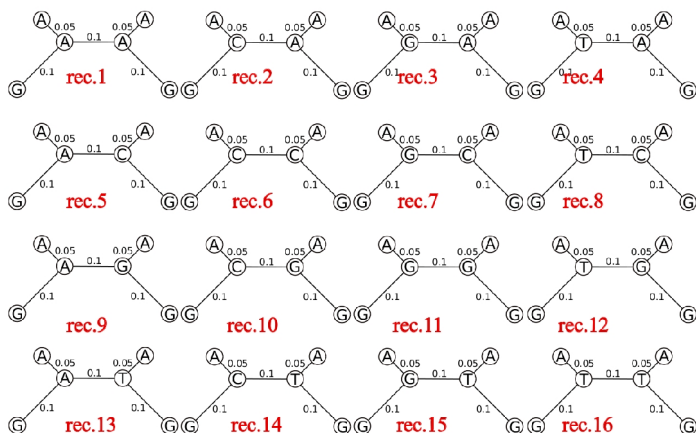
$$\theta = \begin{array}{c} \text{A} \\ \text{0.05} \\ \text{A} \\ \text{0.1} \\ \text{G} \end{array} \begin{array}{c} \text{A} \\ \text{0.05} \\ \text{A} \\ \text{0.1} \\ \text{G} \end{array} + GRT = \begin{bmatrix} P_{AA}(t)P_{AC}(t)P_{AG}(t)P_{AT}(t) \\ P_{CA}(t)P_{CC}(t)P_{CG}(t)P_{CT}(t) \\ P_{GA}(t)P_{GC}(t)P_{GG}(t)P_{GT}(t) \\ P_{TA}(t)P_{TC}(t)P_{TG}(t)P_{TT}(t) \end{bmatrix}$$

Lógica da inferência filogenética

Avaliação e critério de seleção: verossimilhança máxima (ML)

$$\theta = \left[\begin{array}{c} \text{Diagrama de árvore} \end{array} \right] + GRT = \left[\begin{array}{cccc} P_{AA}(t) & P_{AC}(t) & P_{AG}(t) & P_{AT}(t) \\ P_{CA}(t) & P_{CC}(t) & P_{CG}(t) & P_{CT}(t) \\ P_{GA}(t) & P_{GC}(t) & P_{GG}(t) & P_{GT}(t) \\ P_{TA}(t) & P_{TC}(t) & P_{TG}(t) & P_{TT}(t) \end{array} \right]$$

Possíveis reconstruções para HTUs



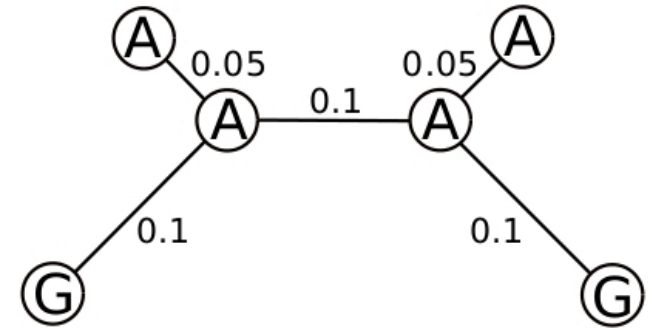
+ Diferentes topologias
Diferentes comprimentos de ramos

$$L = \Pr \left(\begin{array}{c} \text{Diagrama de árvore} \\ \left[\begin{array}{cccc} P_{(AA)} & P_{(AC)} & P_{(AG)} & P_{(AT)} \\ P_{(CA)} & P_{(CC)} & P_{(CG)} & P_{(CT)} \\ P_{(GA)} & P_{(GC)} & P_{(GG)} & P_{(GT)} \\ P_{(TA)} & P_{(TC)} & P_{(TG)} & P_{(TT)} \end{array} \right] \left| \begin{array}{l} \text{sp.X CTGGCTACGT} \\ \text{sp.A TGGAGTAAGT} \\ \text{sp.B CCTAGCAAGT} \\ \text{sp.C CCTGATTGCA} \end{array} \right. \right) + \left[\begin{array}{cccc} P_{(AA)} & P_{(AC)} & P_{(AG)} & P_{(AT)} \\ P_{(CA)} & P_{(CC)} & P_{(CG)} & P_{(CT)} \\ P_{(GA)} & P_{(GC)} & P_{(GG)} & P_{(GT)} \\ P_{(TA)} & P_{(TC)} & P_{(TG)} & P_{(TT)} \end{array} \right]$$

