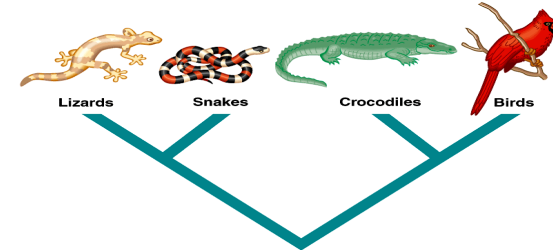


Today

- This morning
 - What is a phylogenetic tree?
 - How to “read” simple phylogenetic trees
 - How can you make a phylogeny?
 - How can you root a phylogeny?
- **Gene Duplications & orthology**

So far: one gene, one individual, one species

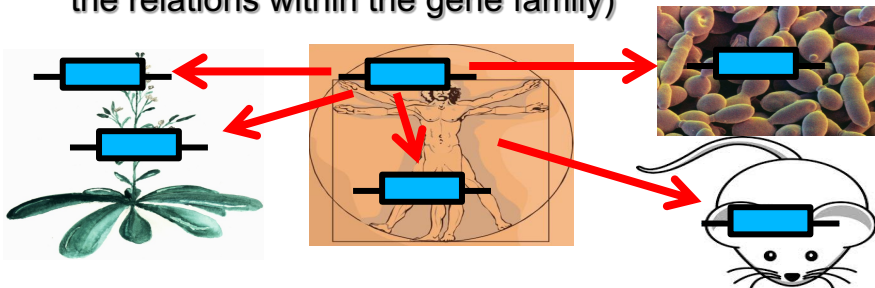


Copyright © Pearson Education, Inc., publishing as Benjamin Cummings.

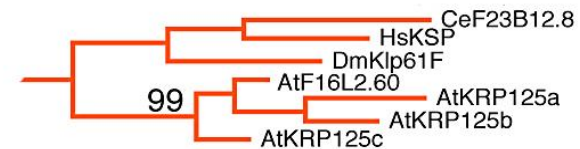
- Biology is more complex
- Not all internal nodes are divergences between species

General idea of today

- Having all the genomes ...
- What is the **relation** of my gene to *homologous* genes in the same and other organisms (what are the relations within the gene family)



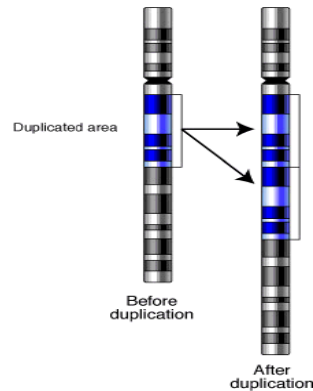
Not just divergence between species; what is up here?



Kinase-5: essential for centrosome separation in mitosis

?b

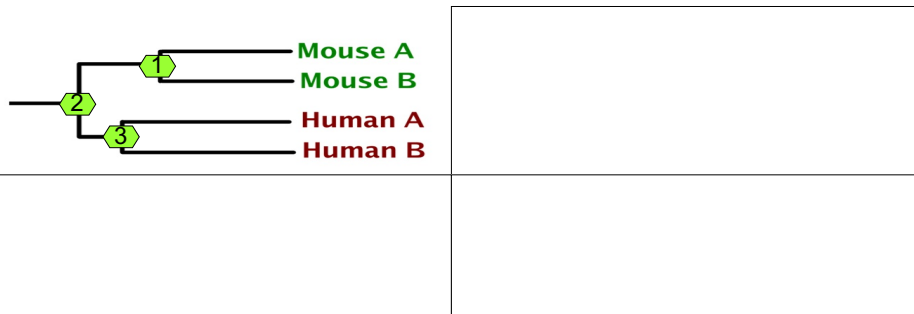
Gene duplication is the duplication of a region of DNA that contains a gene (important force in evolution of gene families / genome evolution)



Two genes per species: how to differentiate between one ancient or two recent duplications?

- Two genes in Human chromosomes (Human A & Human B) & two genes in mouse chromosomes (Mouse A & Mouse B)

