

## ***Conceitos fundamentais desta aula:***

*Cladística:*

*Evidência de relação de parentesco*

*Caráter = Série de transformação*

*Séries binárias e multi-estados*

*Estados de caráter*

*Parcimônia*

*Distância patrística*

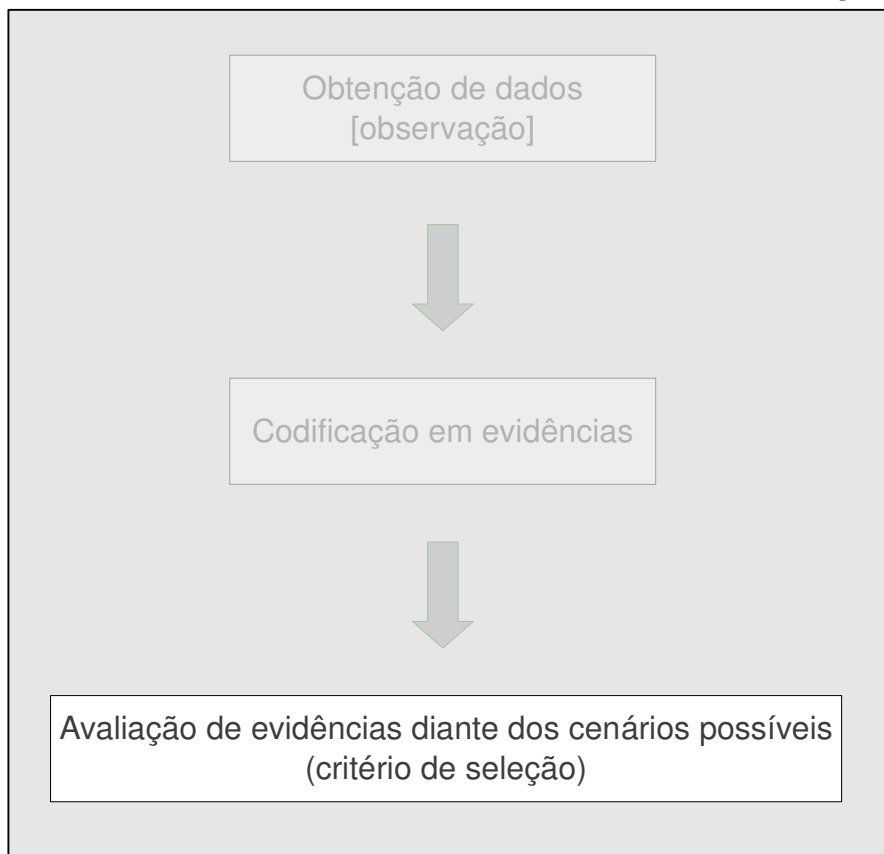
*Otimização*

*Conteúdo informativo de caracteres*

*Justificativa para adoção do critério*

# Lógica da inferência filogenética

Critério de otimalidade: parcimônia



Critério de seleção: < distância patrística



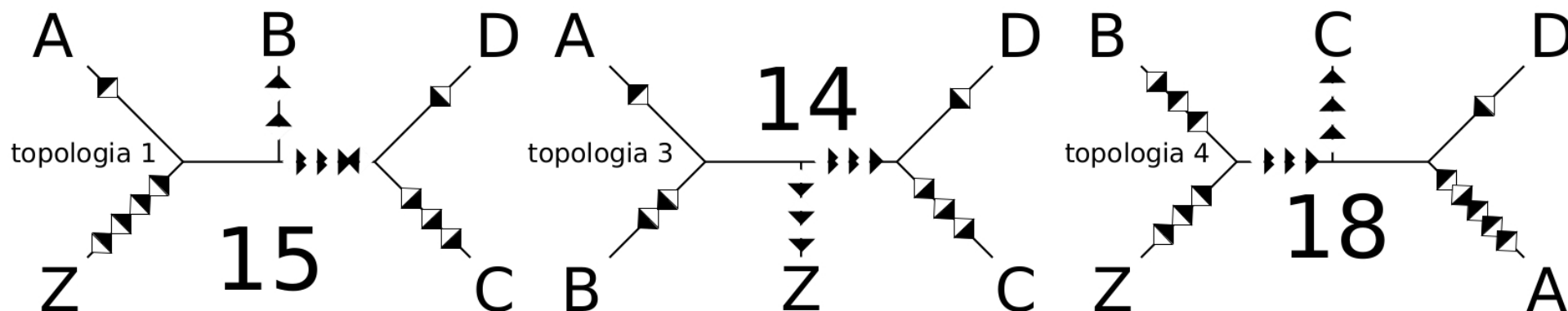
ockham wielding razor

**William of Ockham** (c. 1288 - c. 1348):  
*lex parsimoniae* ou "Occam's Razor"

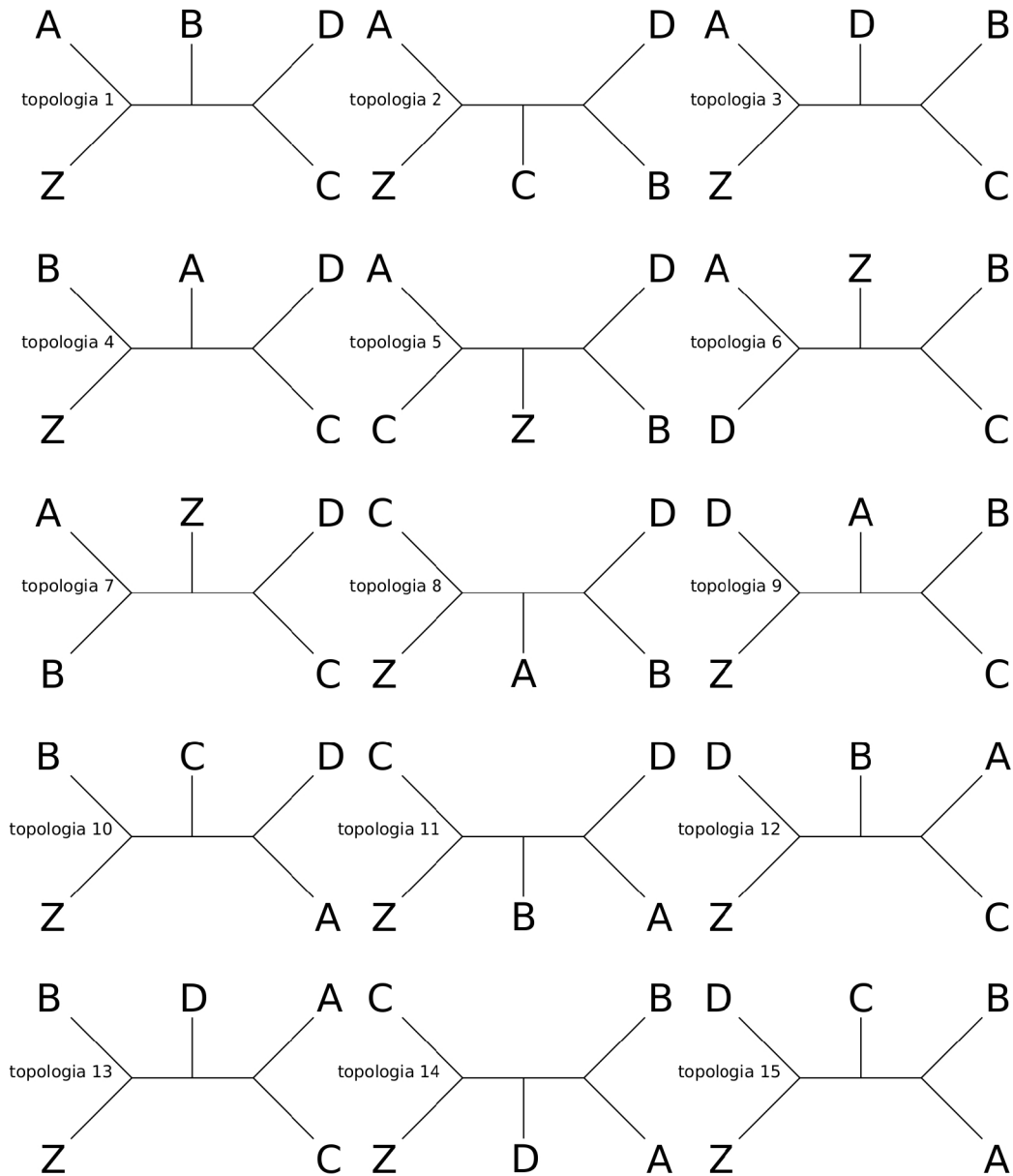
"*entia non sunt multiplicanda praeter necessitatem*"

"Entities should not be multiplied unnecessarily."

"when you have two competing theories which make exactly the same predictions, the one that is simpler is the better."



# *Os possíveis cenários (soluções):*



*Soluções matemáticas, mas ...*

# *Os possíveis cenários (soluções):*

Primeira diferença (trivial): nomes diferentes!

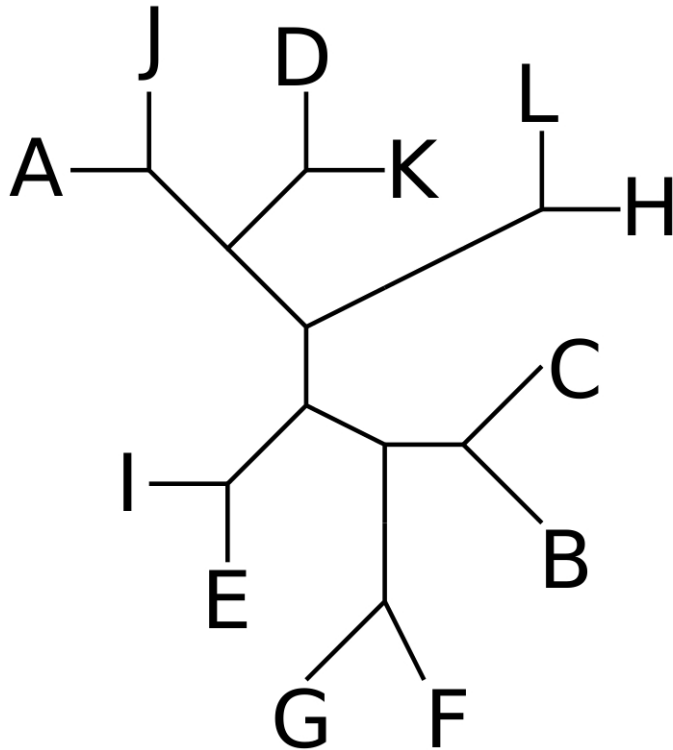


diagrama não-enraizado

(grafo binário acíclico não - direcionado)

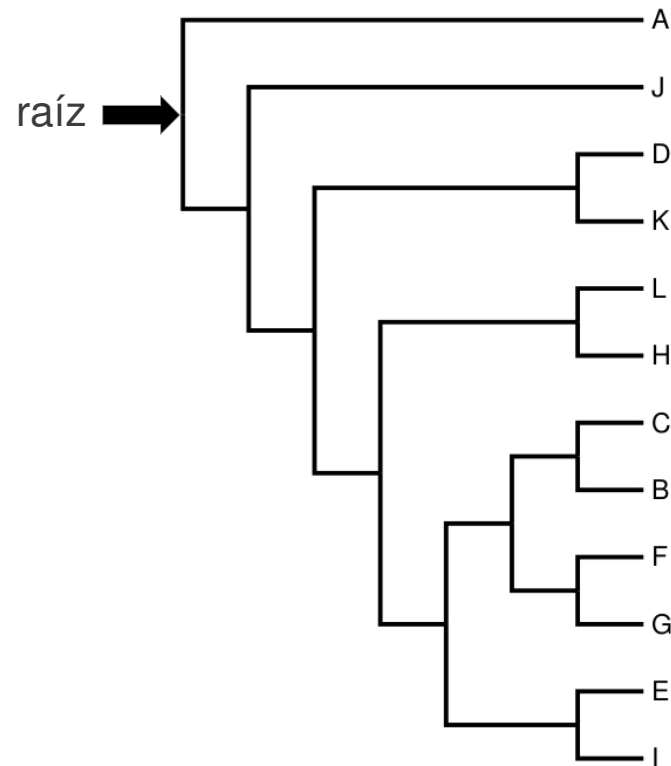


diagrama enraizado

(grafo binário acíclico direcionado)

*Enraizamento: ato de atribuir uma raíz a um diagrama.*

# *Os possíveis cenários (soluções):*

Segunda diferença (fundamental): inserção de um vetor temporal

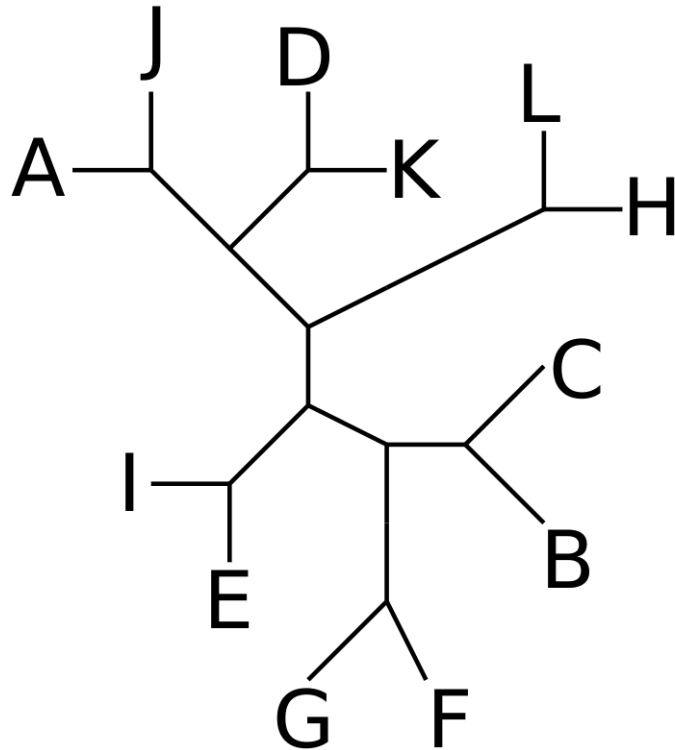


diagrama não-enraizado

(grafo binário acíclico não - direcionado)

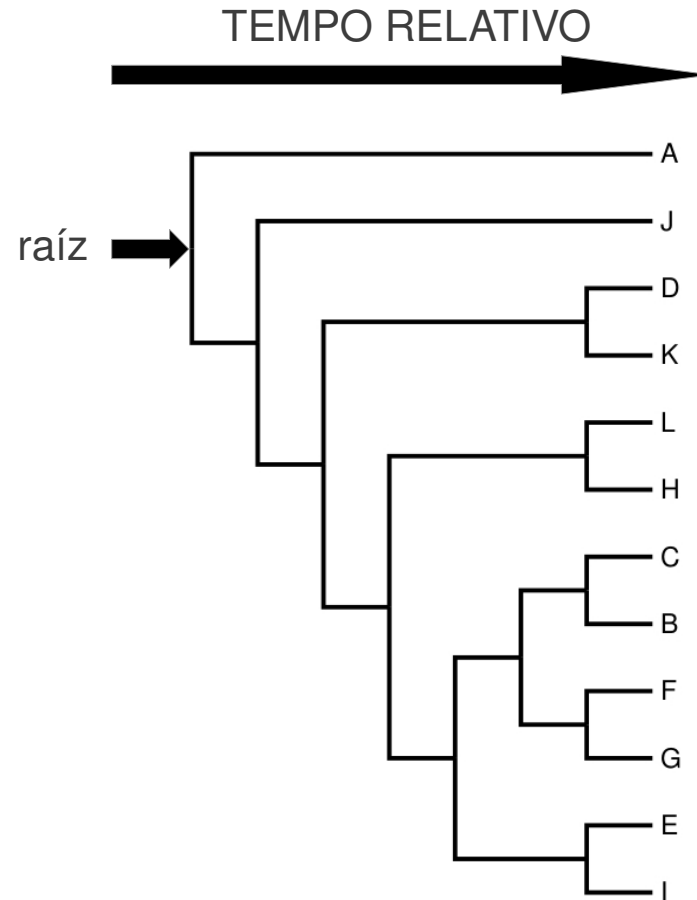


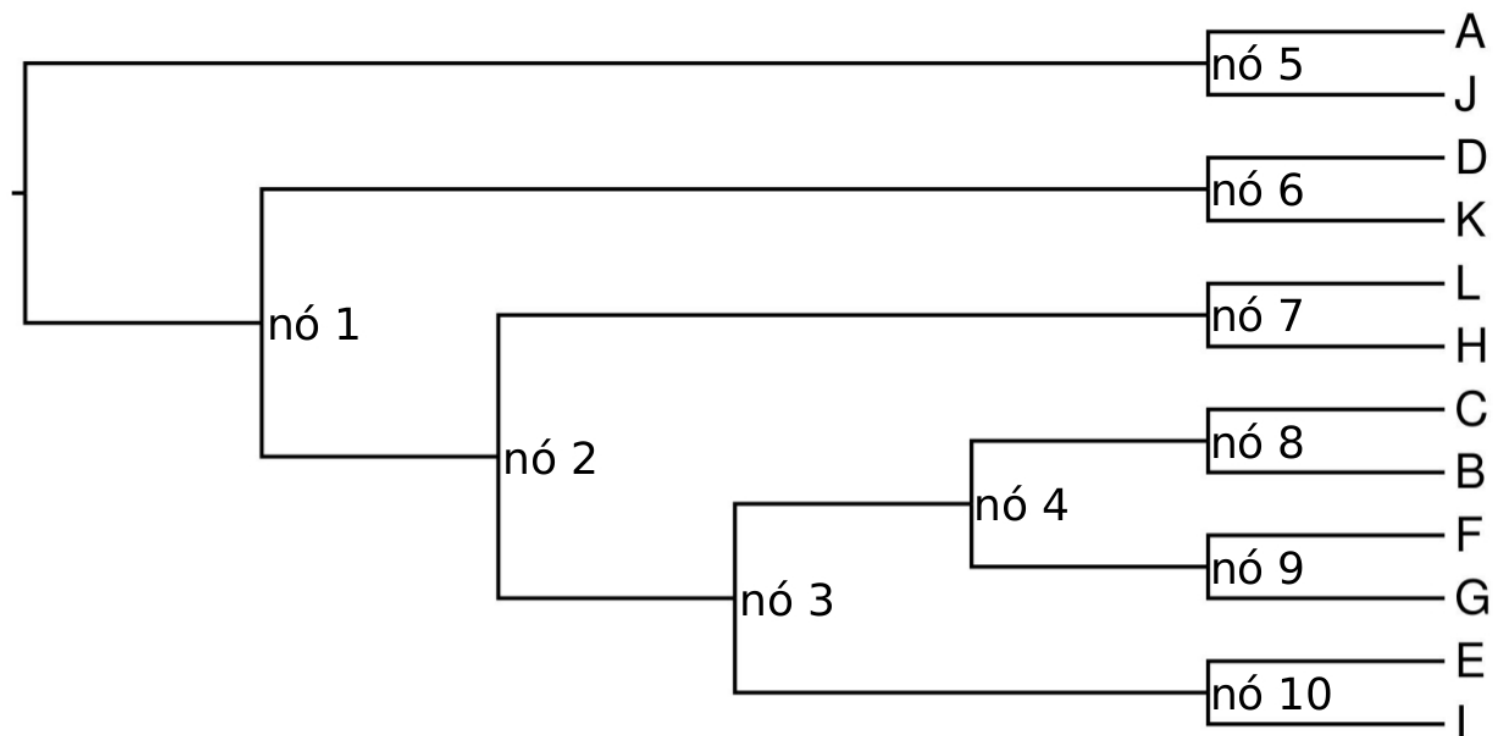
diagrama enraizado

(grafo binário acíclico direcionado)

*Diagramas não-enraizados são destituídos de sentido biológico!*

# *Enraizamento de diagramas:*

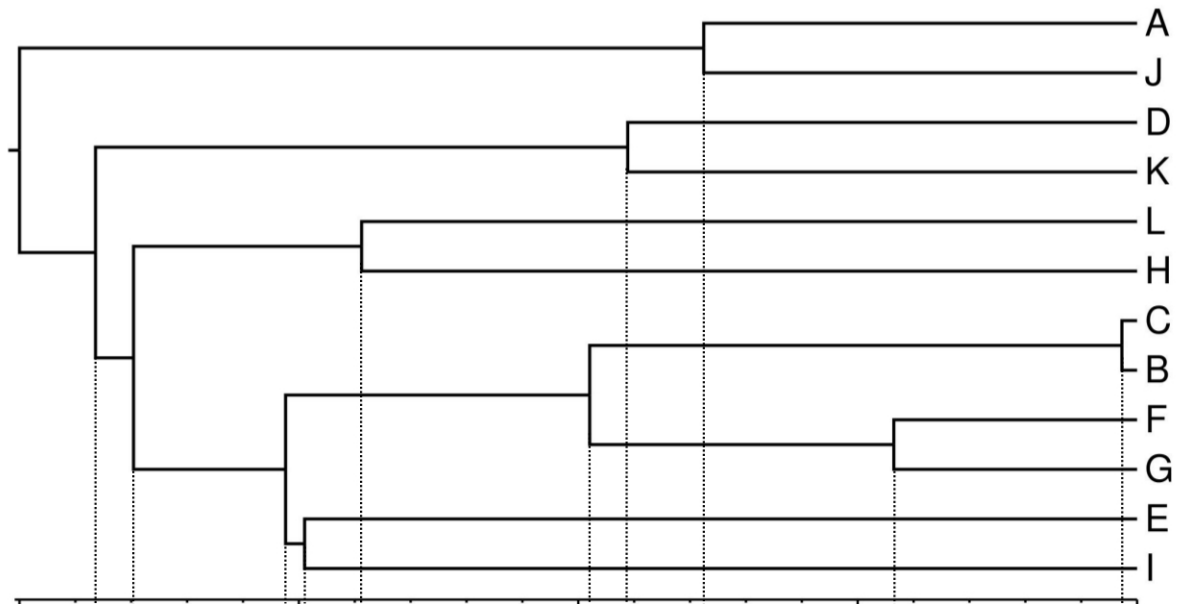
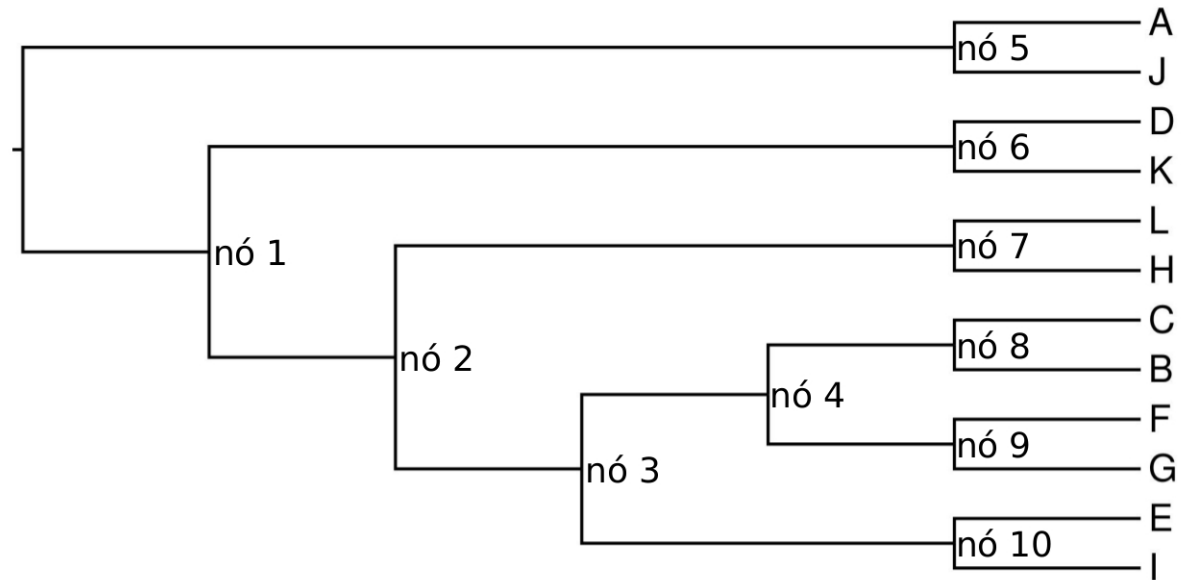
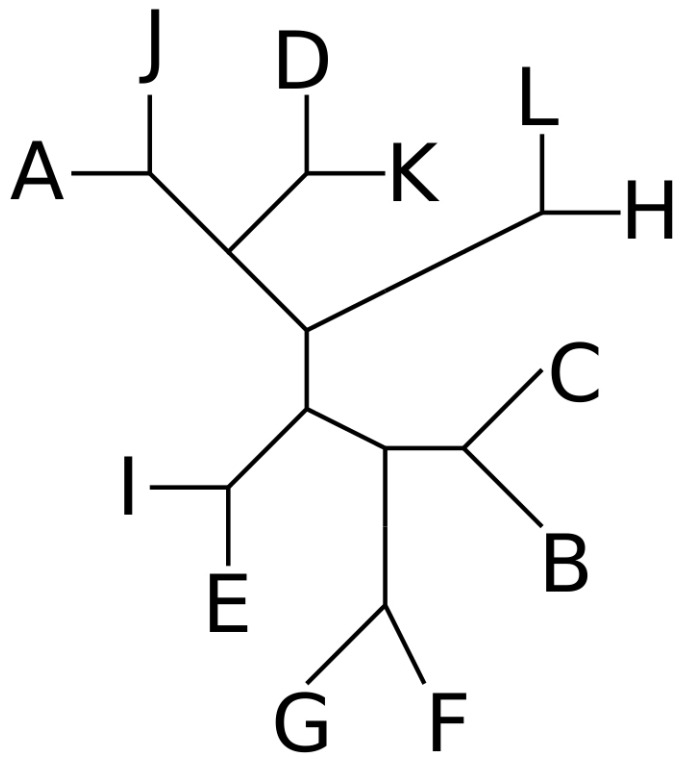
Vetor temporal relativo



*Observe a sequencia hierárquica de determinados nós.*



# *Enraizamento de diagramas:*

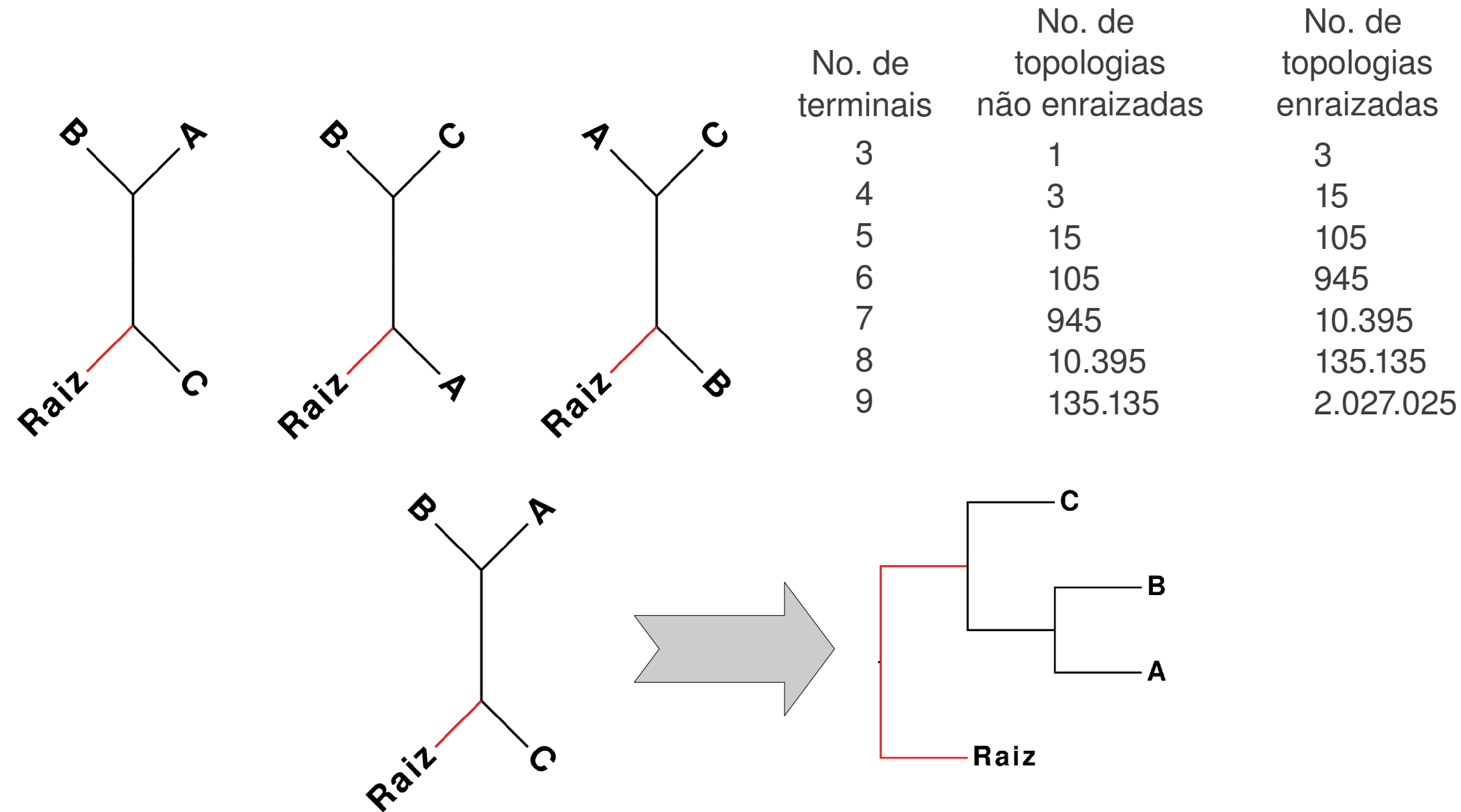


A Biologia é uma disciplina histórica, portanto, diagramas não-enraizados devem assumir simplesmente um caráter operacional em inferência filogenética.



