

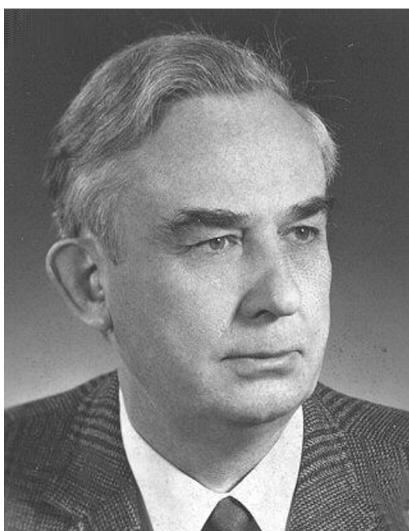
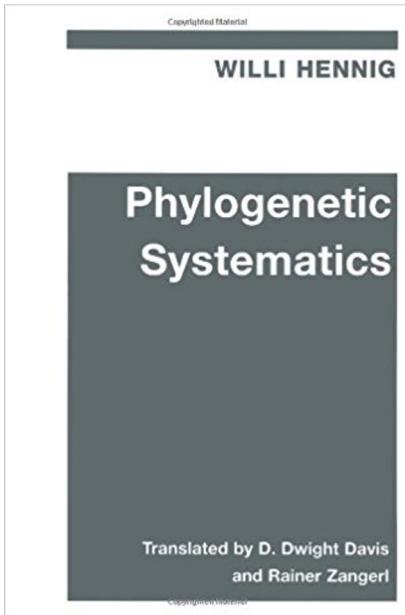
# INFERENCIA FILOGENÉTICA

## USANDO CRITERIO DE OPTIMALIDAD

Julián Aguirre-Santoro  
Instituto de Ciencias Naturales  
Universidad Nacional de Colombia



# SISTEMÁTICA FILOGENÉTICA



**Objetivo:** reconstruir las relaciones de parentesco entre taxones y proveer una clasificación concordante usando clados monofiléticos

## Axioma

**S**

En la naturaleza hay un orden jerárquico

Los caracteres permiten reconstruir ese orden

Se acepta la reconstrucción con menor número de explicaciones

# **FUNDAMENTOS DE LA SISTEMÁTICA DE ACUERDO A HENNIG**

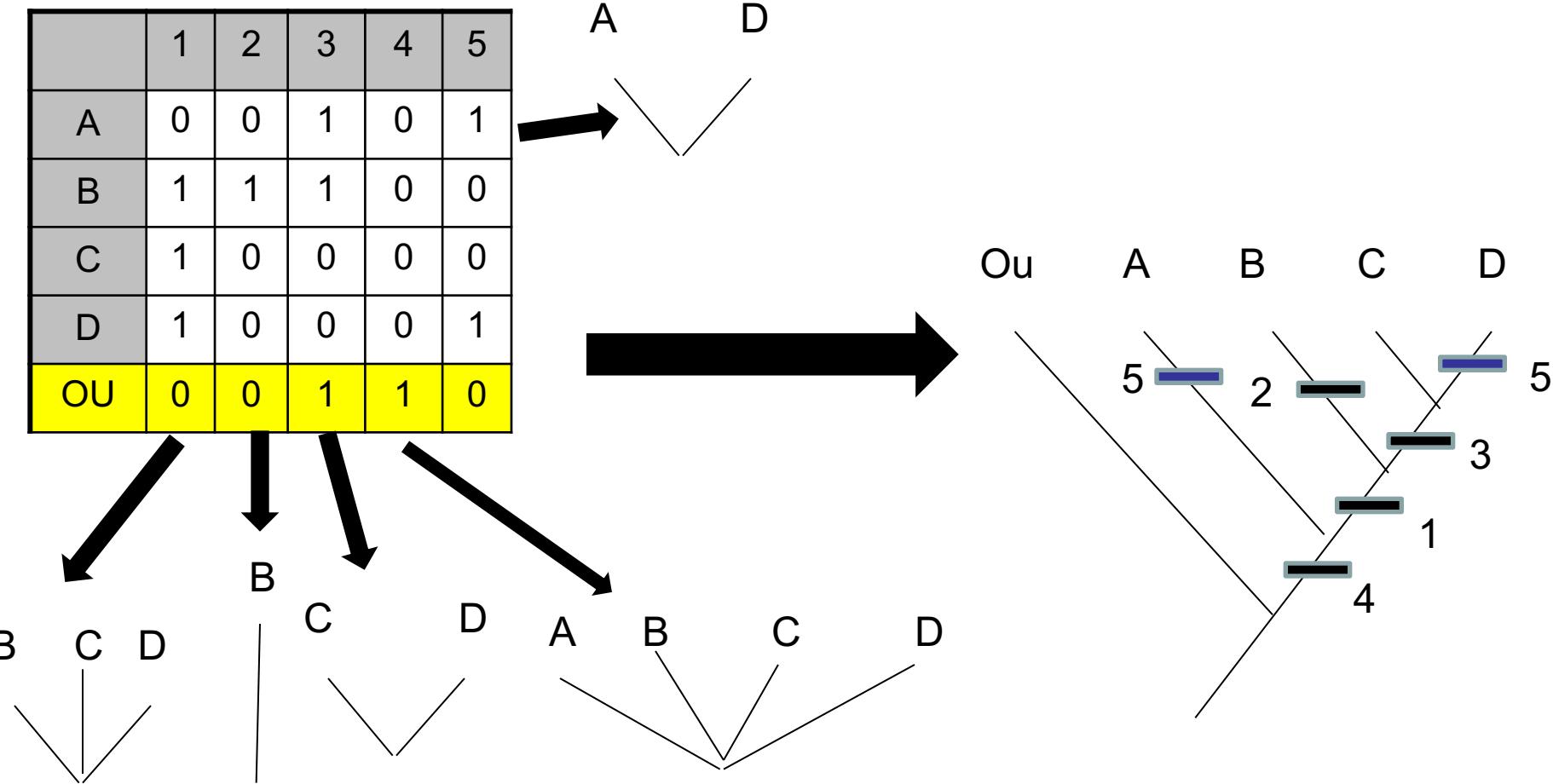
- I. La única base objetiva de relación entre especies es la genealógica
- I. La monofilia, como el grupo compuesto por la especie ancestral y todos sus descendientes, define los taxones objetivamente
- I. Especies y taxones superiores pueden definirse como monofiléticos si y solo si comparten una novedad evolutiva única (sinapomorfía)

*El trabajo de la sistemática es definir grupos monofiléticos*

# ARGUMENTACIÓN HENNIGIANA

- Determinar homología primaria
- Determinar **polaridad** (Grupo ajeno)
- Identifique congruencias
- Construya el o los árboles

# ARGUMENTACIÓN HENNIGIANA



# Argumentación Hennigiana

Infiera las relaciones filogenéticas de las monocotiledóneas con base en la siguiente matriz:

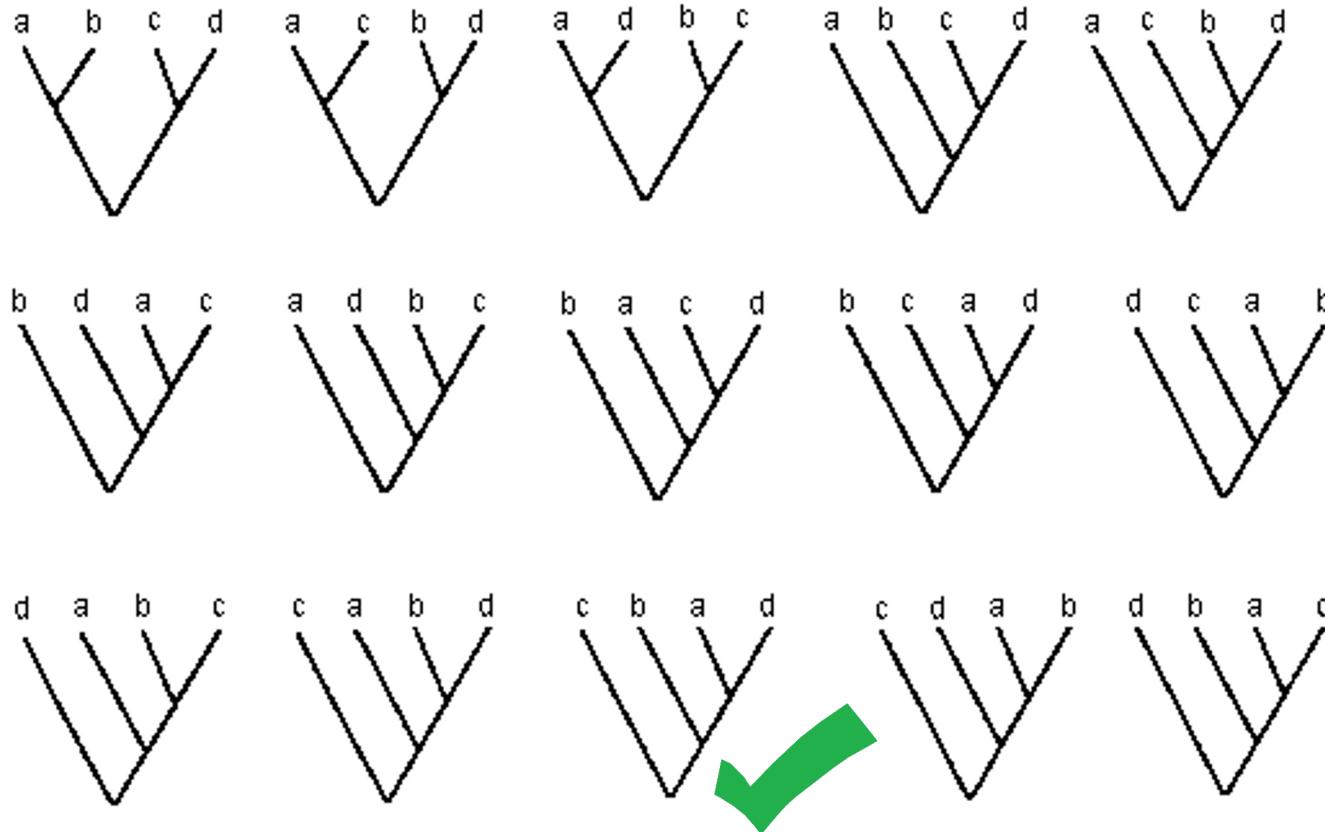
	Flores trímeras	Trepadora	Cuerpos de sílica	Cera tipo "Strelitzia"	Plantas dioicas	Fruto capsular	Venación palmeada	Ovario súpero
Algodón	0	0	0	0	0	0	0	0
Ñame	1	1	0	0	1	0	0	1
Coco	1	0	0	0	1	1	1	0
Piña	1	0	1	1	0	1	1	1
Banano	1	0	1	1	0	1	1	1

