

Preparando para aula

- Materiais necessários estão presentes no site:
 - <http://biodados.icb.ufmg.br/pinguim/filogenia/>
- Roteiro da aula:
 - Aula_Filogenia.pdf
- Crie a sua pasta:
 - Abra o terminal (Ctrl + Alt + T);
 - Crie uma pasta (comando “mkdir aula_filogenia”);
 - Vá para a pasta criada (comando “cd aula_filogenia”).
- Usuários com laptop → Baixe e instale o programa MEGA.
- Baixe na pasta o arquivo VIL1.fasta

Filogenia molecular

Tetsu Sakamoto

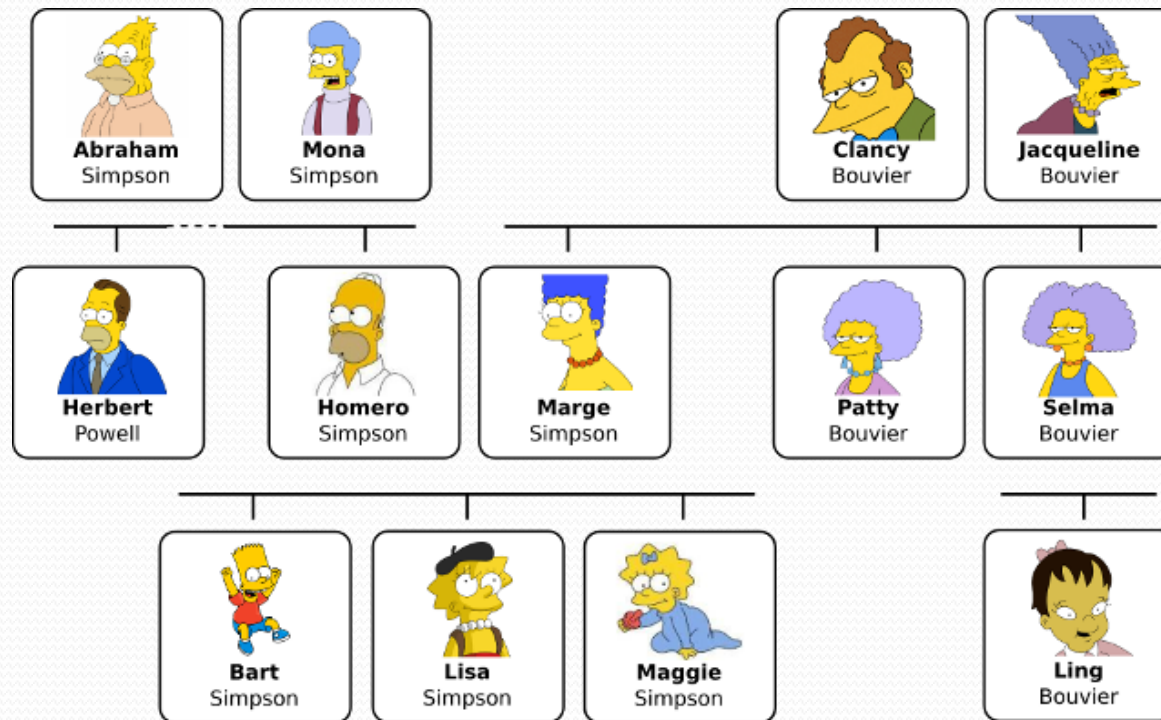
Doutorando em Bioinformática (UFMG)

Roteiro da aula

- Conteúdo:
 - Breve introdução sobre Filogenia Molecular;
 - Construção de uma árvore filogenética;
- Objetivo:
 - Conceitos da filogenia molecular;
 - Saber fazer uma leitura de uma árvore filogenética;
 - Saber construir uma árvore filogenética.