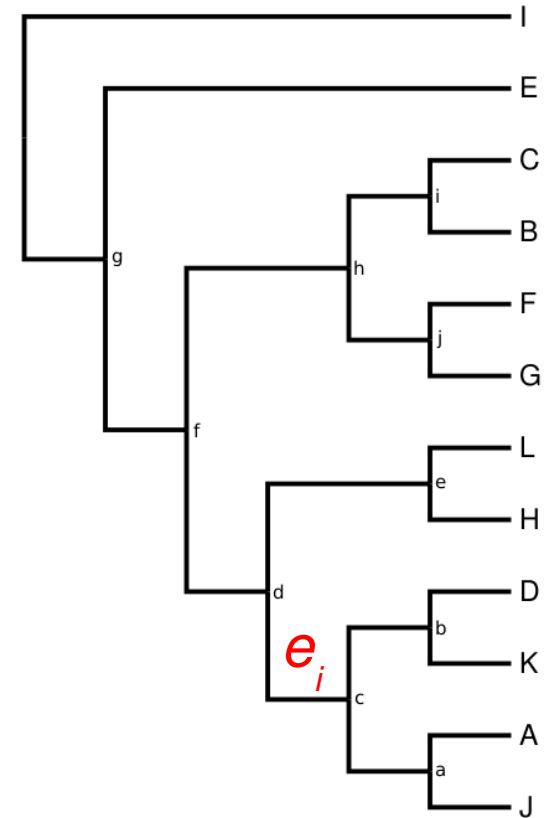
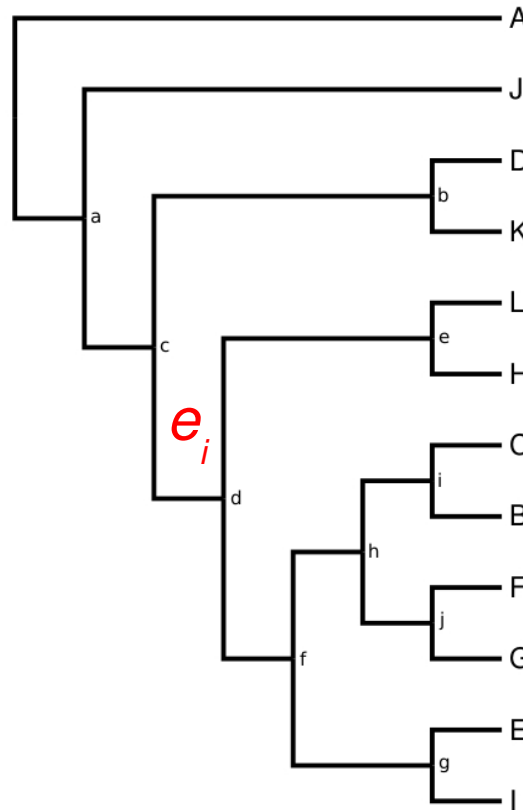
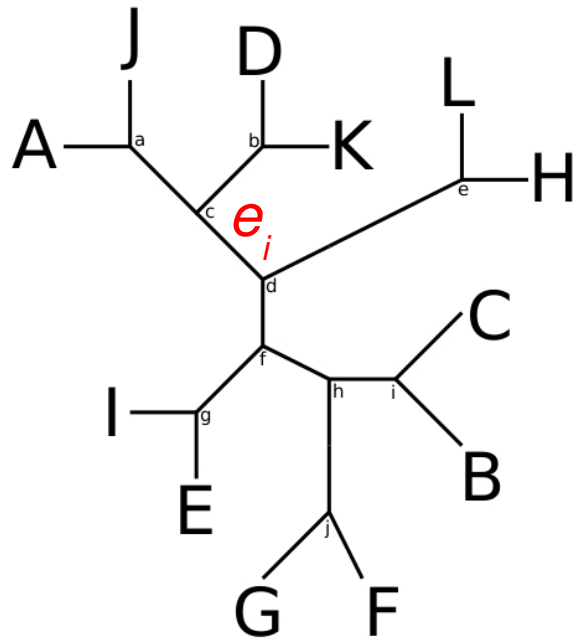
# Enraizamento de diagramas:

Propriedade 1: número de diagramas enraizados é uma ordem de grandeza em comparação aos não-enraizados (enumeração).



# Enraizamento de diagramas:

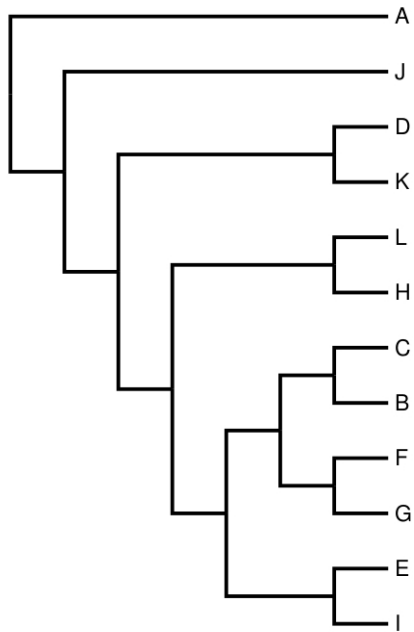
Propriedade 2: a ordem relativa de eventos de cladogênese **muda** com o enraizamento.



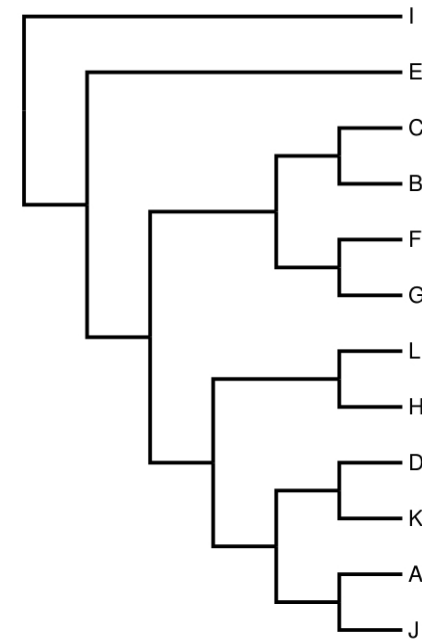
Note que a relação de adjacência destes eventos permanece a mesma.

# Enraizamento de diagramas:

Propriedade 3: a noção de grupos monofiléticos<sup>1</sup> e grupos-irmãos **muda** com o enraizamento.



{I} é monofilético  
{I,E} é monofilético  
{F,G} é monofilético  
{C,B} é monofilético  
{C,B,F,G} é monofilético  
{C,B,F,G,E,I} é monofilético

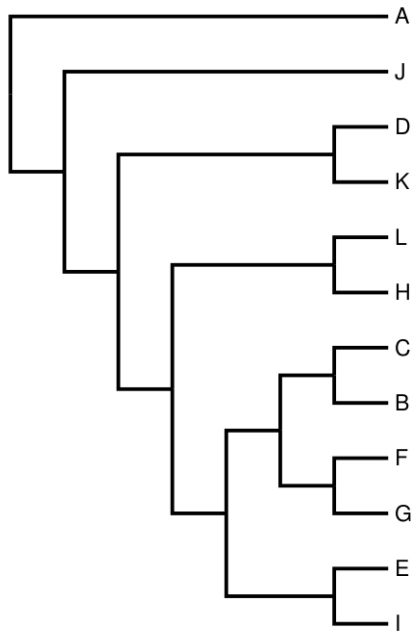


{I} é monofilético  
{I,E} **não** é monofilético  
{F,G} é monofilético  
{C,B} é monofilético  
{C,B,F,G} é monofilético  
{C,B,F,G,E,I} **não** é monofilético

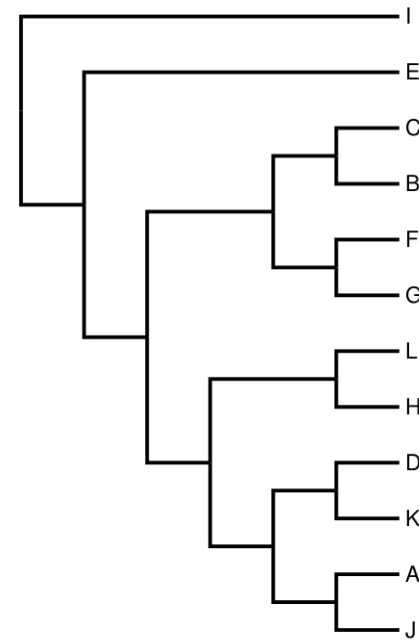
<sup>1</sup>. Formado por terminais que descendem de uma ancestral comum exclusivo.

# Enraizamento de diagramas:

Propriedade 3: a noção de grupos monofiléticos e grupos-irmãos<sup>1</sup> **muda** com o enraizamento.



{C,B} e {F,G} são grupos-irmãos  
{C} e {B} são grupos-irmãos  
{C,B,F,G} e {E,I} são grupos-irmãos  
{C,B,F,G,E,I} e {L,H} são grupos-irmãos

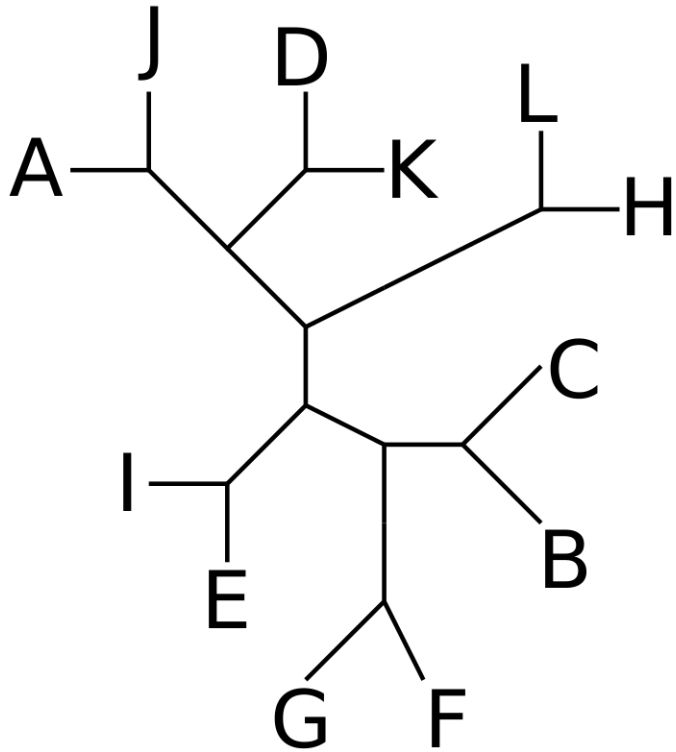


{C,B} e {F,G} são grupos-irmãos  
{C} e {B} são grupos-irmãos  
{C,B,F,G} e {E,I} **não** são grupos-irmãos  
{C,B,F,G,E,I} e {L,H} **não** são grupos-irmãos

<sup>1</sup>. **Dois** grupos que compartilham o mesmo ancestral comum exclusivo.

## *Enraizamento de diagramas:*

Propriedade 3: a noção de grupos monofiléticos e grupos-irmãos **muda** com o enraizamento.



**ATENÇÃO:** *Diagramas não-enraizados NÃO permitem definir grupos monofiléticos e grupos irmãos. Estes diagramas permitem apenas acessar noções de adjacência entre terminais.*



# Enumeração:

3 1  
4 3      Para topologias não direcionadas e  $n \geq 3$ :  
5 15

6 105

7 945

8 10395

9 135135

10 2027025

11 34459425

12 654729075

13 13749310575

14 316234143225

15 7905853580625

16 213458046676875

17 6190283353629375

18 191898783962510625

19 6332659870762850625

20 221643095476699771875

21 8200794532637891559375

22 319830986772877770815625

23 13113070457687988603440625

24 563862029680583509947946875

25 25373791335626257947657609375

26 1192568192774434123539907640625

27 58435841445947272053455474390625

28 2980227913743310874726229193921875

29 157952079428395476360490147277859375

30 8687364368561751199826958100282265625

31 495179769008019818390136611716089140625

32 29215606371473169285018060091249259296875

33 1782151988659863326386101665566204817109375

34 112275575285571389562324404930670903477890625

35 7297912393562140321551086320493608726062890625

36 488960130368663401543922783473071784646213671875

37 33738248995437774706530672059641953140588743359375

38 2395415678676082004163677716234578672981800778515625

39 174865344543353986303948473285124243127671456831640625

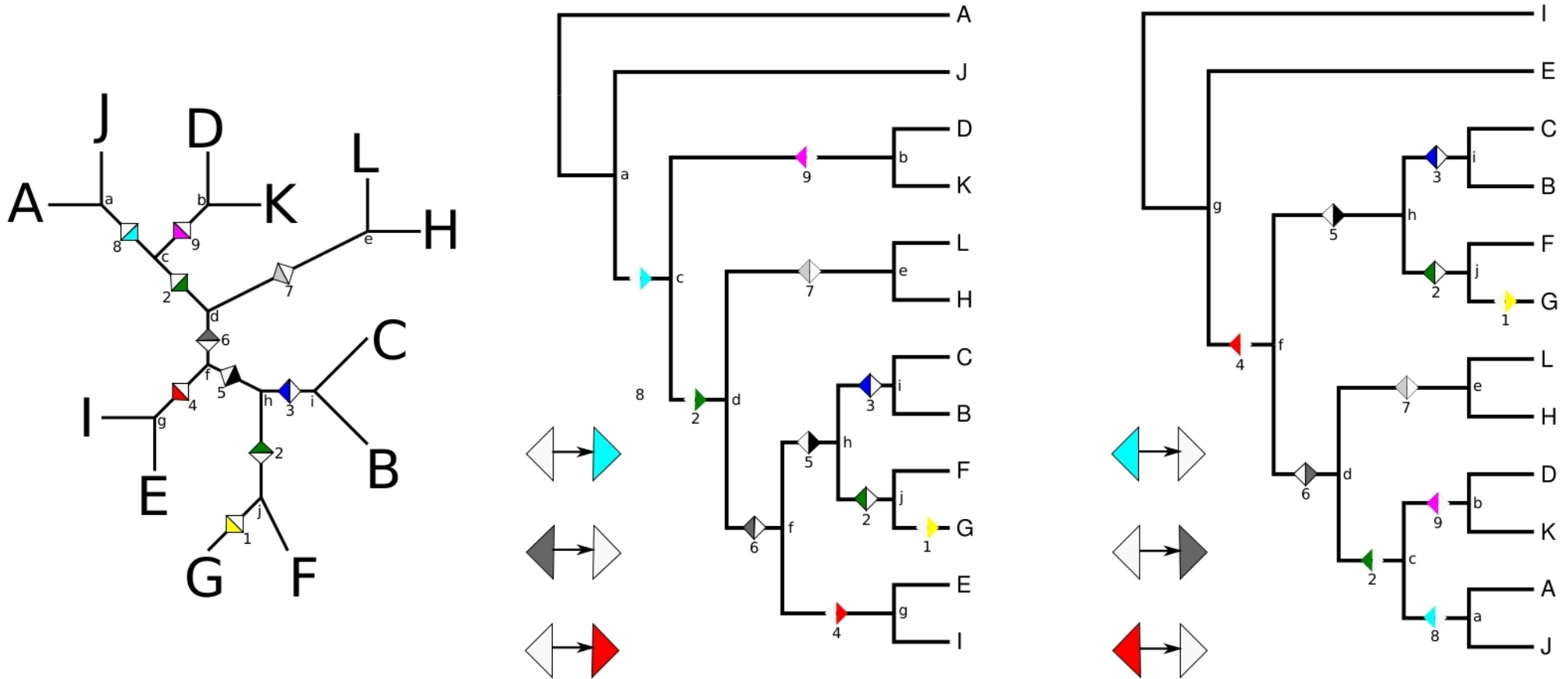
40 13114900840751548972796135496384318234575359262373046875

$$\frac{(2n - 4)!}{(n - 2)! 2^{n-2}}$$

O número de topologias enraizadas pode ser calculado multiplicando a fórmula acima pelo número de ramos  $(2n-3)$  ou incrementando  $+1$  à  $n$ .

# Enraizamento de diagramas:

Propriedade 5: Note que a direção da transformação **pode mudar** dependendo do enraizamento.

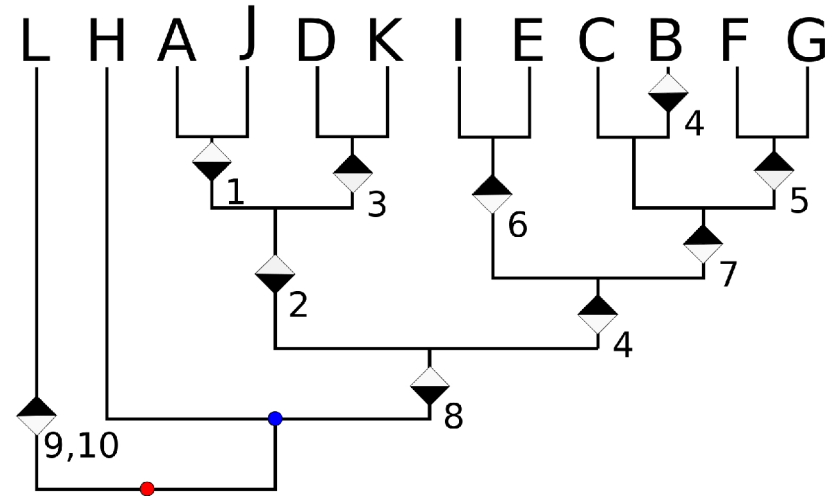
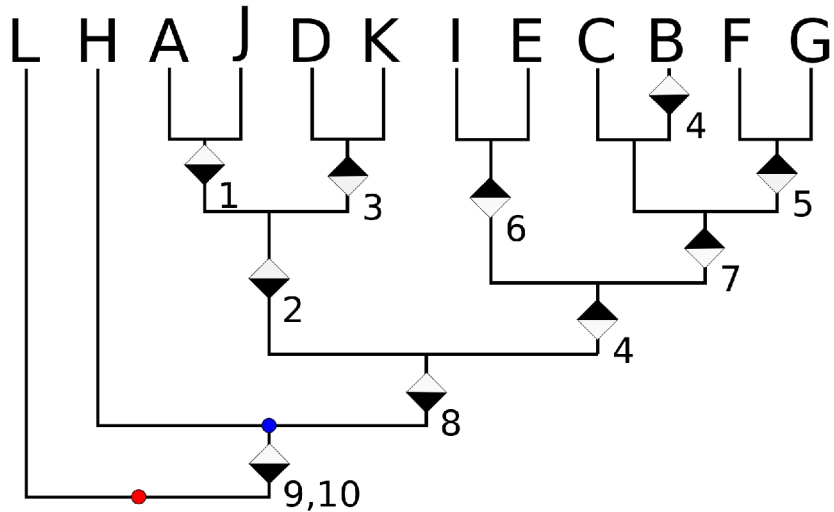


Note os caracteres que sofreram mudanças de **polarização**.



# Enraizamento de diagramas:

Propriedade 5: Ambiguidade na região da raiz.



POLARIZAÇÃO:

- 1: 1  $\Rightarrow$  0
- 2: 1  $\Rightarrow$  0
- 3: 0  $\Rightarrow$  1
- 4: 0  $\Rightarrow$  1  $\Rightarrow$  0
- 5: 0  $\Rightarrow$  1
- 6: 1  $\Rightarrow$  0
- 7: 0  $\Rightarrow$  1
- 8: 1  $\Rightarrow$  0
- 9: 1  $\rightarrow$  0
- 10: 1  $\rightarrow$  0

POLARIZAÇÃO:

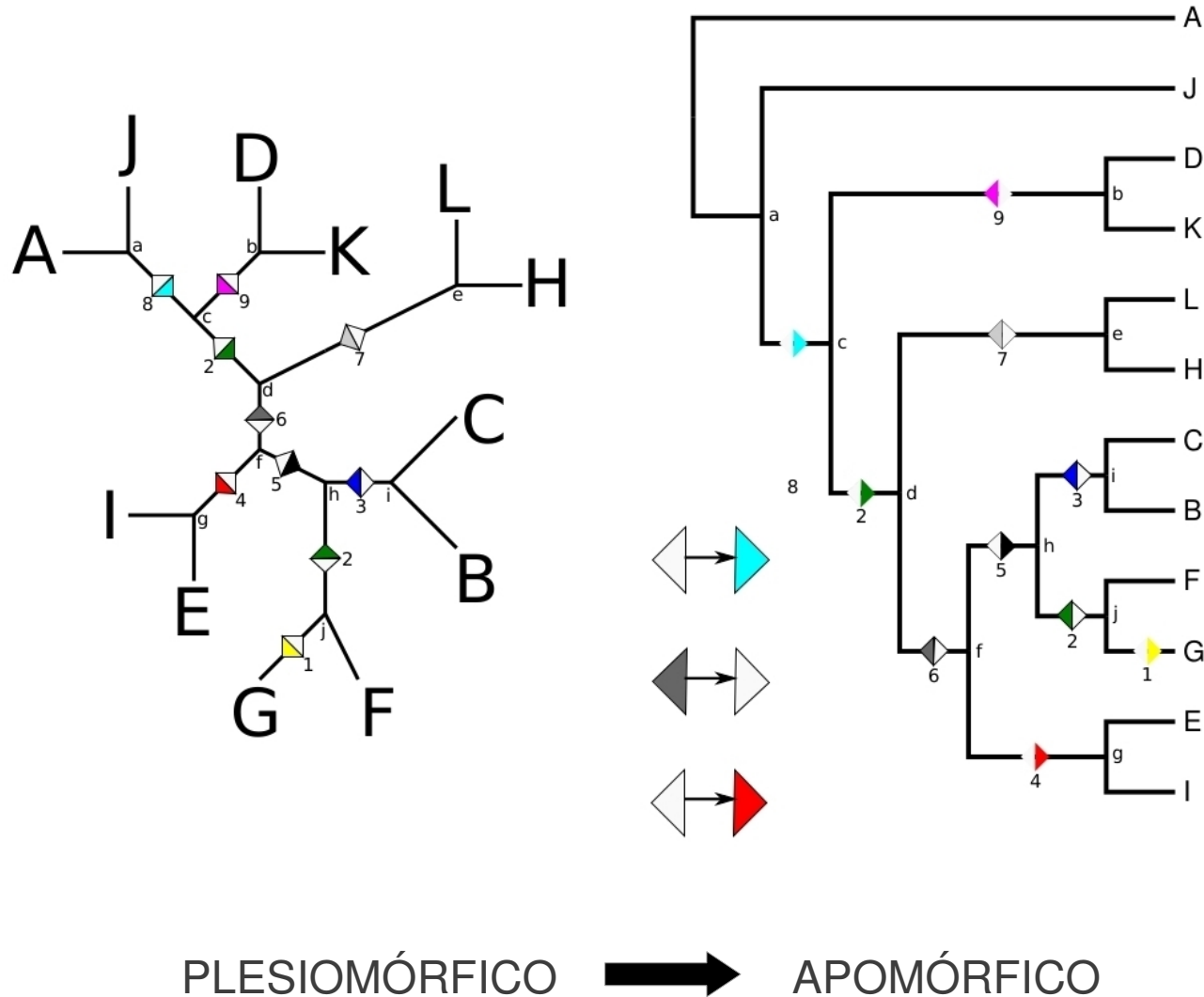
- 1: 1  $\Rightarrow$  0
- 2: 1  $\Rightarrow$  0
- 3: 0  $\Rightarrow$  1
- 4: 0  $\Rightarrow$  1  $\Rightarrow$  0
- 5: 0  $\Rightarrow$  1
- 6: 1  $\Rightarrow$  0
- 7: 0  $\Rightarrow$  1
- 8: 1  $\Rightarrow$  0
- 9: 0  $\rightarrow$  1
- 10: 0  $\rightarrow$  1

$\rightarrow$  = ambiguidade

$\Rightarrow$  = sem ambiguidade

# *Enraizamento e termos para estados de caráter:*

Em diagramas enraizados, todas as transformações envolvem a mudança de um estado **plesiomórfico** para um estado **apomórfico**.



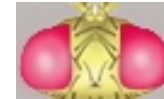
# Codificação & Evidência de grupos

## Fenética:

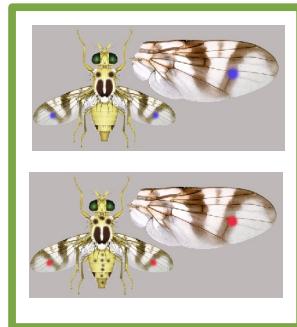
EVIDÊNCIAS: similaridade global

CRITÉRIO DE OTIMALIDADE: Distância (fenética)

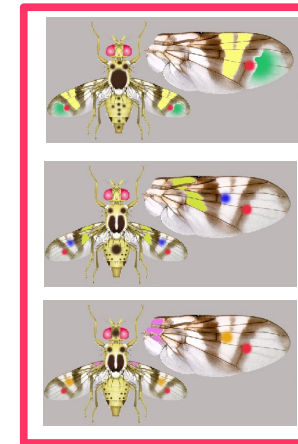
OBSERVAÇÃO: olhos apresentam duas cores:



Este caráter sugere dois grupos:



E



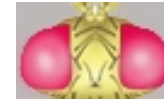
# Codificação & Evidência de grupos

## Cladística:

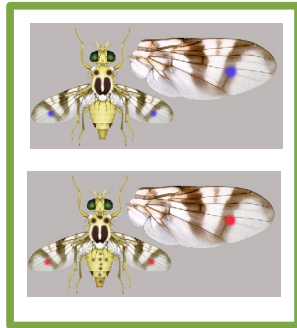
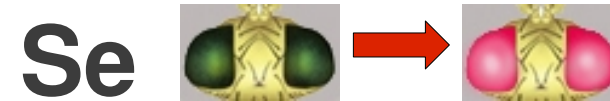
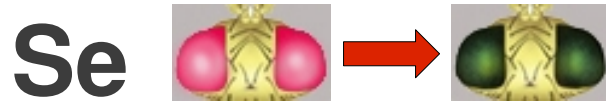
EVIDÊNCIAS: séries de transformações

CRITÉRIO DE OTIMALIDADE: Número mínimo de transformações

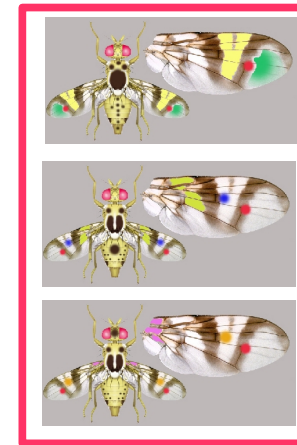
OBSERVAÇÃO: olhos apresentam duas cores:



Este caráter sugere um desses grupos:



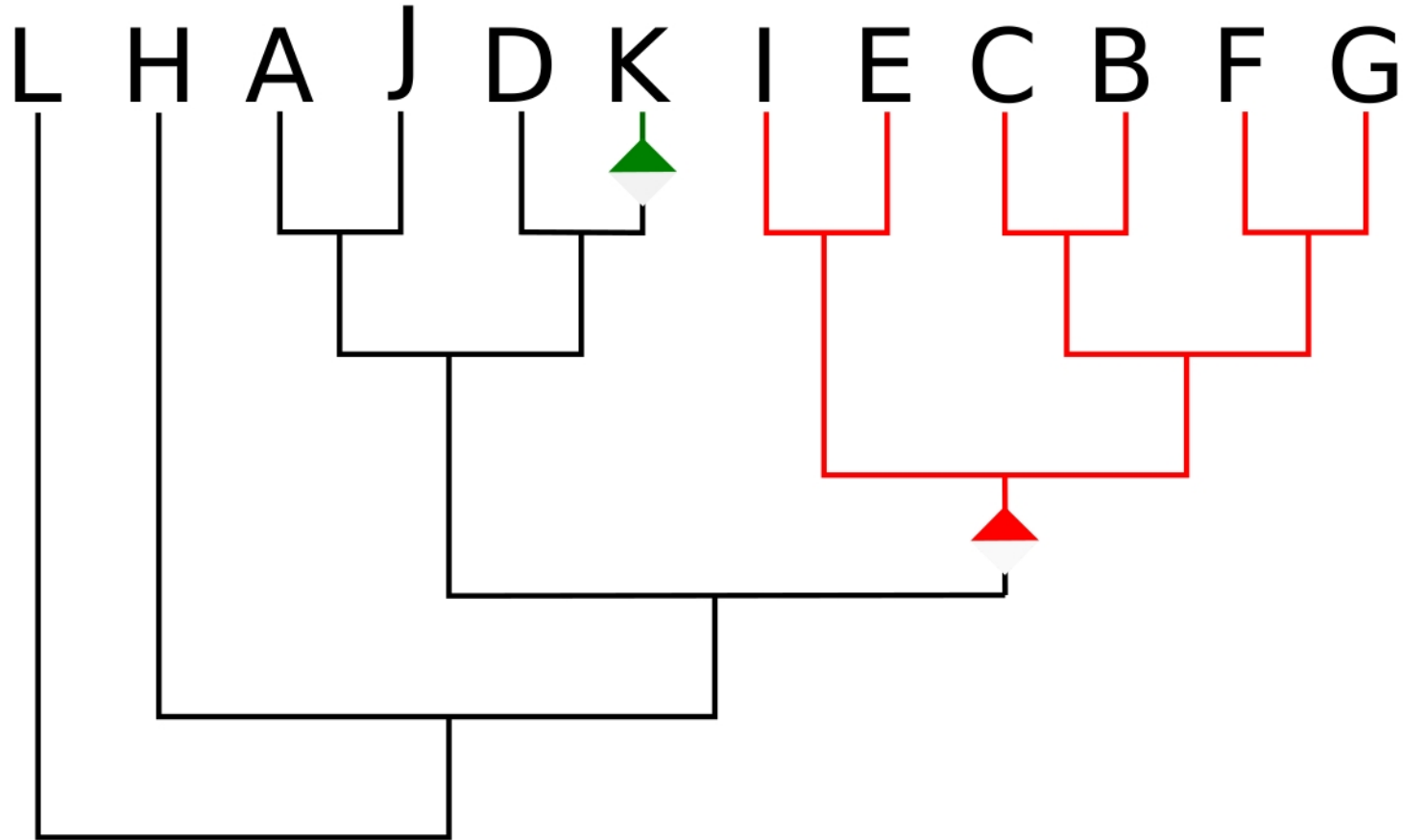
**ou**



# *Enraizamento e termos para estados de caráter:*

Apomorfias: compartilhadas e não compartilhadas.