# PROBLEMAS DE LA ARGUMENTACION HENNIGIANA

- Polarización *a priori* de caracteres (grupo ajeno = ancestral)
- El método asume que **no hay homoplasia** (poco realista)
  - Caracteres inconsistentes con otros violan el modelo Hennigiano
- Imposible evitar errores o malas interpretaciones al codificar caracteres

# CRITERIO DE OPTIMALIDAD

Medida que permite decidir, con base en un conjunto de datos, cuáles árboles (hipótesis) son mejores y cuales son peores



# MÁXIMA PARSIMONIA

**Cuchilla de Occam:** la mejor hipótesis para explicar un proceso es aquella que requiere el menor número de suposiciones

**En inferencia filogenética:** el mejor árbol es aquel que explica los datos observados con la menor cantidad de homoplasia posible (menos transformaciones)

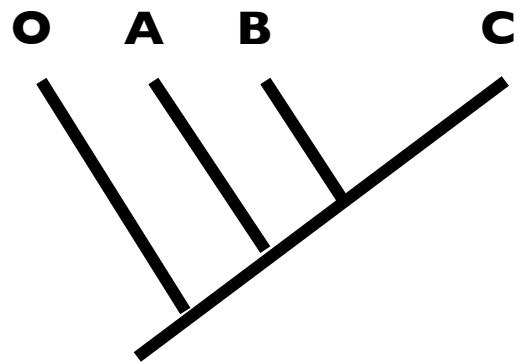
# MÁXIMA PARSIMONIA

- I. Contar el **mínimo** número de cambios (pasos) de cada carácter en un árbol determinado
- I. Sumar todos los números de pasos para determinar la **LONGITUD DEL ÁRBOL**
- I. Repetir en los otros árboles alternativos y escoger aquel con la menor longitud como el **ÁRBOL MÁS PARSIMONIOSO**

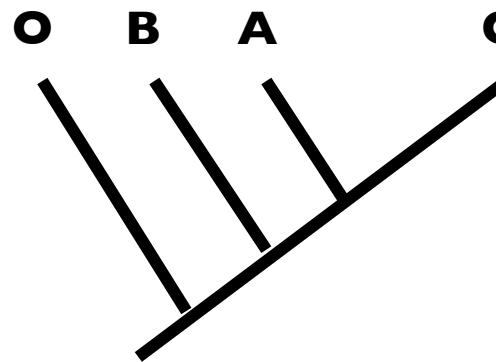
# MÁXIMA PARSIMONIA

## EJEMPLO

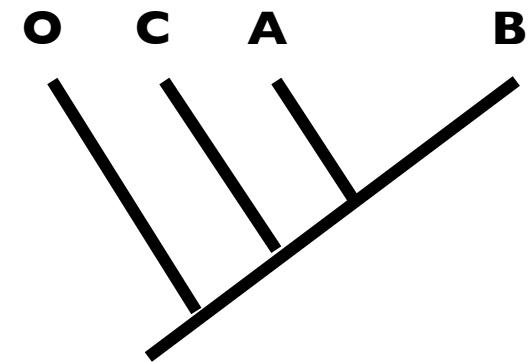
	1	2	3	4	5	6	7	8
O	0	0	0	0	0	0	0	0
A	0	1	0	0	0	1	1	0
B	1	1	0	1	1	1	1	1
C	0	0	1	1	0	0	0	0