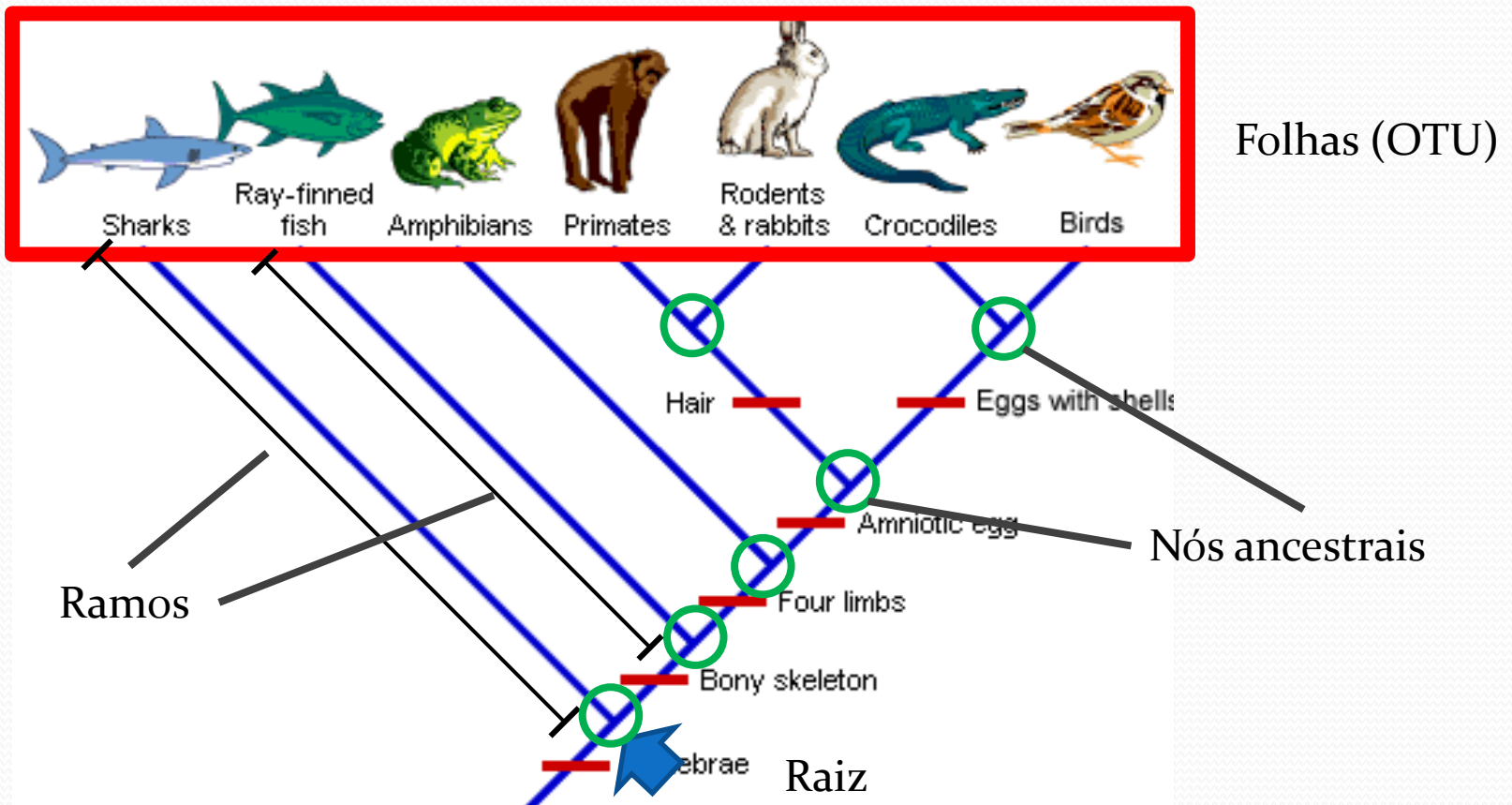
O que é Filogenia...

- Estudo das relações evolutivas entre grupos de organismos;
- Semelhante a uma árvore genealógica;