PLESIOMÓRFICO  $\longrightarrow$  APOMÓRFICO



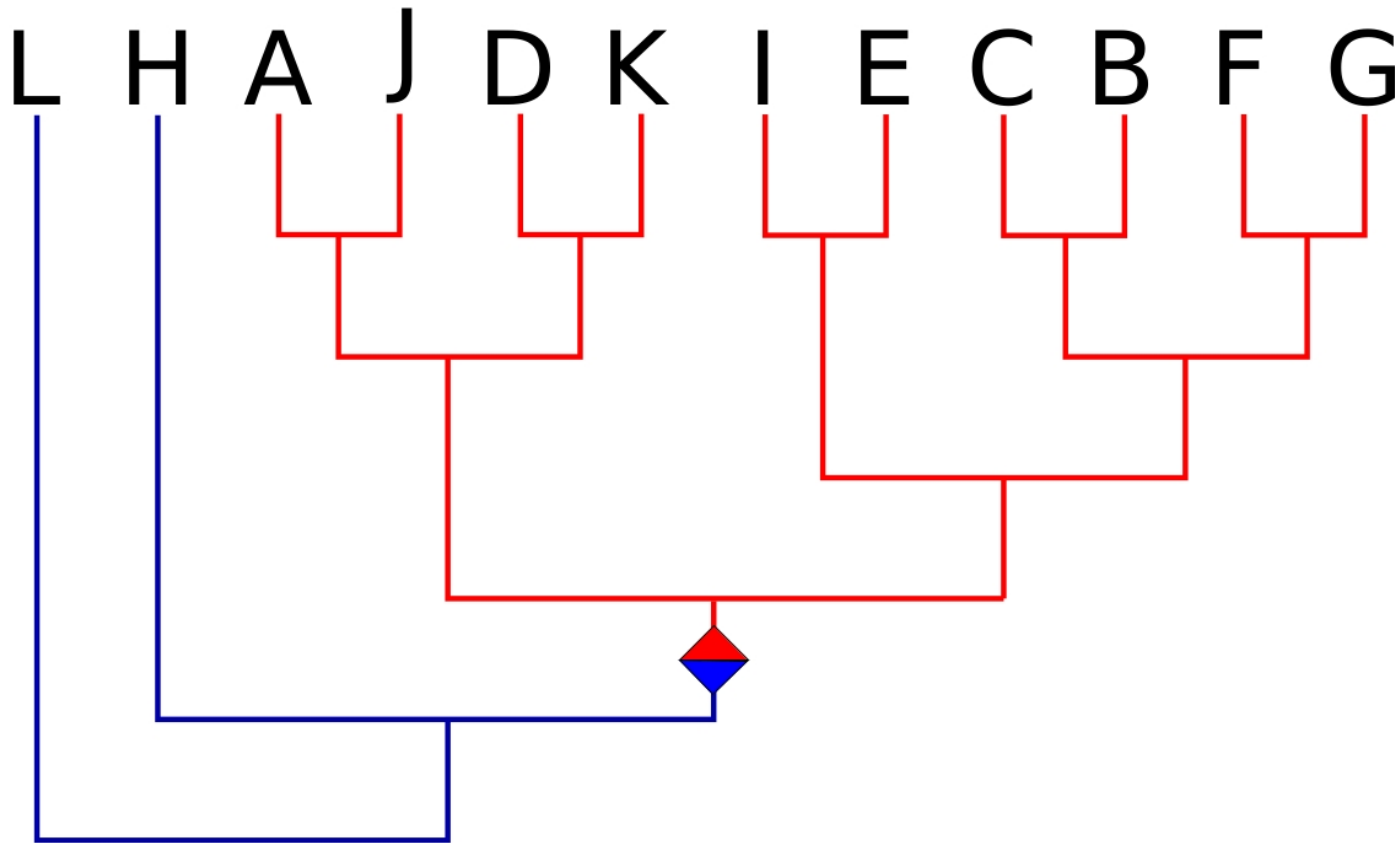
**SINAPOMORFIAS:** Apomorfias compartilhadas.

**AUTAPOMORFIAS:** Apomorfias não compartilhadas.

# *Enraizamento e termos para estados de caráter:*

Plesiomorfias compartilhadas.

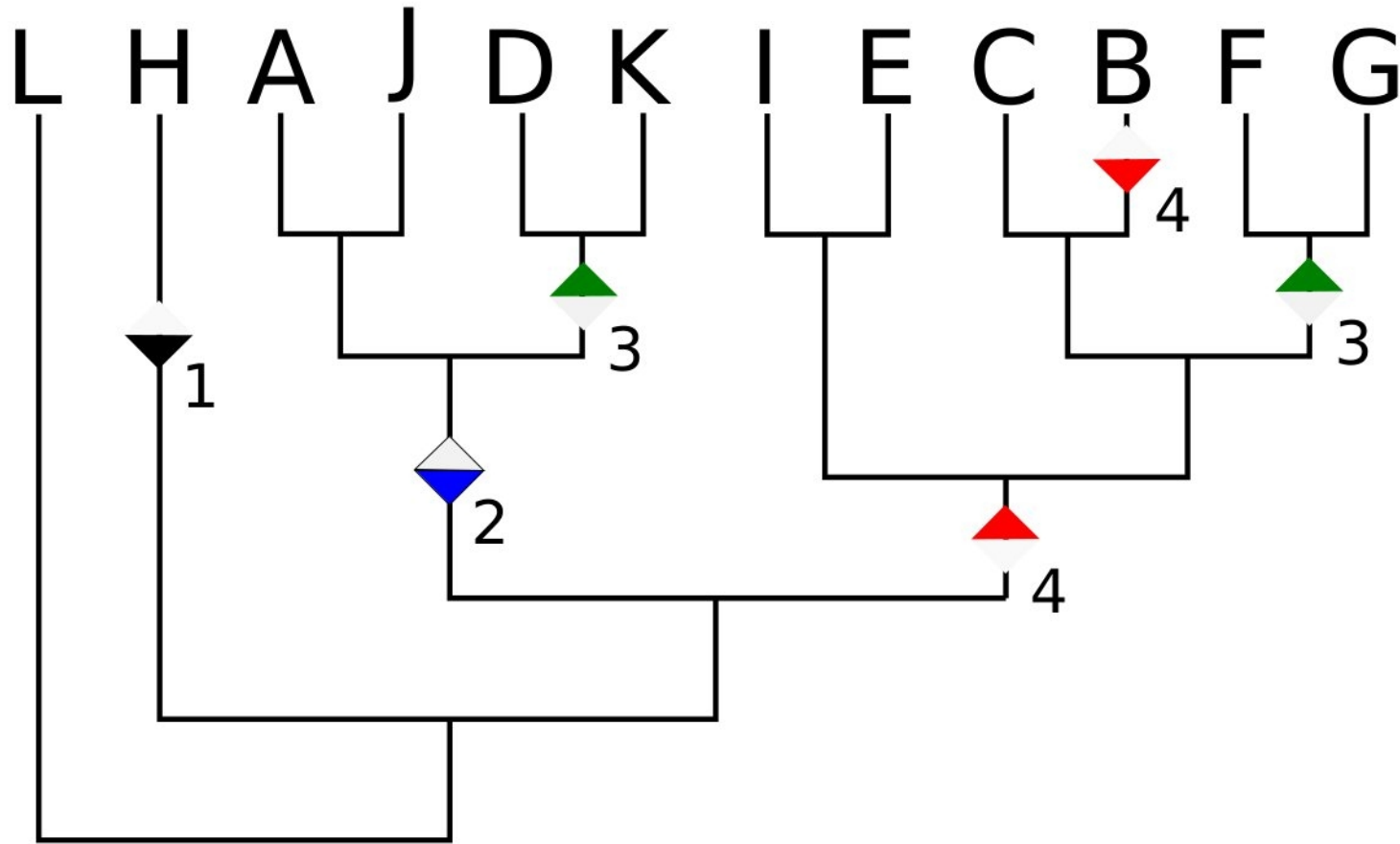
PLESIOMÓRFICO  $\longrightarrow$  APOMÓRFICO



**SIMPLESIOMORFIAS:** Plesiomorfias compartilhadas pelos membros do grupo {L H}.

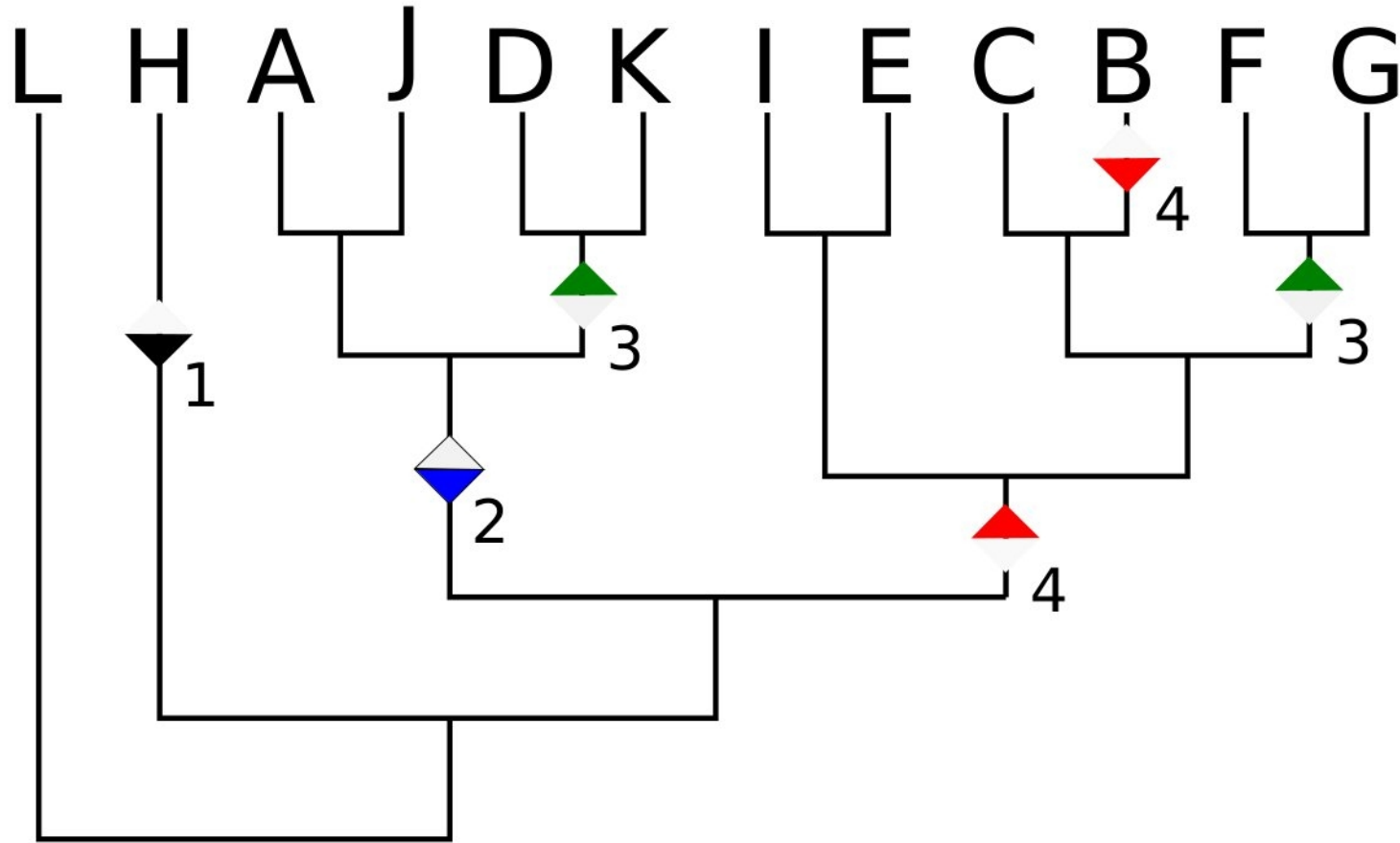
# *Enraizamento e termos para estados de caráter:*

Quais seriam os termos aplicáveis aos caracteres abaixo?



# *Enraizamento e termos para estados de carácter:*

Quais seriam os termos aplicáveis aos caracteres abaixo?



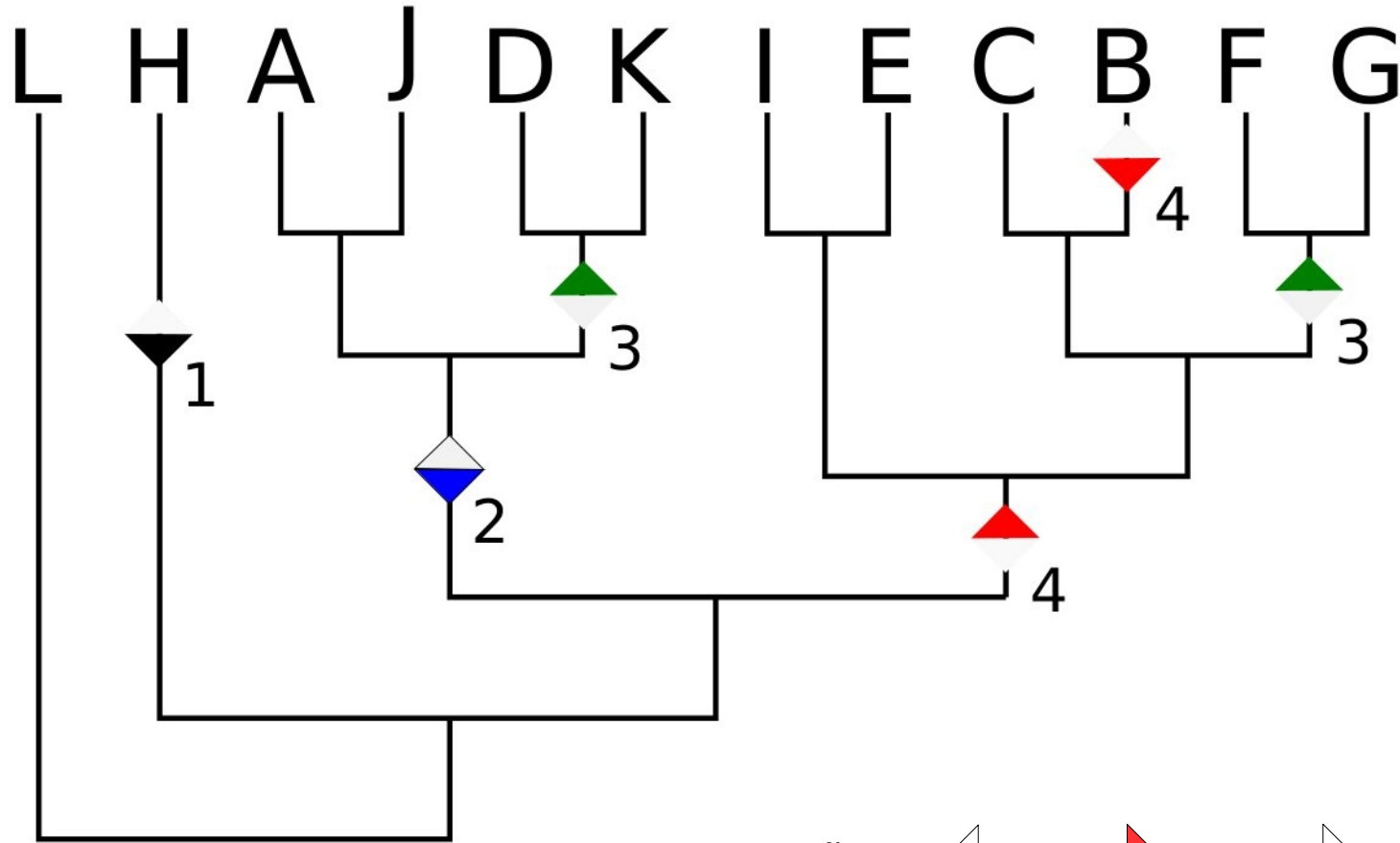
## **Exemplo:**

Carácter 3: homoplástico, sinapomorfia para D e K e para F e G.



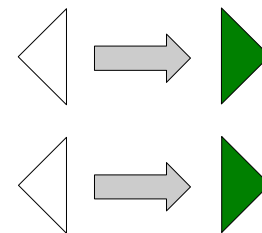
# *Enraizamento e termos para estados de caráter:*

Quais seriam os termos aplicáveis aos caracteres abaixo?



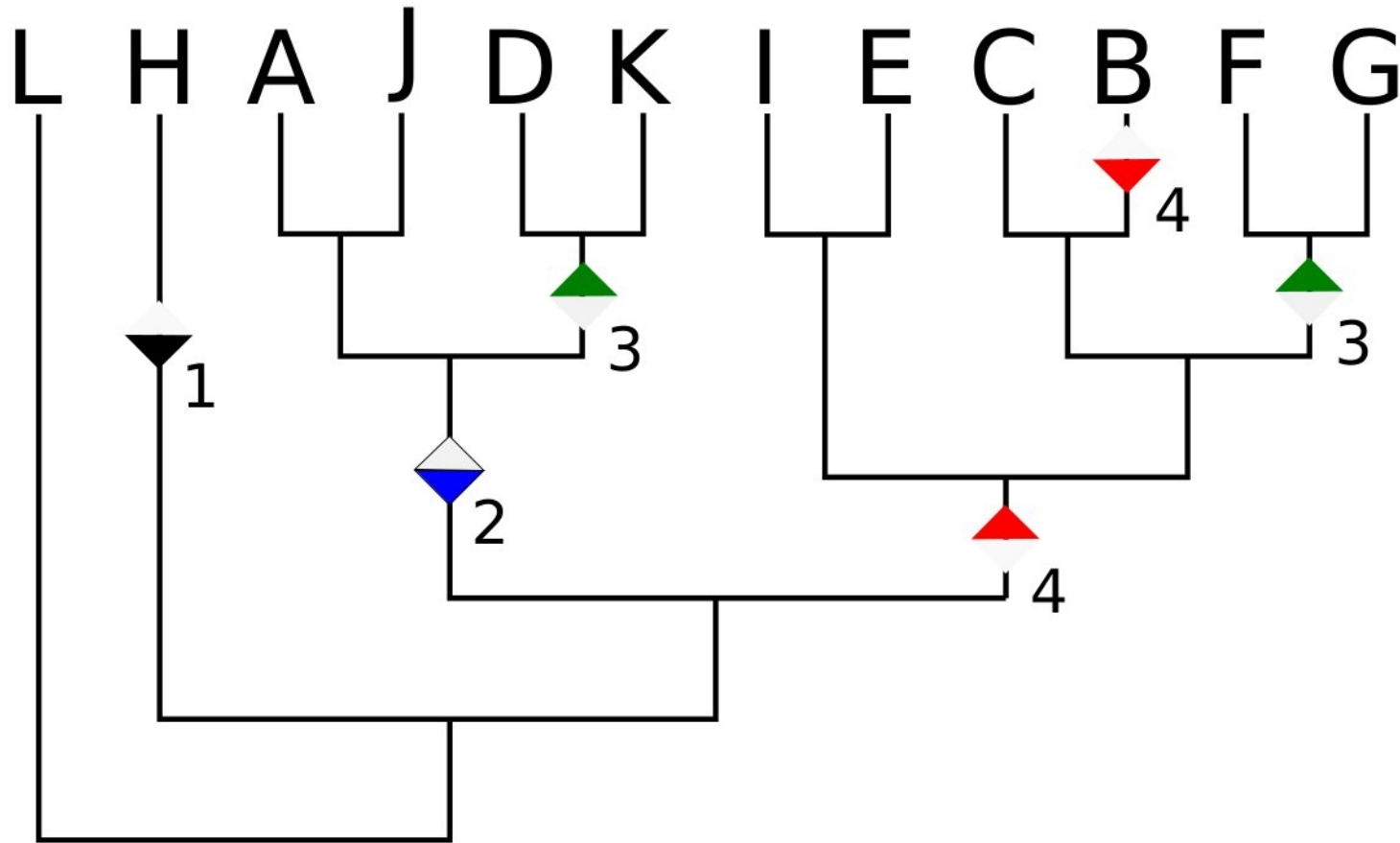
REVERSÃO:

CONVERGÊNCIA:



# *Enraizamento e termos para estados de caráter:*

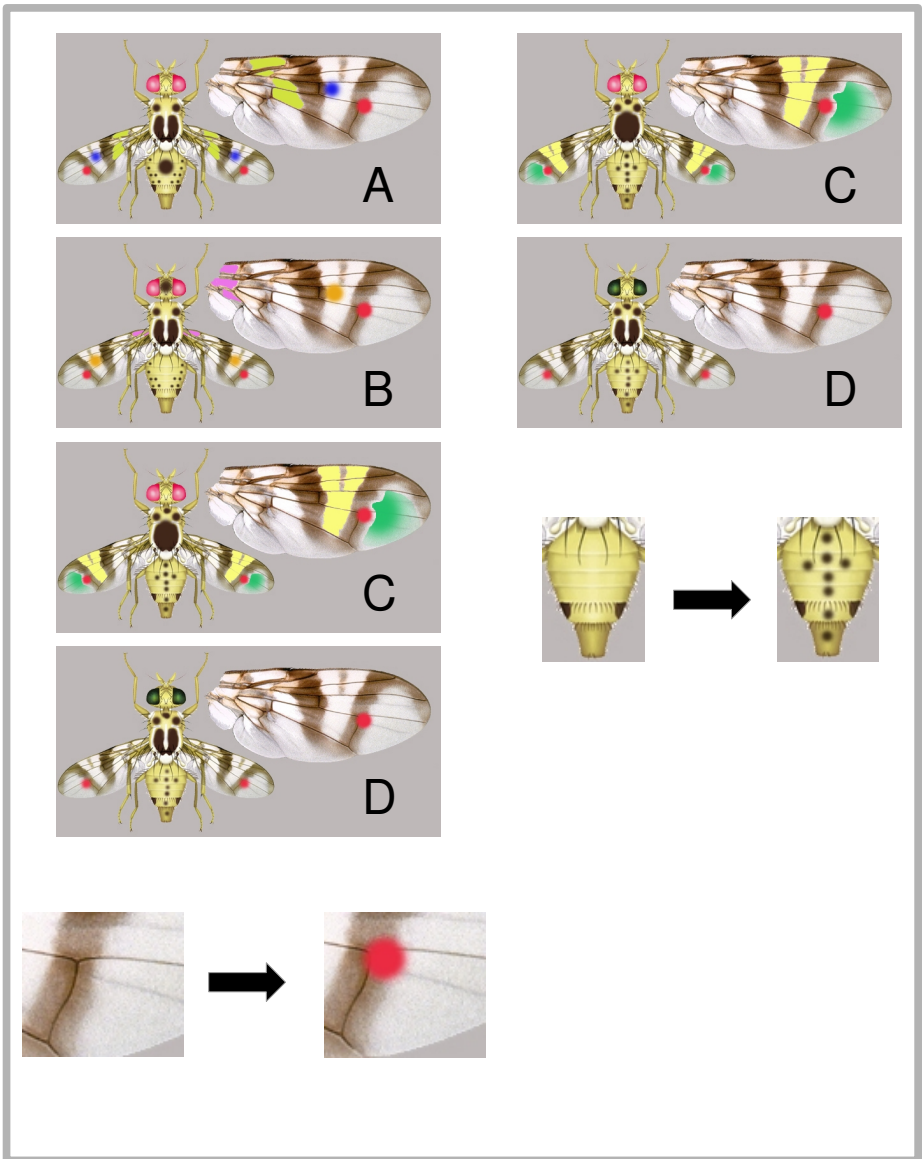
*Qual a relação entre homologia, sinapomorfias, simplesiomorfias e homoplasia?*



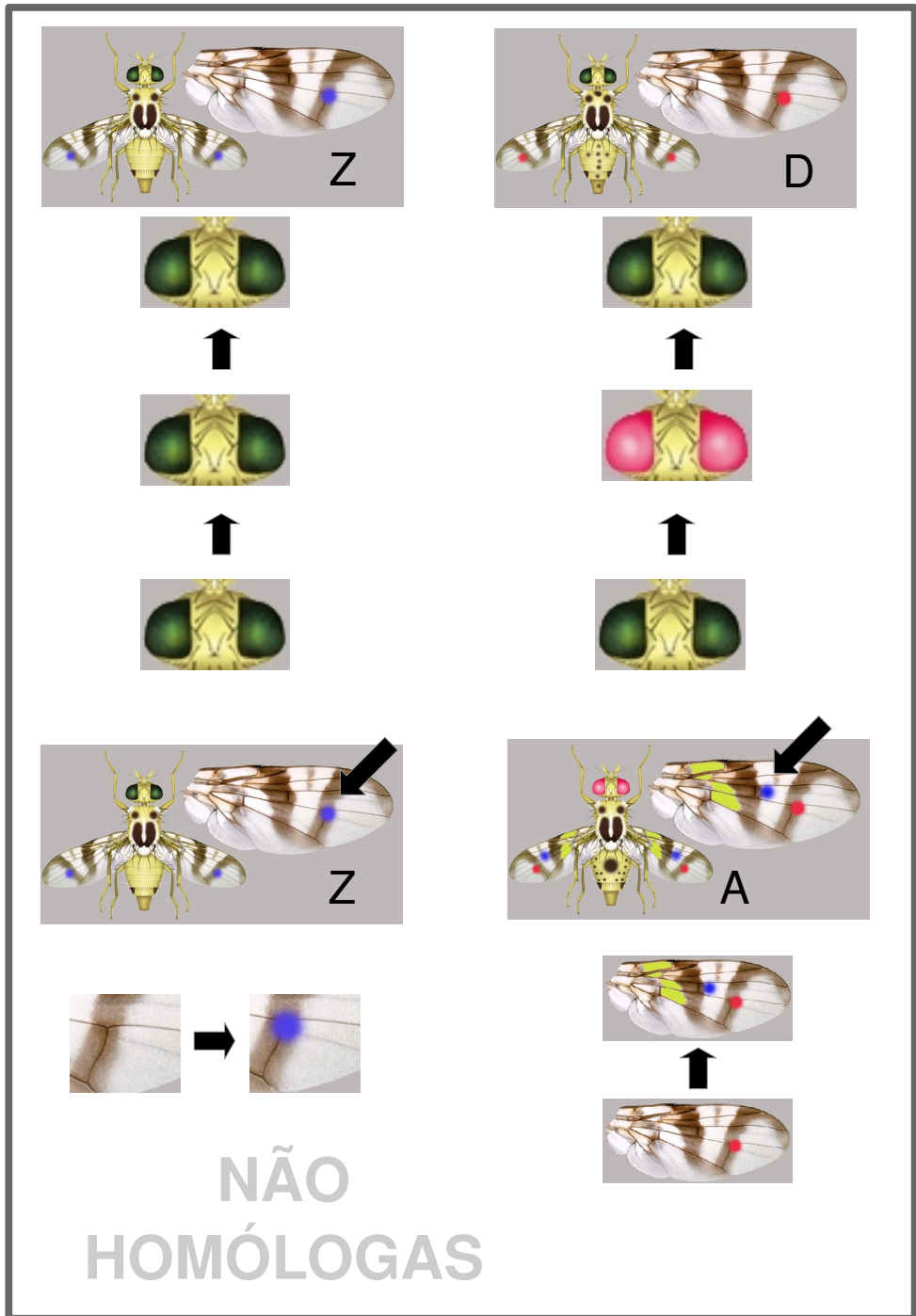
# Compartilhamento de semelhanças:



# Compartilhamento de semelhanças:



Características HOMÓLOGAS



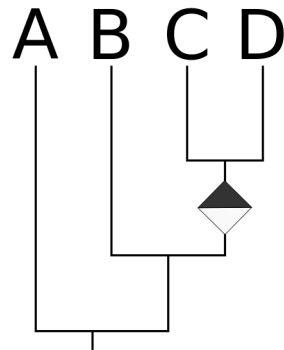
NÃO HOMÓLOGAS

# *Enraizamento e termos para estados de caráter:*

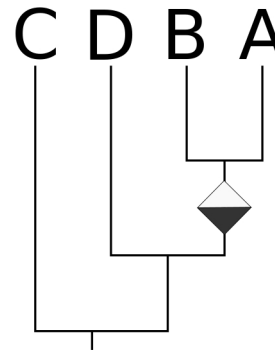
*Qual a relação entre homologia, sinapomorfias, simplesiomorfias e homoplasia?*

Considere:

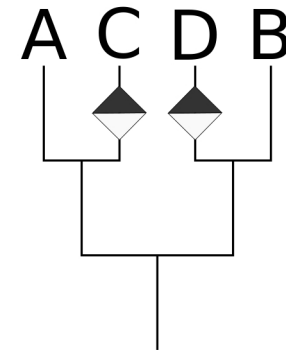
	$c_1$
A	0
B	0
C	1
D	1



Hipótese 1



Hipótese 2



Hipótese 3

**Grant & Kluge (2004):**

**Homologia** refere-se à relação entre partes que resultaram do mesmo evento de transformação.

**Sinapomorfia** refere-se ao compartilhamento de estados de caráter derivados (apomorfias), independentemente deste compartilhamento ser o resultado de uma mesmo evento de transformação (homologia) ou eventos distintos (homoplasia)

**Tautologia:**

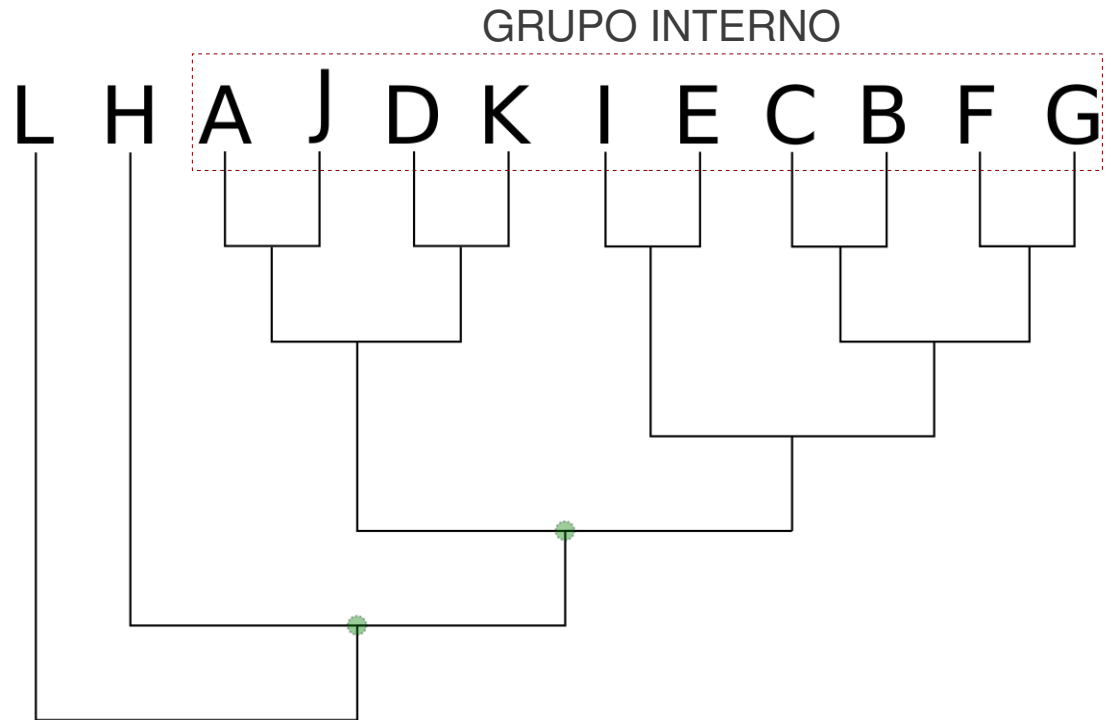
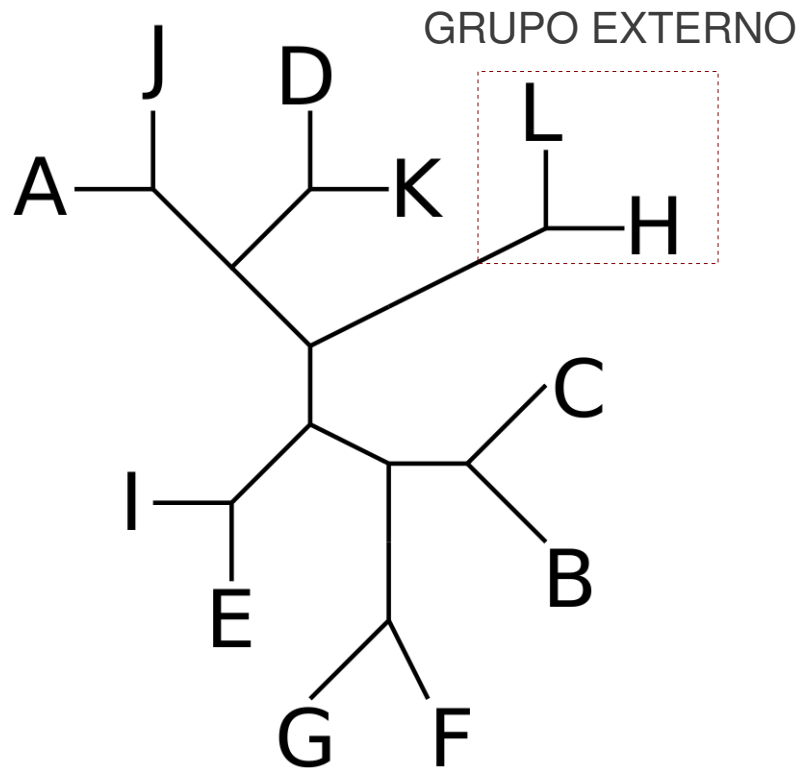
“homologia“ = “sinapomorfia“

“sinapomorfia“ = “caracteres apomórficos compartilhados“

“estados de caráter compartilhados devido à homologia“ = “homologia devido à homologia“

# Método de enraizamento:

Grupos: externos e interno

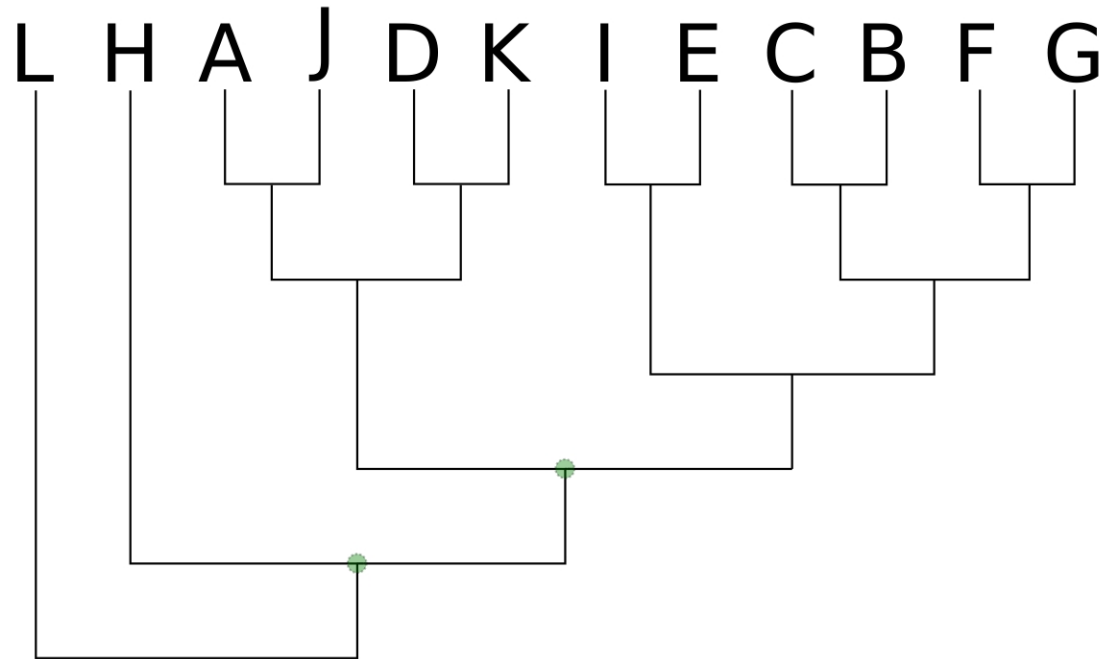
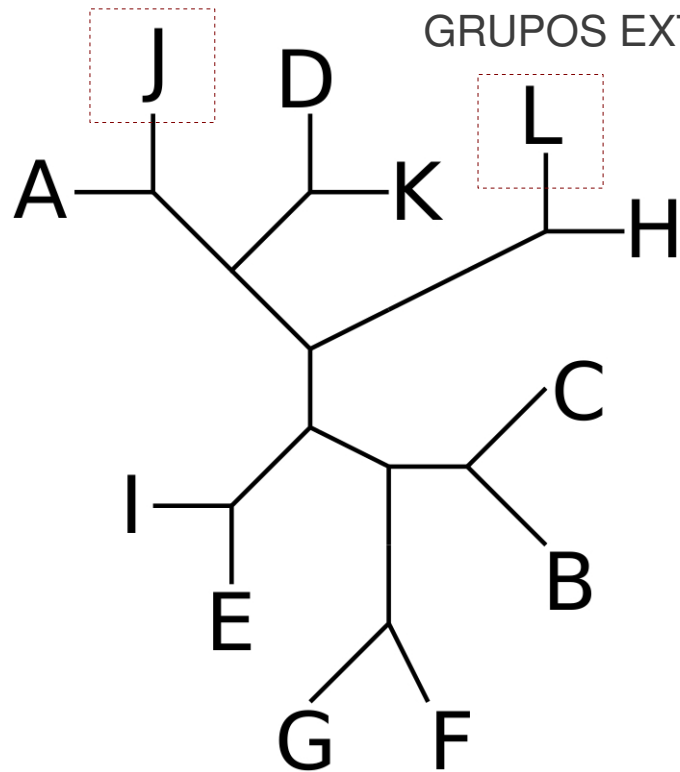


Mito 1: O grupo externo não precisa ser o grupo-irmão do grupo interno.

Mito 2: O(s) grupo(s) externo(s) não precisa(m) ser monofilético em relação ao grupo interno.

# Método de enraizamento:

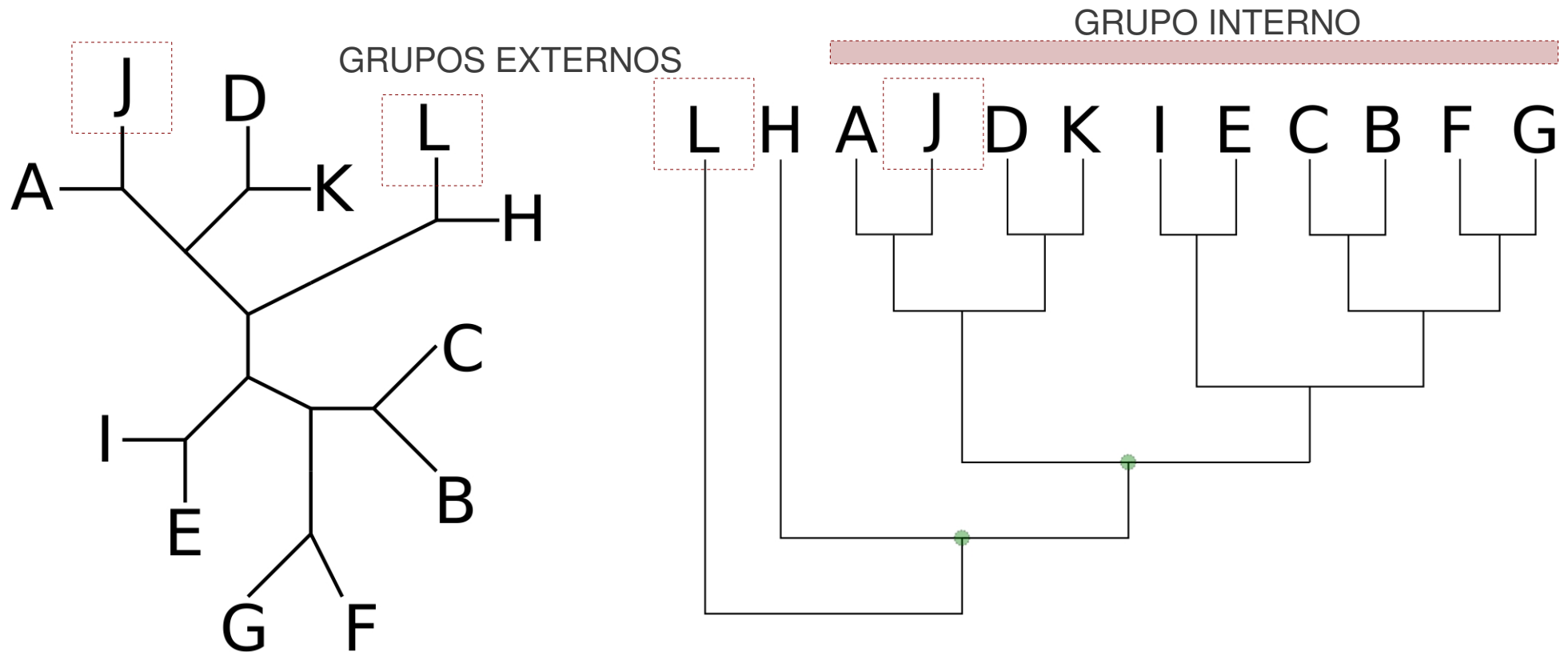
Grupos: externos e interno



*O que aconteceria se escolhêssemos estes terminais como grupos externos?*

# Método de enraizamento:

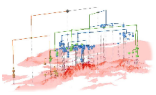
Grupos: externos e interno



*O que aconteceria se escolhêssemos estes terminais como grupos externos?*

- Violariamos a premissa de que o grupo interno é monofilético.





# ***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*

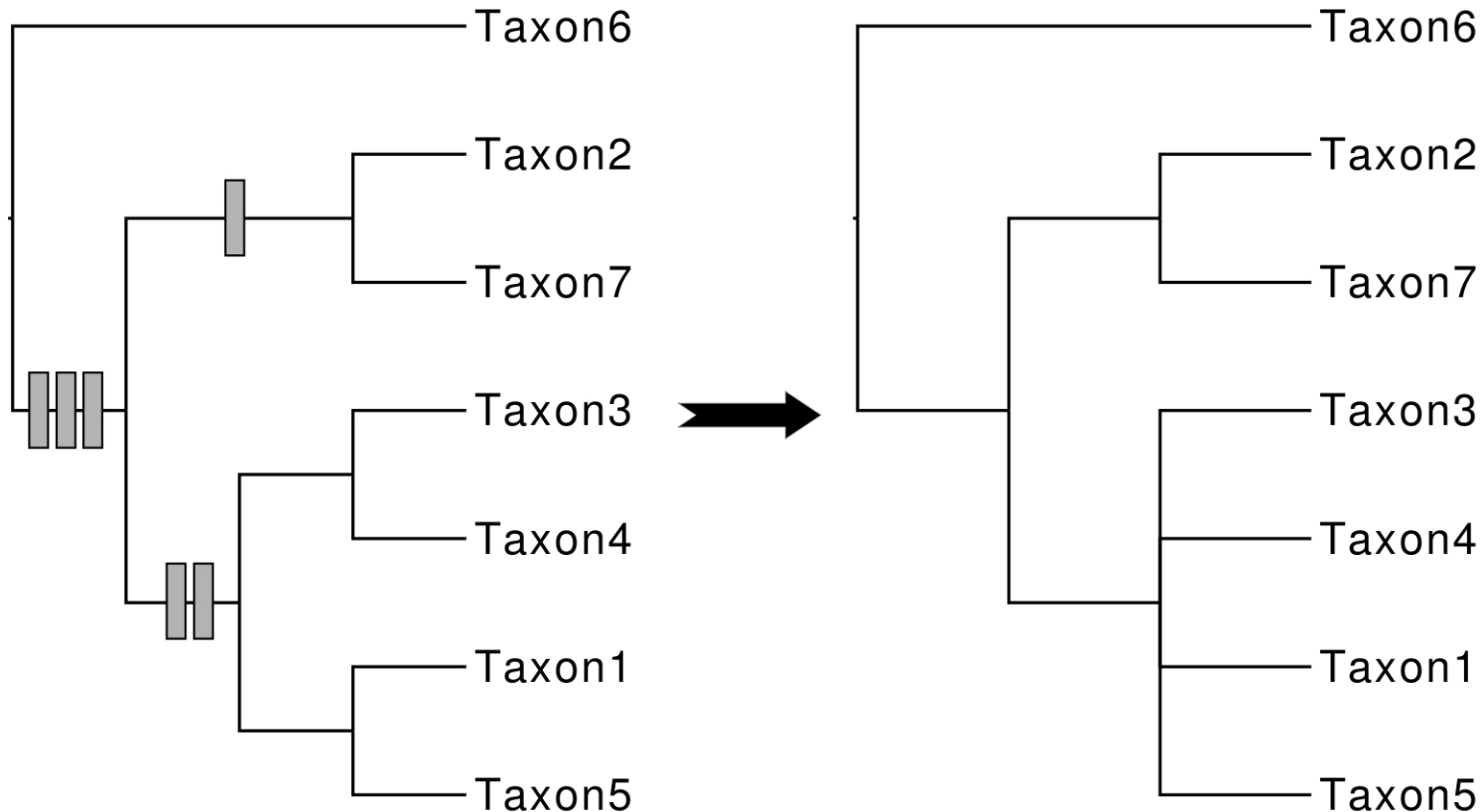
*Sinapomorfias*

*Simplesiomorfias*

*Autapomorfias*

# *Leitura de cladograma:*

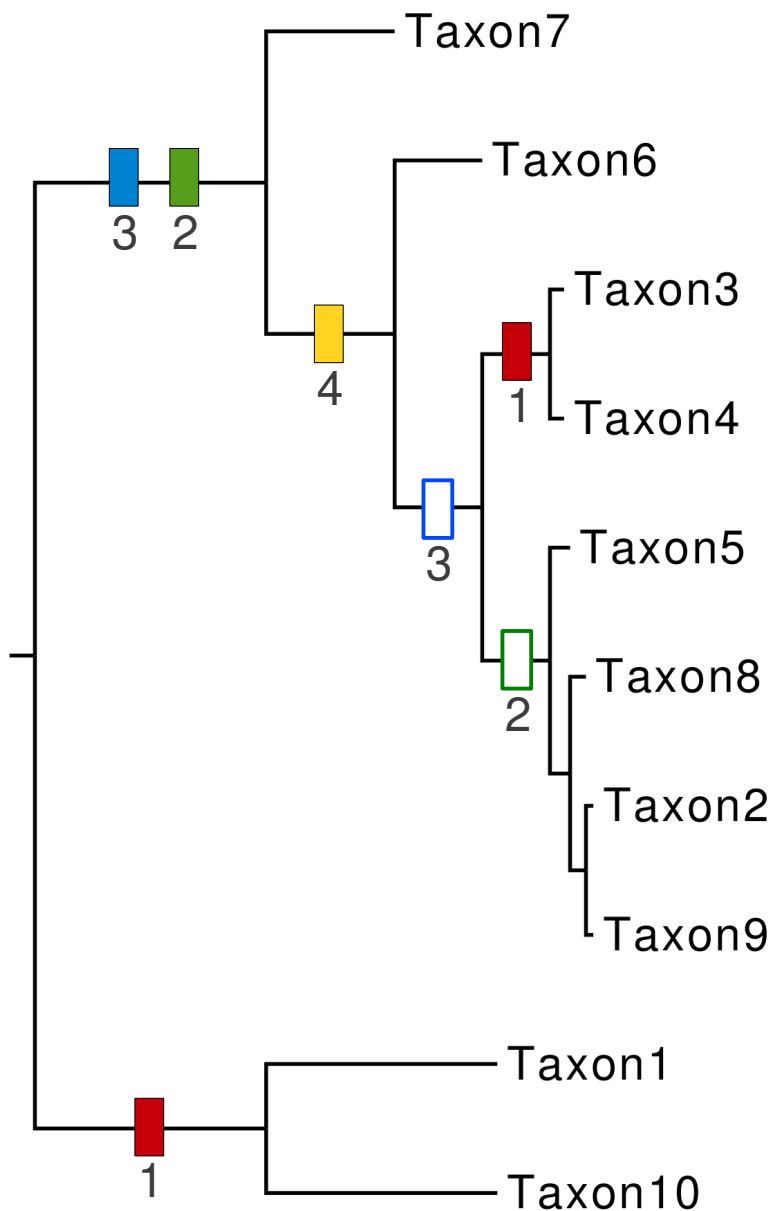
## Suporte e resolução



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

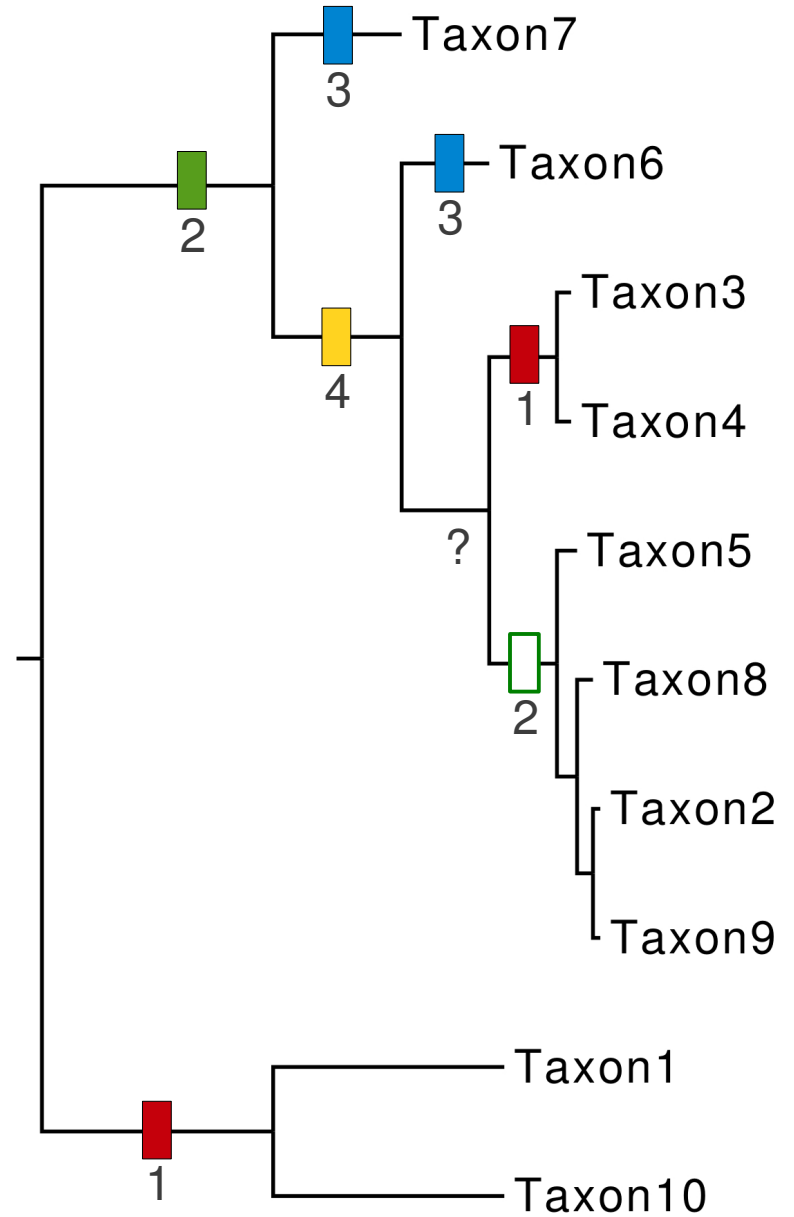
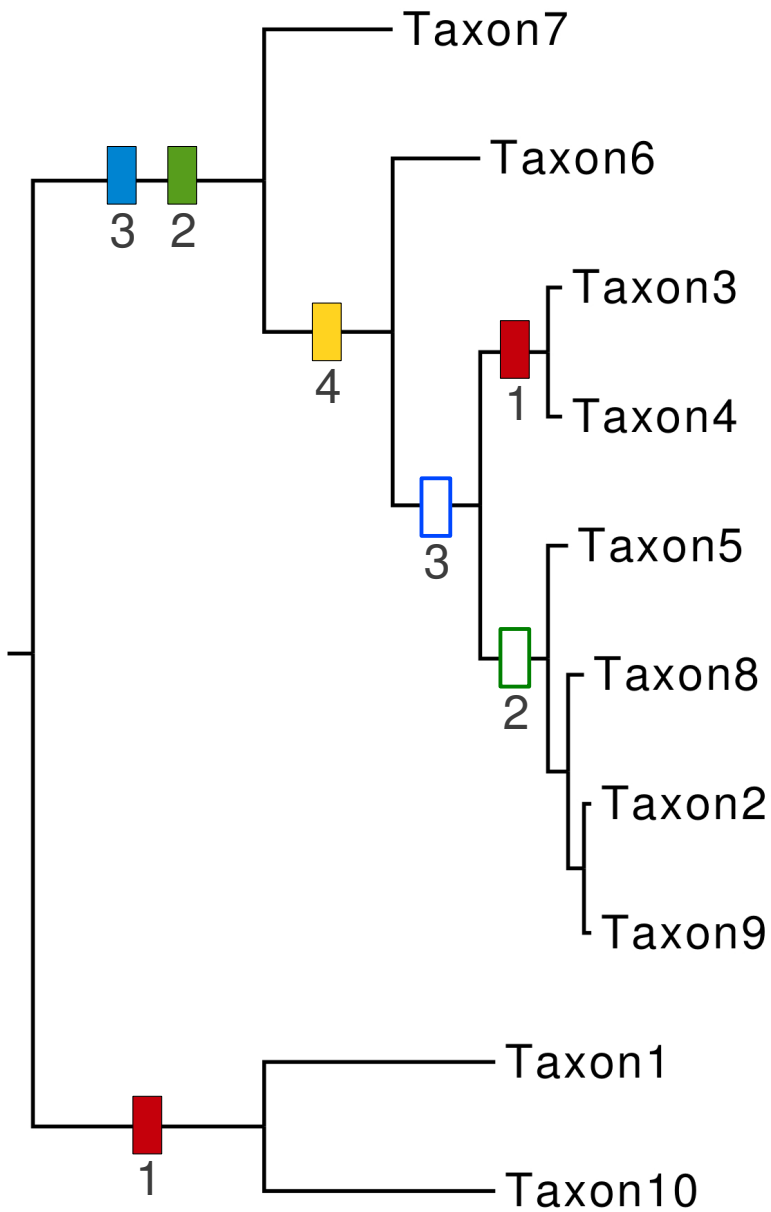


**Considere os nós que estão sendo sustentados por transformações e responda:**

Existe algum nó com suporte ambíguo neste cladograma?