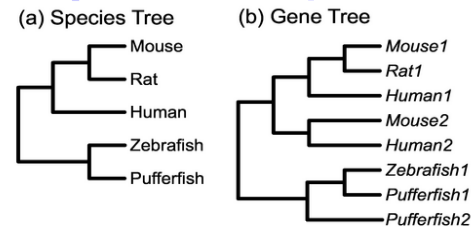
Duplications, Speciations



● Speciation ■ Gene Duplication

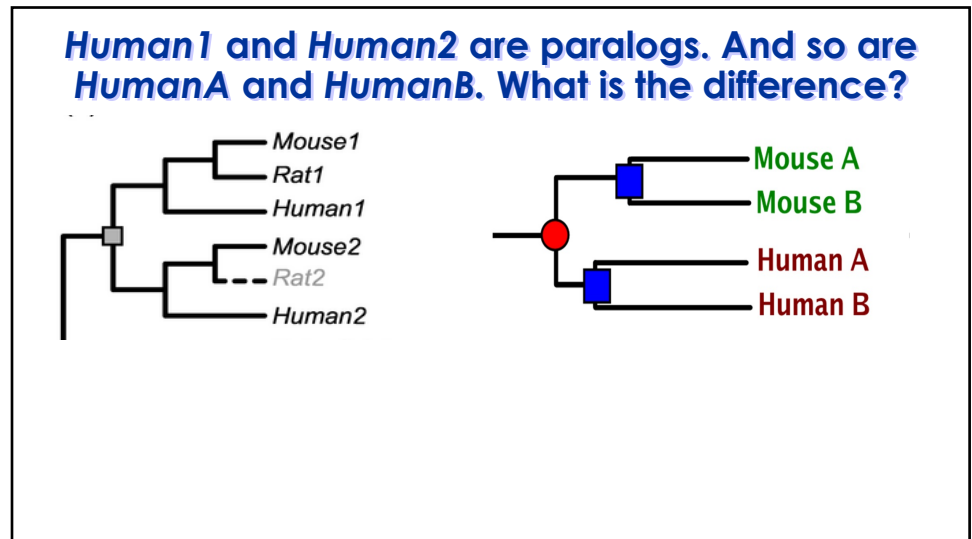
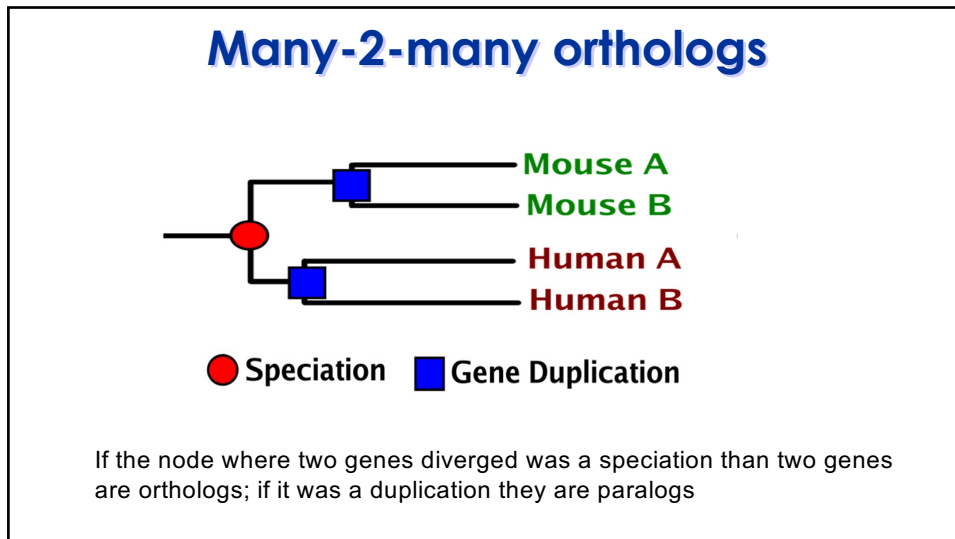
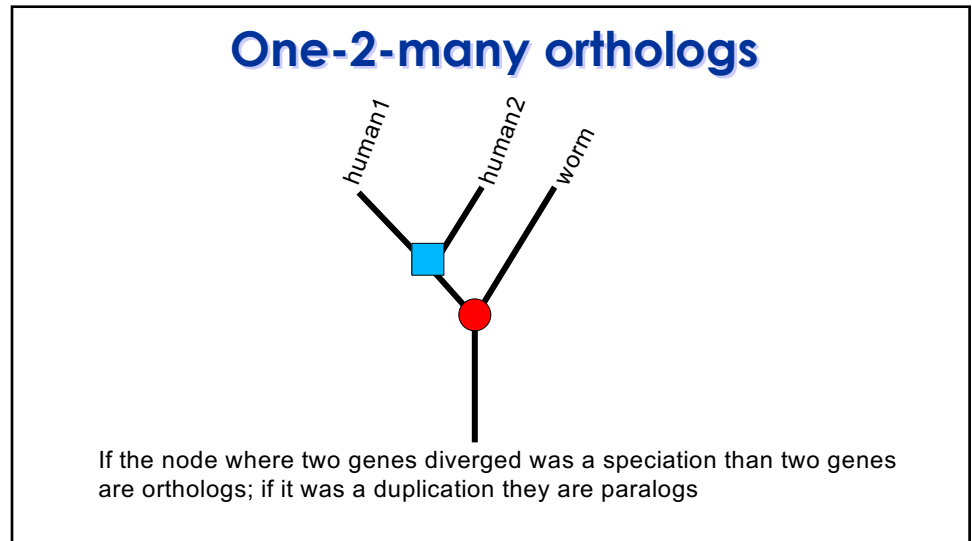
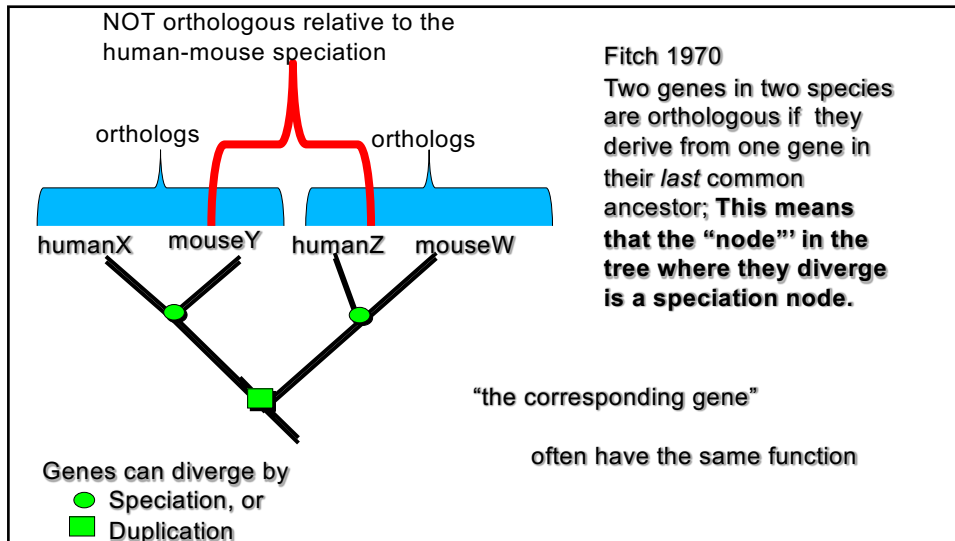
?

