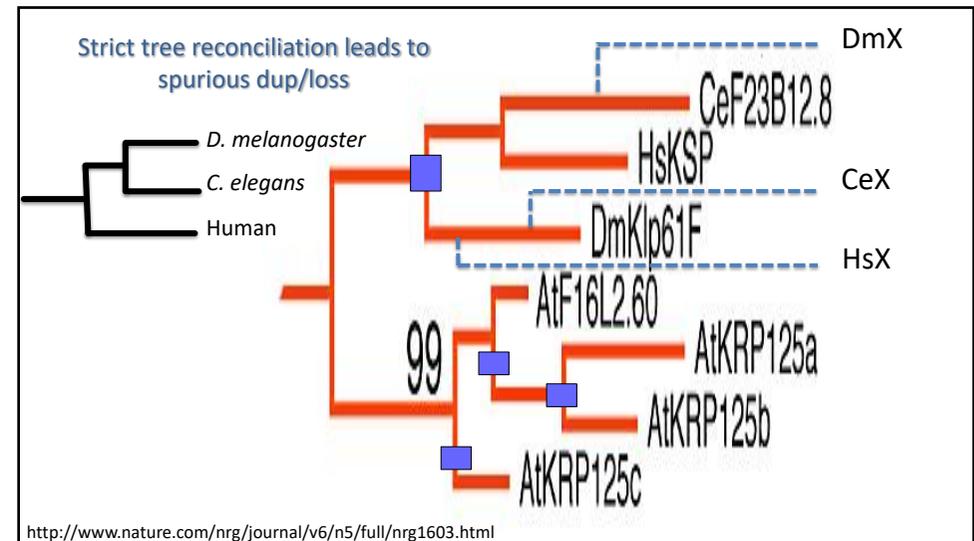
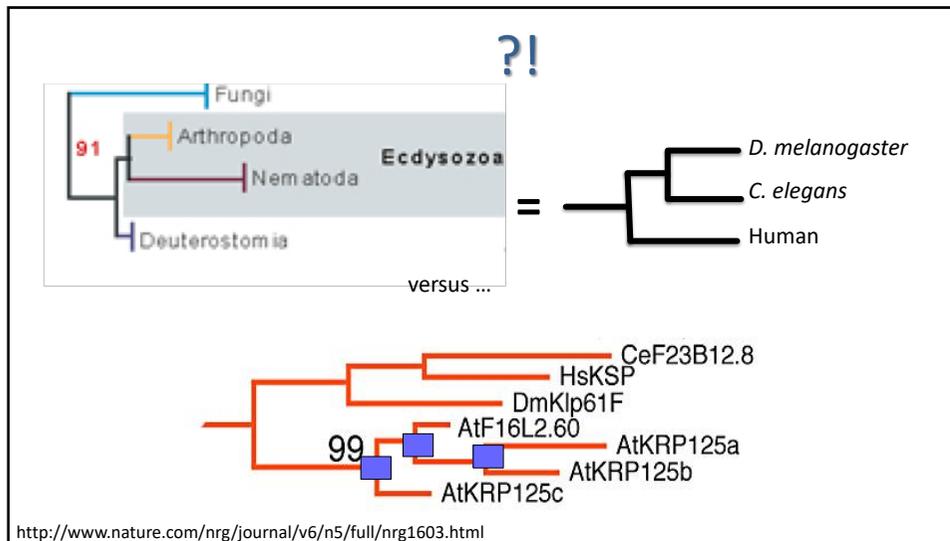


Interpreting/annotating gene phylogenies. Beyond intro bioinformatics.

- [Species phylogeny](#)
- [Timing duplications, outparalogs/inparalogs](#)
- [Whole genome duplications](#)
- [Duplications in FECA-2-LECA, endosymbiosis & HGT](#)

Given problems with single gene trees where does the species tree comes from?