Árbol 1



Árbol 2

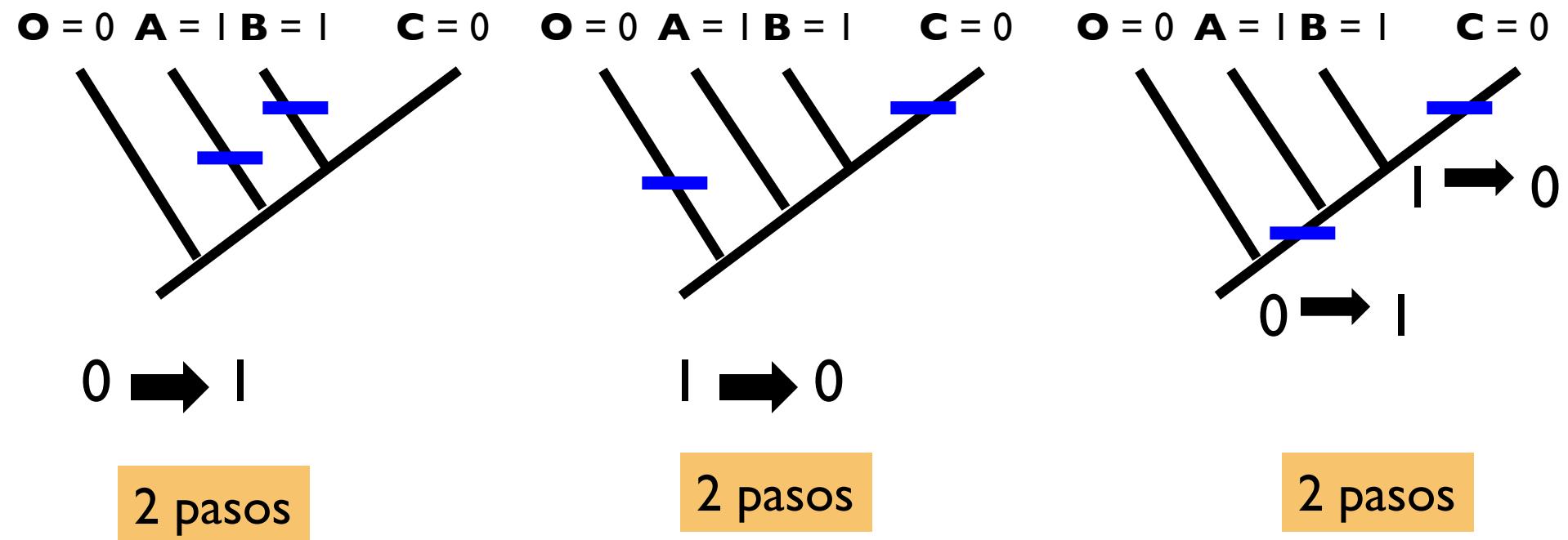


Árbol 3

# MÁXIMA PARSIMONIA

Árbol I

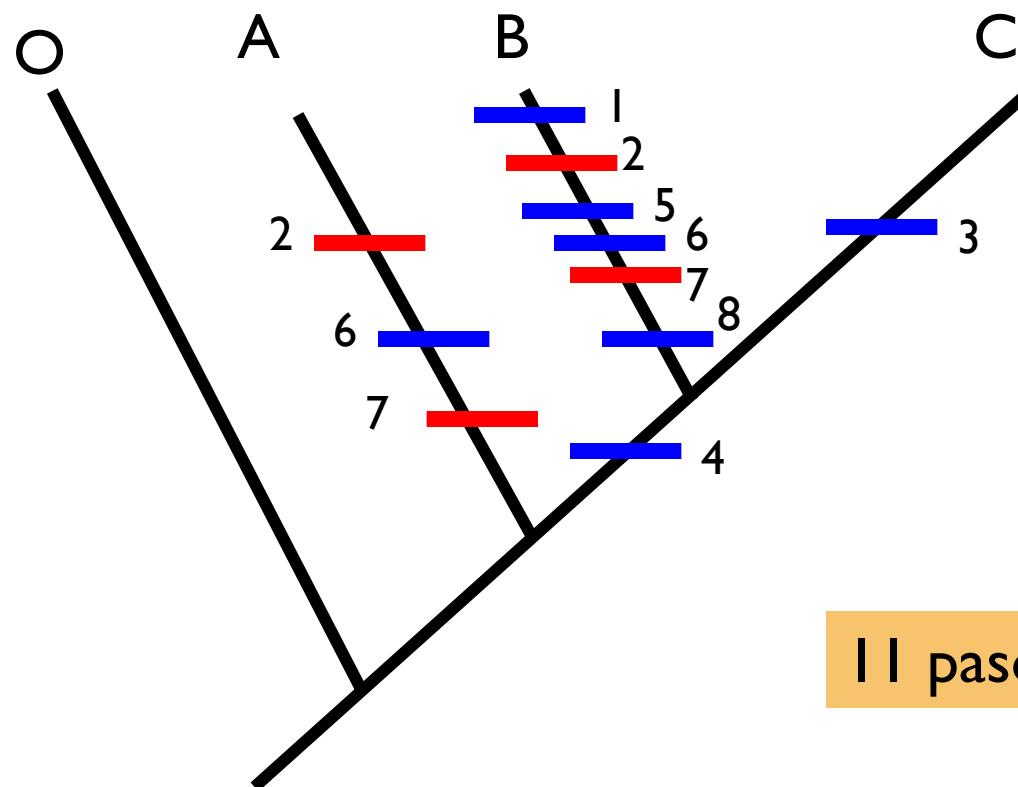
	1	2	3	4	5	6	7	8
O	0	0	0	0	0	0	0	0
A	0	1	0	0	0	1	1	0
B	1	1	0	1	1	1	1	1
C	0	0	1	1	0	0	0	0



# MÁXIMA PARSIMONIA

Árbol I

	1	2	3	4	5	6	7	8
O	0	0	0	0	0	0	0	0
A	0	1	0	0	0	1	1	0
B	1	1	0	1	1	1	1	1
C	0	0	1	1	0	0	0	0

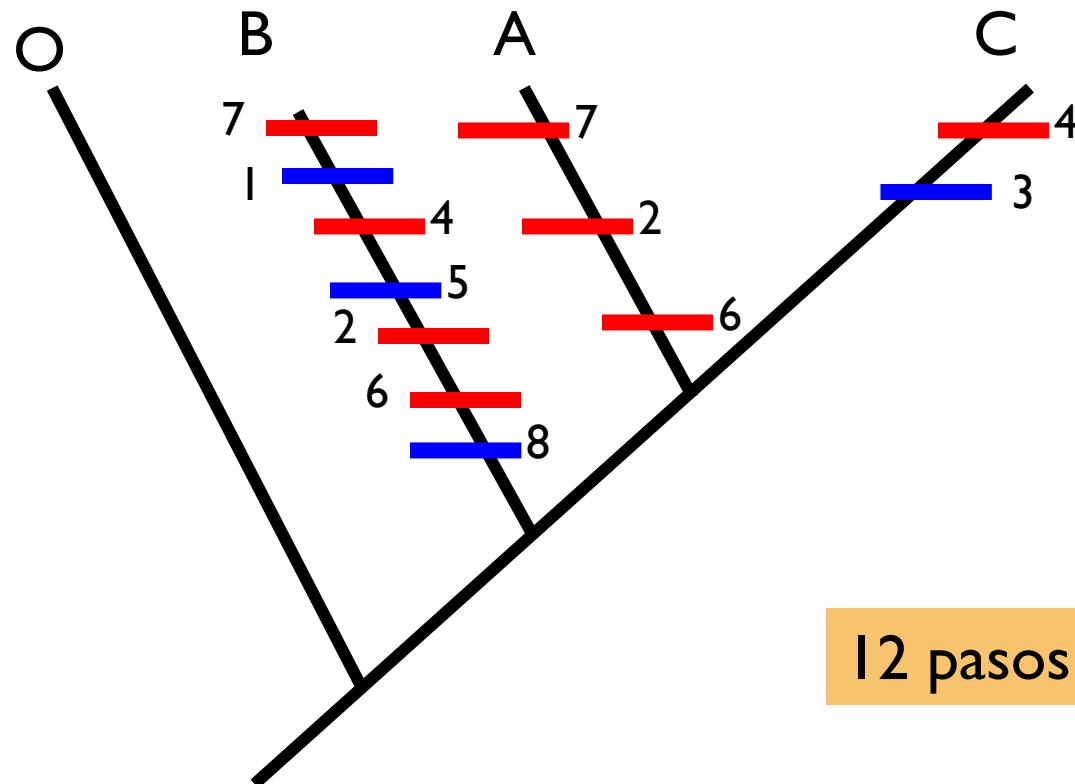


II pasos

# MÁXIMA PARSIMONIA

Árbol 2

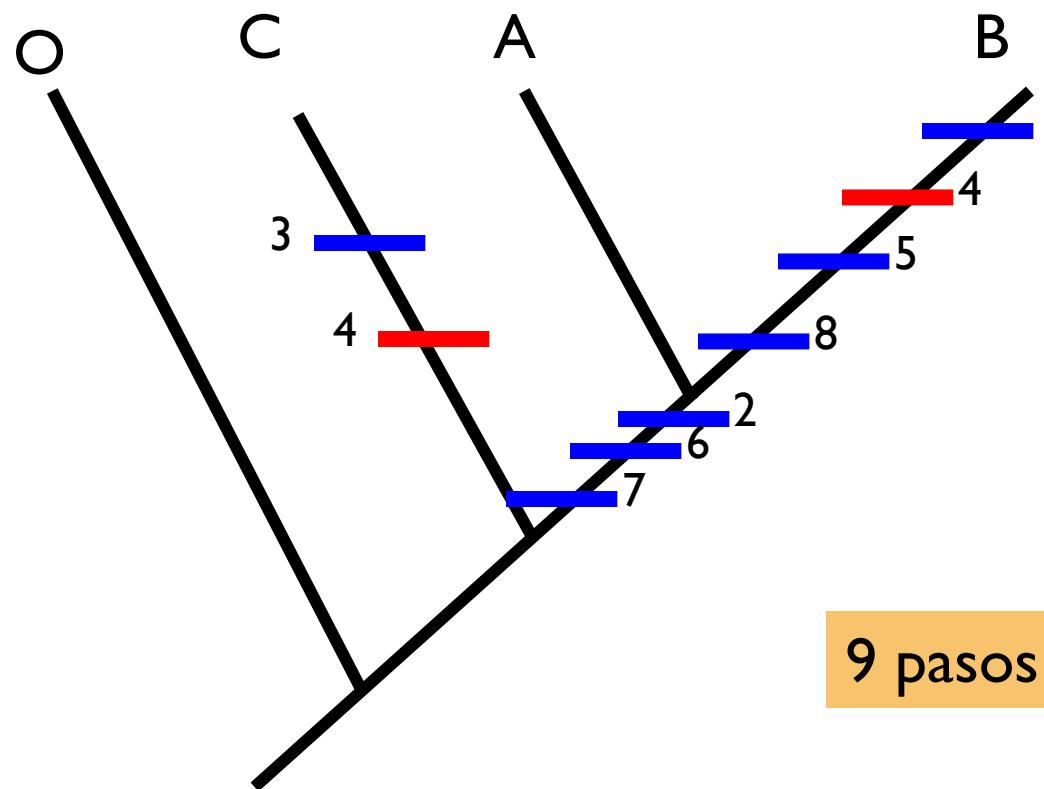
	1	2	3	4	5	6	7	8
O	0	0	0	0	0	0	0	0
A	0	1	0	0	0	1	1	0
B	1	1	0	1	1	1	1	1
C	0	0	1	1	0	0	0	0



# MÁXIMA PARSIMONIA

Árbol 3

	1	2	3	4	5	6	7	8
O	0	0	0	0	0	0	0	0
A	0	1	0	0	0	1	1	0
B	1	1	0	1	1	1	1	1
C	0	0	1	1	0	0	0	0



# MÁXIMA PARSIMONIA

	1	2	3	4	5	6	7	8	
O	0	0	1	0	1	1	0	0	
A	0	1	1	0	1	0	1	0	
B	1	1	1	1	0	0	1	1	
C	0	0	0	1	1	1	0	0	Total length
Length on tree 1	1	2	1	1	1	2	2	1	11
Length on tree 2	1	2	1	2	1	2	2	1	12
Length on tree 3	1	1	1	2	1	1	1	1	9