Lógica da inferência filogenética: Likelihood

Função objetiva:

$$C \Rightarrow P(\text{sp. X CTGGCTACGT} \\ \text{sp. A TGGAGTAAGT} \\ \text{sp. B CCTAGCAAGT} \\ \text{sp. C CCTGATTGCA} \mid \theta)$$



Modelo de substituição:

	A	C	G	T
A	$P_{(A,A)}$	$P_{(A,C)}$	$P_{(A,G)}$	$P_{(A,T)}$
C	$P_{(C,A)}$	$P_{(C,C)}$	$P_{(C,G)}$	$P_{(C,T)}$
G	$P_{(G,A)}$	$P_{(G,C)}$	$P_{(G,G)}$	$P_{(G,T)}$
T	$P_{(T,A)}$	$P_{(T,C)}$	$P_{(T,G)}$	$P_{(T,T)}$

$$P_{ij}(t) = \frac{1}{4} - \frac{1}{4}e^{-4\nu/3}$$

probabilidade de mudança de estado

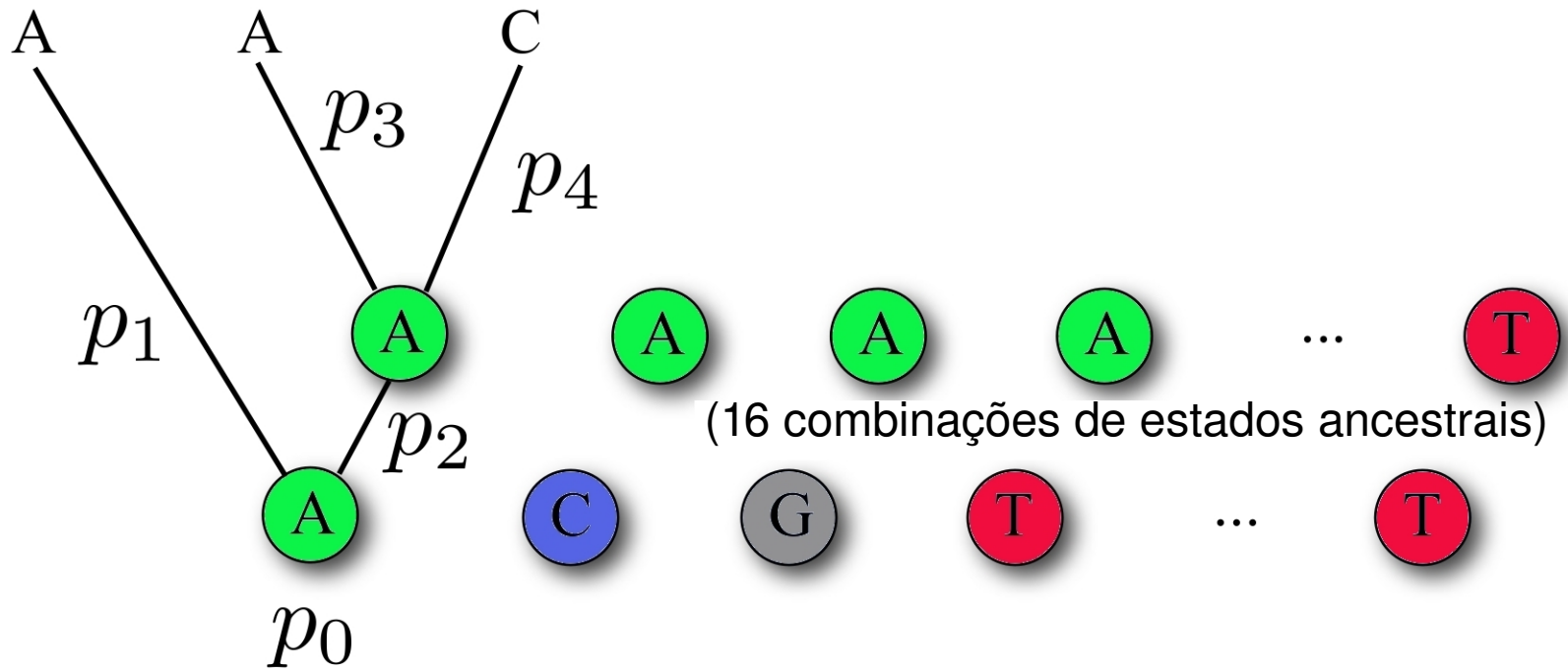
$$P_{ii}(t) = \frac{1}{4} + \frac{3}{4}e^{-4\nu/3}$$

probabilidade de não mudança de estado

		$\nu = 0.1$			
		A	C	G	T
A		0.9064	0.0312	0.0312	0.0312
C		0.0312	0.9064	0.0312	0.0312
G		0.0312	0.0312	0.9064	0.0312
T		0.0312	0.0312	0.0312	0.9064