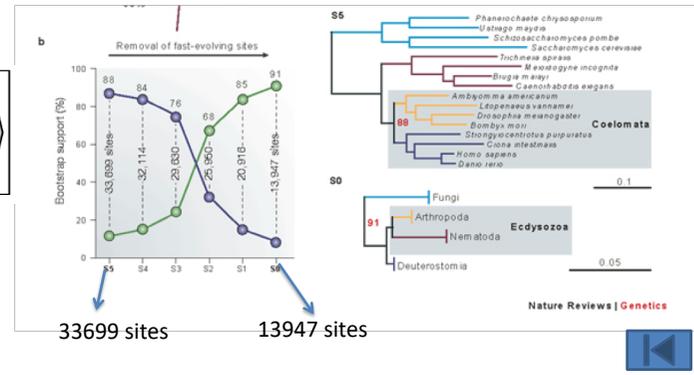
- Noisiness of gene trees necessitates use of species tree guided gene tree construction as e.g. TREEBEST as used by ENSEMBL
- Wait a second? If gene trees are often wrong, how do we know the species tree?
- Trees based on massive amounts of data: concatenated alignments, multi locus trees



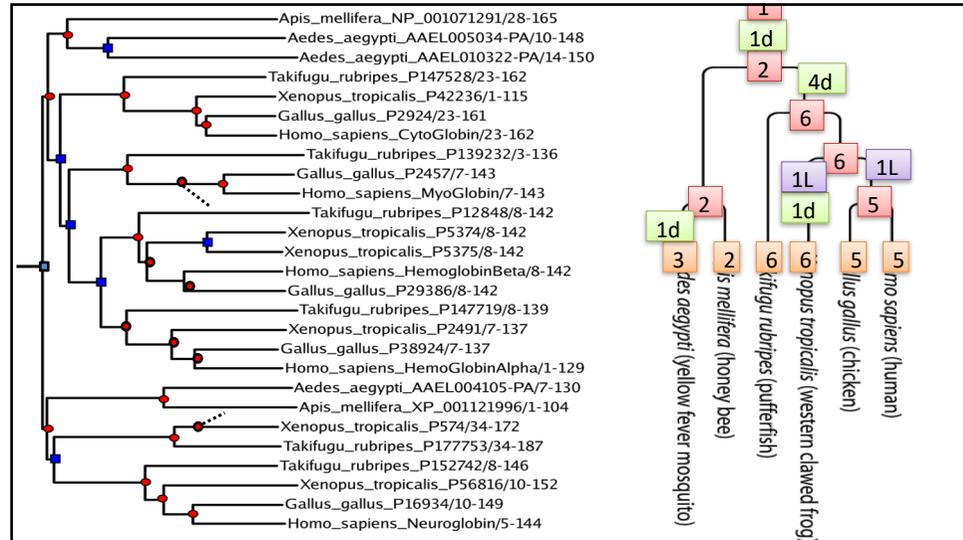
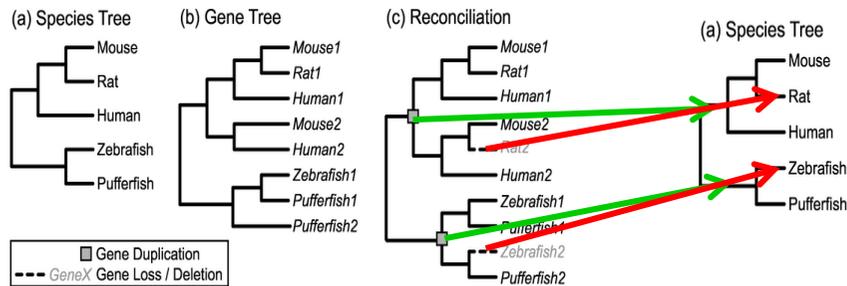
So why are gene trees often wrong?

- Too little data (too short alignments) for events that happened a long time ago, i.e. lack of resolution
- So how do we normally solve that for the *tree of life*, e.g. the *C. elegans* / *D. melanogaster* / human relations?

Concatenating genes / multimarker and very strict data selection



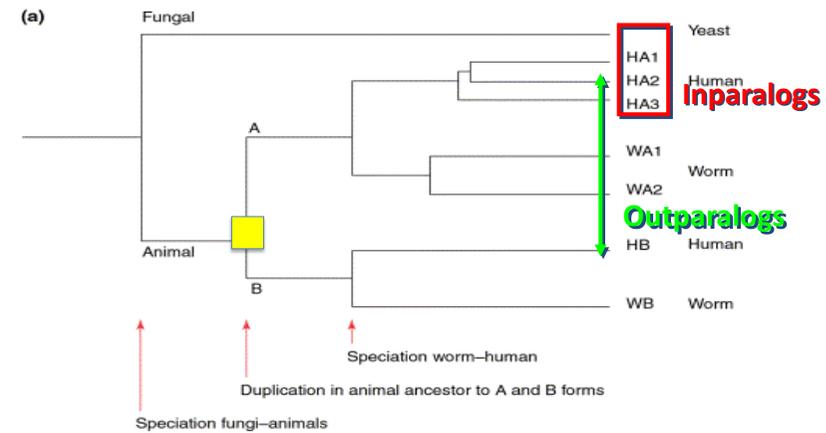
Tree reconciliation gives gene family dynamics on a species tree