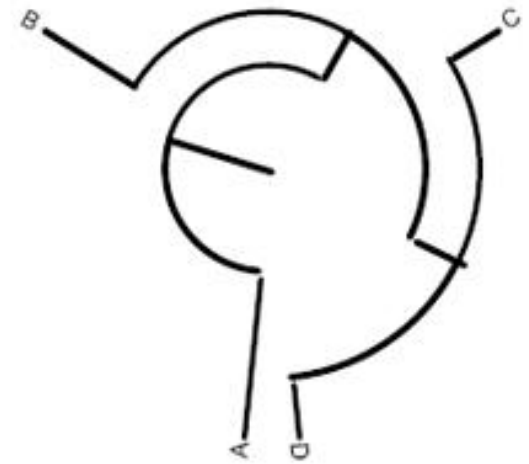
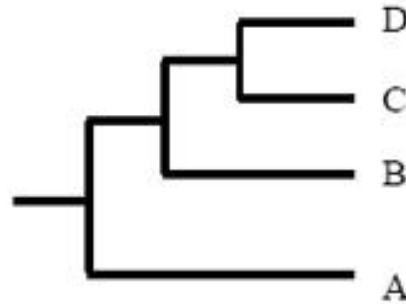
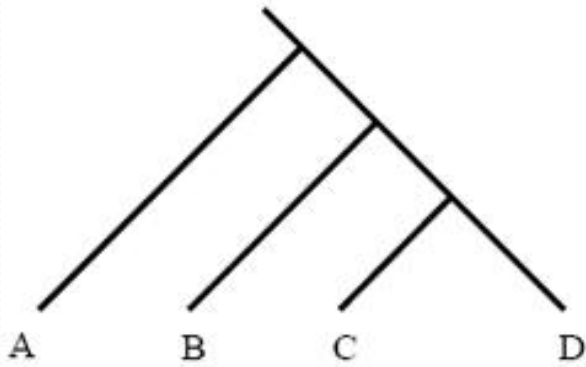


O que é Filogenia...

- Mas aplicada às espécies ou a entidades taxonômicas.



Tipos de árvores filogenéticas



Formas de estudar Filogenia...

- Dados morfológicos
 - Crânio, dentição...
 - Morphological Image Database;
 - Vantagens:
 - Aplicação em materiais de museu;
 - Aplicação em fósseis;
- Dados de sequências
 - DNA, Proteínas;
 - Vantagens:
 - Tamanho do banco de dados (genoma de humano: 3×10^9 bases);
 - Extensão de organismos que podem ser analisados ao mesmo tempo;

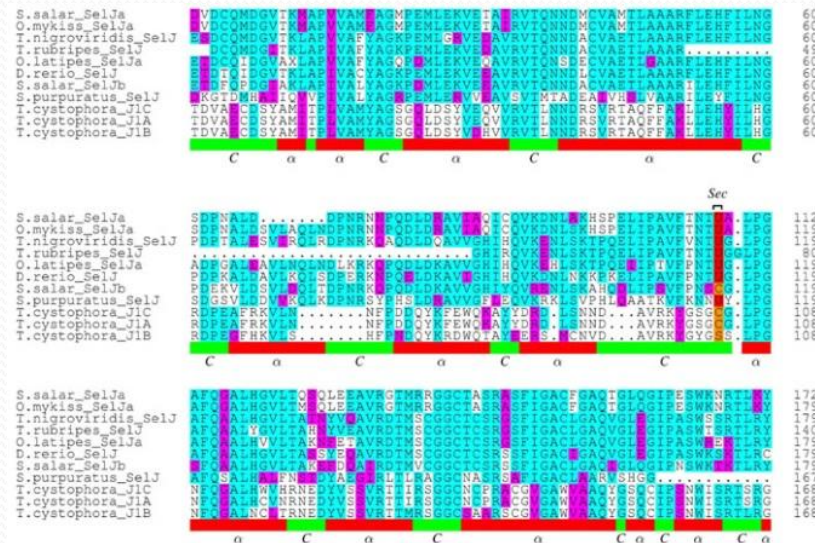
Por que fazer Filogenia?

- Filogenia = estudos de relações entre organismos;
- Nível de espécie:
 - Estudo de sistemática (árvore de espécie);
 - Datação dos eventos (Relógio Molecular).
- Nível de gene:
 - Comportamento do gene ao longo da evolução;
 - Agrupamento (Formação de famílias gênicas);