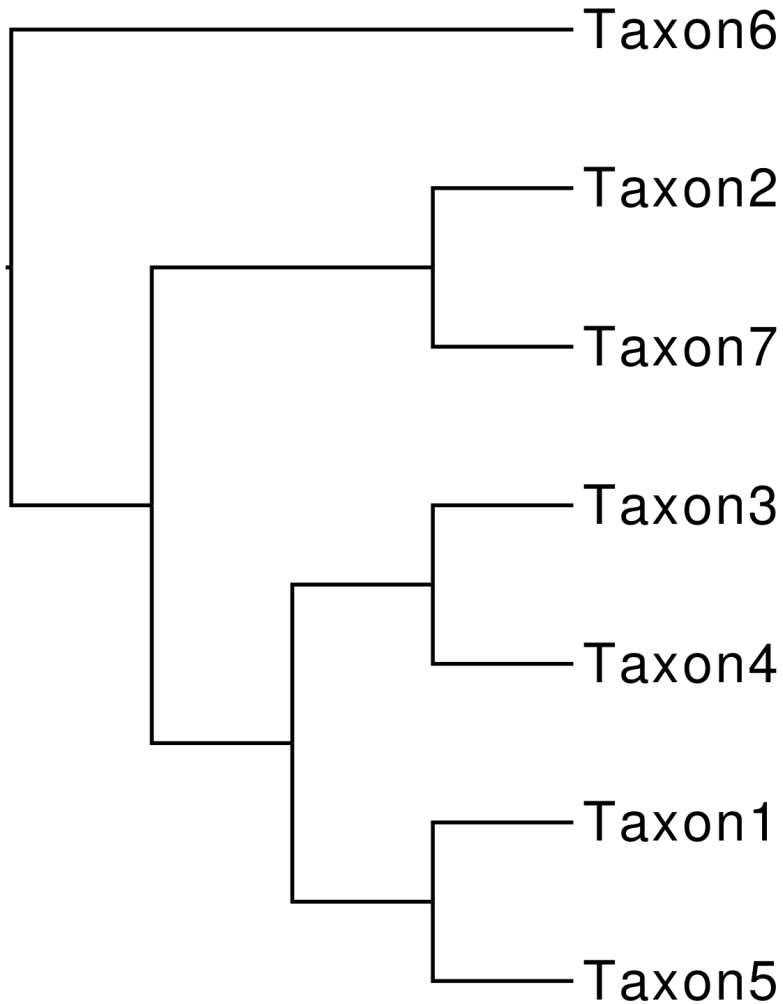
# Leitura de cladograma:

## 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

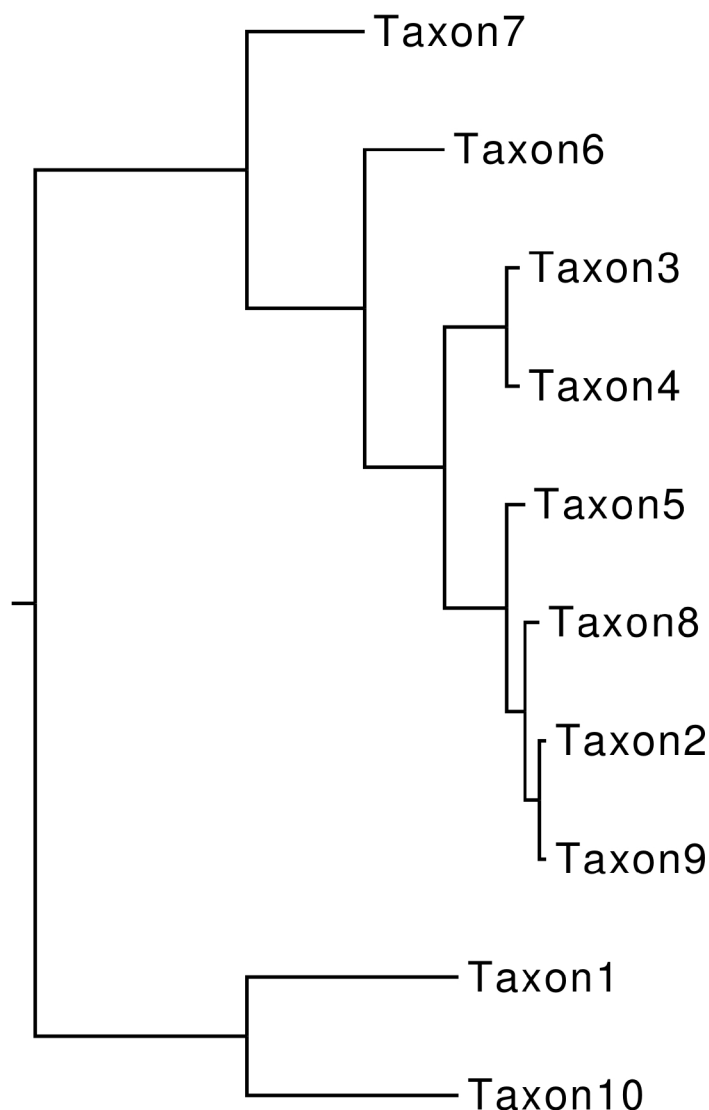
Considere a descrição abaixo e desenhe o cladograma:

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:

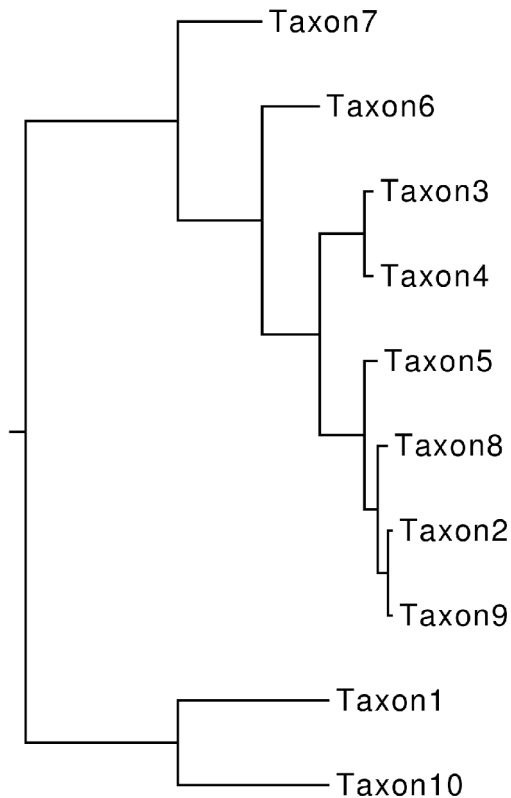


O cladograma com 10 terminais (T1-T10) contém dois grupos monofiléticos: um formado por T1 e T10 e o outro formado pelos demais. Dentro deste grande clado, T3 é grupo-irmão de T4, clado este que compartilha o mesmo nó que o grupo monofilético formado por T5, T8, T2 e T9. A resolução destes quatro terminais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

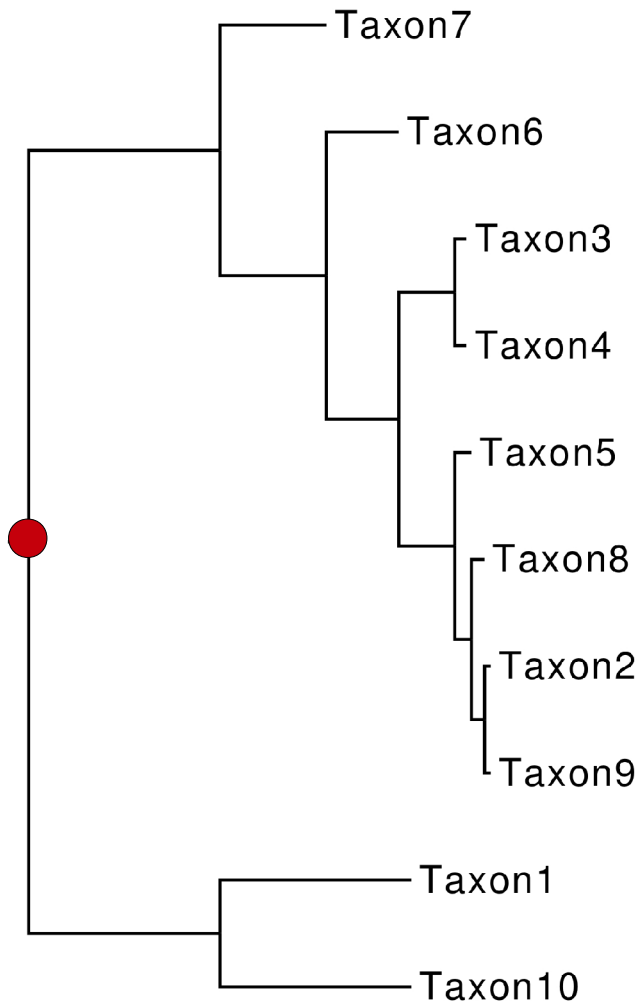
Dito de outra maneira, a relação destes terminais pode ser expressa da seguinte forma:

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



# *Leitura de cladograma:*

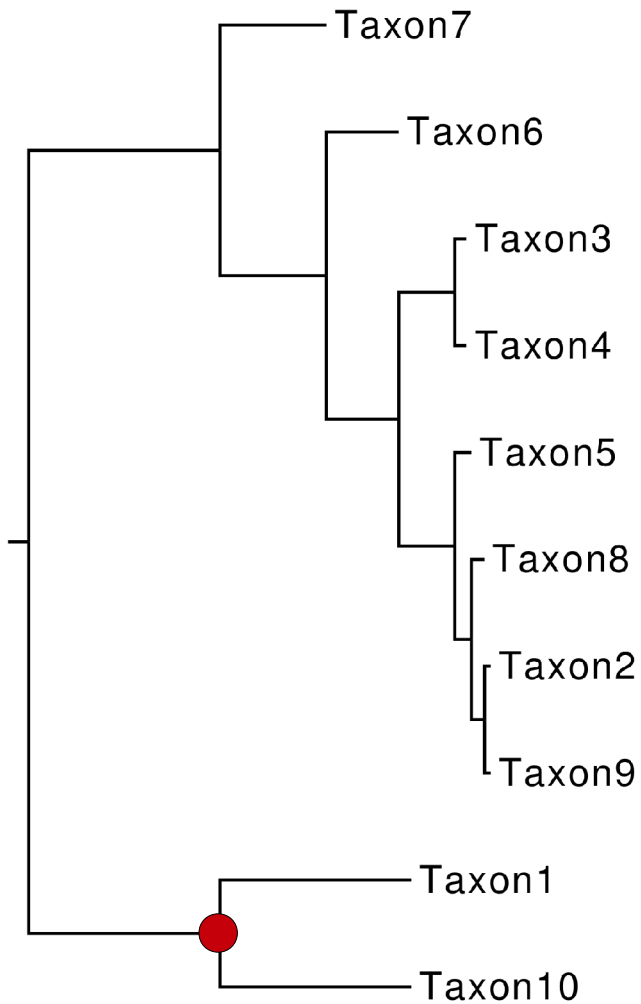
Descrição das relações de parentesco  
Notação parentética



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

# *Leitura de cladograma:*

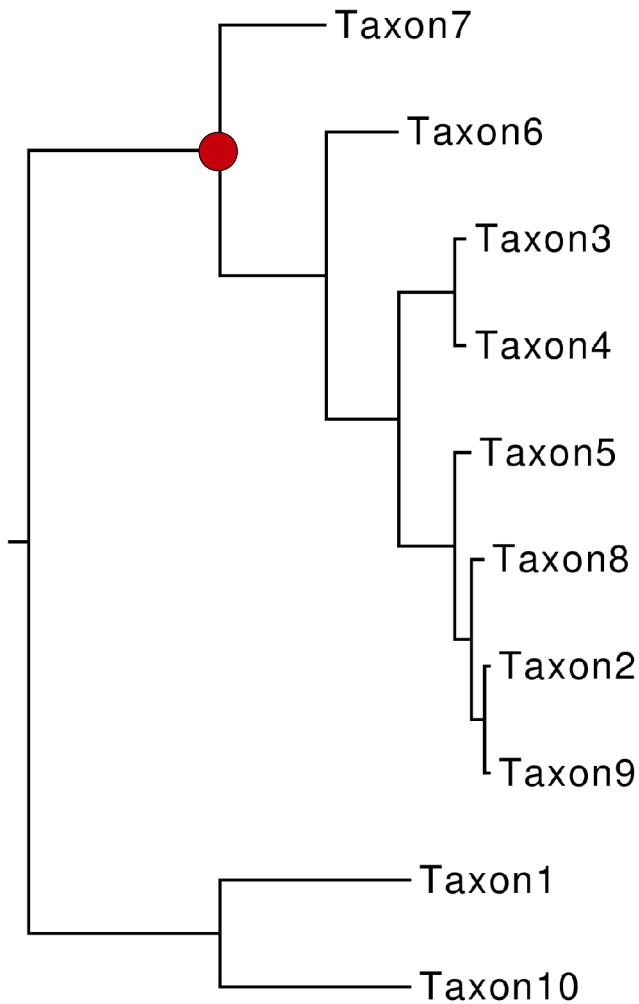
Descrição das relações de parentesco  
Notação parentética



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

# *Leitura de cladograma:*

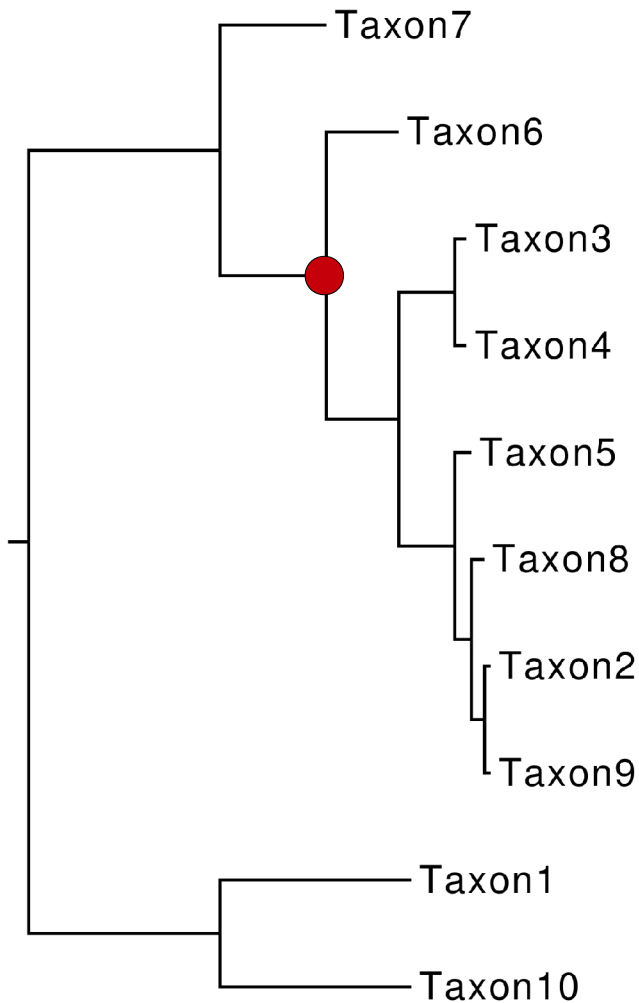
Descrição das relações de parentesco  
Notação parentética



(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

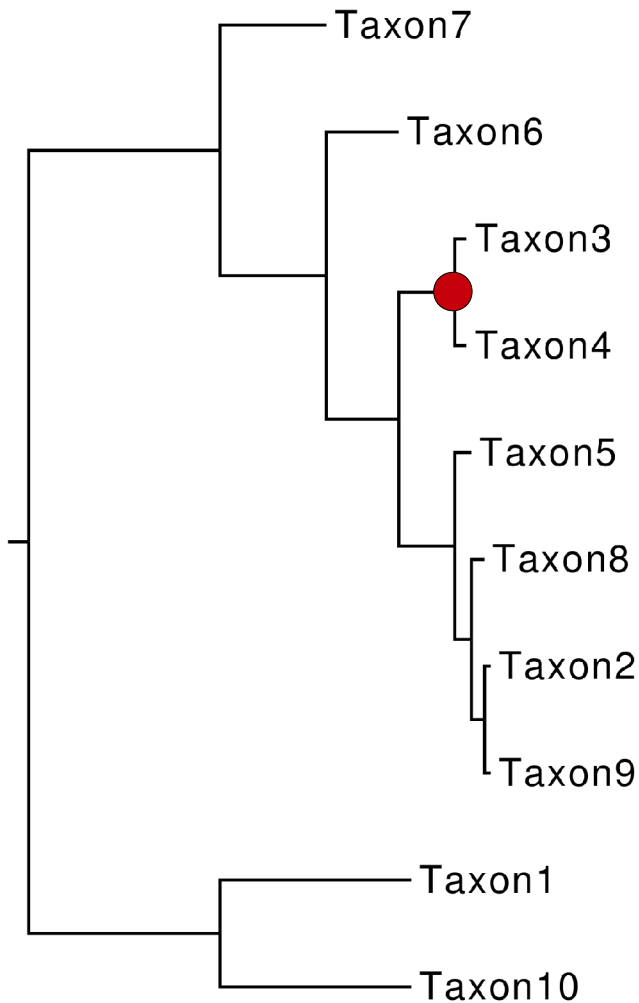


(T10,T1,T7,T6,T3,T4,T5,T8,T2,T9)  
((T10,T1),T7,T6,T3,T4,T5,T8,T2,T9)  
((T10,T1),(T7,T6,T3,T4,T5,T8,T2,T9))  
((T10,T1),(T7,(T6,T3,T4,T5,T8,T2,T9)))



# Leitura de cladograma:

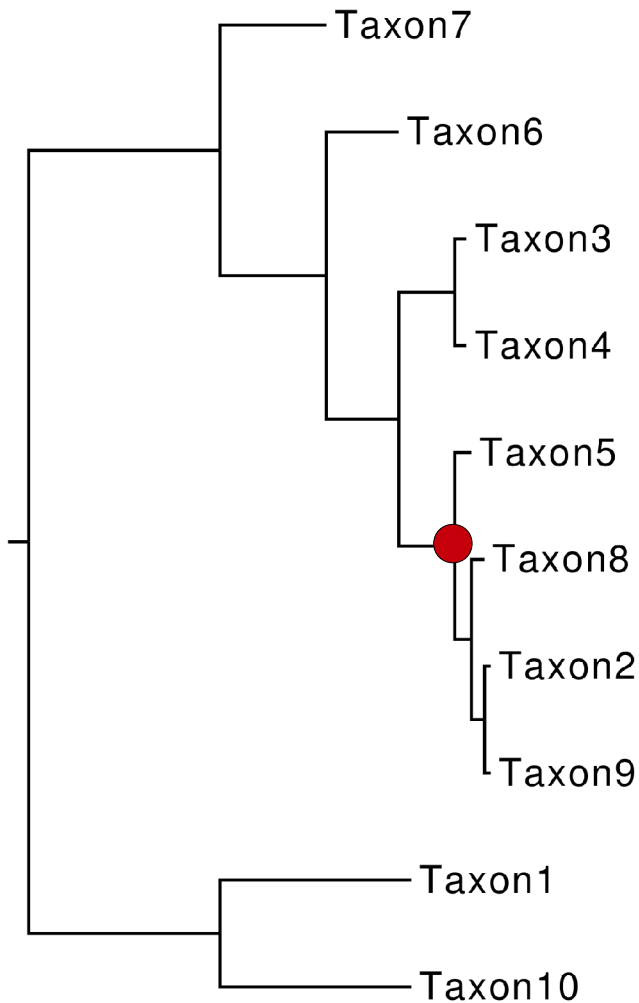
Descrição das relações de parentesco  
Notação parentética



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

# Leitura de cladograma:

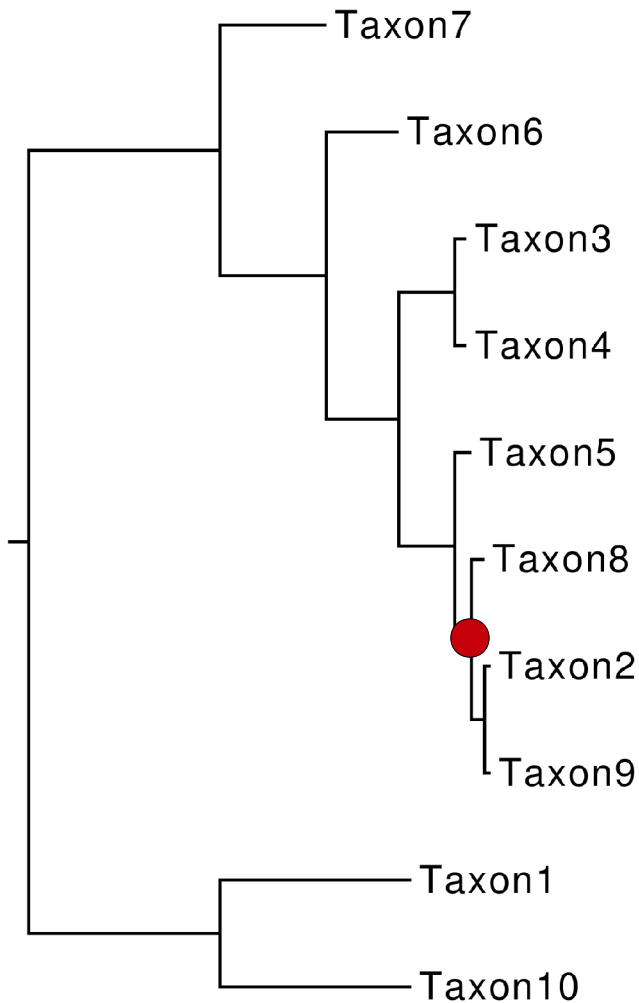
Descrição das relações de parentesco  
Notação parentética



(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



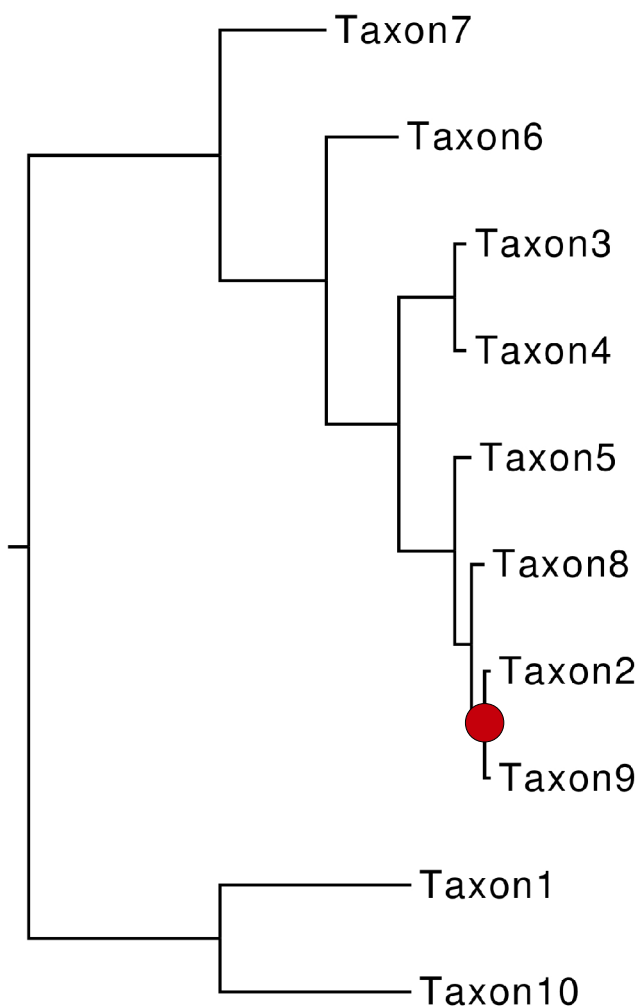
(T10,T1,T7,T6,T3,T4,T5,T8,T2,T9)  
((T10,T1),T7,T6,T3,T4,T5,T8,T2,T9)  
((T10,T1),(T7,T6,T3,T4,T5,T8,T2,T9))  
((T10,T1),(T7,(T6,T3,T4,T5,T8,T2,T9)))  
((T10,T1),(T7,(T6,(T3,T4,T5,T8,T2,T9))))  
((T10,T1),(T7,(T6,((T3,T4),T5,T8,T2,T9))))  
((T10,T1),(T7,(T6,((T3,T4),(T5,T8,T2,T9)))))  
((T10,T1),(T7,(T6,((T3,T4),(T5,(T8,T2,T9))))))



# Leitura de cladograma:

Descrição das relações de parentesco

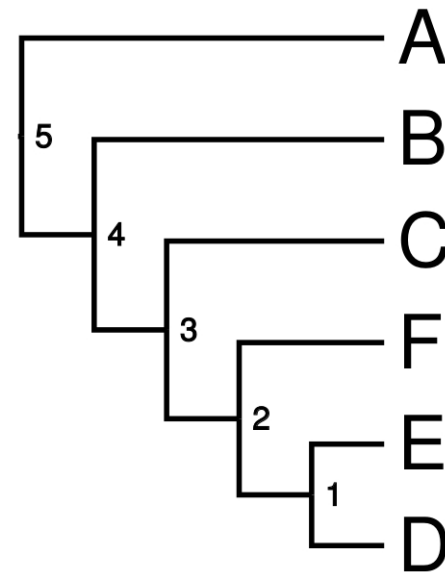
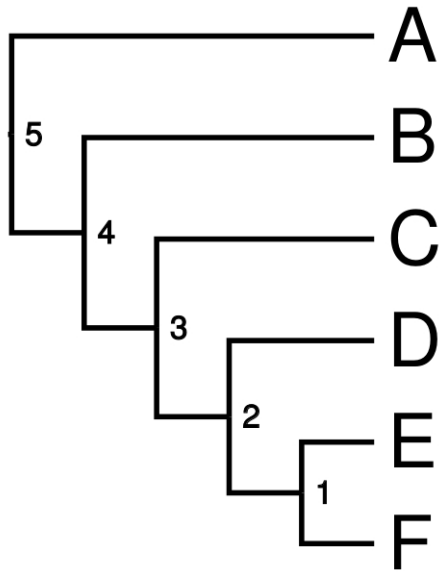
Notação parentética



(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?*

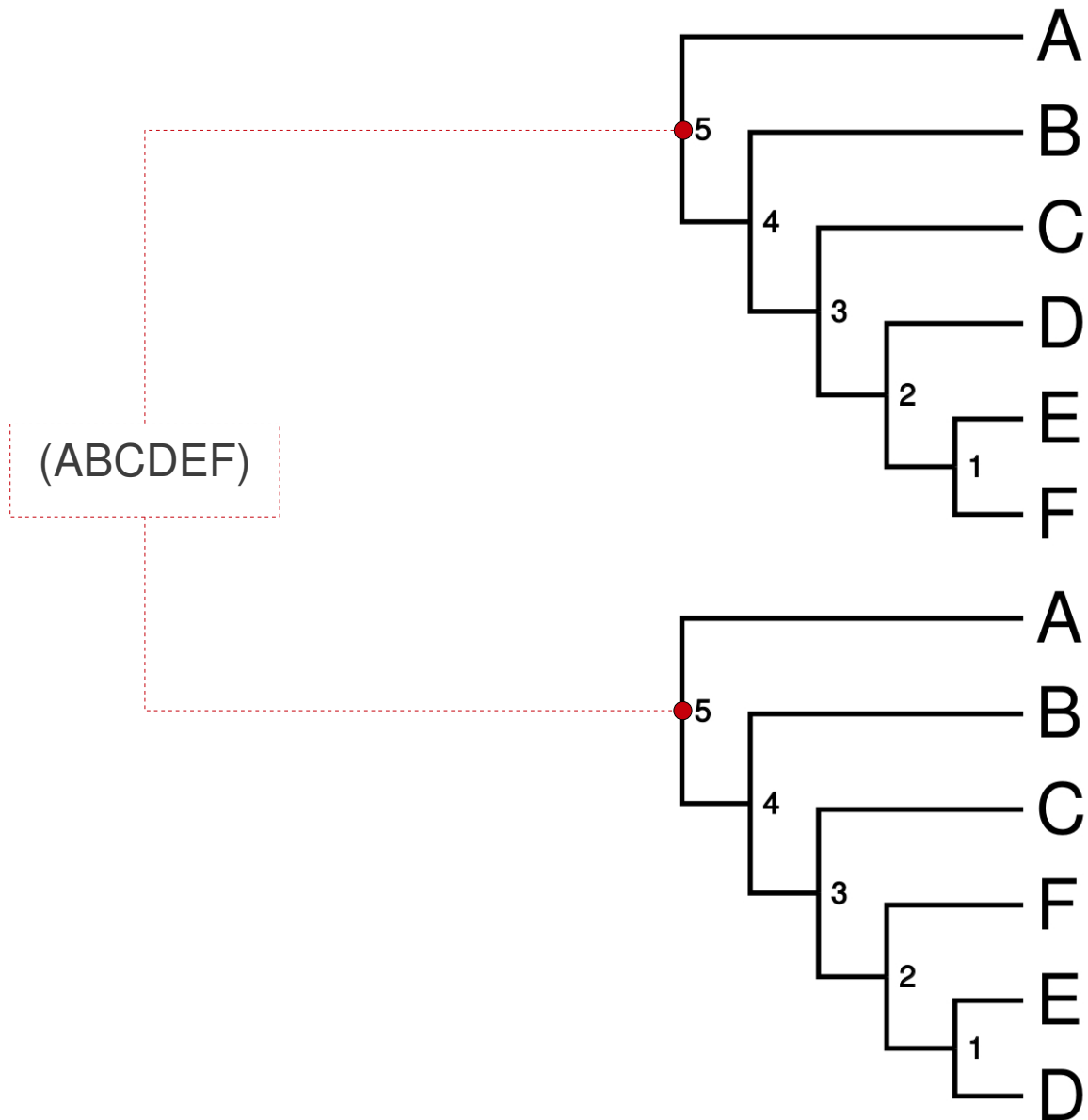
*Quais componentes são comuns entre estas duas topologias?*