		$\nu = 0.05$			
		A	C	G	T
A		0.9516	0.0161	0.0161	0.0161
C		0.0161	0.9516	0.0161	0.0161
G		0.0161	0.0161	0.9516	0.0161
T		0.0161	0.0161	0.0161	0.9516

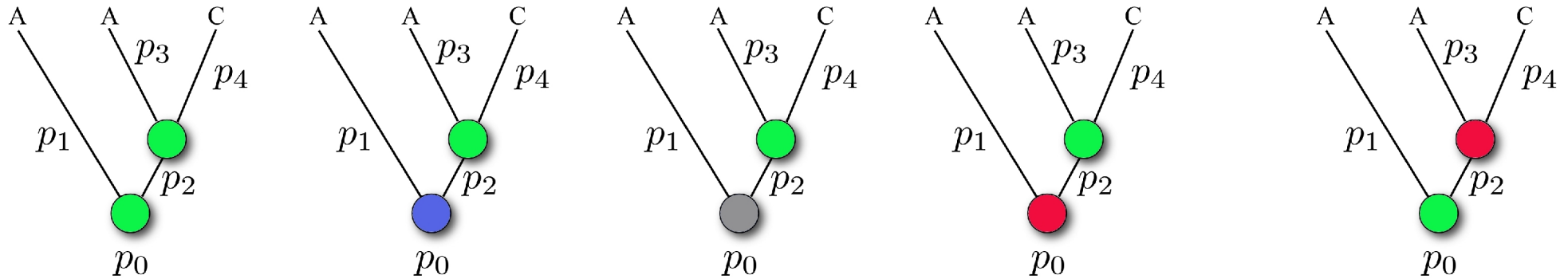
Cálculo de verossimilhança para topologias:



Regra 5: lintersecção de eventos, ou probabilidade conjunta: $P_{(A \text{ e } B)} = P_{(A)} * P_{(B)}$.

$$\text{Reconstrução 1} = P_0 * P_1 * P_2 * P_3 * P_4$$

Cálculo de verossimilhança para topologias:



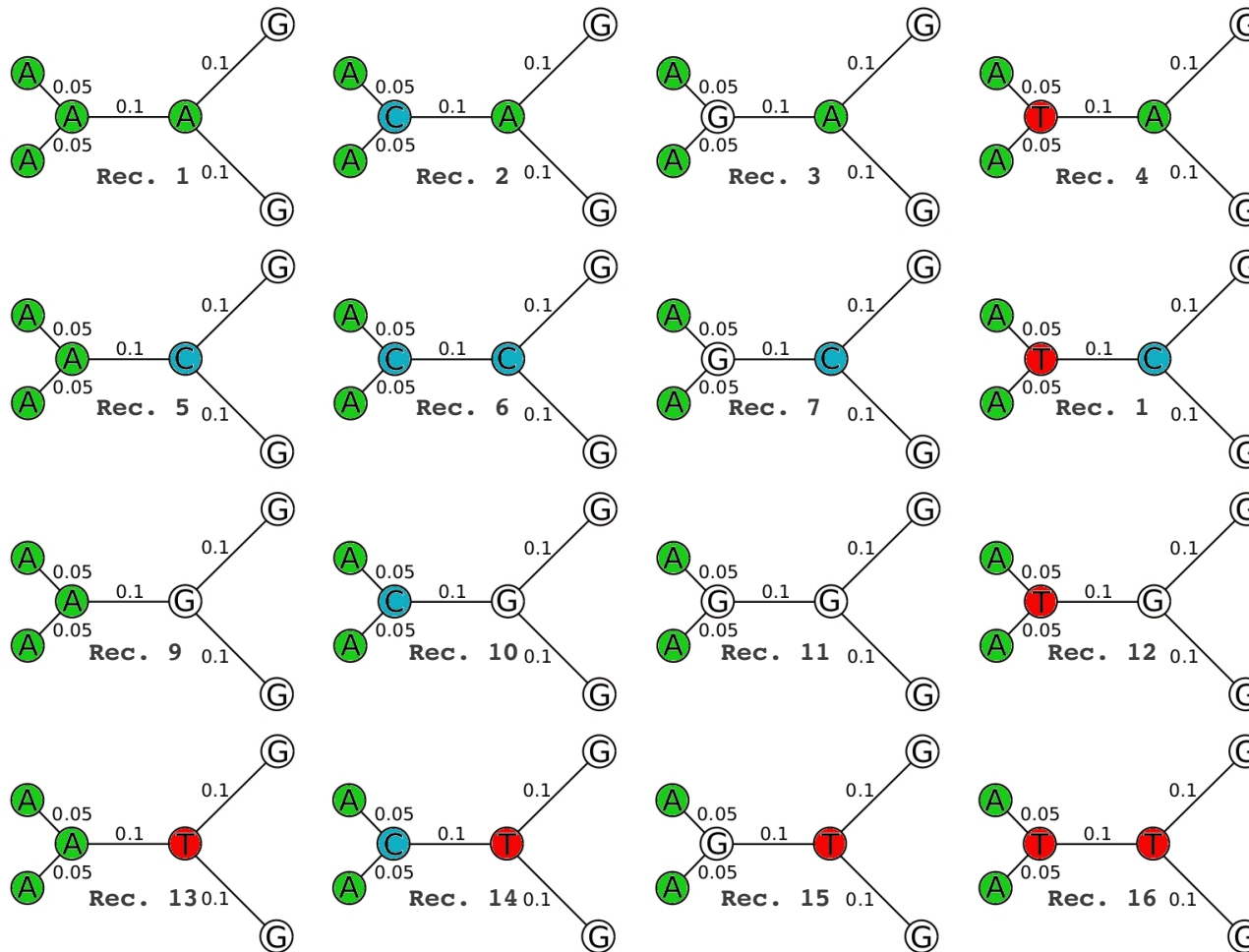
Reconstrução 1 + Reconstrução 2 + Reconstrução 3 + Reconstrução 4 + ... + Reconstrução 16