Árbol más  
parsimonioso

# Máxima Parsimonia

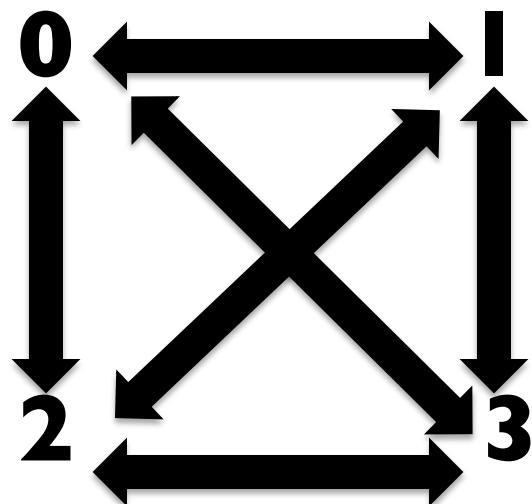
Infiera las relaciones filogenéticas de las monocotiledóneas con base en la siguiente matriz:

	Flores trímeras	Trepadora	Cuerpos de sílica	Cera tipo "Strelitzia"	Plantas dioicas	Fruto capsular	Venación palmeada	Ovario súpero
Algodón	0	0	0	0	0	0	0	0
Ñame	1	1	0	0	1	0	0	1
Coco	1	0	0	0	1	1	1	0
Piña	1	0	1	1	0	1	1	1

# MÁXIMA PARSIMONIA

## Variaciones de optimización de caracteres en árboles

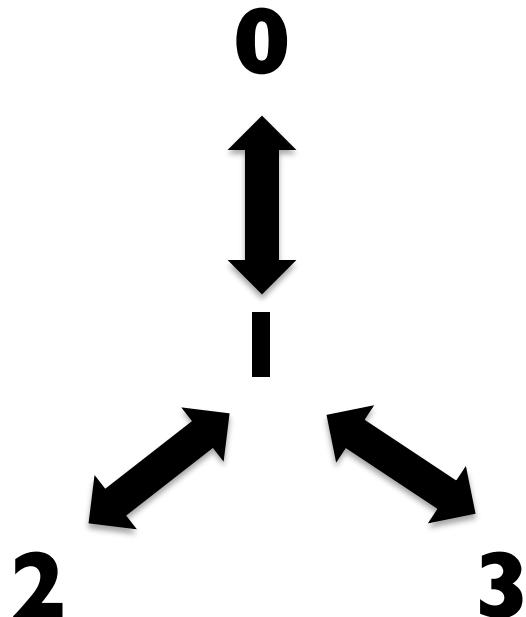
- Caracteres no ordenados o no aditivos (pesos iguales) = **Parsimonia de Fitch**



# MÁXIMA PARSIMONIA

## Variaciones de optimización de caracteres en árboles

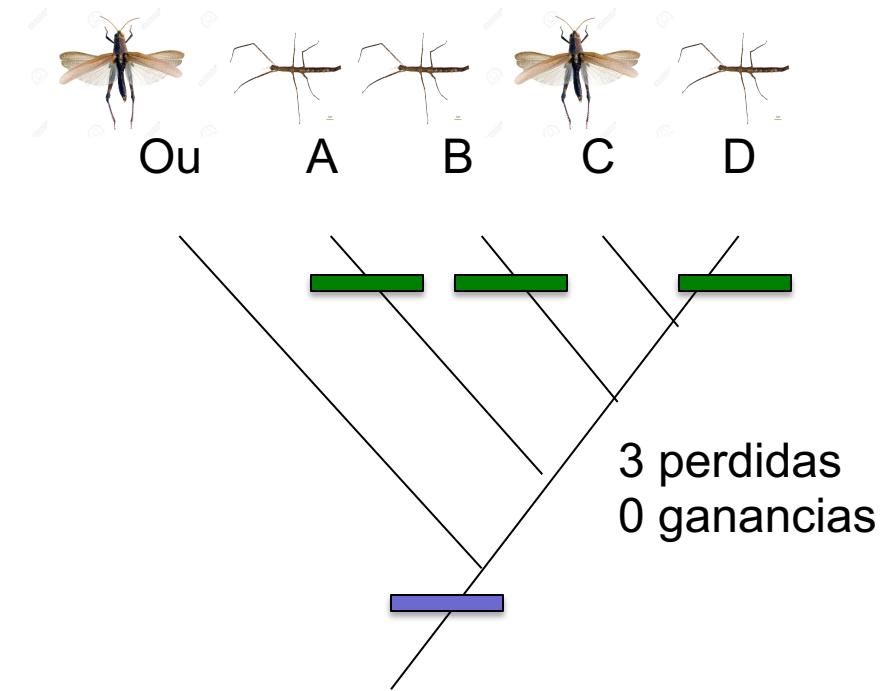
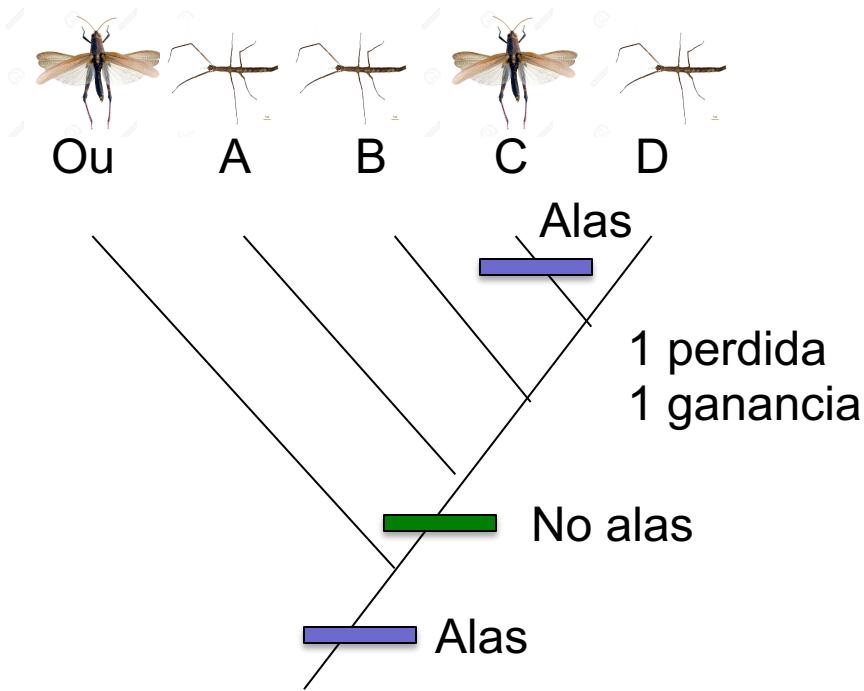
- Caracteres ordenados o aditivos = **Parsimonia de Wagner**