One more genome evolution process / gene family evolution process: what is up here? (besides duplication & speciation)

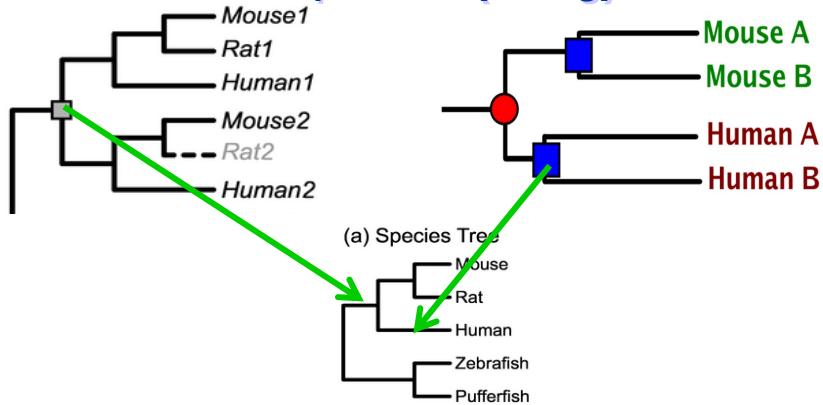


Gene loss

b



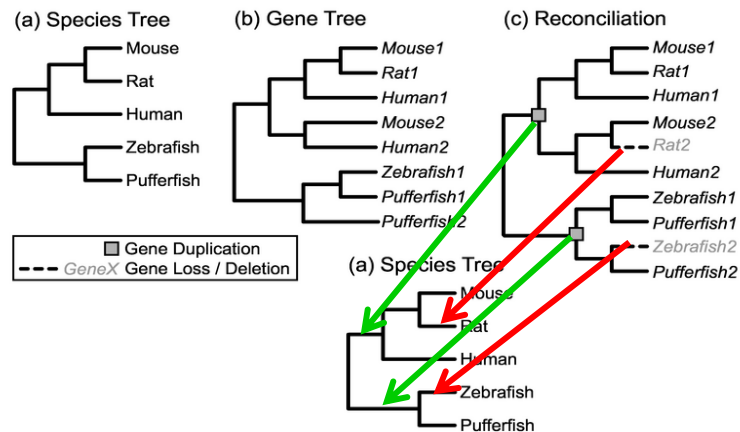
Human1 and Human2 are paralogs. And so are HumanA and HumanB. They differ in when they duplicated (timing)



Paralogs

- can have different timing: they can be ancient ($>3.8 \times 10^9$) or very recent (10^6)
- Tree reconciliation (*interpreting / annotating a gene tree in terms of speciation and duplication*) allows to time this by placing the events on a *timed* species tree.