Regra 3: Eventos disjuntos: $P_{(A \text{ ou } B)} = P_{(a)} + P_{(b)}$

Lógica da inferência filogenética: Likelihood

$$L = \Pr \left(\begin{matrix} P_{(AA)} & P_{(AC)} & P_{(AG)} & P_{(AT)} \\ P_{(CA)} & P_{(CC)} & P_{(CG)} & P_{(CT)} \\ P_{(GA)} & P_{(GC)} & P_{(GG)} & P_{(GT)} \\ P_{(TA)} & P_{(TC)} & P_{(TG)} & P_{(TT)} \end{matrix} \right) + \left(\begin{matrix} A & 0.05 & 0.05 & A \\ & A & 0.1 & \\ G & 0.1 & & G \end{matrix} \mid \begin{matrix} \text{sp. X CTGGCTACGT} \\ \text{sp. A TGGAGTAAGT} \\ \text{sp. B CCTAGCAAGT} \\ \text{sp. C CCTGATTGCA} \end{matrix} \right)$$

Avaliação de todas as reconstruções possíveis



Lógica da inferência filogenética: o cálculo

Para,	1	2..	j	N							
sp.X	C	T	G	G	C	T	A	...	C	G	T
sp.A	T	G	G	A	G	T	A	...	A	G	T
sp.B	C	C	T	A	G	C	A	...	A	G	T
sp.C	C	C	T	G	A	T	T	...	G	C	A

A Verossimilhança Máxima de um **determinado caráter** (i.e., sítio) é:

$$L_{(j)} = P_{(rec. 1)} + P_{(rec. 2)} + P_{(rec. 3)} + \dots + P_{(rec. n)}$$

A Verossimilhança Máxima (L^*) de uma **determinada hipótese** é dada por:

$$L = L_{(1)} * L_{(2)} * L_{(3)} * \dots * L_{(n)} = \prod_{j=1}^N L_{(j)}$$

* tradicionalmente ela é avaliada pela soma dos logaritmos neperianos das probabilidades de cada caráter:

$$\ln L = \ln L_{(1)} + \ln L_{(2)} + \ln L_{(3)} + \dots + \ln L_{(n)} = \sum_{j=1}^N \ln L_{(j)}$$