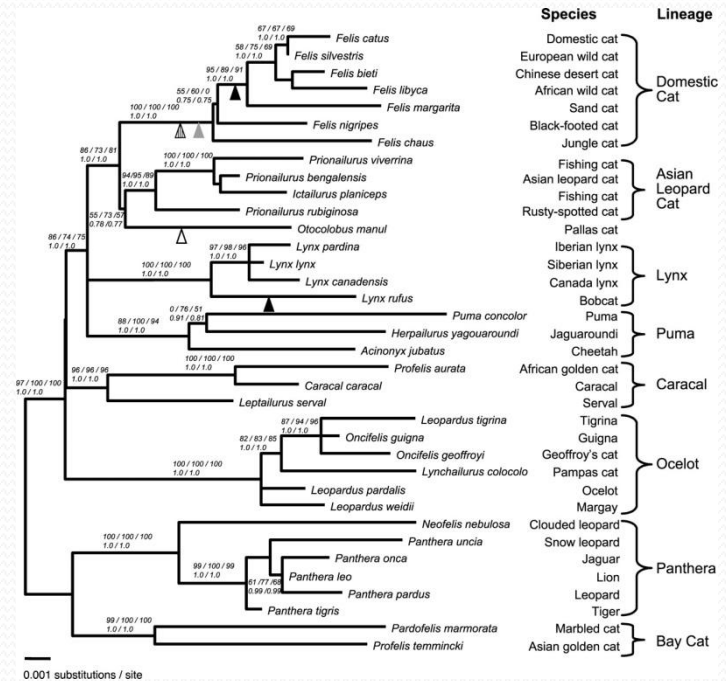
Filogenia - Passo a passo

1) Gene (Estudo de Caso)

2) Alinhamento de sequência



3) Construção da árvore filogenética

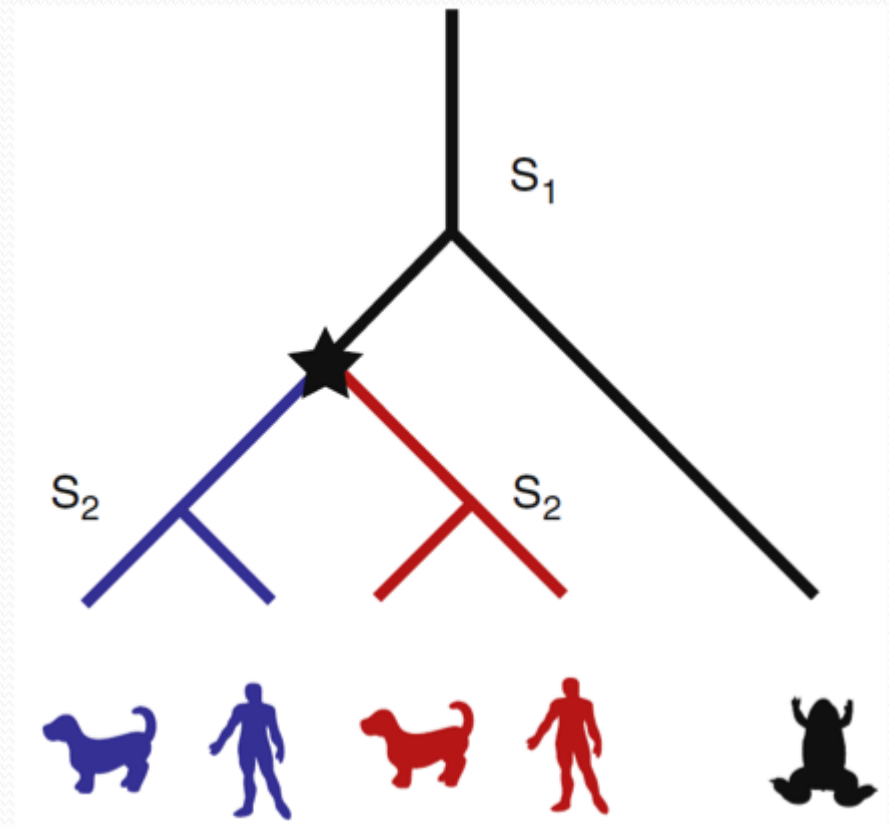


(1) Estudo de caso

- Materiais:
 - Sequências dos organismos vivos;
- Objetivo:
 - Inferir a relação evolutiva entre as sequências utilizadas no estudo.
- Algumas condições:
 - Pelo menos três sequências;
 - Sequências devem ser homólogas;

Homologia, Ortologia, Paralogia...

- **Homologia** – Duas sequências que possuem a mesma origem;
- **Ortologia** – homologia por especiação;
- **Paralogia** – homologia por duplicação.