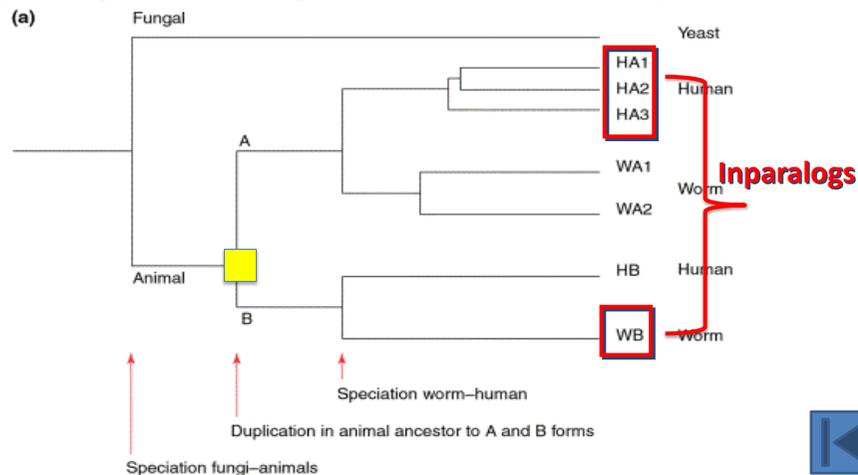
You can not always point to a tree while writing or talking so terminology: inparalogs & outparalogs

- *Outparalogs* duplicated **before** the speciation of the two species between which you define orthologs
- *Inparalogs* duplicated **after** the speciation of the two species between which you define orthologs
- orthologs, inparalogs and outparalogs are \pm shorthand to discuss for a set of proteins how they are related / when they duplicated

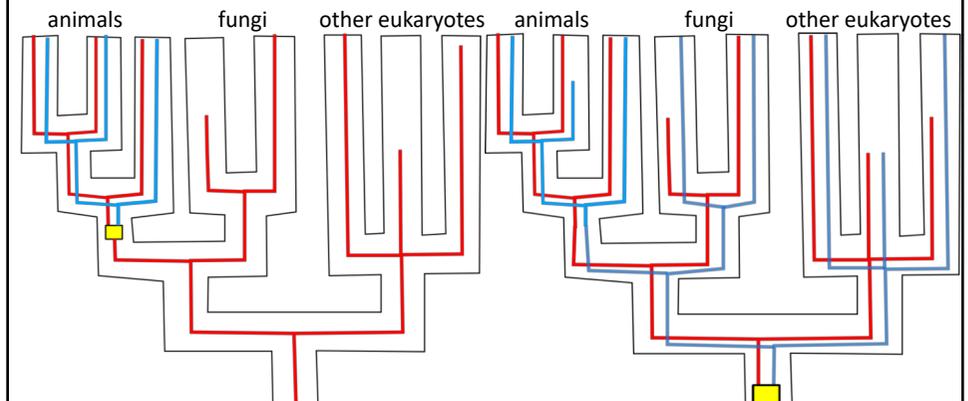
inparalogs, outparalogs relative to speciation worm-human



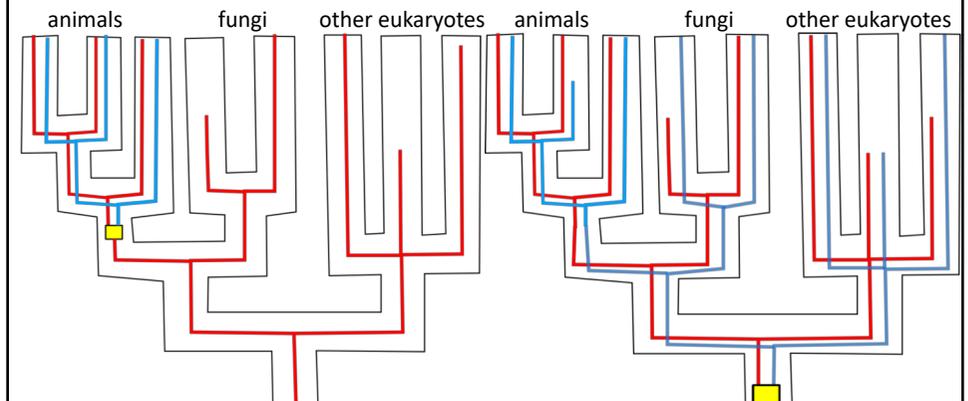
Inparalogs, outparalogs relative to speciation yeast-human