# Consenso

Considere os seguintes cladogramas fundamentais:

Componentes  
comuns:

(ABCDEF)



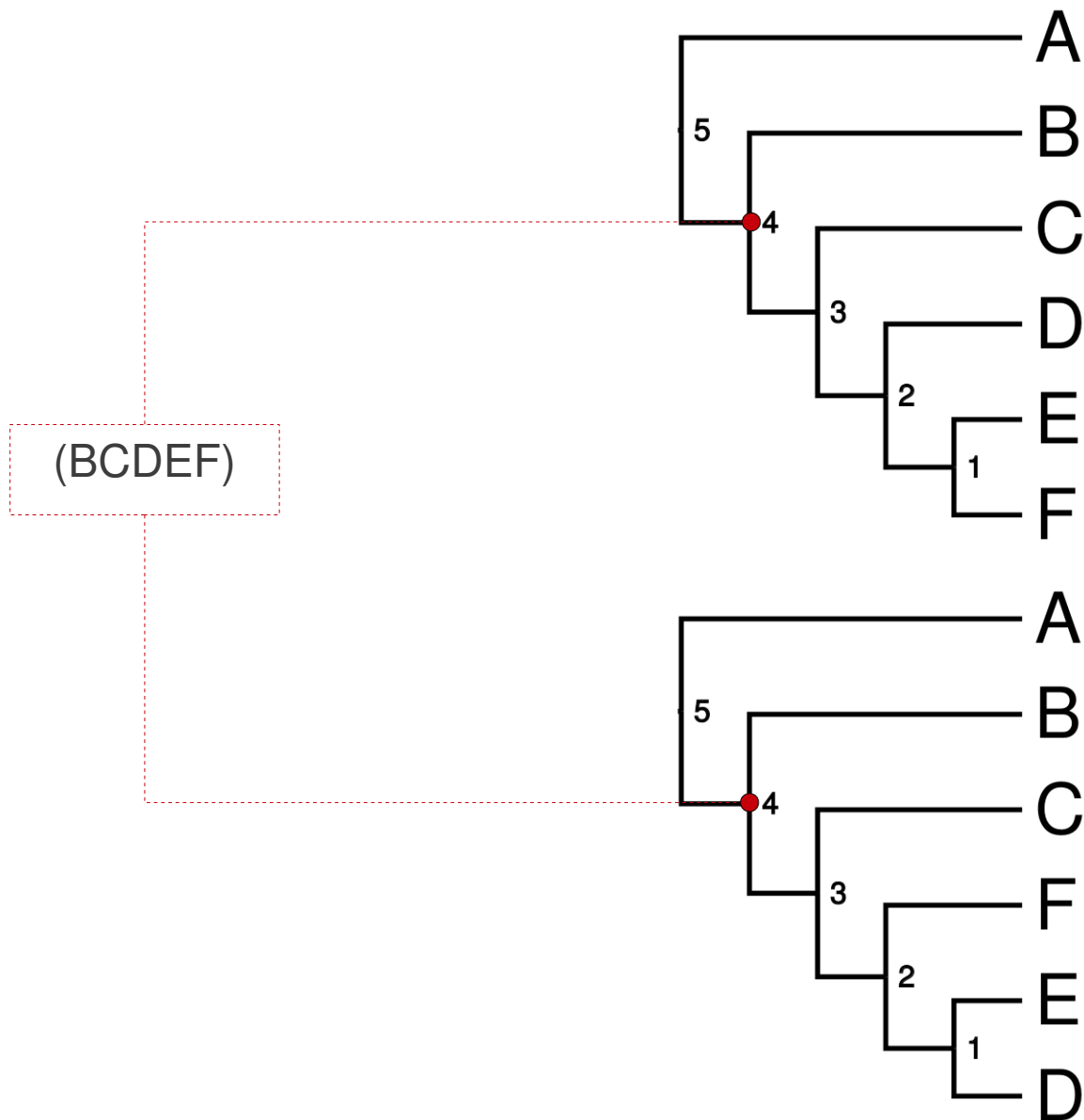
# Consenso

Considere os seguintes cladogramas fundamentais:

Componentes  
comuns:

(ABCDEF)

(BCDEF)

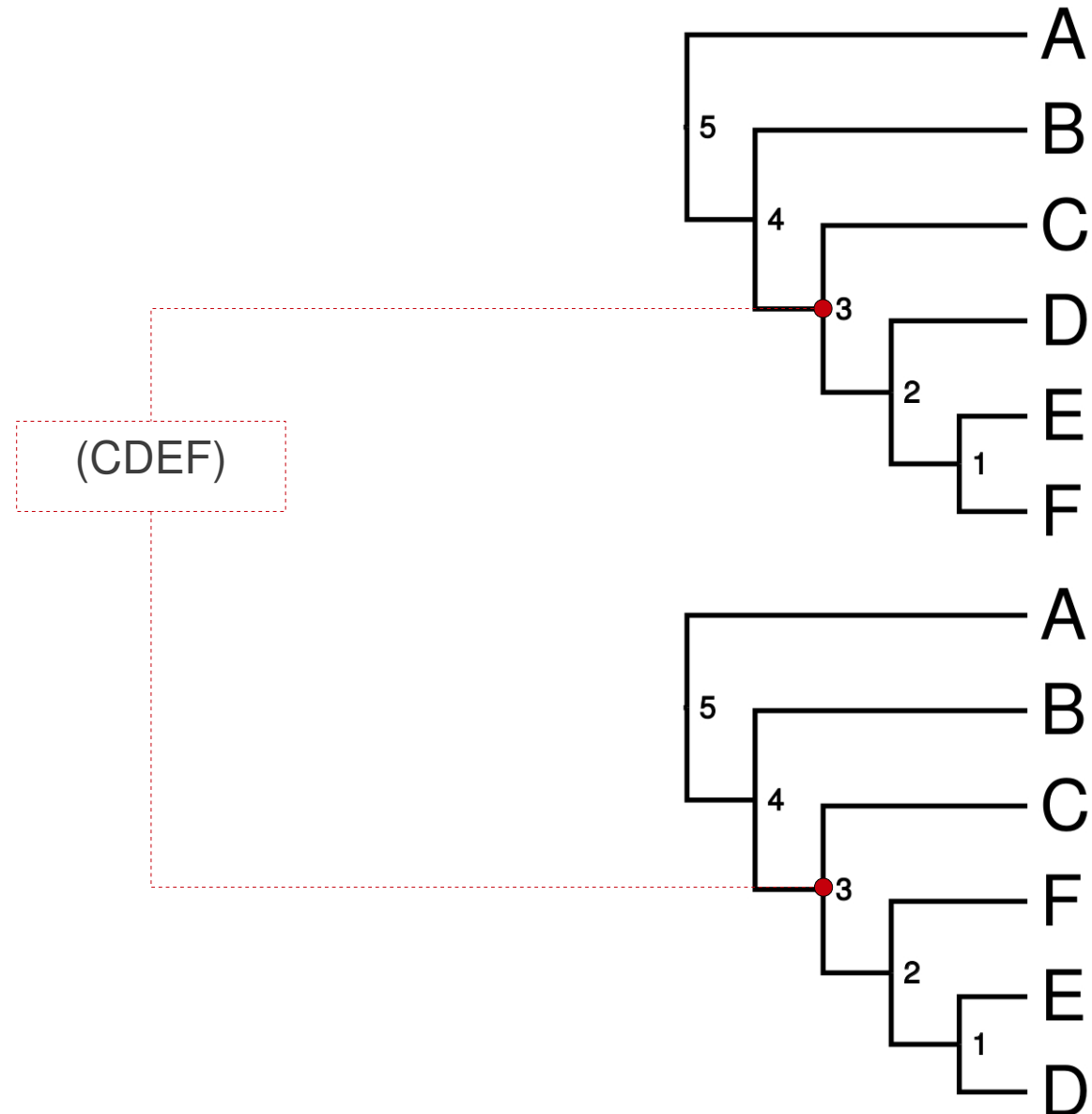


# Consenso

Considere os seguintes cladogramas fundamentais:

Componentes  
comuns:

(ABCDEF)  
(BCDEF)  
(CDEF)

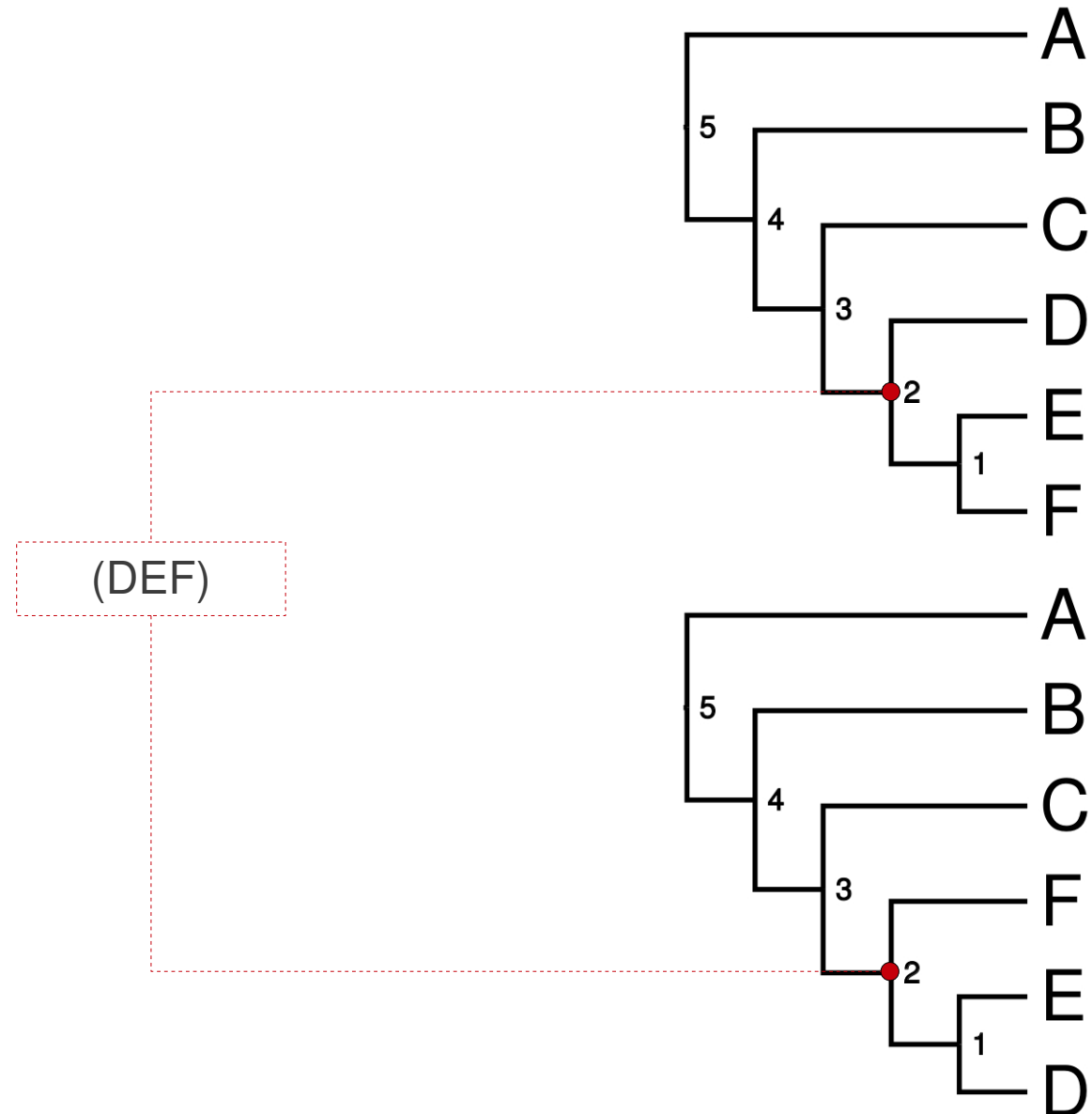


# Consenso

Considere os seguintes cladogramas fundamentais:

Componentes  
comuns:

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

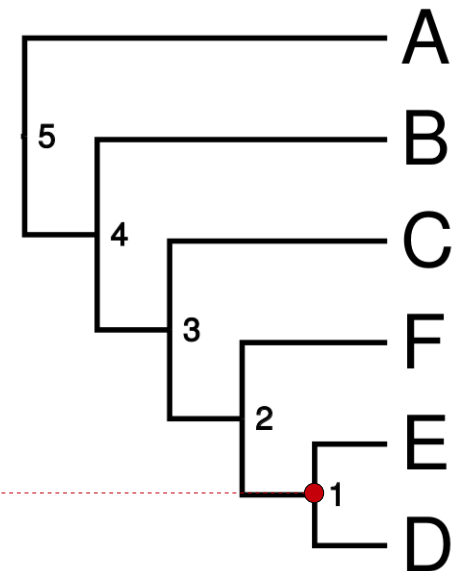
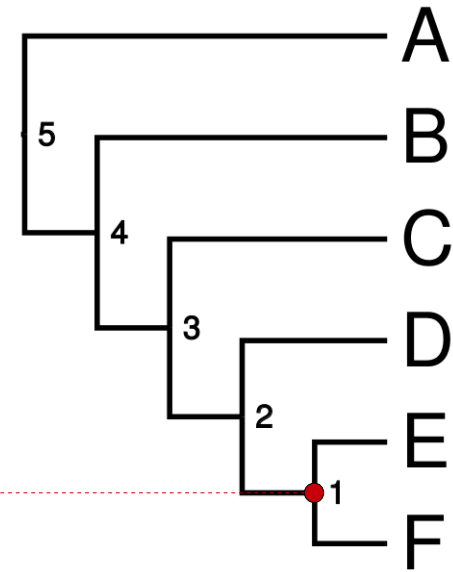
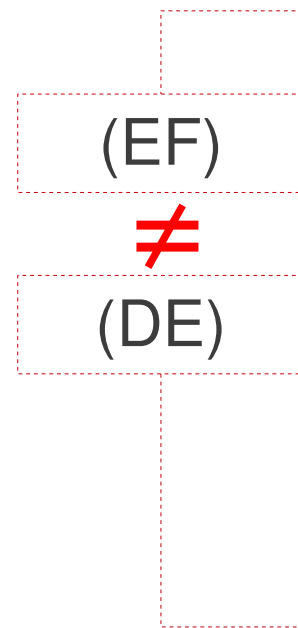


# Consenso

Considere os seguintes cladogramas fundamentais:

Componentes  
comuns:

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



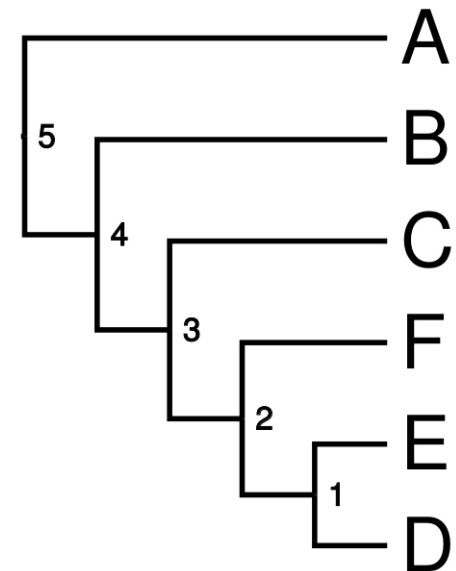
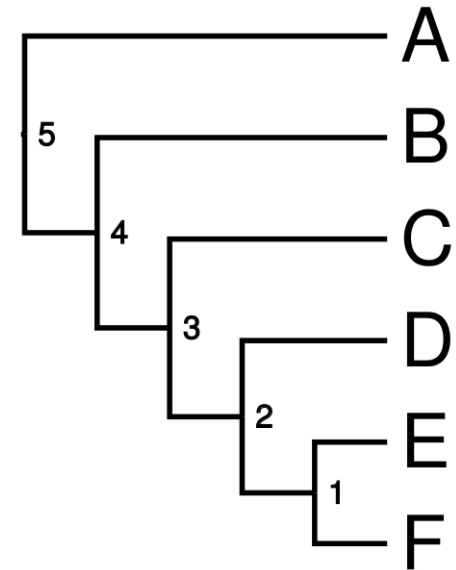
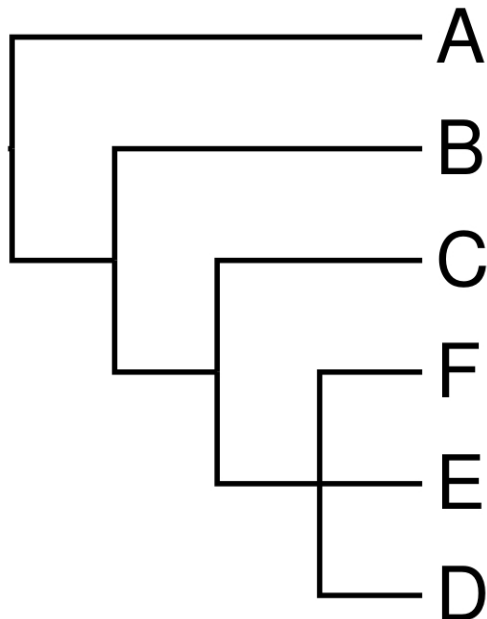
# Consenso

Considere os seguintes cladogramas fundamentais:

Componentes  
comuns:

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

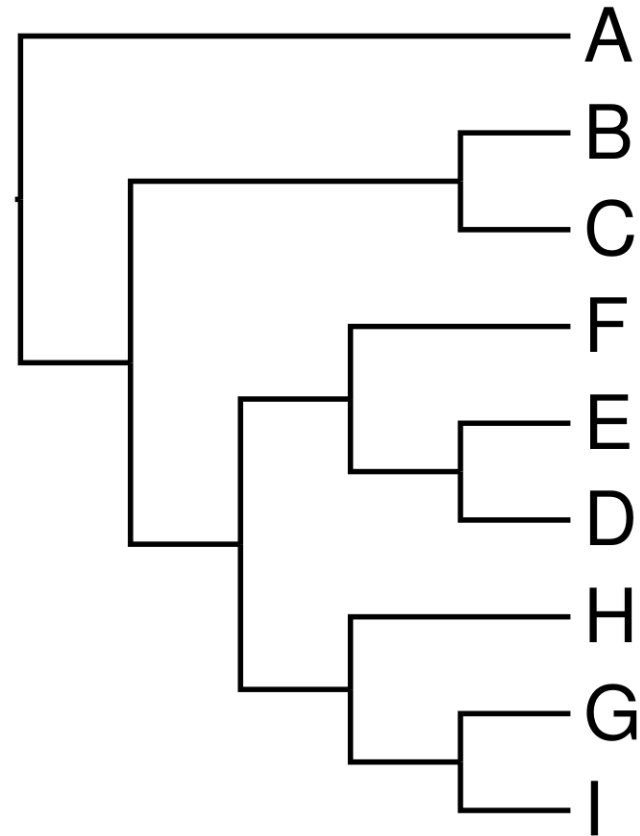
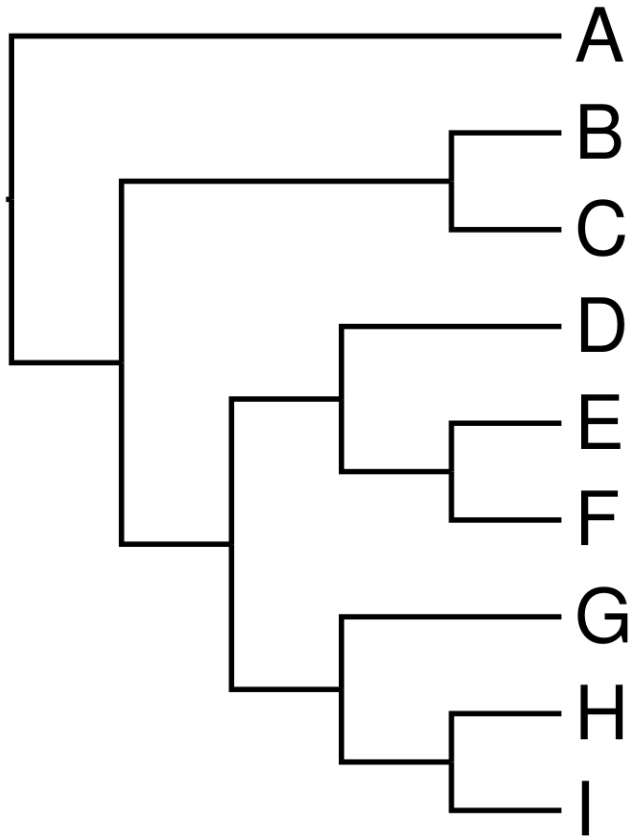
Consenso estrito:





# Consenso

## Exemplo 1

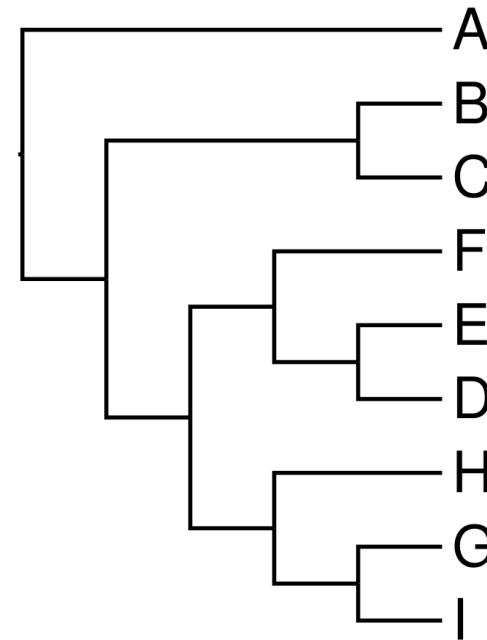
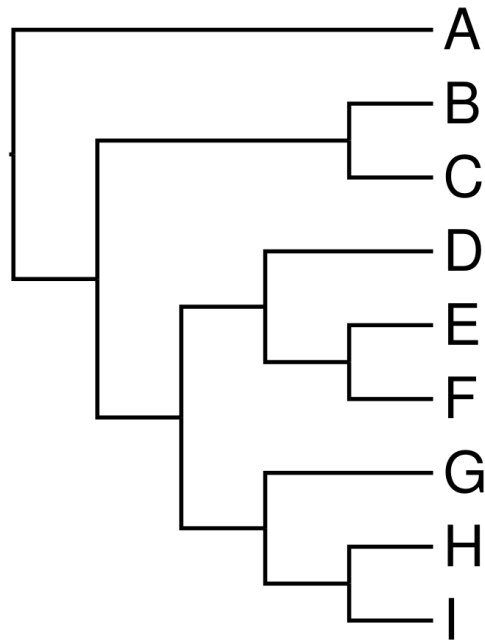


*Quais componentes são comuns entre estas duas topologias?*

*Qual é a topologia de consenso?*

# Consenso

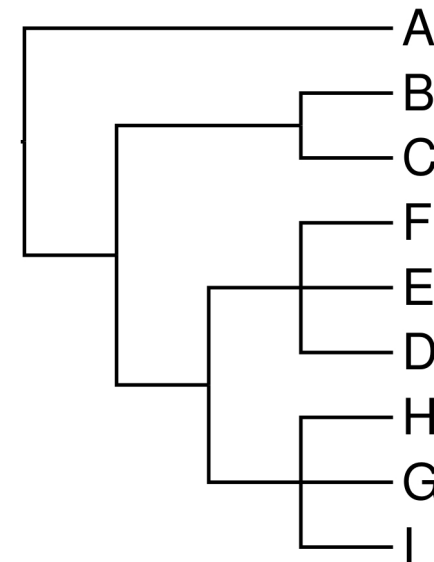
Exemplo



*Quais componentes são comuns entre estas duas topologias?*

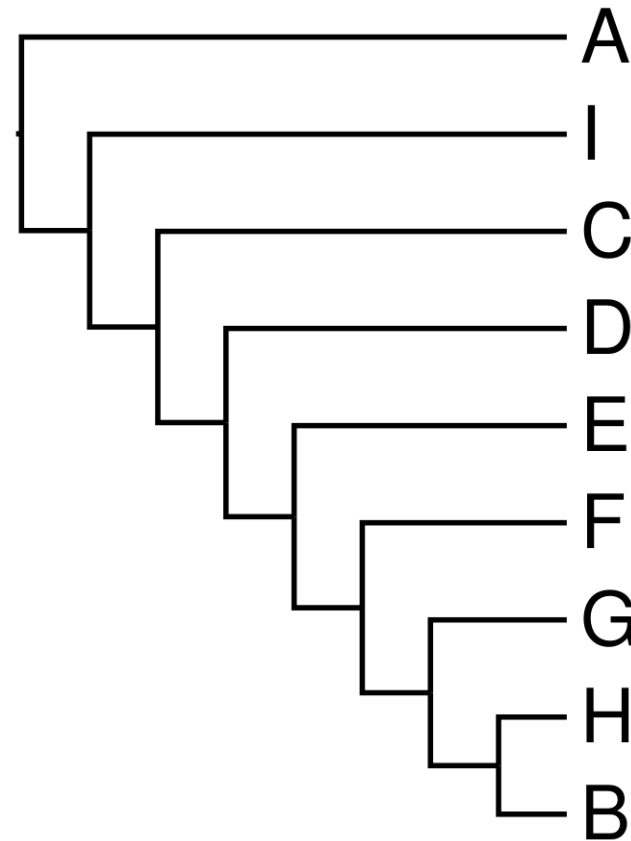
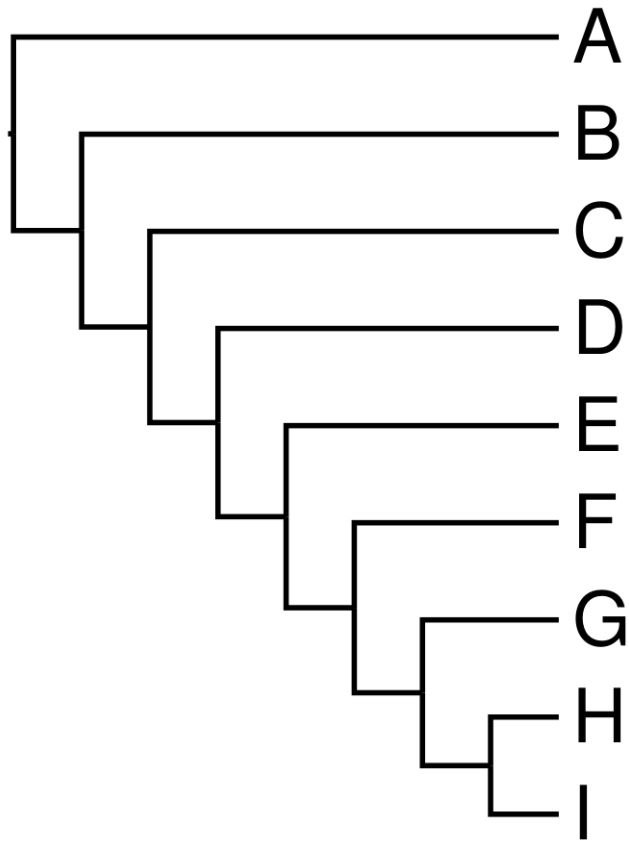
- 1: (ABCDEFGHI)
- 2: (BCDEFGHI)
- 3: (DEFGHI)
- 4: (DEF)
- 5: (GHI)
- 6: (BC)