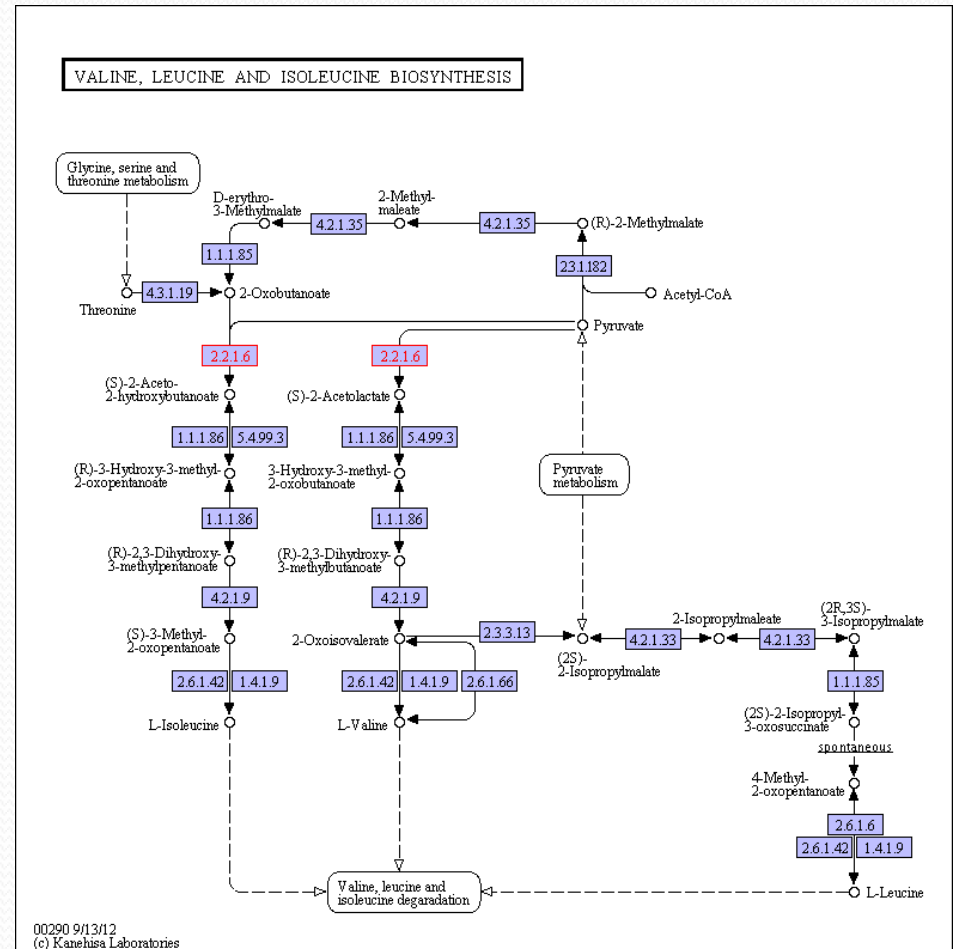
(1) Estudo de caso:

- **Acetolactate synthase**

- Biossíntese de valina, leucina e isoleucina;
- A via não ocorre em metazoários.
- **O que essa proteína está fazendo nesses organismos?**

- **Serine--pyruvate aminotransferase**

- Biossíntese de glicina.
- A via ocorre em todos os eucariotos.



(2) Alinhamento múltiplo de sequências

- Local (BLAST) X Global;
- Par a par X Múltiplo;
- Alinhamento Múltiplo:
 - Alinhamento de 3 ou mais sequências de DNA ou de proteína;
 - Tenta maximizar o número de bases (ou de aminoácidos) que ocorre na mesma posição relativa em todas as sequências;
 - Insere gaps (inserção ou deleção) quando este melhora o alinhamento como um todo.

(3) Alinhamento múltiplo de sequências

- Seq1 ASSGYCAWPACYCYGLPDSASVWDSATNKCGKK
- Seq2 GSSGYCAWPACWCYGLPDNVKIFDSNNNKCGK
- Seq3 GSSGYCAWPACWCYGLPDSVRYNNKCHK
- Seq4 GTDGYCYAWLACYCYNMPNWWKTWD RATNTCGRGK
- Seq5 GSSGYCAWPACYCYGLPDSVKIWTSETNKCGKK



- Seq1 ASSGYC-AWPACYCYGLPDSASVWDSATNKCGKK-
- Seq2 GSSGYC-AWPACWCYGLPDNVKIFDSNNNKCGK--
- Seq3 GSSGYC-AWPACWCYGLPDSVR---YNNKCHK--
- Seq4 GTDGYCYAWLACYCYNMPNWWKTWD RATNTCGRGK
- Seq5 GSSGYC-AWPACYCYGLPDSVKIWTSETNKCGKK-

Programas que realizam o alinhamento múltiplo:

- ClustalW;
- MUSCLE;
- T-Coffee;
- PRANK;
- MAFFT;
- MultAlin;
- Etc.