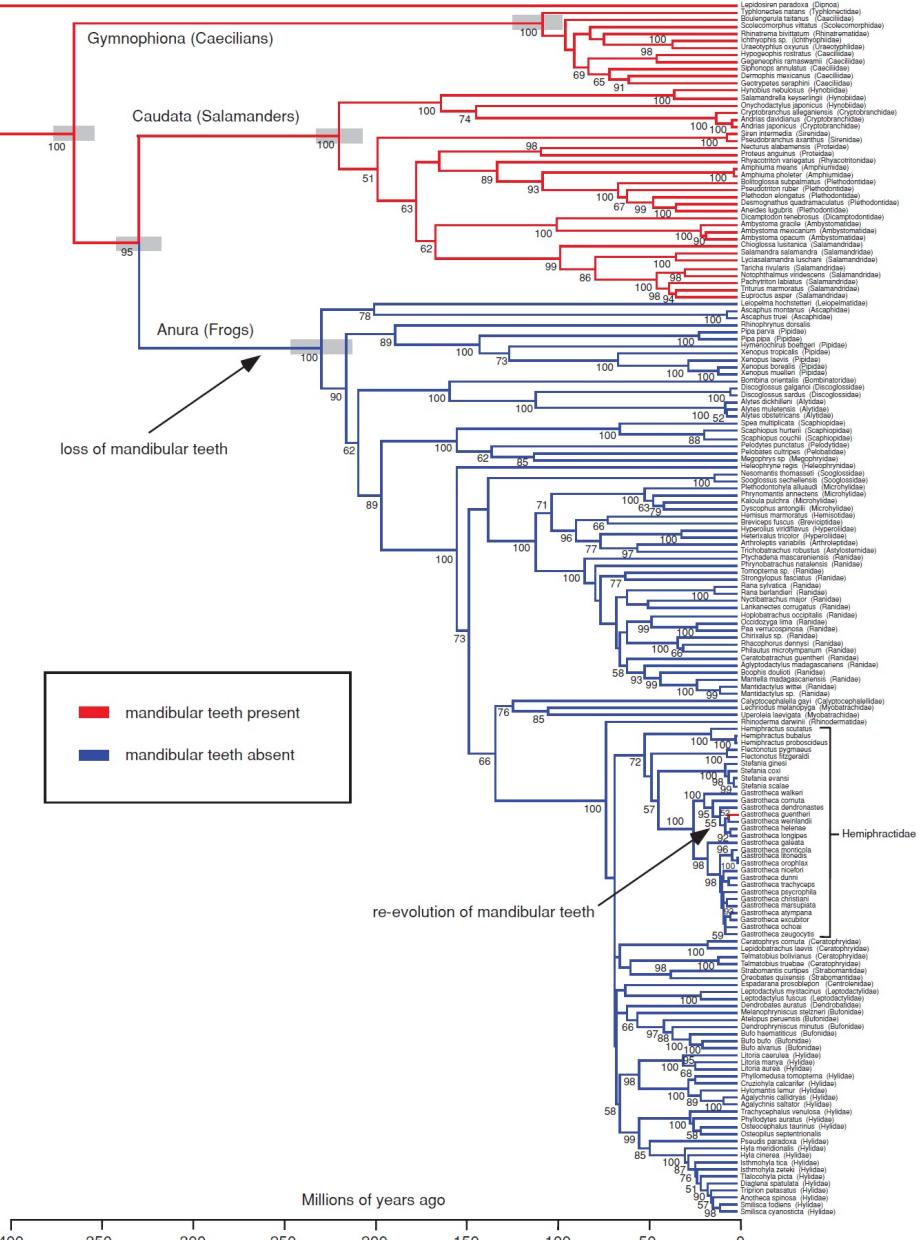


# MÁXIMA PARSIMONIA

## Parsimonia de Dollo

- Una vez un carácter complejo se puede perder muchas veces pero no podrá evolucionar de nuevo





# Dientes mandibulares en *Gastrotheca guentheri*



# MÁXIMA PARSIMONIA

## Variaciones de optimización de caracteres en árboles

- PESAJE DE CARACTERES
  - **Pesaje a priori (Parsimonia generalizada)**
    - Asignación de peso a criterio del investigador
    - Ej.: transversiones vs. transiciones, codones, caracteres diagnósticos
  - **Pesos a posteriori**
    - Pesaje a posteriori después de un análisis de pesos iguales (homólogos pesan más)

# MÁXIMA PARSIMONIA

	1	2	3	4	5	6	7	8		
O	0	0	1	0	1	1	0	0		
A	0	1	1	0	1	0	1	0		
B	1	1	1	1	0	0	1	1		
C	0	0	0	1	1	1	0	0		
Peso	1	1	1	5	1	1	1	1	Longi- tud	Costo total
Costo árbol 1	1	2	1	5	1	2	2	1	11	15
Costo árbol 2	1	2	1	10	1	2	2	1	12	20
Costo árbol 3	1	1	1	10	1	1	1	1	9	17

# El problema de encontrar árboles óptimos...

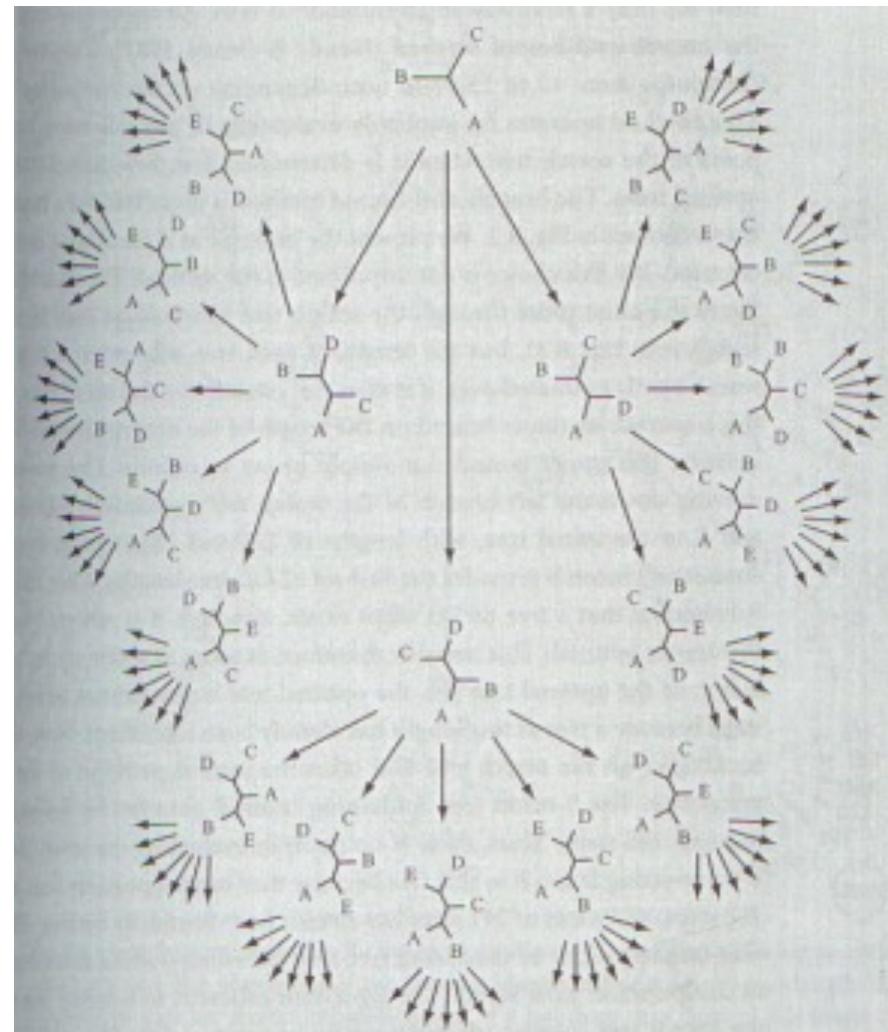
$$= (2n - 5)! / ((n-3)!2^{n-3})$$

Taxones	Árboles resueltos
1	--
2	1
3	3
4	15
5	105
6	945
7	10395
8	135135
9	2027025
10	34459425
11	654729075
12	13749310575
13	316234143225
14	7905853580625
15	213458046676875
16	6190283353629370
17	191898783962510000
18	6332659870762850000
19	221643095476699000000
20	6,66409461 x 10 E 98
62	> 10 E 100