Lógica da inferência filogenética: modelos complexos

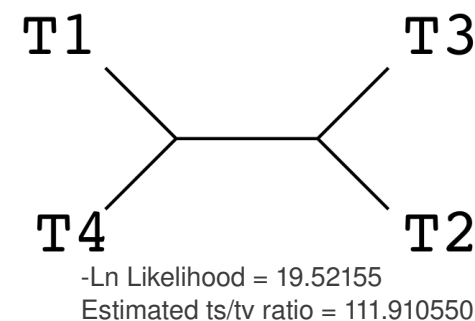
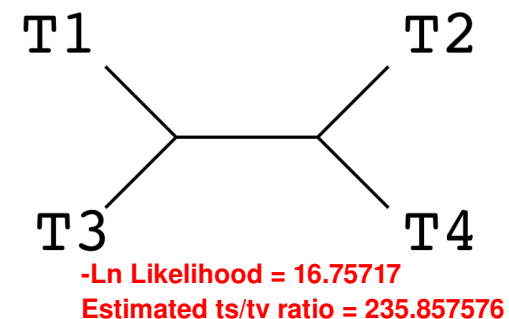
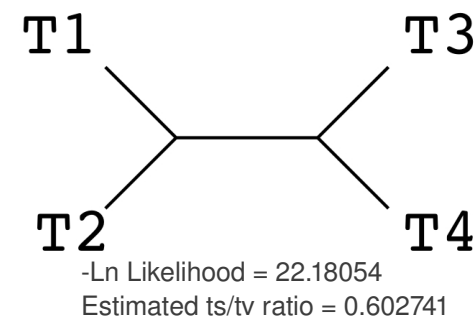
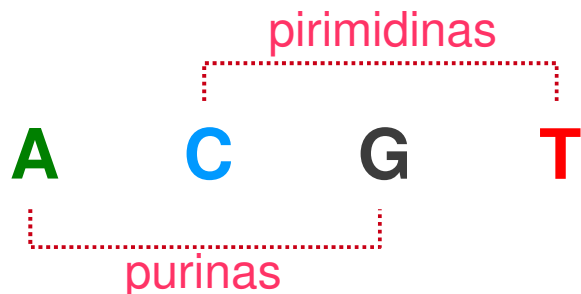
K2P

Kimura Two Parameters

	A	C	G	T
A	$P_{(A,A)}$	$P_{(A,C)}$	$P_{(A,G)} *k$	$P_{(A,T)}$
C	$P_{(C,A)}$	$P_{(C,C)}$	$P_{(C,G)}$	$P_{(C,T)} *k$
G	$P_{(G,A)} *k$	$P_{(G,C)}$	$P_{(G,G)}$	$P_{(G,T)}$
T	$P_{(T,A)}$	$P_{(T,C)} *k$	$P_{(T,G)}$	$P_{(T,T)}$

Considere:

T1 **A C G T**
 T2 **C G T A**
 T3 **G T A C**
 T4 **T A C G**



... onde k é a razão entre eventos de transição e transversão (valor estimado neste exemplo).

Note que utilizando o critério de parcimônia, nenhum caráter seria considerado informativo!

Escolha de critérios de otimalidade

Qual é o melhor método?

Qual é a melhor árvore (topologia)?

Não há uma única resposta para a pergunta e para ambas perguntas é necessário adotar um critério para avaliação objetivo!

MÉTODOS DE DISTÂNCIA E DE TRANSFORMAÇÃO DE CARACTERES

Problemas centrais com o métodos de distância:

Impossibilidade de identificar transformações

Impossibilidade de reconstruir estados ancestrais hipotéticos

Impossibilidade de postular hipóteses de homologia

Escolha de critérios de otimalidade

Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)

Como nós "sabemos" as coisas?

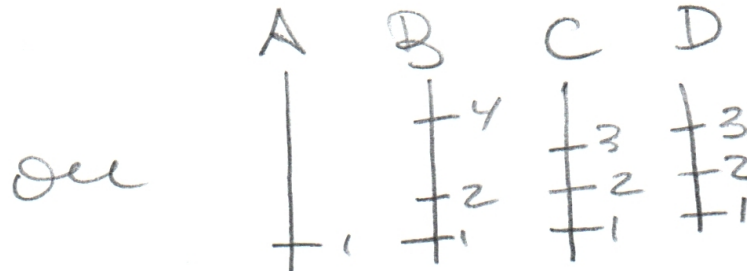
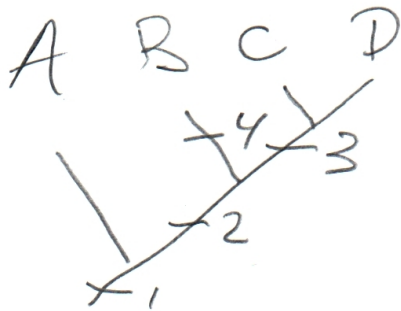
Como hipóteses científicas são propostas e testadas?