Projection on the species tree also allows timing of gene loss

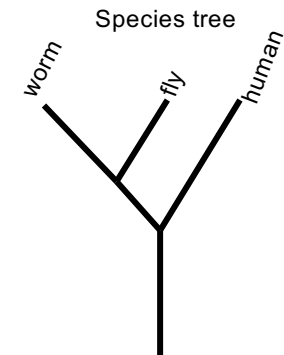
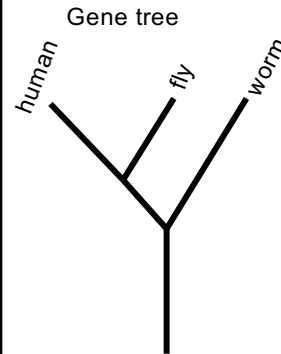


Exercise 4

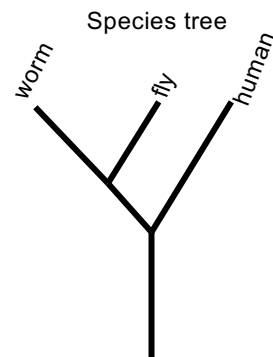
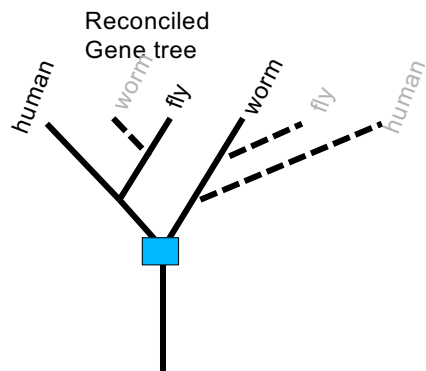
The problem of automating gene tree-species tree reconciliation

- Do you have to do the tree reconciliation yourself?
 - (today yes)
 - Quite often yes, although many algorithms exist
- Algorithms have trouble with tree reconciliation because gene trees are often *wrong* (because they are based on very little data inferring events that happened a very long time ago and species/genes have very different rates of sequences evolution)

A very common case: do you see the problem if you start doing the tree reconciliation?

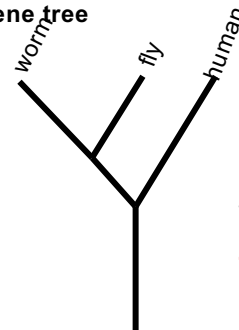


Strict tree reconciliation gives spurious duplications and losses

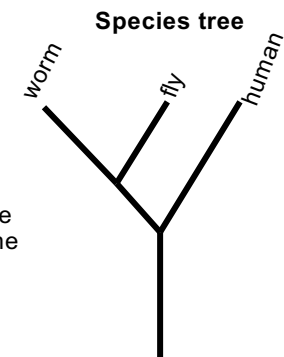


How to perform automatic tree reconciliation given wrong gene trees? ... Species tree guided tree building

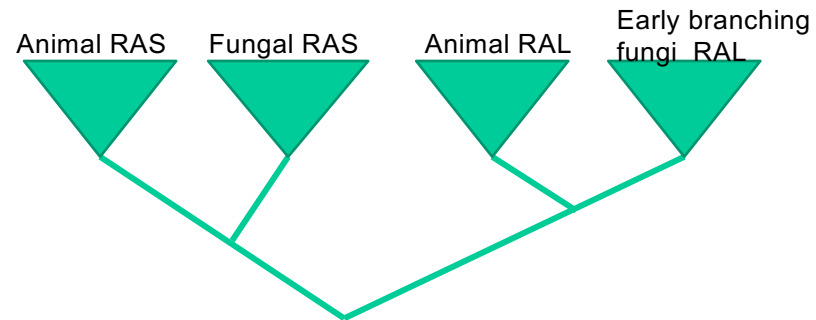
Species tree guided gene tree



Is there significant signal in aspects of the data (the alignment) to support a gene tree that implies less genome evolution events? e.g. **TreeBEST** as used in **ENSEMBL COMPARA**
Prevents spurious duplications and losses



However more data (better taxon sampling) >>
tree reconciliation: the case of RAL evolution?



Old duplication **and** loss. The hypothesis of animal specific duplication and accelerated evolution & wrong gene tree can be rejected

Exercise 5