# MÁXIMA PARSIMONIA

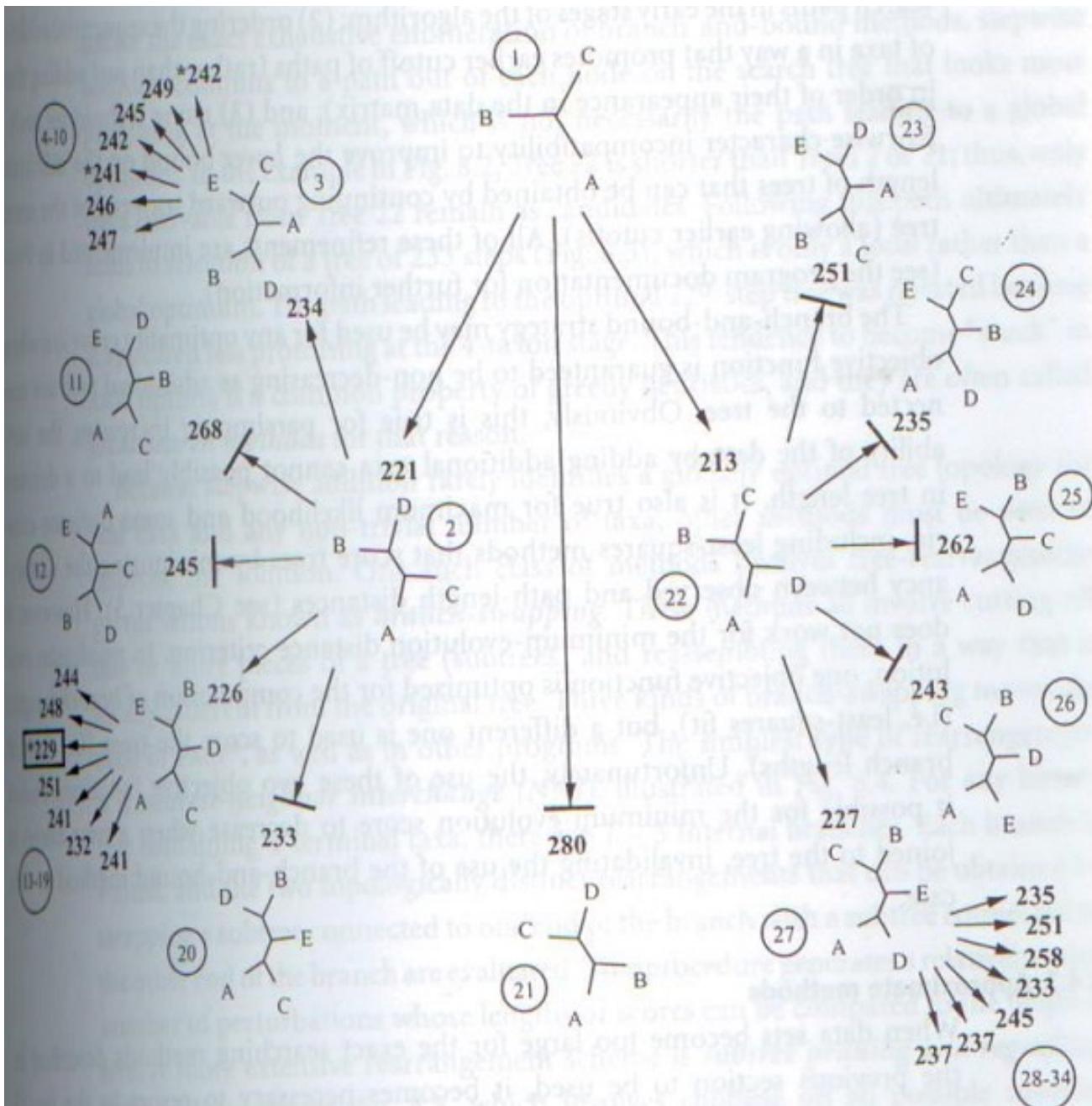
**Métodos exactos:**

I. Búsqueda exhaustiva



# MÁXIMA PARSIMONIA

**Métodos exactos:**  
2. Branch & Bound

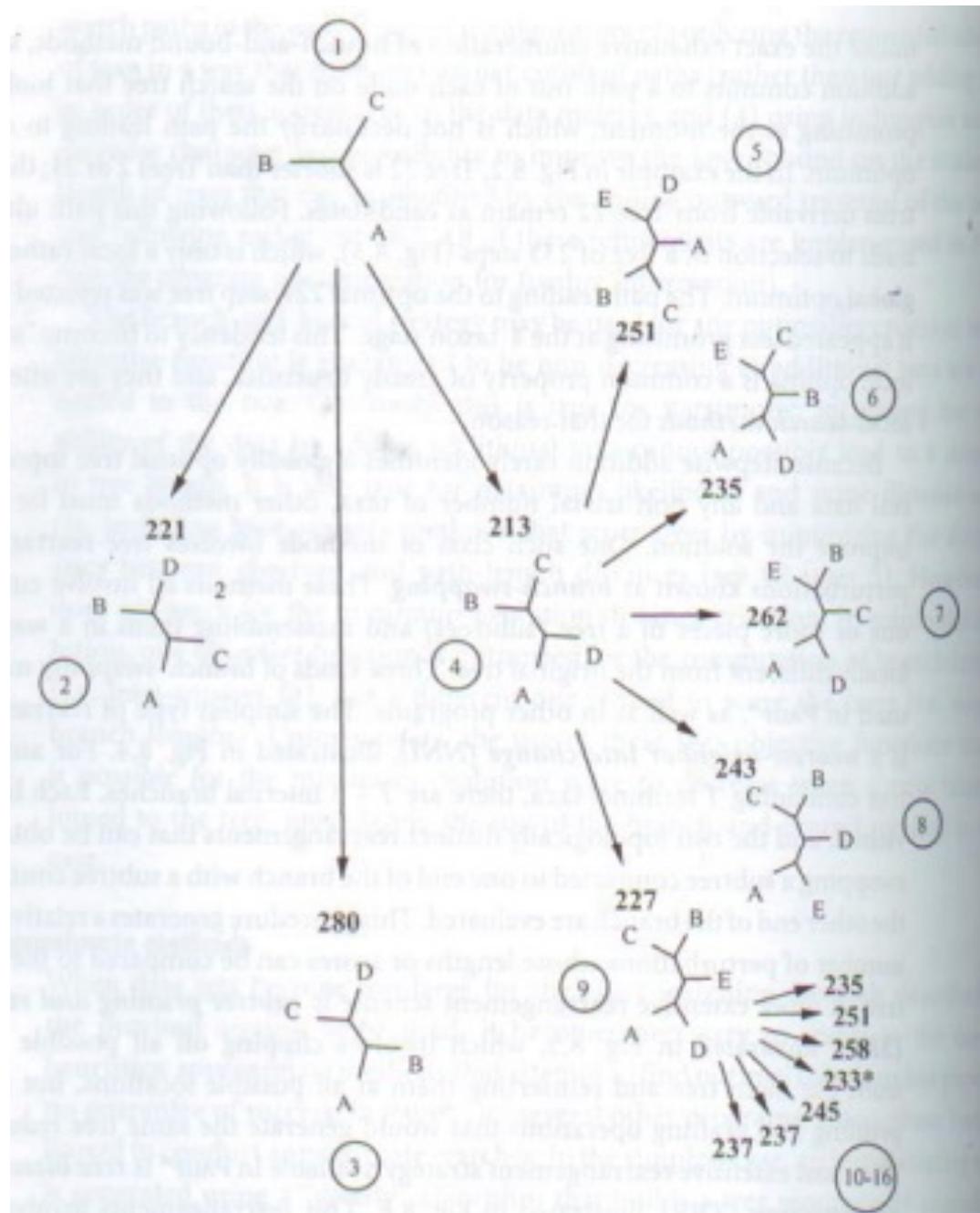


# MÁXIMA PARSIMONIA

## Métodos heurísticos:

### I. Buscar árbol inicial:

- Adición paso a paso  
(Stepwise addition)
- Aleatorio

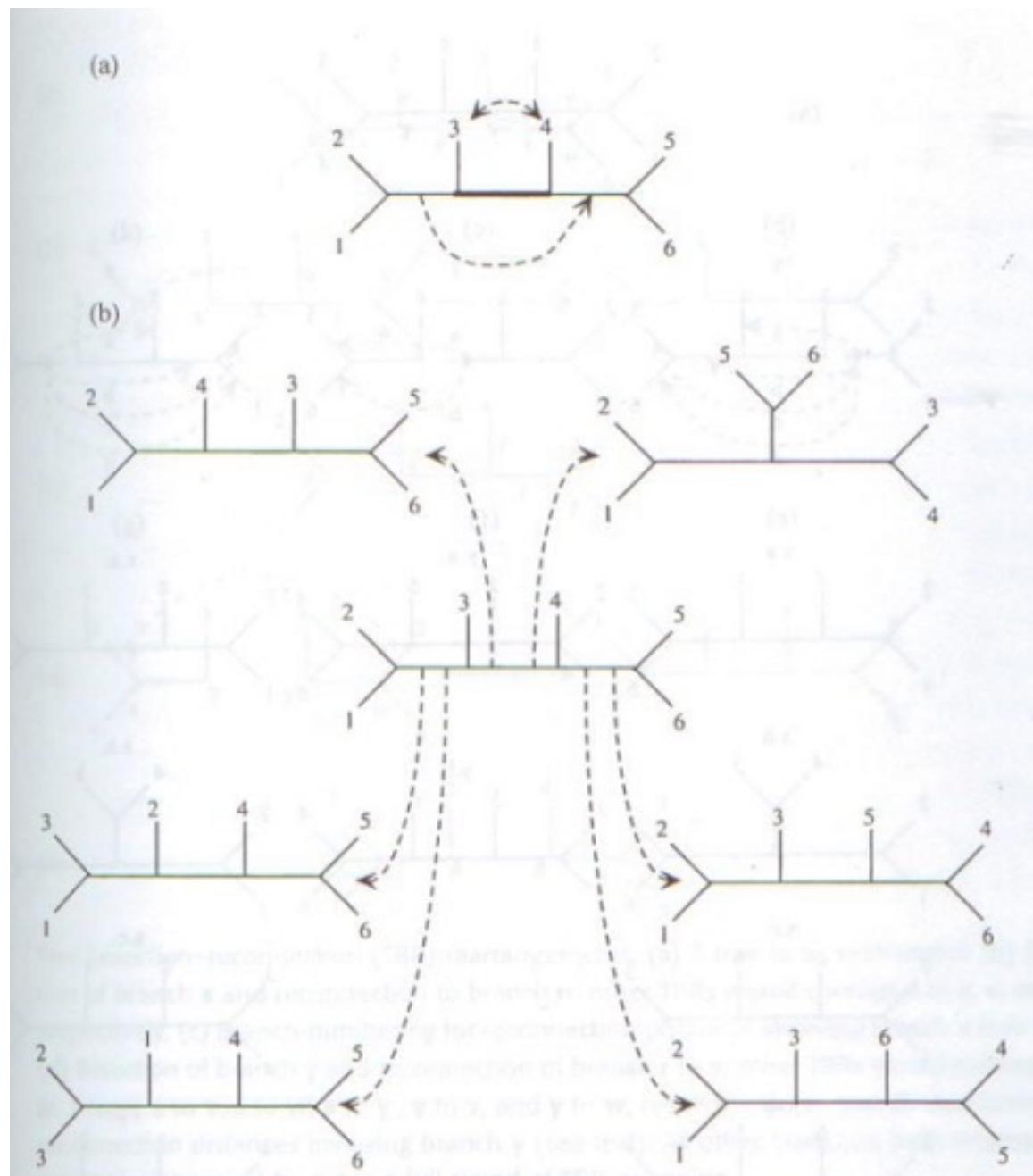


# MÁXIMA PARSIMONIA

## Métodos heurísticos:

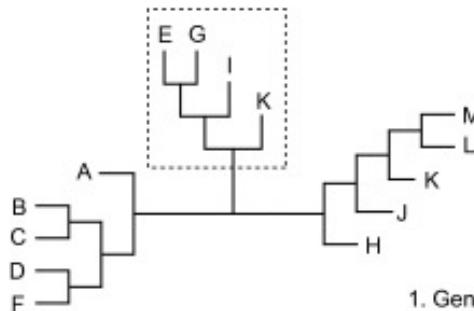
### 2. Perturbar árbol inicial:

- Nearest Neighbor Interchange (NNI)

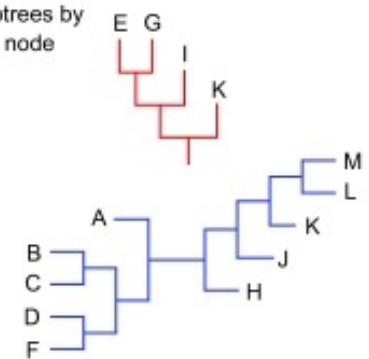


# MÁXIMA PARSIMONIA

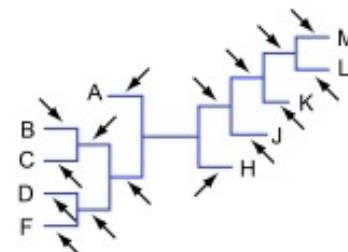
0. Starting tree



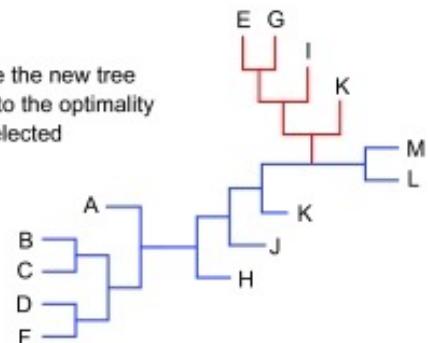
1. Generate two subtrees by breaking an internal node



2. Try to insert the red subtree at each node of the blue subtree



3. Evaluate the new tree according to the optimality criterion selected



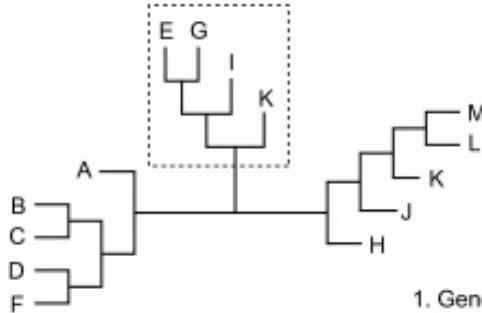
## Métodos heurísticos:

### 2. Perturbar árbol inicial:

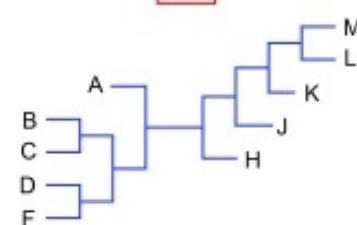
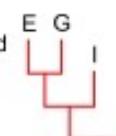
- Subtree Pruning & Regrafting (SPR)

# MÁXIMA PARSIMONIA

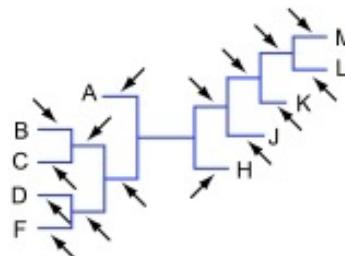
0. Starting tree



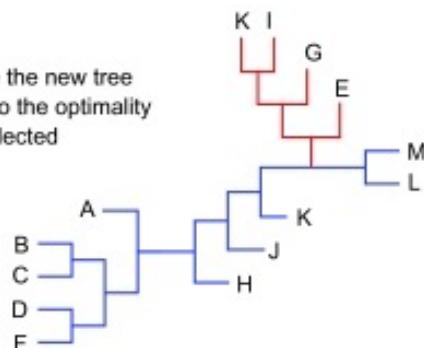
1. Generate two subtrees by breaking an internal node and re-rooting the subtree



2. Try to insert all possible rooted red subtrees at each node of the blue subtree



3. Evaluate the new tree according to the optimality criterion selected



## Métodos heurísticos:

### 2. Perturbar árbol inicial:

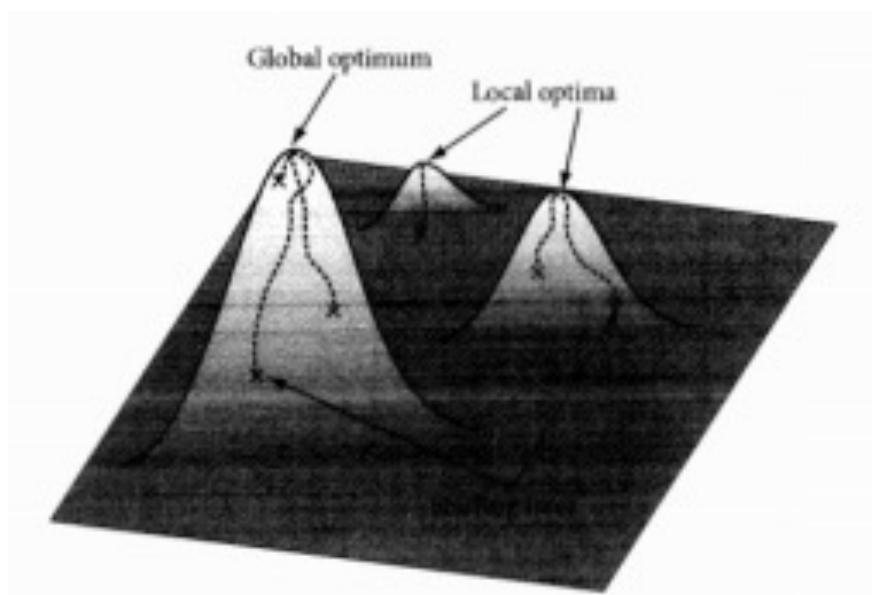
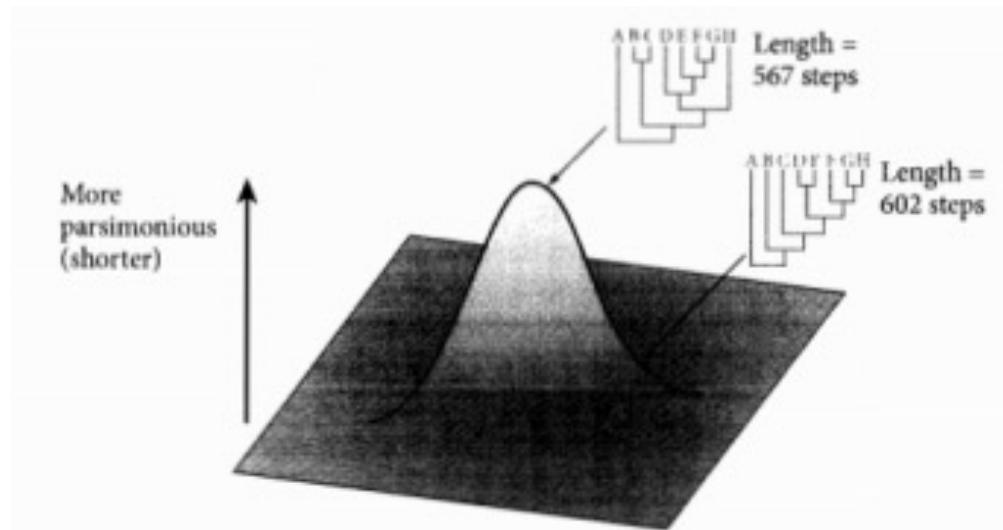
- Tree Bisection & Reconnection (TBR)