Qual é a relação entre observação e inferência?

1. Navalha de Ockham e argumentação Popperiana

"The minimization approach of parsimony offers a specific method to achieve the base-line goal of choosing genealogical explanation of variation, and of identifying non-conforming observations that require additional modes of causality."

(Wheeler 2012:271)



Escolha de critérios de otimalidade

Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)

Como nós "sabemos" as coisas?

Como hipóteses científicas são propostas e testadas?

Qual é a relação entre observação e inferência?

1. Navalha de Ockham e argumentação Popperiana

Popper --> hipóteses devem ser testadas pela busca de evidências contraditórias

> número de observação

> número de oportunidades de falsificar a hipótese

> severamente a hipótese foi testada

"Parsimonious trees have maximal explanatory power with regard to observed variation, and they do this without the accessory explanatory elements of other methods (e.g., model and tree parameters)."

(Wheeler 2012:271)

Escolha de critérios de otimalidade

Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)

Como nós "sabemos" as coisas?

Como hipóteses científicas são propostas e testadas?

Qual é a relação entre observação e inferência?

2. Parcimônia e processos evolutivos

procedimento de inferência

vs.

Modelo de evolução

- i. Não assume que evolução ocorre parcimoniamamente (ex. Homoplasia)
- ii. Em ciência utilizada como um processo de descobrimento e teste
- lii. Pode ser defendida sem apelo ao autoritarismo a apriorismo

Escolha de critérios de otimalidade

Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)

Como nós "sabemos" as coisas?

Como hipóteses científicas são propostas e testadas?

Qual é a relação entre observação e inferência?

3. Indução e estimativa estatística

Sistemática, assim como muitas outras disciplinas em Biologia, é uma ciência histórica na qual os eventos que buscamos explicar são únicos no tempo e no espaço.

"If each character observation is a unique object, it cannot be a random sample drawn from a parameterized distribution expressed as a model of change (e.g., GTR+I+G)."

"By this rationale, parsimonious trees have explanatory power in minimizing ad hoc hypotheses, and increased observation will increase the severity of test, but would not reduce any measure of 'sampling error', since there is no sample distribution that is measured."

(Wheeler 2012:272)

Escolha de critérios de otimalidade

Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)

Como nós "sabemos" as coisas?

Como hipóteses científicas são propostas e testadas?

Qual é a relação entre observação e inferência?

4. Teste de hipóteses e critério de otimização

Inferência hipotético-dedutiva pode ser feita via etapa de falsificação baseada em qualquer função de otimização objetiva:

Requer: hipóteses devem competir matematicamente de forma objetiva.

Questão central: a função de otimização é apropriada?

Escolha de critérios de otimalidade

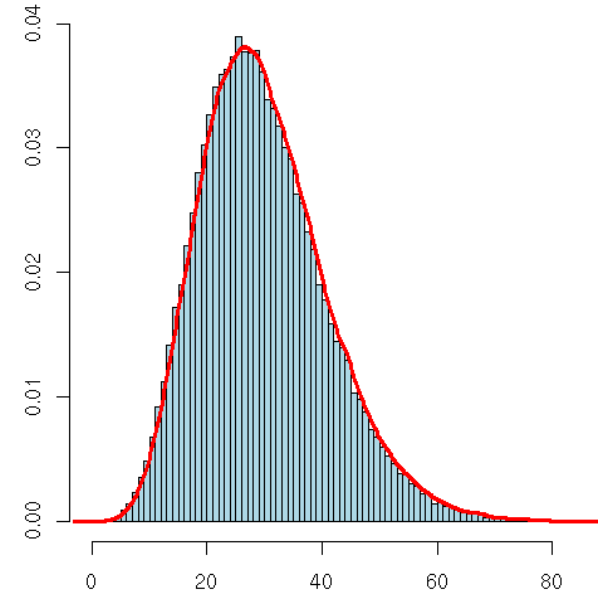
Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)



Probabilidades: duas interpretações

i. Grau de crença:

Confiança relativa em determinado resultado, considere

"Esta moeda tem 80% de chance de ser honesta."

Mas a moeda é ou não é honesta!