Inparalogs relative to LECA & a single "orthologous group"



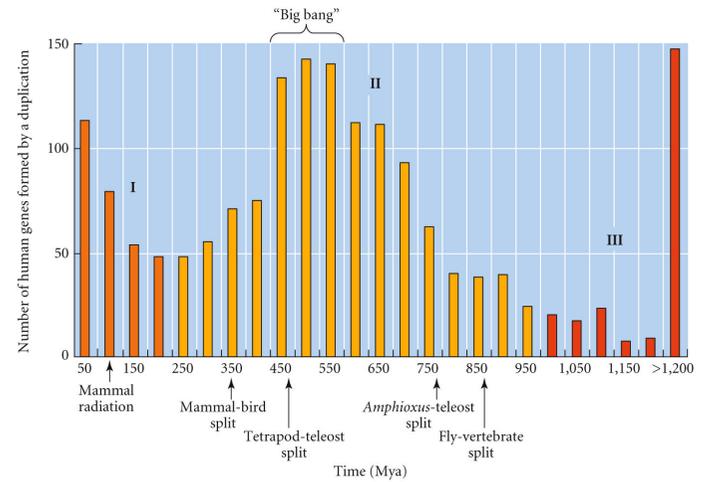
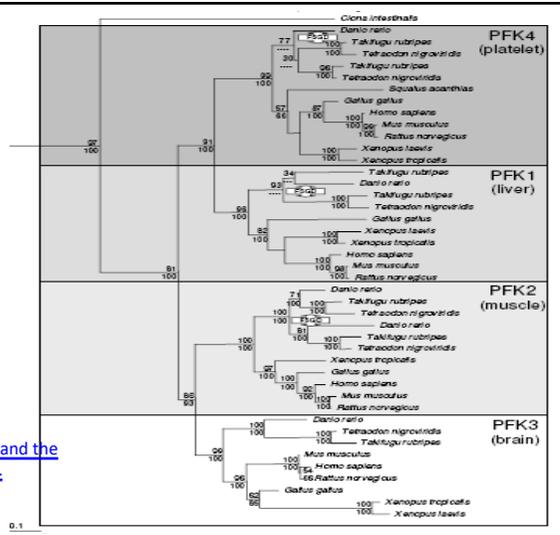
Outparalogs relative to LECA & two "orthologous groups"



Interpreting the tree Example: vertebrate duplications

- Tetraploidy?

Three rounds (1R/2R/3R) of genome duplications and the evolution of the glycolytic pathway in vertebrates.
Steinke D, Hoegg S, Brinkmann H, Meyer A.
BMC Biol. 2006 Jun 6;4:16.