CADA PROGRAMA POSSUI SUAS PARTICULARIDADES E UMA PODE SER MAIS APROPRIADA QUE O OUTRO EM CADA CASO.

MultAlin

(3) Construção da árvore filogenética

- Existem vários métodos:
 - Distância;
 - UPGMA;
 - **Neighbor-Joining**;
 - Evolução Mínima;
 - Máxima Parcimônia;
 - Procura pela topologia com menor número de substituições;
 - Máxima Verossimilhança;
 - Procura por uma topologia com melhor adequação a um determinado modelo de substituição;
 - Inferência Bayesiana;

Forma mais simples de Filogenia

Seq1 AAATGCGAAAGCAGGT

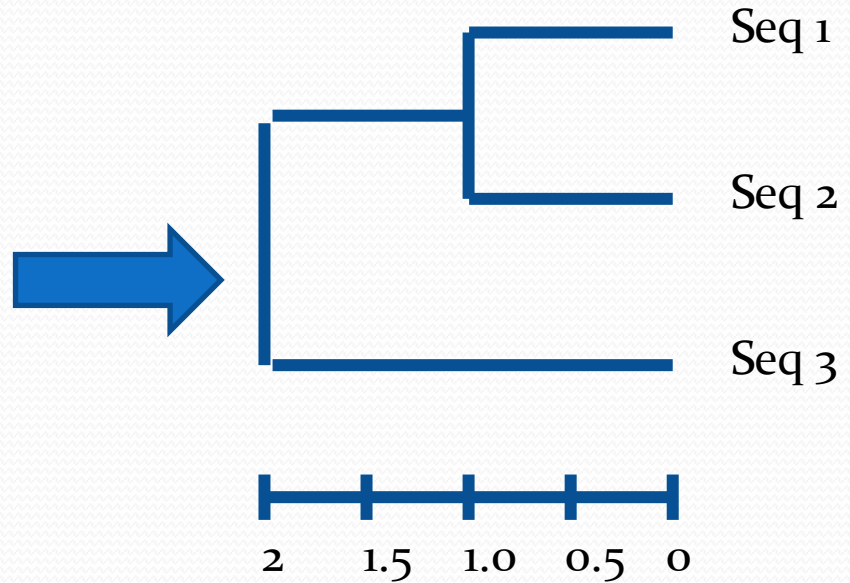
Seq2 AAAGGCGAAAACAGGT

Seq3 AATGGCGAGAGGAGGT

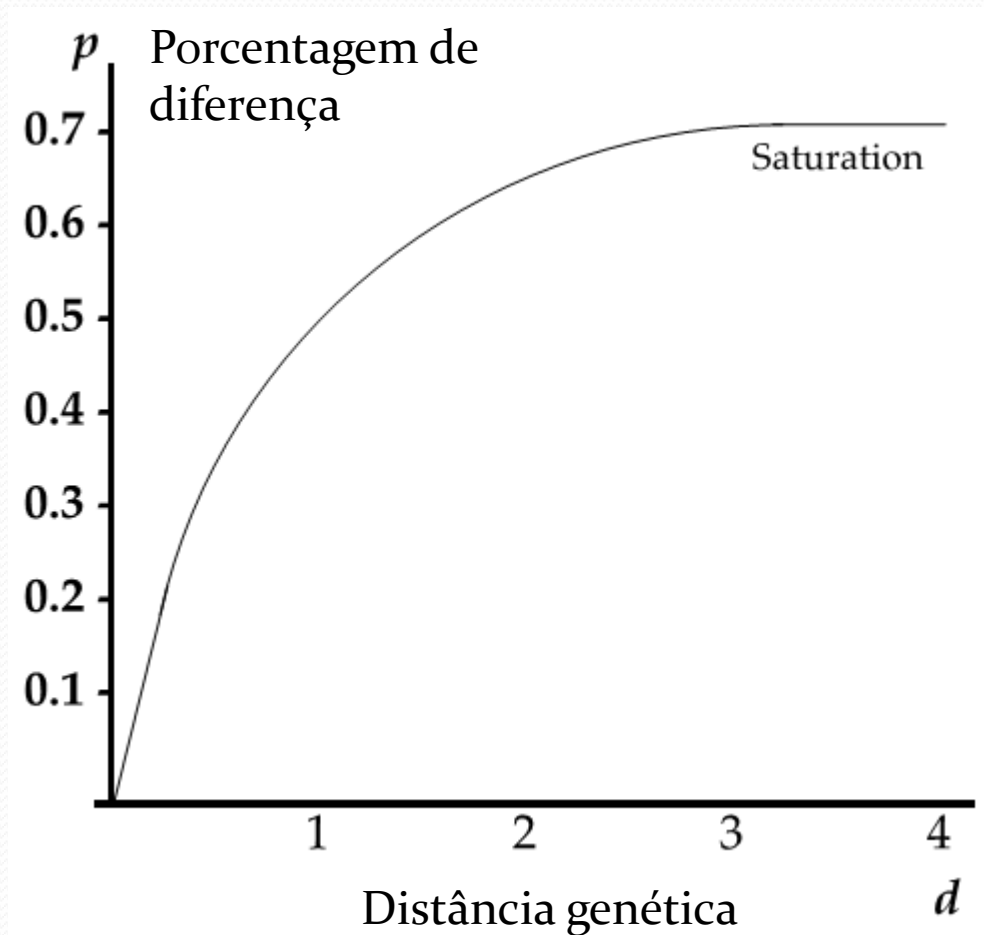
Forma mais simples de Filogenia

Número de Diferenças:

	Seq 1	Seq 2	Seq 3
Seq 1	0		
Seq 2	2	0	
Seq 3	4	4	0

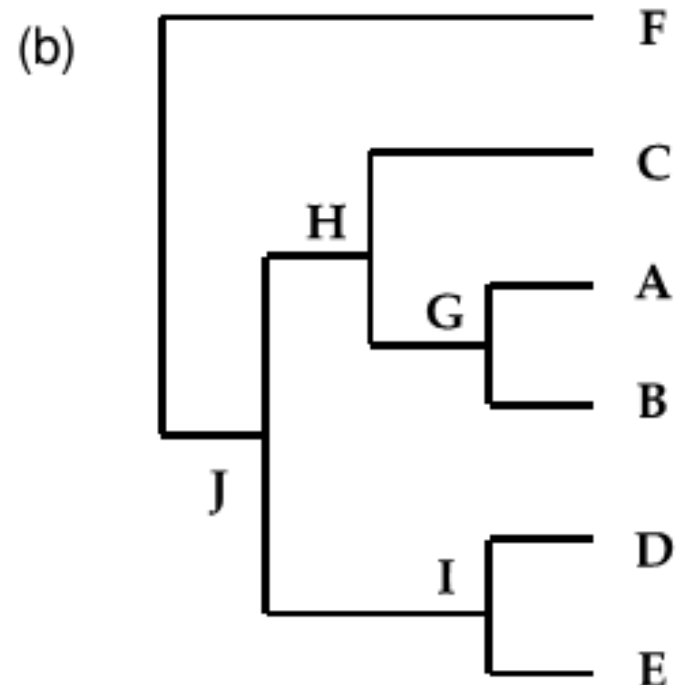
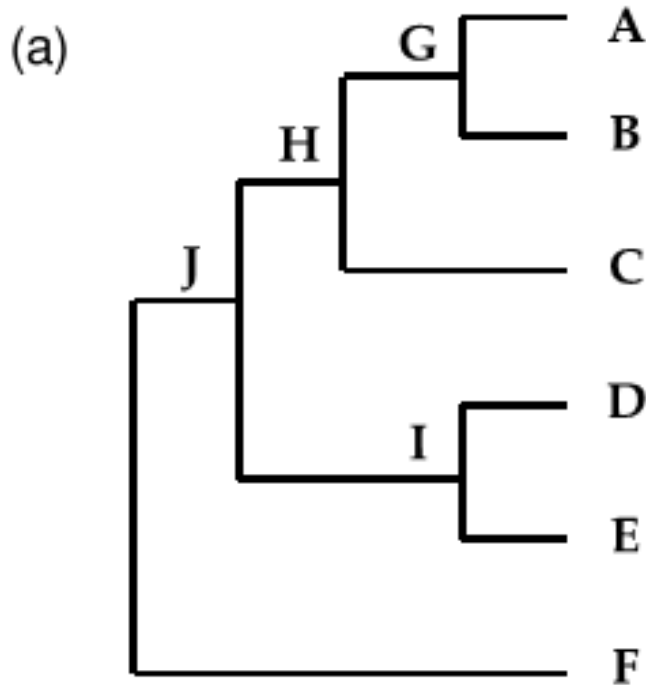