Escolha de critérios de otimalidade

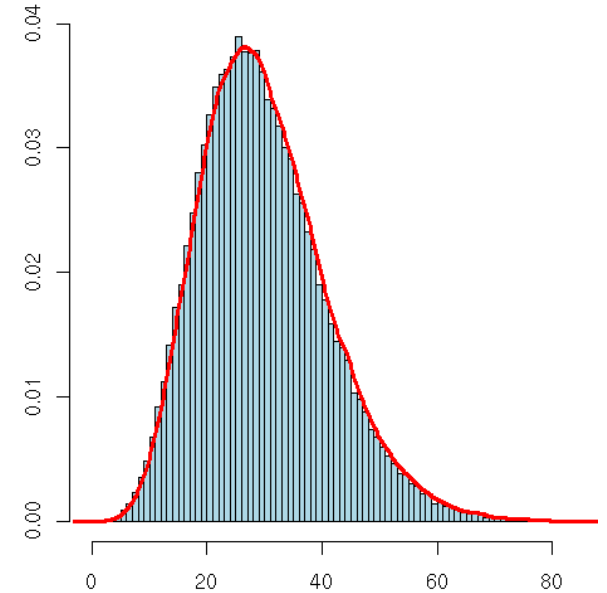
Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)



Probabilidades: duas interpretações

ii. Uma proposição verdadeira sobre o universo natural semelhante à eventos repetidos:

Implica em uma proposição sobre o que é ontologicamente real, considere:

"Esta moeda tem 80% de chance de dar coroa em uma jogada."



Oferece uma propriedade da moeda e uma proposição relacionada à sua realidade física.

Escolha de critérios de otimalidade

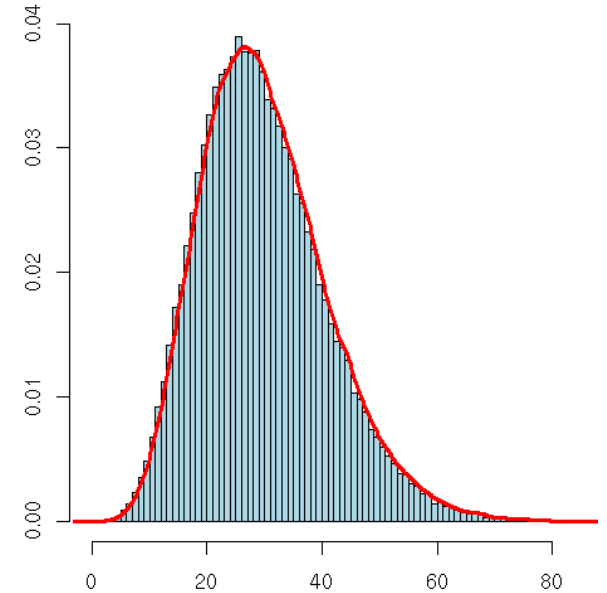
Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)



Probabilidades:

Ciências históricas não permitem facilmente a interpretação frequencista (i.e., ii)

Considere:

"Qual é a probabilidade de que Tiradentes e Joaquim José da Silva Xavier eram a mesma pessoa?"

Escolha de critérios de otimização

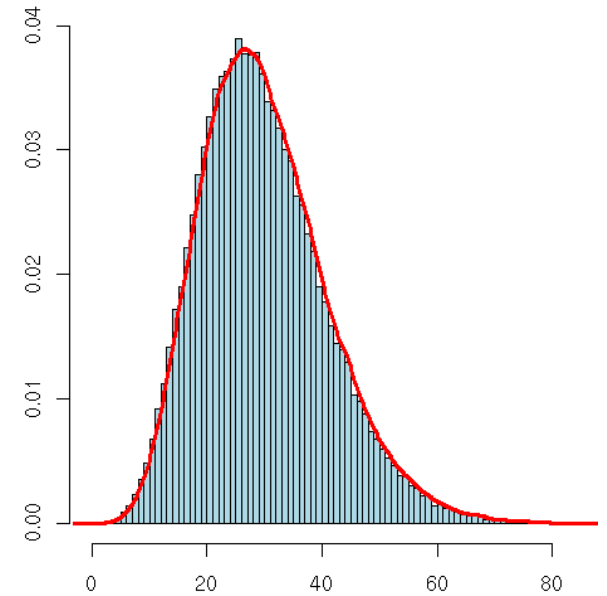
Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)



Probabilidades:

Ciências históricas não permitem facilmente a interpretação frequencista (i.e., ii)

Considere:

"A e B são grupos-irmãos?"