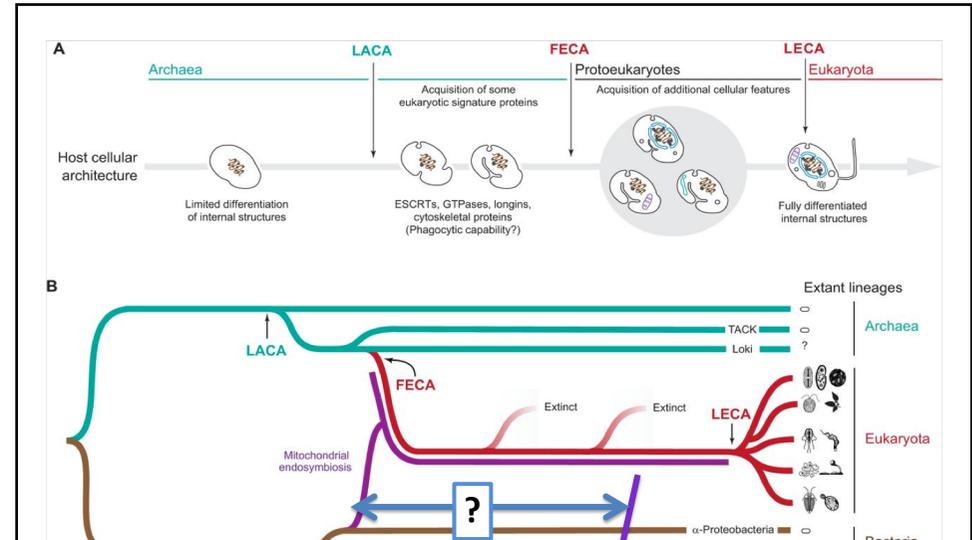
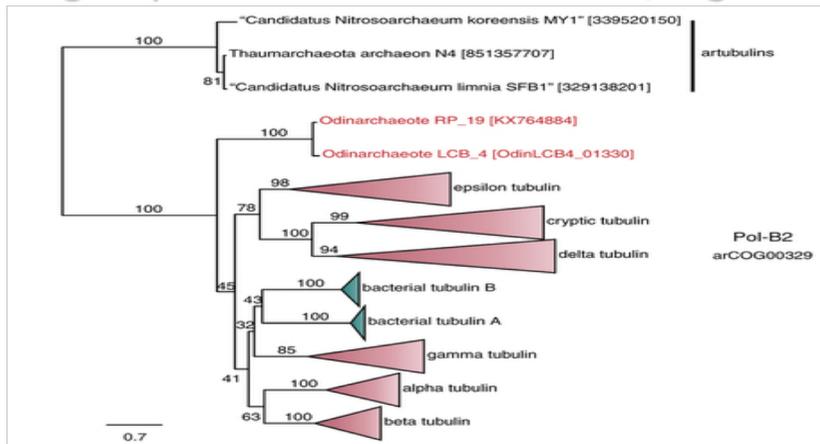


22

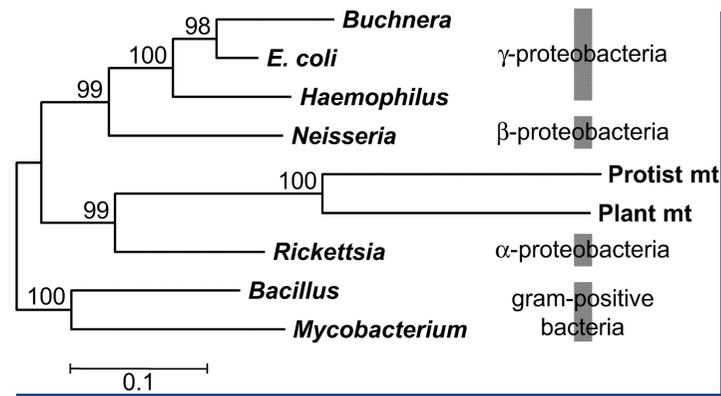
EVOLUTION 2e, Figure 20.19

© 2009 Sinauer Associates, Inc.

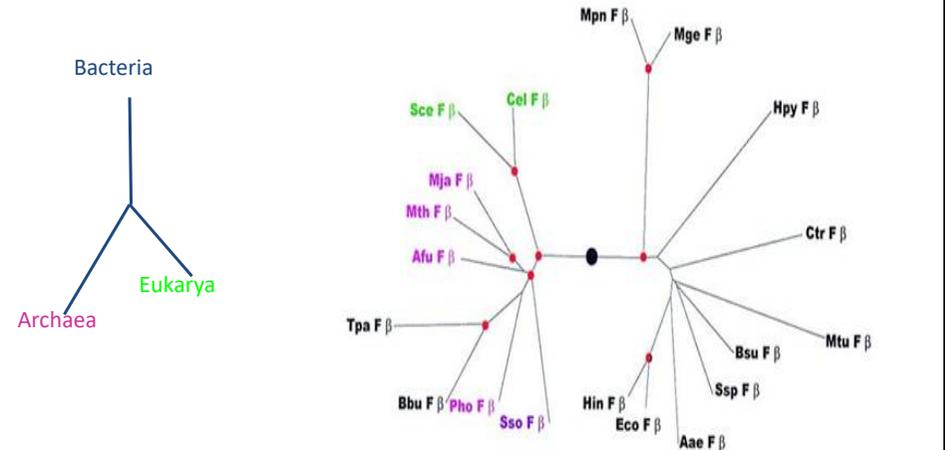
Timing Duplications to FECA-2-LECA, e.g. tubulin



Interpreting the tree: endosymbiosis



Interpreting the tree: Horizontal Gene Transfer (HGT)



Interpreting the tree: HGT between eukaryotes