# MÁXIMA PARSIMONIA

## Métodos heurísticos:

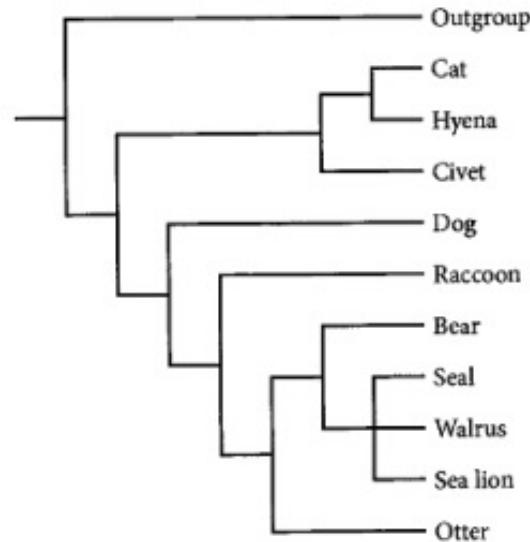
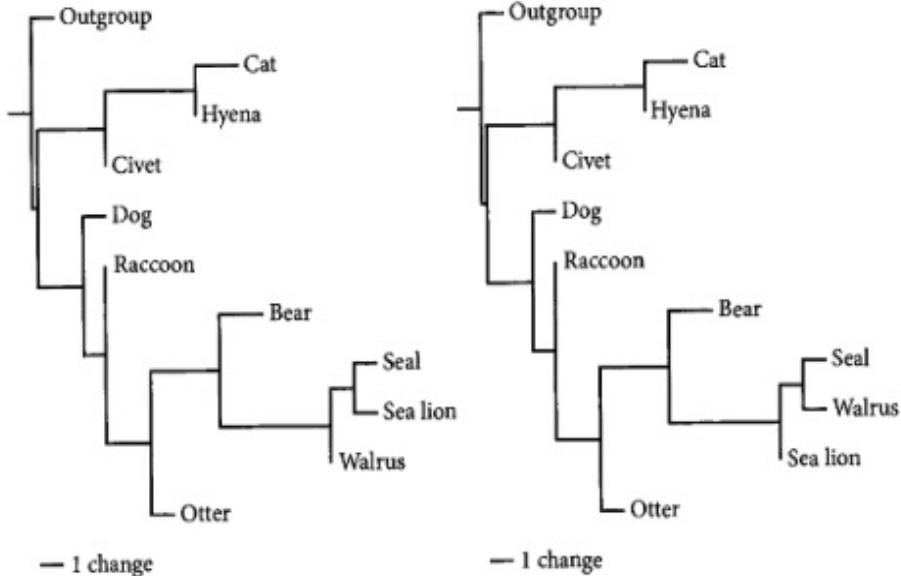
3. Visitar óptimos locales para tener óptimo global:

- Réplicas
- Stepwise-random-addition



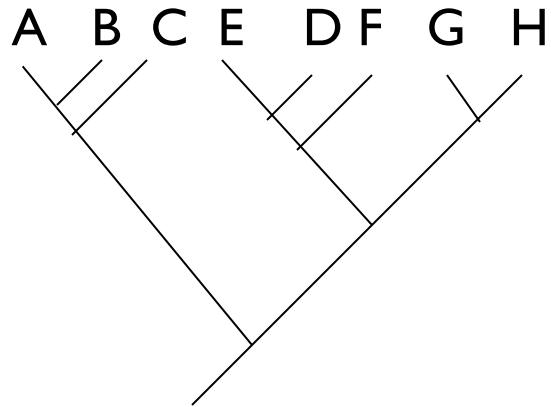
# MÁXIMA PARSIMONIA

¿Qué pasa si hay más de un árbol más parsimonioso?

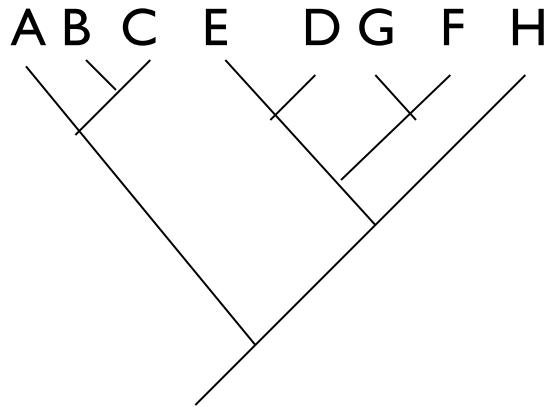


# Ejercicio

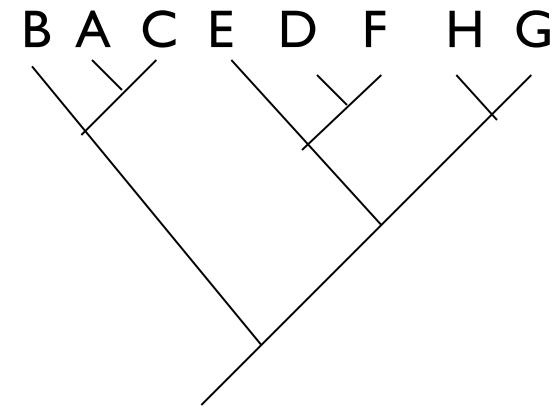
DNA gen COI



DNA gen 28S



DNA gen NADH

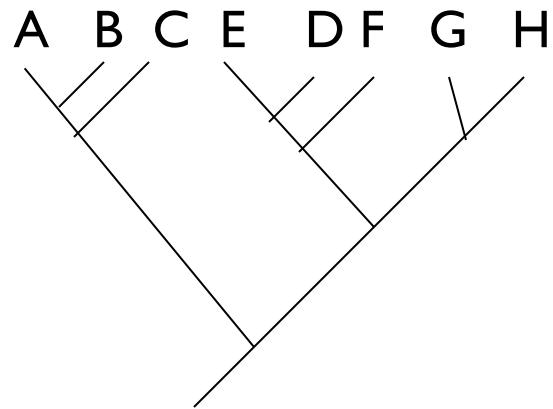


Generar árbol de consenso estricto

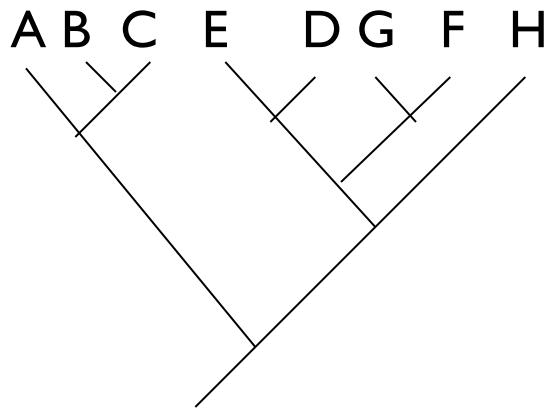
Generar árbol de consenso de mayoría

# Ejercicio

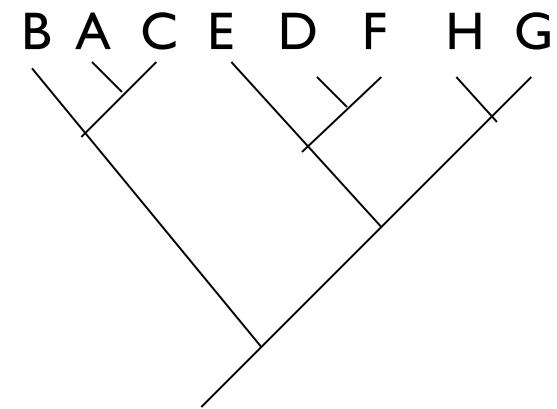
DNA gen COI



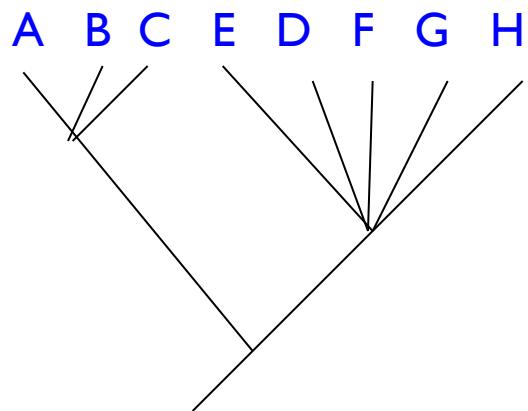
DNA gen 28S



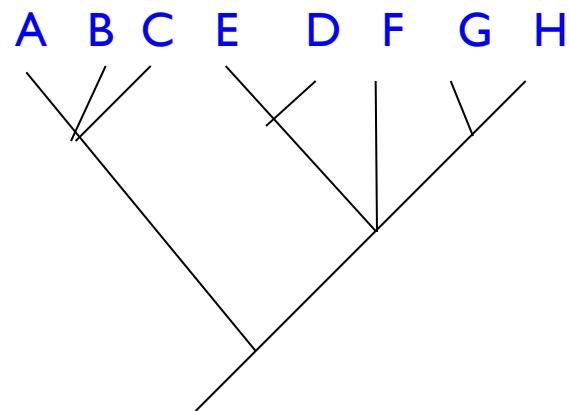
DNA gen NADH



Estricto



Mayoría



# MÁXIMA PARSIMONIA

## Problemas

- Se relaja la búsqueda de homologías primarias
- Longitud de ramas no se toma en cuenta (se ignora la tasa de evolución de los caracteres en cada rama)
- Atracción de ramas largas (¡¡entre más caracteres, peor!!)
- Pesaje de caracteres es necesario (aún si son pesos iguales)
  - No hay métodos formales para decidir pesos