Escolha de critérios de otimalidade

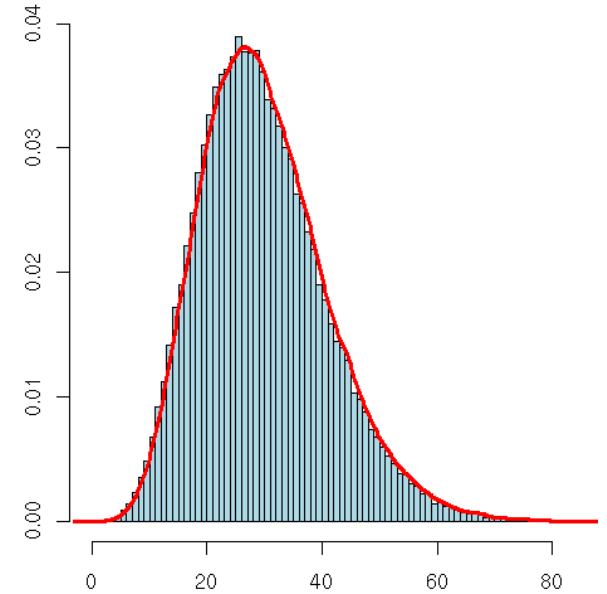
Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)



Probabilidades:

Ciências históricas não permitem facilmente a interpretação frequencista (i.e., ii)

Considere:

"Asas de morcego e asas de passarinho são homólogas?"

Escolha de critérios de otimalidade

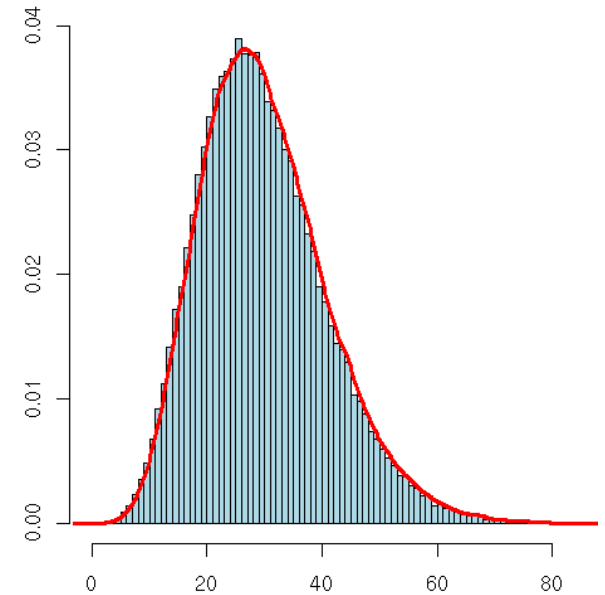
Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)



Probabilidades:

Ciências históricas não permitem facilmente a interpretação frequencista (i.e., ii)

Nesse contexto, qualquer interpretação de probabilidade está restrita ao nosso grau de crença. Portanto probabilidades associadas a eventos históricos conferem:

- i.* grau de suporte entre hipóteses e não sua realidade;
- ii.* elas não descrevem o universo natural, mas sim nosso entendimento dele.

"We cannot say whether a given set of historical statements (i.e., a tree) is true, but we can say we have relative degree of belief based on empirical observation and an optimality criterion"

(Wheeler 2012:274)

Escolha de critérios de otimização

Como escolher?

Epistemologia não oferece caminho inequívoco

- i. Todos os métodos podem participar do processo hipotético-dedutivo.
- ii. Todos estão sujeitos à comportamentos inconsistentes.

Uma possibilidade: sincretismo

- i. Convergência de resultados é trivial, embora rara.
- ii. Na maioria dos casos há incongruências, o que fazer?
 - a. critérios de otimização possuem diferentes motivações e epistemologias subjacentes.
 - b. Consenso é um caminho ilógico.

Escolha de critérios de otimização

Como escolher?

Critérios de otimização são premissas analíticas utilizadas para interpretar eventos historicamente únicos. Como tal, eles não podem ser testados empiricamente por acurácia. Por outro lado, congruência pode ser um caminho a ser seguido. No entanto, em seu cerne, critérios de otimização são pressupostos que requer suporte e justificativa quando empregados. Pressupostos diferentes levam a resultados diferentes. Somente justificando nossas premissas nós podemos defender nossas conclusões.

Conceitos fundamentais desta aula:

Cladogramas s/c comprimentos de ramo

Descrição de hipóteses filogenéticas

Suporte e ambiguidade

Consenso

Estrito

Semi-estricto

Regra de Maioria

Probabilismo

Verossimilhança e Probabilidades

Escolha de critérios de otimalidade