*Qual é a topologia de consenso?*



# Consenso

## Exemplo 2



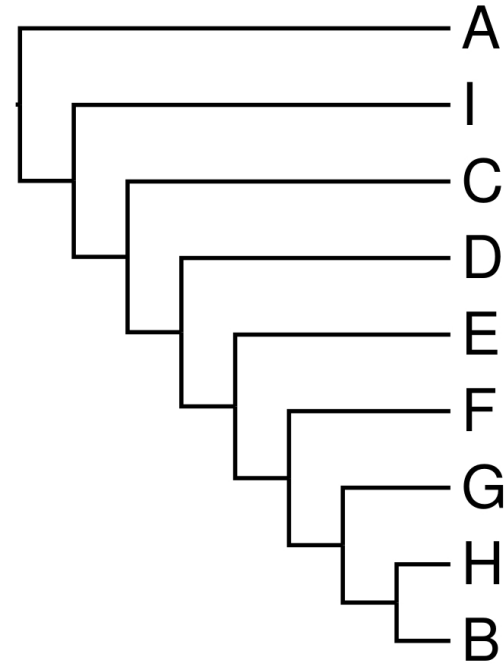
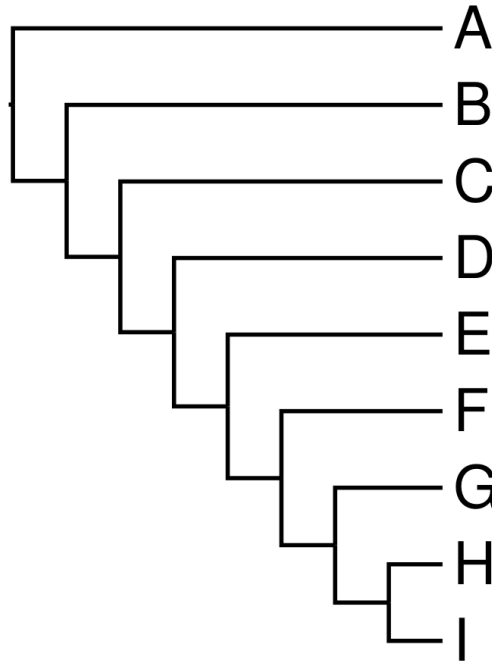
*Quais componentes são comuns entre estas duas topologias?*

*Qual é a topologia de consenso?*

*Existe algum terminal que está forçando a topologia de consenso?*

# Consenso

Exemplo 2

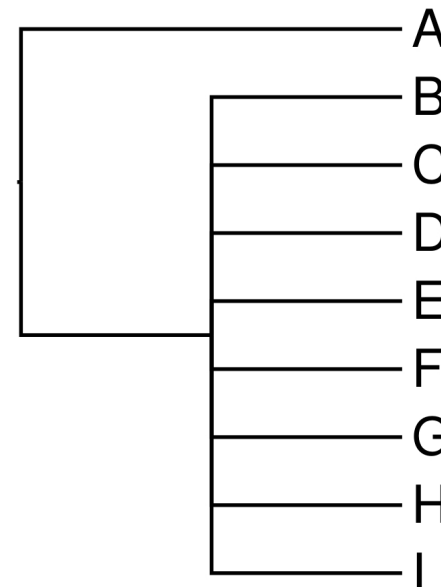


*Quais componentes são comuns entre estas duas topologias?*

1: (ABCDEFGHI)

2: (BCDEFGHI)

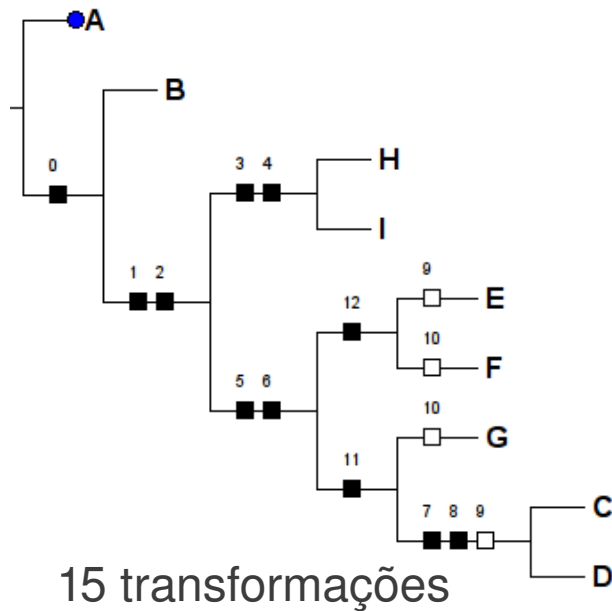
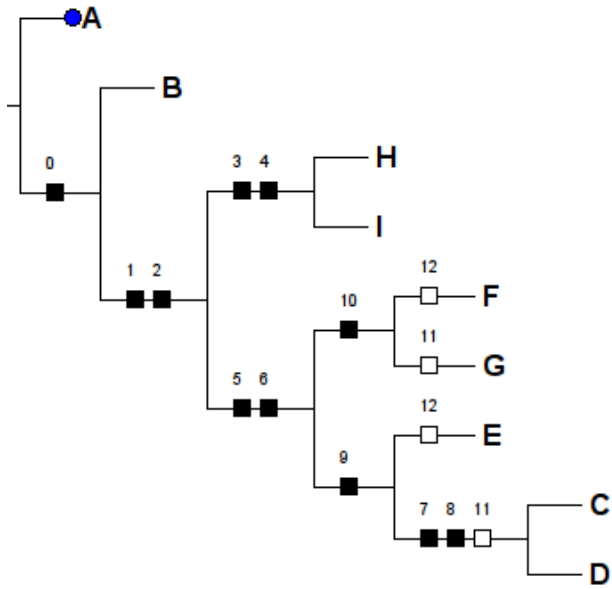
*Qual é a topologia de consenso?*



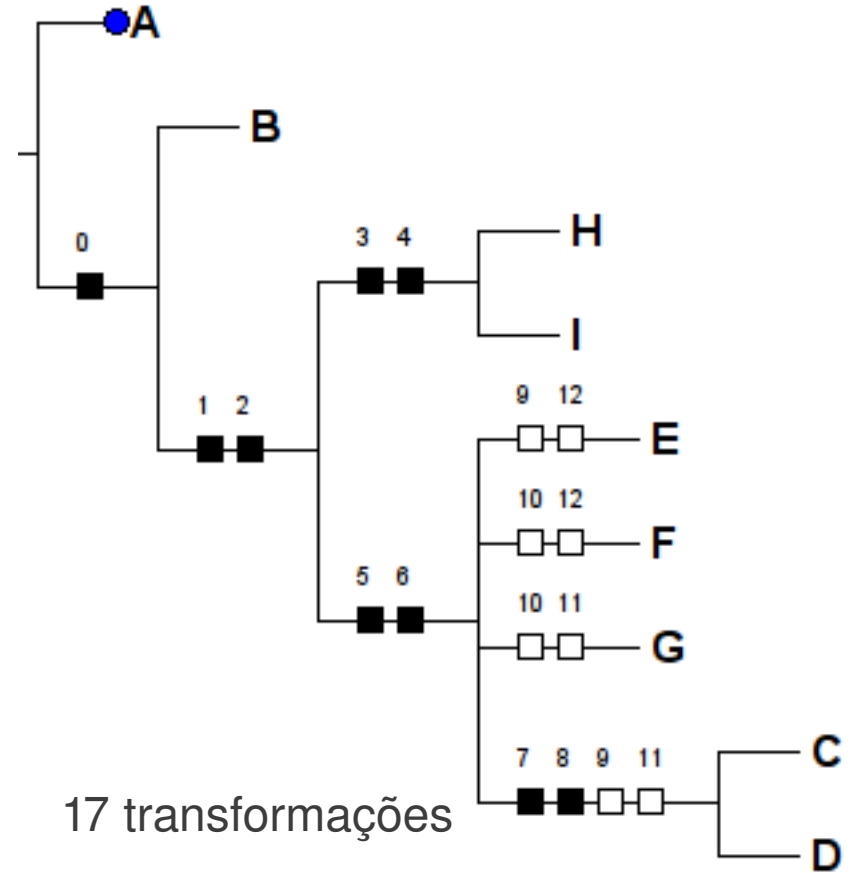


# Consenso

## Exemplo 3



15 transformações



17 transformações

# O desenvolvimento da Cladística

## Dados moleculares e modelos probabilísticos

Período essencialista

Mundo dinâmico

Resistência e Nova Síntese

Sistemática Evolutiva

1960's

Fenética

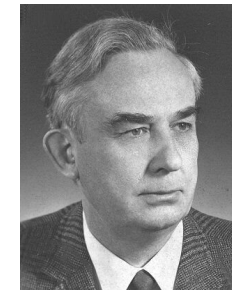
1970's

Cladística

1990's

Probabilismo

CRITÉRIO DE OTIMALIDADE	# DE PUBLICAÇÕES
UPGMA	4340
Neighbor Joining (NJ)	17000
Distância Patrística (PARCIMÔNIA)	15900
Probabilidade Posterior (BAYESIANA)	17000
Verossimilhança Máxima (ML)	28100



Willi Hennig  
1913 - 1976



Joe Felsenstein



Steve Farris



David Hillis

**Fonte:** Google Scholar, 8 maio 2017. Parâmetros da busca:

“phylogenetics” AND “<method>”, restrição “Since 2016”.

[Compilado pelo Prof. Daniel Lahr]

# Motivação para o uso de Verossimilhança:

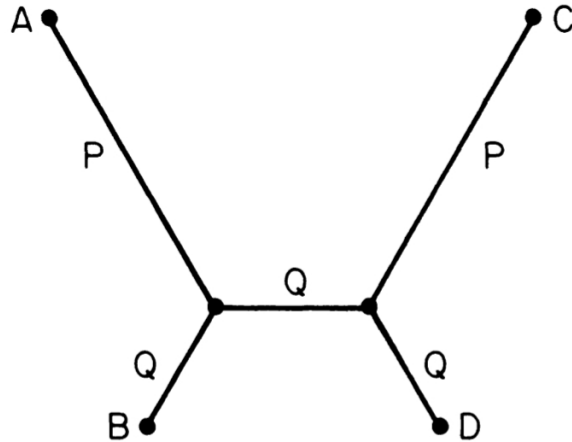
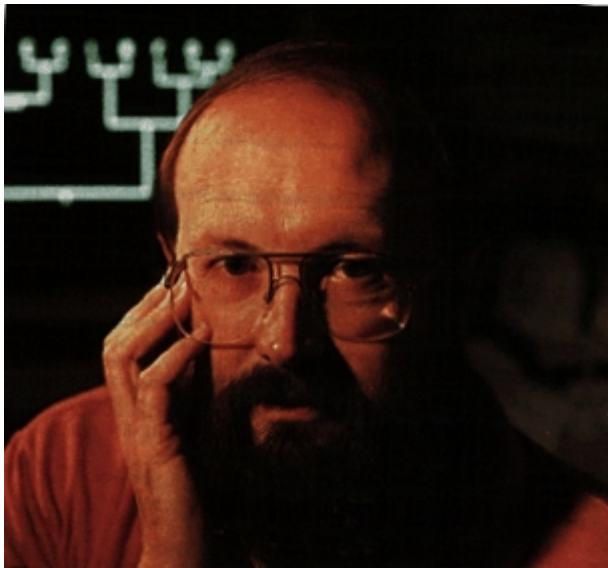


FIG. 3.—True unknown phylogeny (with root omitted) used to find cases in which unrooted Wagner tree parsimony methods will be inconsistent.



Joe Felsenstein  
University of  
Washington

## CASES IN WHICH PARSIMONY OR COMPATIBILITY METHODS WILL BE POSITIVELY MISLEADING<sup>1</sup>

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.

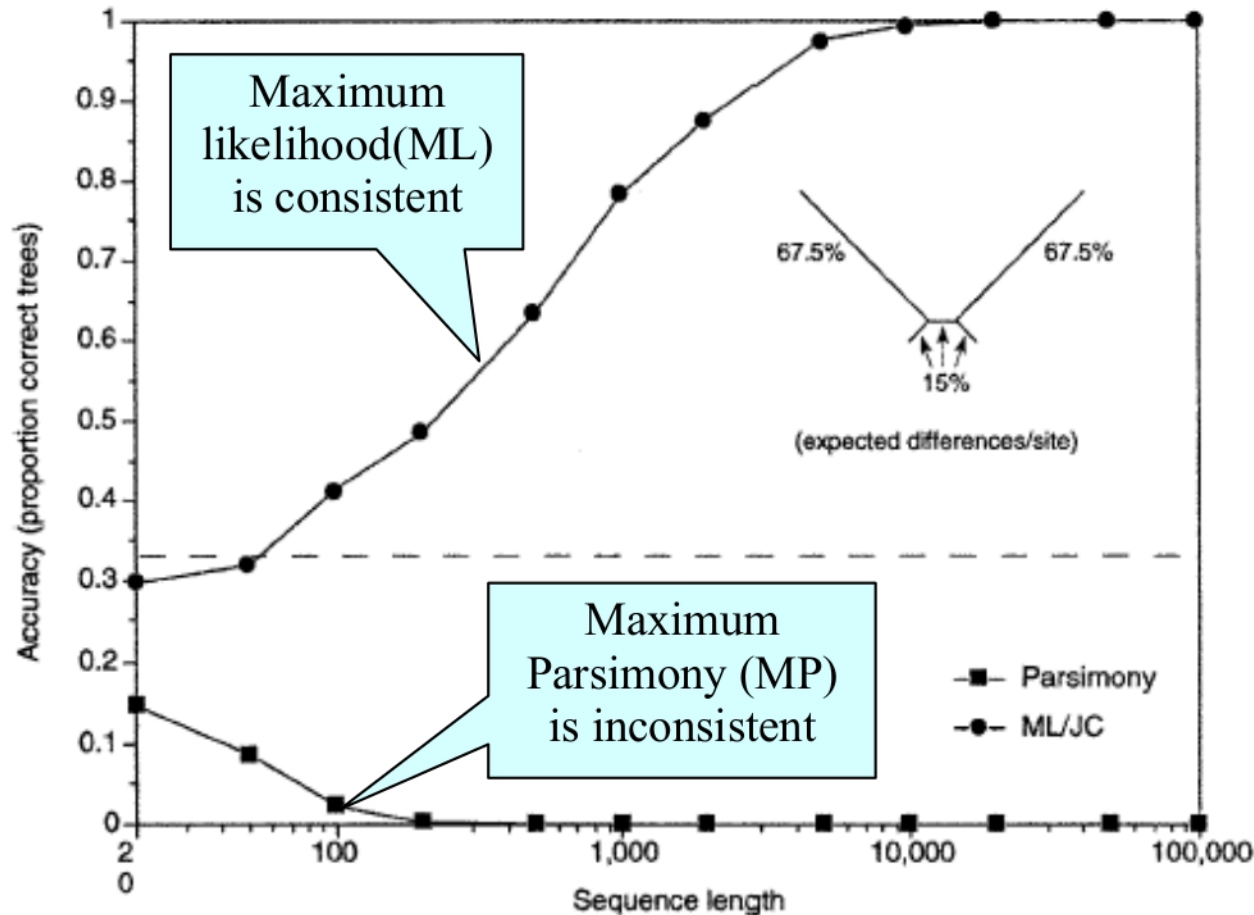
An alternative methodology for phylogenetic inference is the compatibility method, introduced by Le Quesne (1969, 1972). He suggested that phylogenetic inference be based on finding the largest possible set of characters which could simultaneously have all states be uniquely derived on the same tree. 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,

<sup>1</sup> 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.

# Cálculo de probabilidades:

**Regra 1a:** A probabilidade de um evento **A** ocorrer é  $P_{(A)} = A/S$ , onde S é o espaço de soluções possíveis.

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

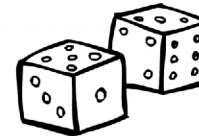


Copyright © Ron Leishman

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

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

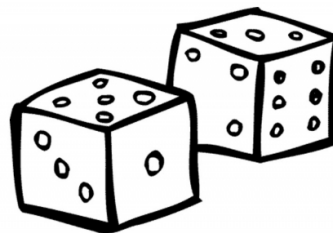
Se dois eventos não possuem soluções em comum, eles são chamados de **disjuntos** e suas probabilidades obedecem à seguinte regra:



**Regra 2:** Se dois eventos A e B são disjuntos, então a probabilidade de qualquer um destes eventos e a soma das probabilidades dos dois eventos:  $P_{(A \text{ ou } B)} = P_{(a)} + P_{(b)}$

# Probabilidades e modelos

Considere que você jogue dois dados:



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











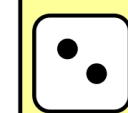
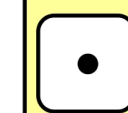
**Regra 3:** Se dois eventos A e B são independentes, então a probabilidade dos dois eventos ocorrerem um após o outro (i.e., intersecção de eventos, ou **probabilidade conjunta**) é o produto das probabilidades de cada evento:  $P_{(A \text{ e } B)} = P_{(A)} * P_{(B)}$ .

## 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

# *O que é verossimilhança?*



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

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

A verossimilhança ( $L$ ) de um conjunto de parâmetros ( $\theta$ ), dado alguma observação ( $x$ ) é proporcional 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.

# *Estimativa de parâmetros:*

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



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

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

Qual é a probabilidade de uma estudante do sexo feminino fazer uma pergunta durante a aula?

# *Estimativa de parâmetros:*

## Modelo probabilístico:

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.

Em teoria das probabilidades e estatística, a **distribuição binomial** é a distribuição de probabilidade discreta do número de sucessos numa sequência de tentativas tais que:

Cada tentativa tem exclusivamente como resultado duas possibilidades, sucesso ou fracasso (binomial, a que se chama de tentativa de Bernoulli), e;

Cada tentativa é independente das demais, e;

A probabilidade de sucesso a cada tentativa permanece constante independente das demais, e;

A variável de interesse, ou pretendida, é o número de sucessos nas tentativas

# *Estimativa de parâmetros:*

Qual é a probabilidade de uma estudante do sexo feminino fazer uma pergunta durante a aula?

## Modelo probabilístico:

Em teoria das probabilidades e estatística, a **distribuição binomial** é a distribuição de probabilidade discreta do número de sucessos numa sequência de tentativas.

$$f(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}$$

onde  $k$  é no número de eventos de sucesso (i.e., pergunta de aluna),  $n$  é o número de perguntas feitas, e  $p$  é a probabilidade deste evento ocorrer, neste caso, ele é o parâmetro ( $\theta$ ) que queremos estimar.

# *Estimativa de parâmetros:*

Qual é a probabilidade de uma estudante do sexo feminino fazer uma pergunta durante a aula?

## Modelo probabilístico:

Em teoria das probabilidades e estatística, a **distribuição binomial** é a distribuição de probabilidade discreta do número de sucessos numa sequência de tentativas .

$$f(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}$$

## Observações de perguntas em PSB 2016:

104 perguntas em 3 aulas  
21 provenientes de alunas

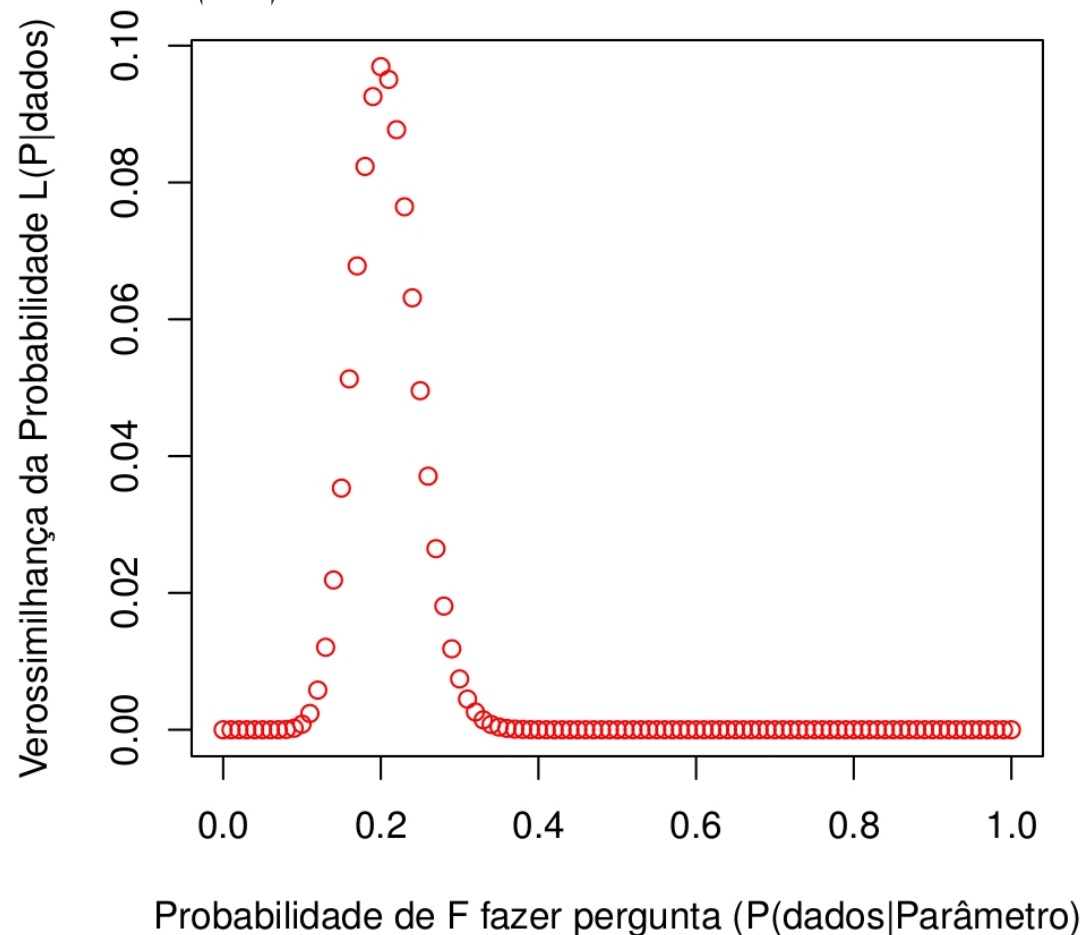


# Estimativa de parâmetros:

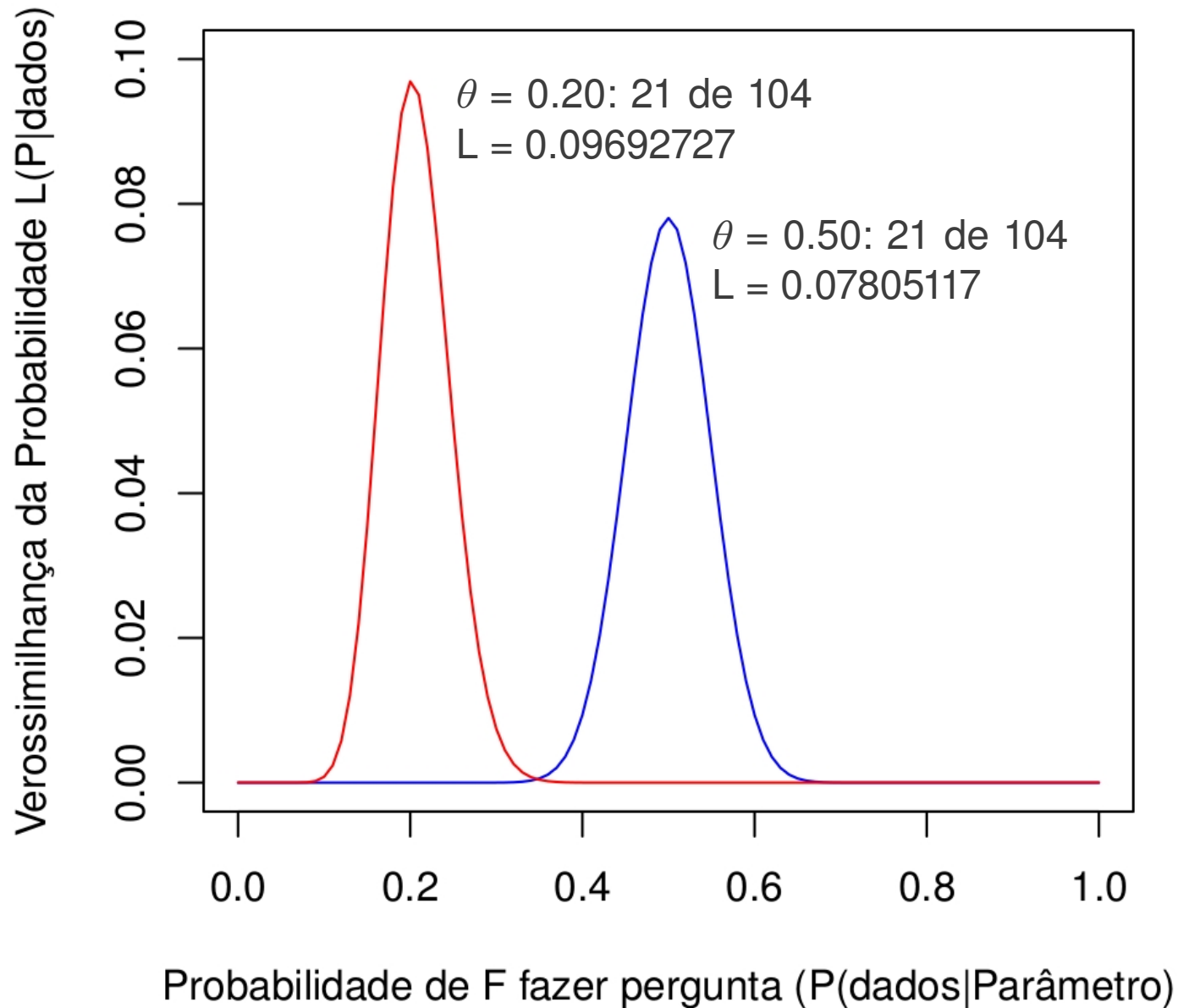
Modelo probabilístico:

$$f(21; 104, \theta) = \binom{n}{k} p^k (1-p)^{n-k}$$

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



# Verossimilhanças e seleção de modelos:



# Verossimilhanças e seleção de modelos:



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

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

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

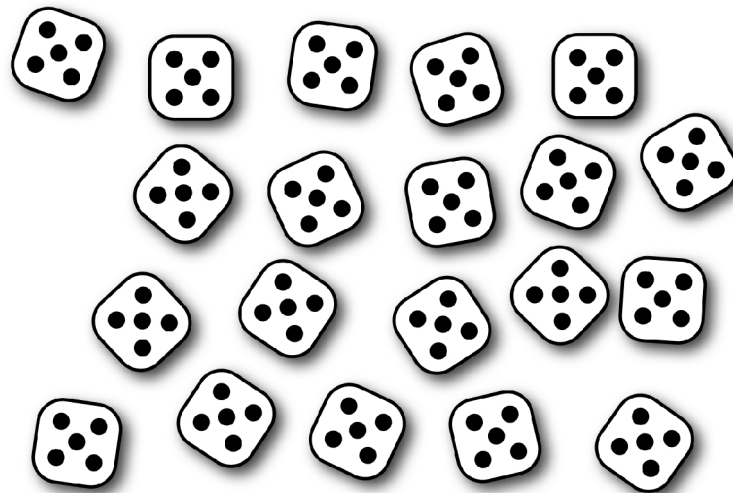
# *Verossimilhanças e seleção de modelos:*

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

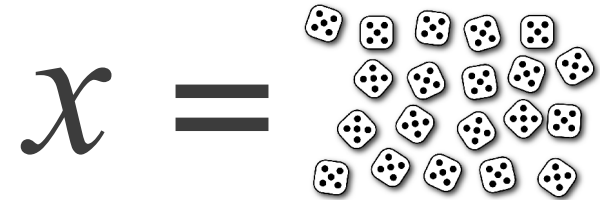
Considere:

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

$x =$



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



Considere que você deseja avaliar a  $L$  de dois modelos ( $\theta_1$  e  $\theta_2$ ):

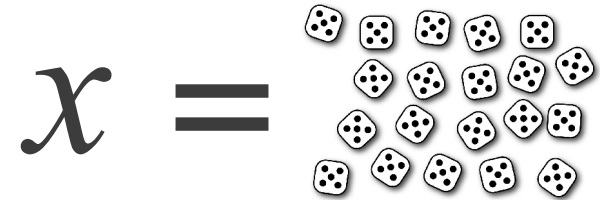
**Modelo 1** ( $\theta_1$ ) – dado honesto – assume que cada face do dado possui um número de 1 a 6 cuja a probabilidade conjunta (i.e.,  $P_{(A \text{ e } B)} = P_{(A)} \times P_{(B)}$ ) de obter 20 dados apresentando 5 é igual a  $(1/6)^{20}$ .

**Modelo 2** ( $\theta_2$ ) – dado duvidoso (Feito em **Brasilia/DF**) – 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.

# Verossimilhanças e seleção de modelos:

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



Considere que você deseja avaliar a  $L$  de dois modelos ( $\theta_1$  e  $\theta_2$ ):

**Modelo 1** ( $\theta_1$ ) – 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  $(1/6)^{20}$ .

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

**Modelo 2** ( $\theta_2$ ) – dado duvidoso – 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  $(1)^{20}$ .

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

# Verossimilhança vs. Probabilidade

$$x = \text{[img alt="10 dice showing various faces]" data-bbox="180 200 320 320]} = \text{observação} = O$$

**Modelo 1** ( $\theta_1$ ) – dado honesto = Hipótese 1 =  $H_1$

**Modelo 2** ( $\theta_2$ ) – dado duvidoso = Hipótese 2 =  $H_2$

$$L(\theta|x) \approx P(x|\theta) \quad x = \text{[img alt="10 dice showing various faces]" data-bbox="760 600 910 730]}$$

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

# Verossimilhança vs. Probabilidade

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



Considere:

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

$H$  = há *minions* jogando futebol no seu forro.