- Múltiplas mutações:
 - $A \rightarrow T \rightarrow G$
- Mutações reversas:
 - $A \rightarrow T \rightarrow A$



Observações:

- Qual a diferença entre estas árvores?



Observações:

- Qual a diferença entre estas árvores?

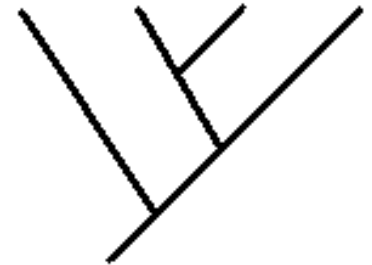
MOSS FERN PINE ROSE



MOSS FERN ROSE PINE

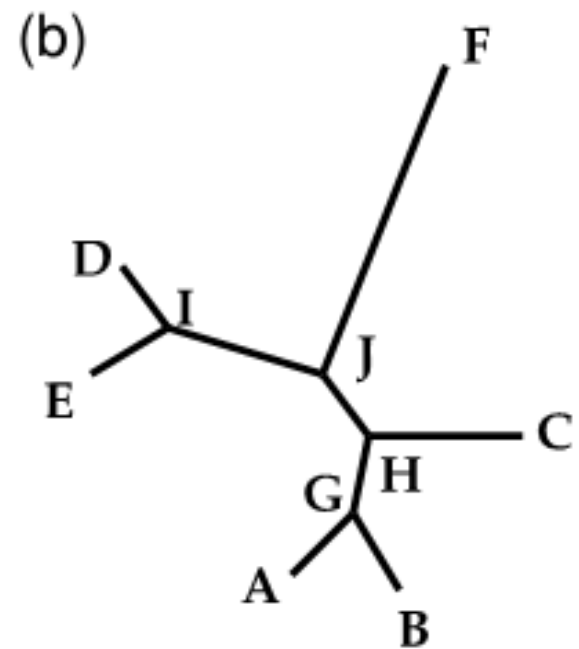
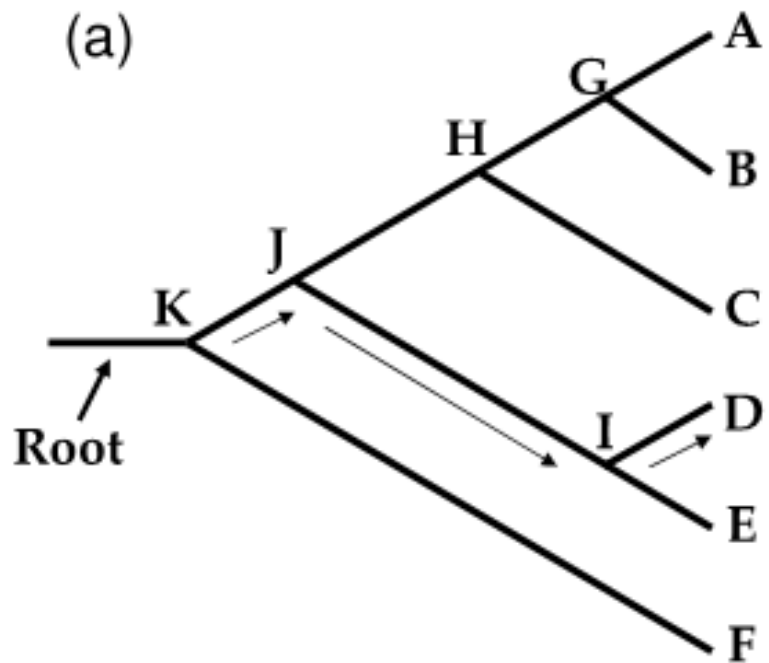


MOSS PINE ROSE FERN



Observações:

- Qual a diferença entre estas duas árvores?





Vamos seguir o roteiro da aula