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

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

A verossimilhança desta hipótese é muito alta, pois se há *minions* jogando futebol 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 *minions* jogando futebol no seu forro.



# Voltemos a ideia original:

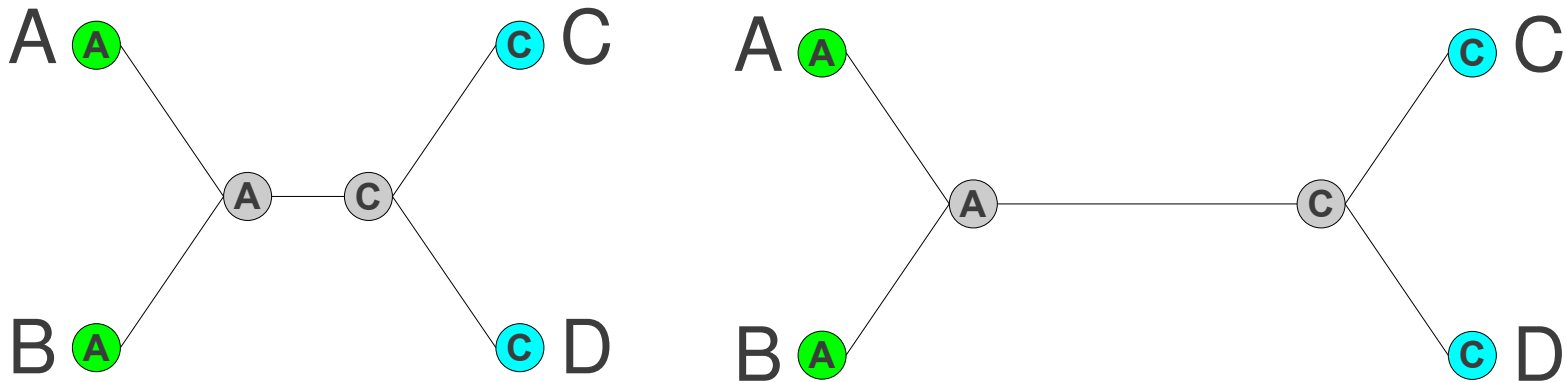


Sir Ronald Aylmer Fisher  
(17/02/1890 – 29/07/1962)

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

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

## Estimativa de máxima verossimilhança: TOPOLOGIA 1



Relevância da estimativa de comprimento de ramos

# Voltemos a ideia original:



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

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

Qual o valor de  $\theta$  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:

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

Qual o valor de  $\theta$  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\*

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

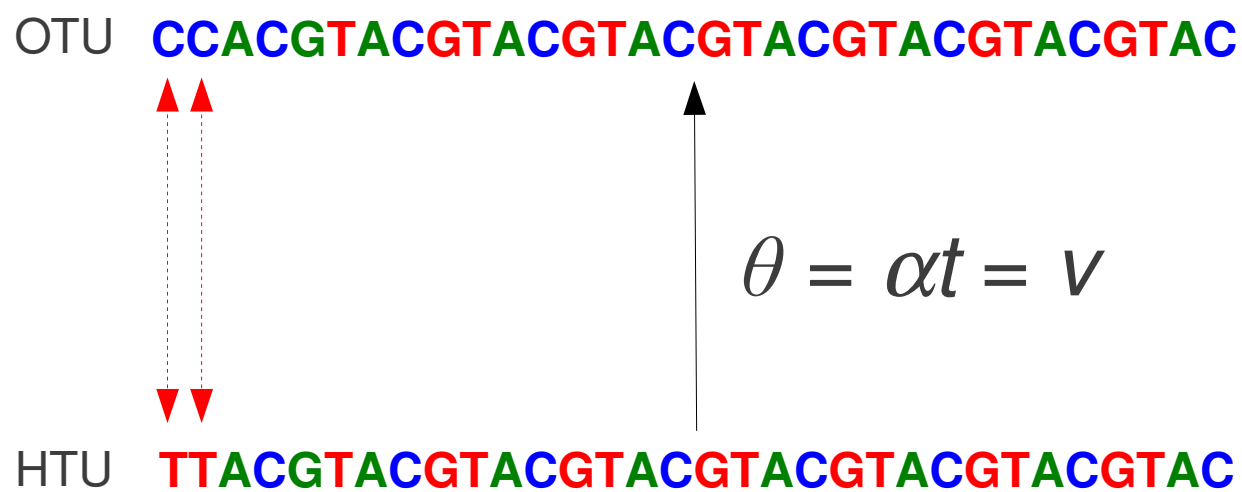
probabilidade de mudança de estado\*

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

\* 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:



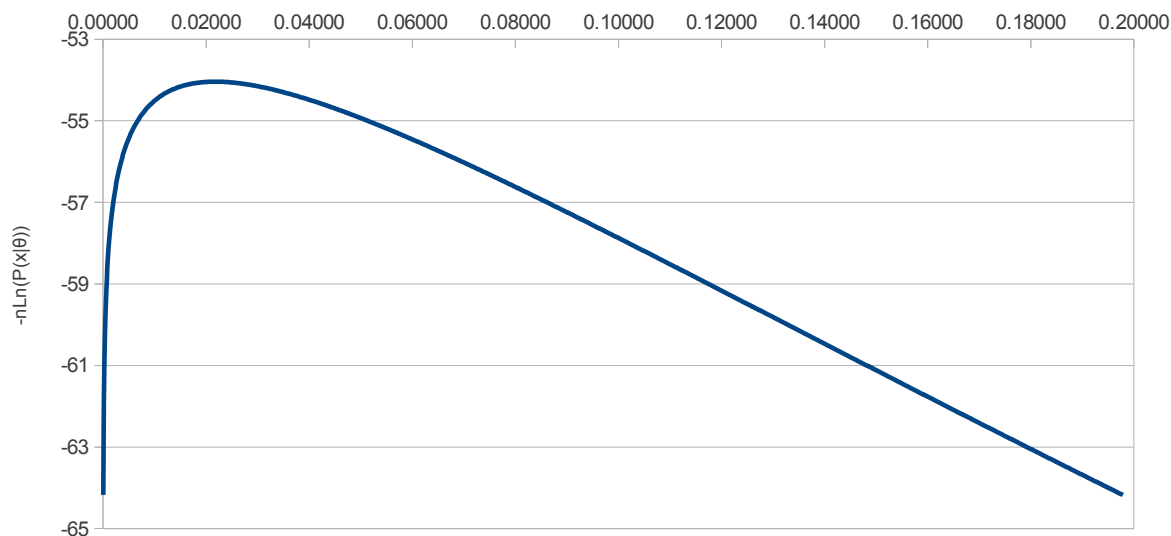
Sir Ronald Aylmer Fisher  
(17/02/1890 – 29/07/1962)

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

Qual o valor de  $\theta$  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)

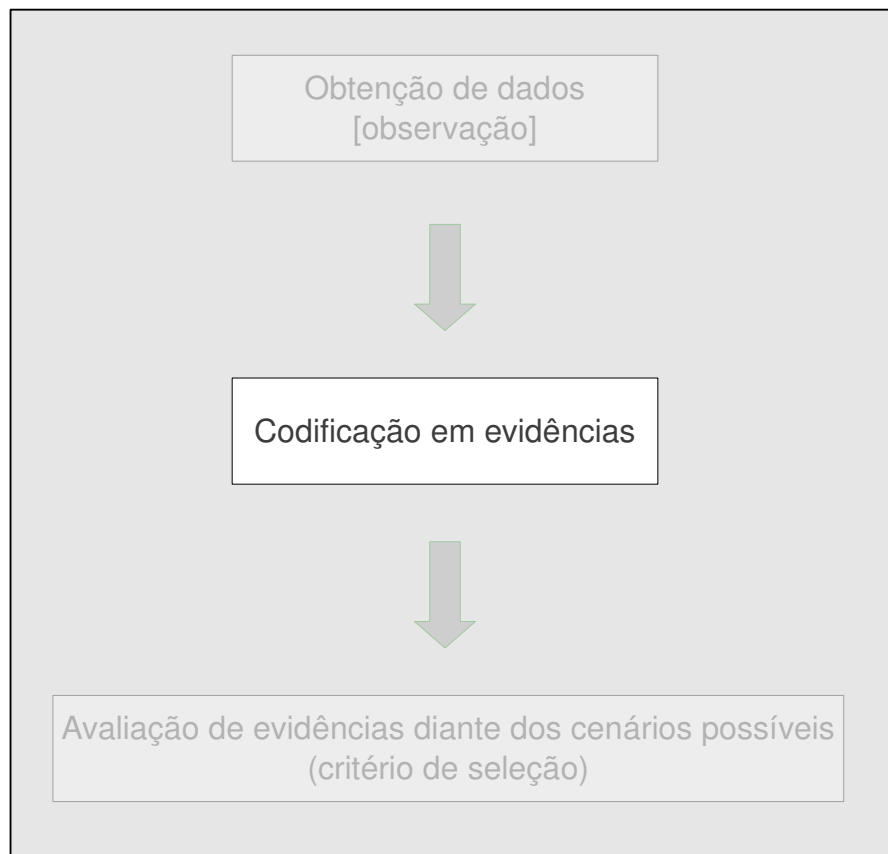


Estimativa de distância evolutiva (v)

Valor estimado para v: 0.021752

# Lógica da inferência filogenética

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



## Parcimônia:

EVIDÊNCIAS: transformação de estados de caracteres

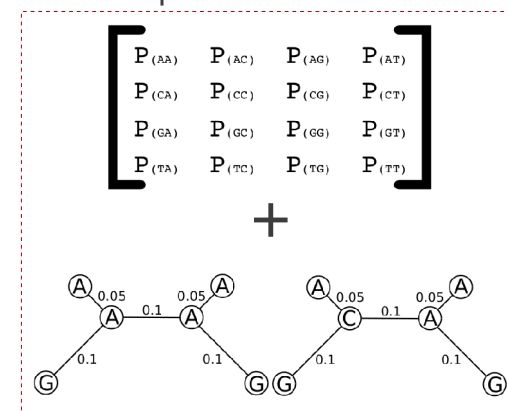
- 1a. posição: C ↔ T
- 3a. posição: G ↔ T
- 10a. posição: T ↔ A

CODIFICAÇÃO: matriz de dados

	c <sub>1</sub>	c <sub>2</sub>	c <sub>3</sub>	c <sub>4</sub>	c <sub>5</sub>	c <sub>6</sub>	c <sub>7</sub>	c <sub>8</sub>	c <sub>9</sub>	c <sub>10</sub>
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):

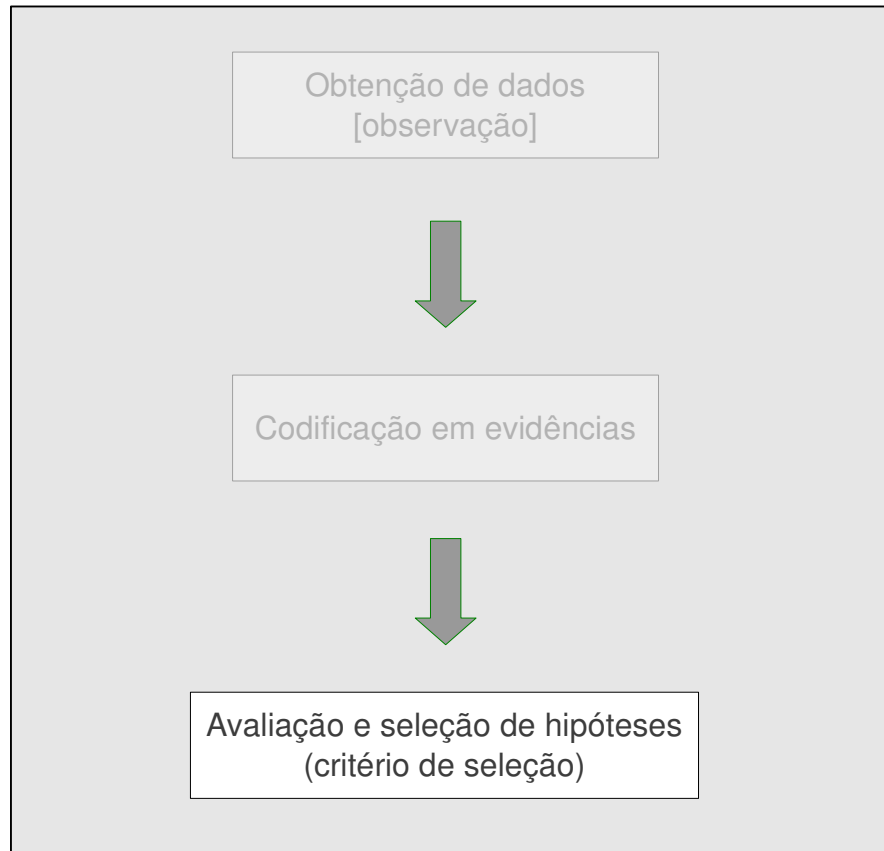
EVIDÊNCIAS: modelo de transformações + topologia que melhor explicam seus dados.



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

# Lógica da inferência filogenética

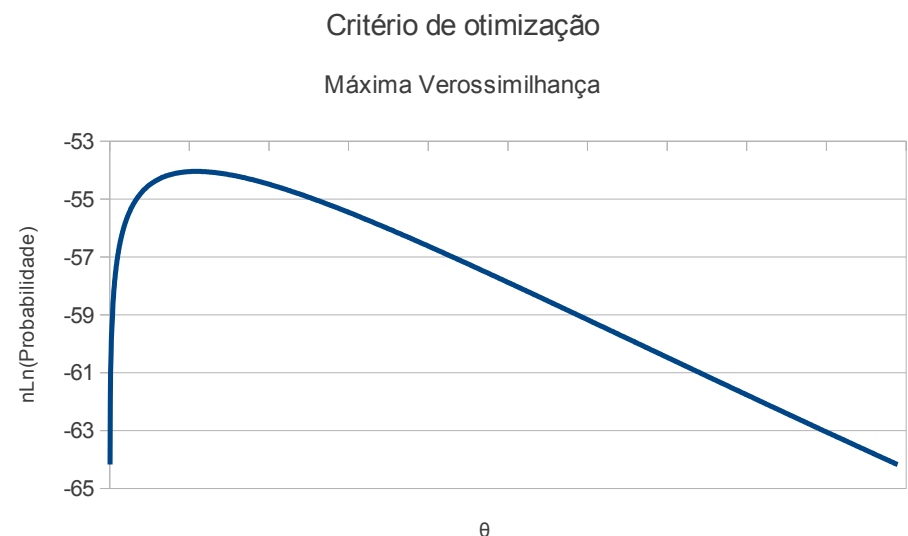
Avaliação e critério de seleção: verossimilhança máxima (ML)



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

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

Qual o valor de  $\theta$  que maximiza a probabilidade dos dados observados?



# Lógica da inferência filogenética

Avaliação e critério de seleção: verossimilhança máxima (ML)

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

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

Maximização do parâmetro  $\theta$ :

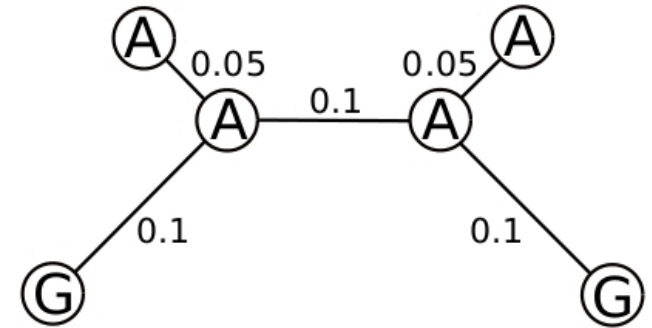
$$\theta = \begin{array}{c} \text{A} \\ \text{0.05} \\ \text{A} \\ \text{0.1} \\ \text{G} \end{array} \begin{array}{c} \text{A} \\ \text{0.05} \\ \text{A} \\ \text{0.1} \\ \text{G} \end{array} + GRT = \begin{bmatrix} P_{AA}(t)P_{AC}(t)P_{AG}(t)P_{AT}(t) \\ P_{CA}(t)P_{CC}(t)P_{CG}(t)P_{CT}(t) \\ P_{GA}(t)P_{GC}(t)P_{GG}(t)P_{GT}(t)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



# Lógica da inferência filogenética: 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:

	<b>A</b>	<b>C</b>	<b>G</b>	<b>T</b>
<b>A</b>	$P_{(A,A)}$	$P_{(A,C)}$	$P_{(A,G)}$	$P_{(A,T)}$
<b>C</b>	$P_{(C,A)}$	$P_{(C,C)}$	$P_{(C,G)}$	$P_{(C,T)}$
<b>G</b>	$P_{(G,A)}$	$P_{(G,C)}$	$P_{(G,G)}$	$P_{(G,T)}$
<b>T</b>	$P_{(T,A)}$	$P_{(T,C)}$	$P_{(T,G)}$	$P_{(T,T)}$

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

probabilidade de  
mudança de estado

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

probabilidade de não  
mudança de estado

$\nu = 0.1$				
	<b>A</b>	<b>C</b>	<b>G</b>	<b>T</b>
<b>A</b>	0.9064	0.0312	0.0312	0.0312
<b>C</b>	0.0312	0.9064	0.0312	0.0312
<b>G</b>	0.0312	0.0312	0.9064	0.0312
<b>T</b>	0.0312	0.0312	0.0312	0.9064

$\nu = 0.05$				
	<b>A</b>	<b>C</b>	<b>G</b>	<b>T</b>
<b>A</b>	0.9516	0.0161	0.0161	0.0161
<b>C</b>	0.0161	0.9516	0.0161	0.0161
<b>G</b>	0.0161	0.0161	0.9516	0.0161
<b>T</b>	0.0161	0.0161	0.0161	0.9516

# Lógica da inferência filogenética: modelos complexos

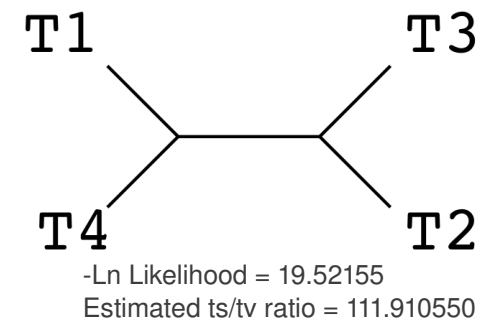
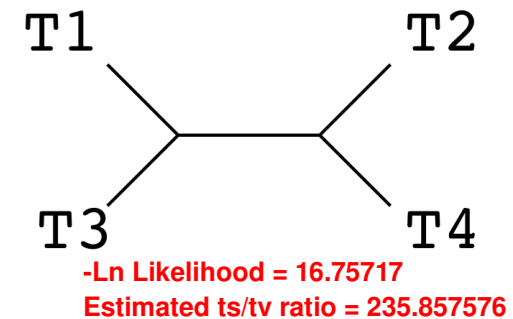
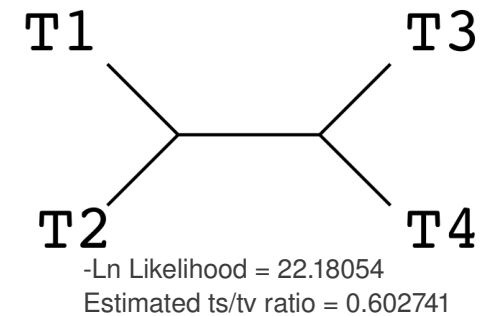
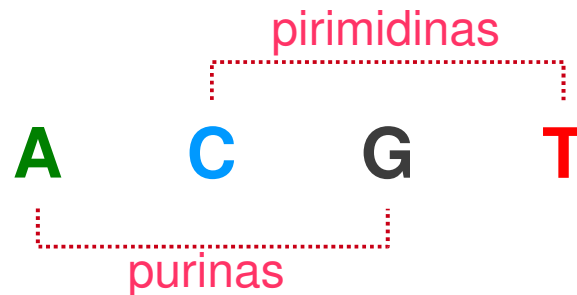
## K2P

Kimura Two Parameters

	<b>A</b>	<b>C</b>	<b>G</b>	<b>T</b>
<b>A</b>	$P_{(A,A)}$	$P_{(A,C)}$	$P_{(A,G)} *k$	$P_{(A,T)}$
<b>C</b>	$P_{(C,A)}$	$P_{(C,C)}$	$P_{(C,G)}$	$P_{(C,T)} *k$
<b>G</b>	$P_{(G,A)} *k$	$P_{(G,C)}$	$P_{(G,G)}$	$P_{(G,T)}$
<b>T</b>	$P_{(T,A)}$	$P_{(T,C)} *k$	$P_{(T,G)}$	$P_{(T,T)}$

Considere:

<b>T1</b>	<b>A</b>	<b>C</b>	<b>G</b>	<b>T</b>
<b>T2</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>A</b>
<b>T3</b>	<b>G</b>	<b>T</b>	<b>A</b>	<b>C</b>
<b>T4</b>	<b>T</b>	<b>A</b>	<b>C</b>	<b>G</b>



... 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

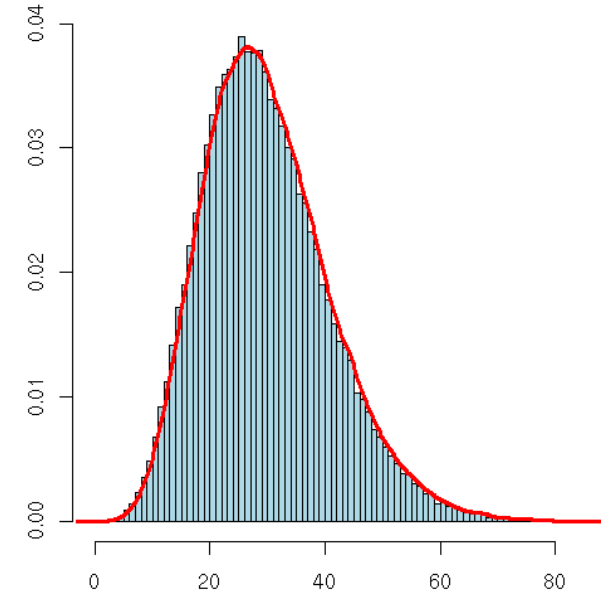
Méritos relativos dos métodos:

## Epistemologia

Comportamento estatístico

Performance

Congruência (?)



## 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

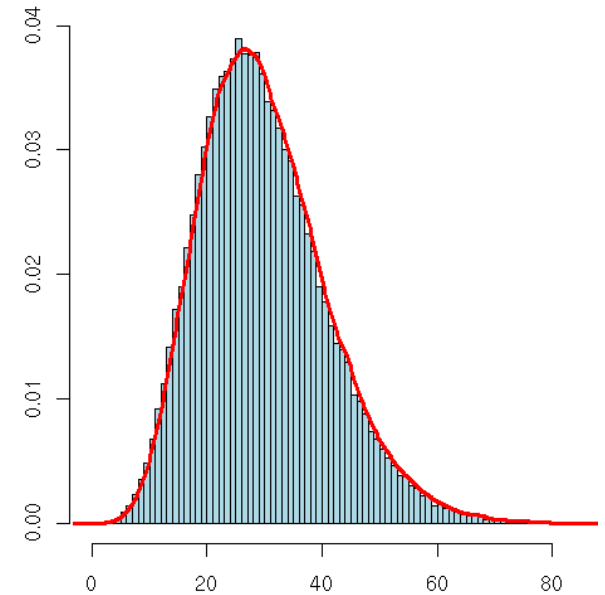
Méritos relativos dos métodos:

## 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*

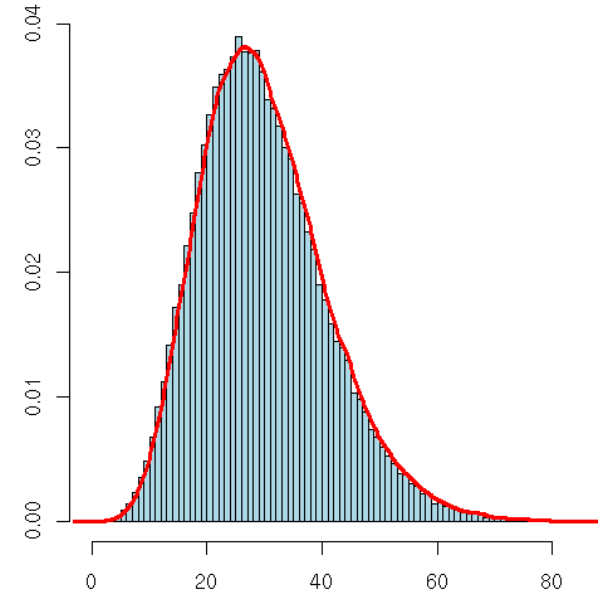
Méritos relativos dos métodos:

**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*

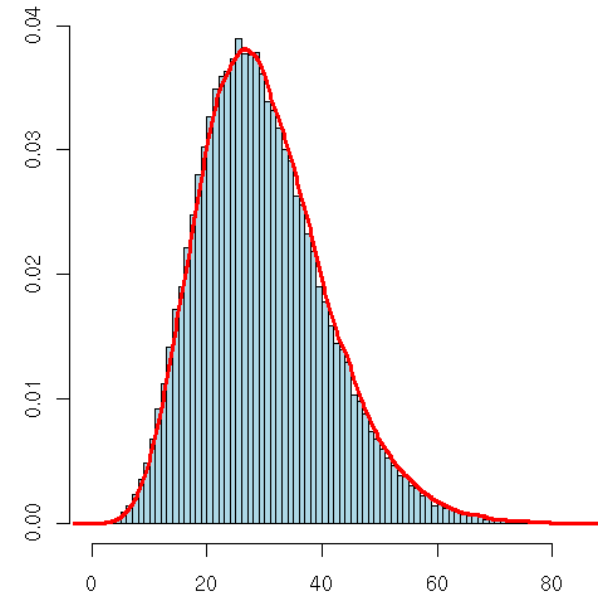
Méritos relativos dos métodos:

**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*

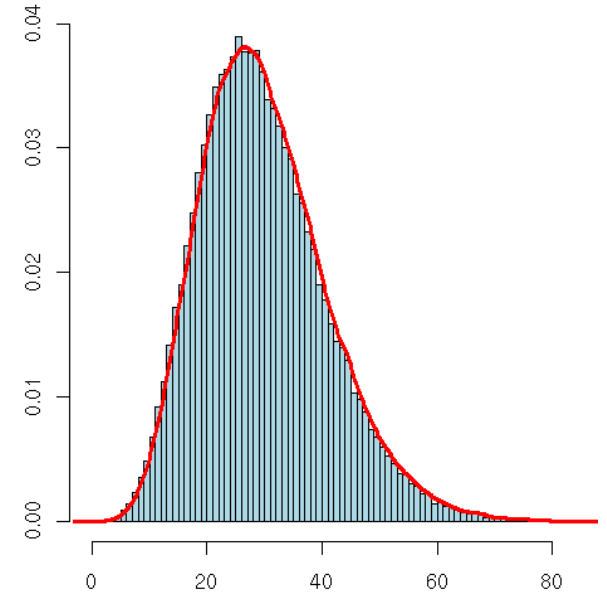
Méritos relativos dos métodos:

**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

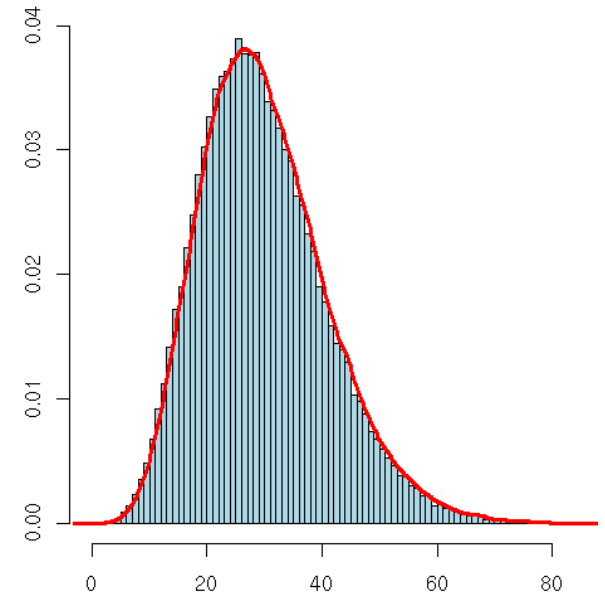
Méritos relativos dos métodos:

## 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*

*Probabilismo*

*Verossimilhança e Probabilidades*

*Escolha de critérios de otimalidade*