

Tree-Thinking Skills

KIZ, Fall 2024

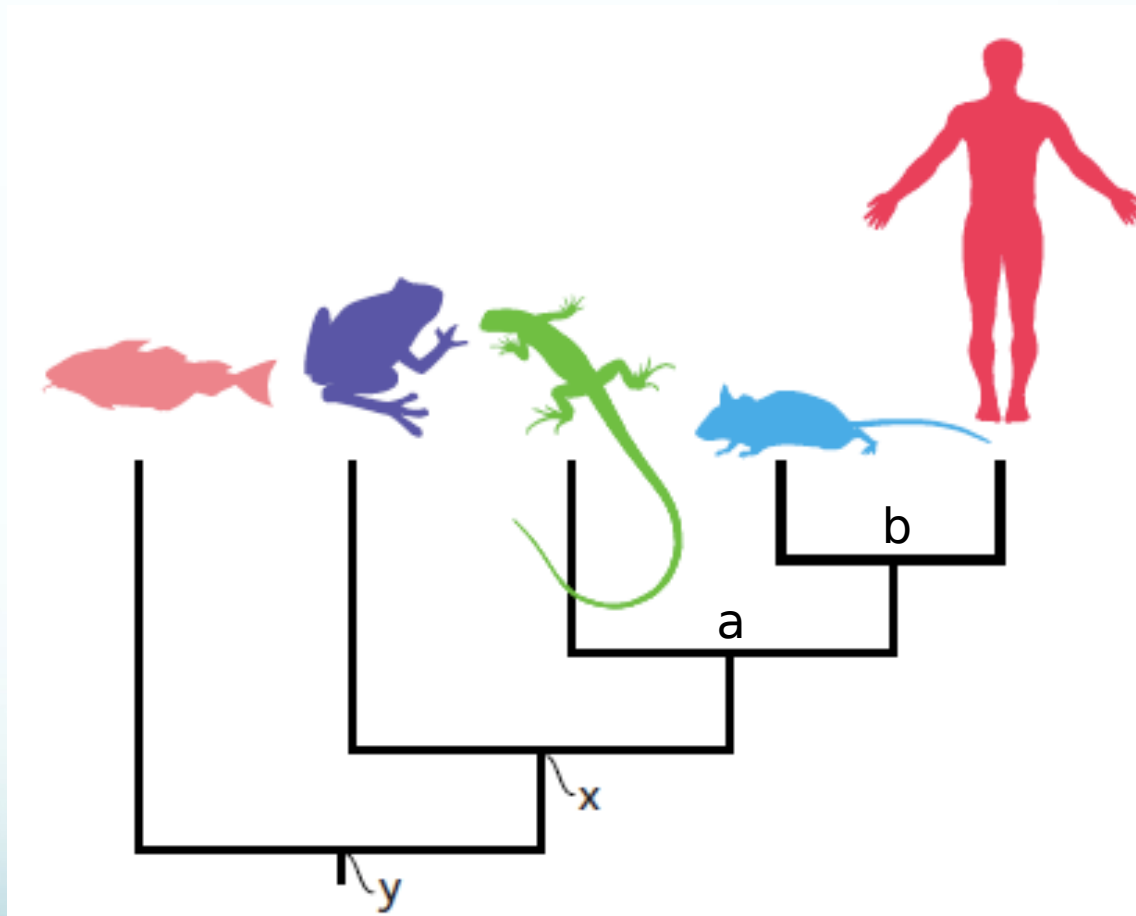
Prof Weigang Qiu

Department of Biological Sciences

Hunter College of City University of New York

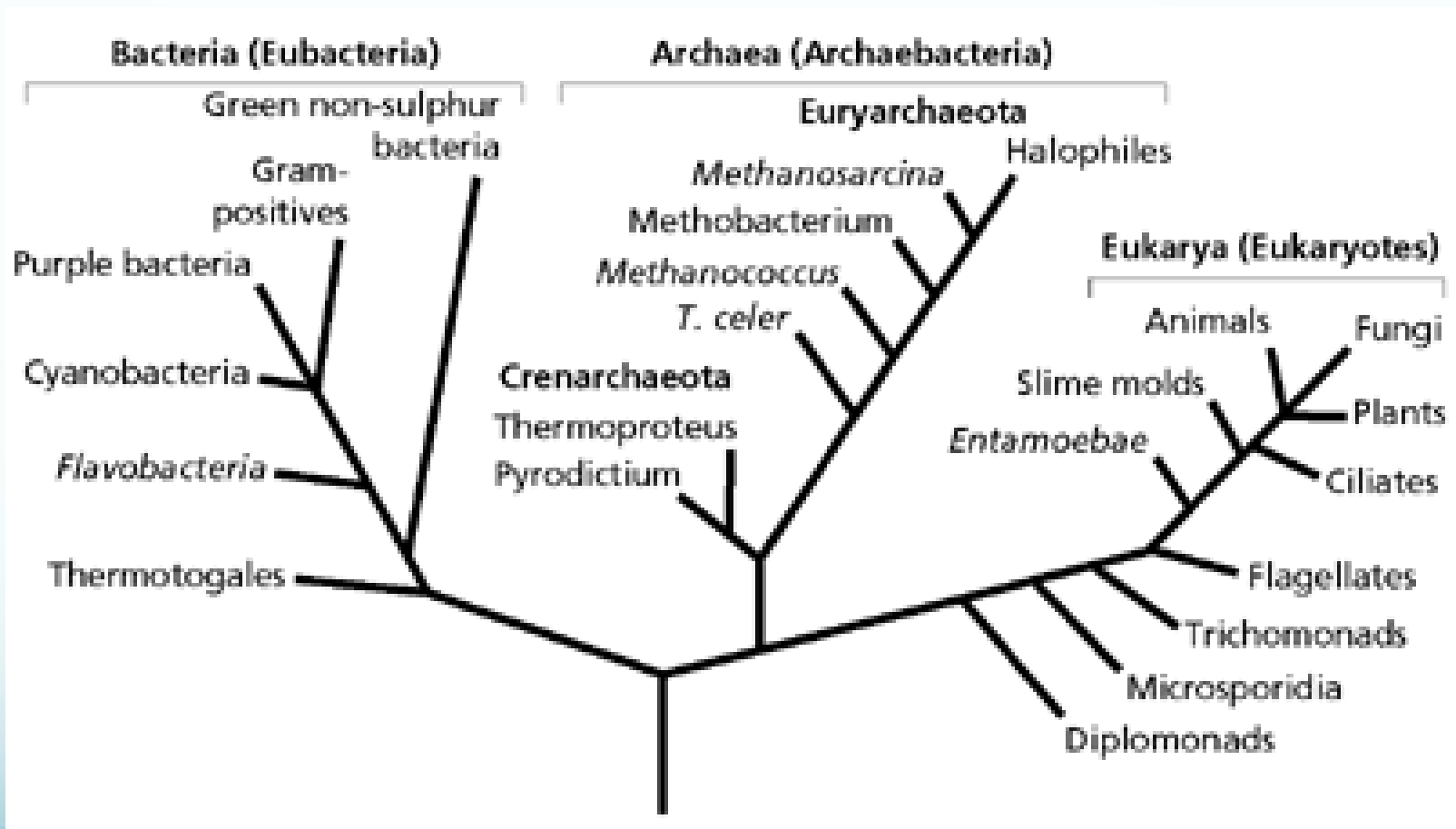
Session I. How to Read a Tree?

Determine Relatedness (**solely**) by Level of Last Common Ancestors



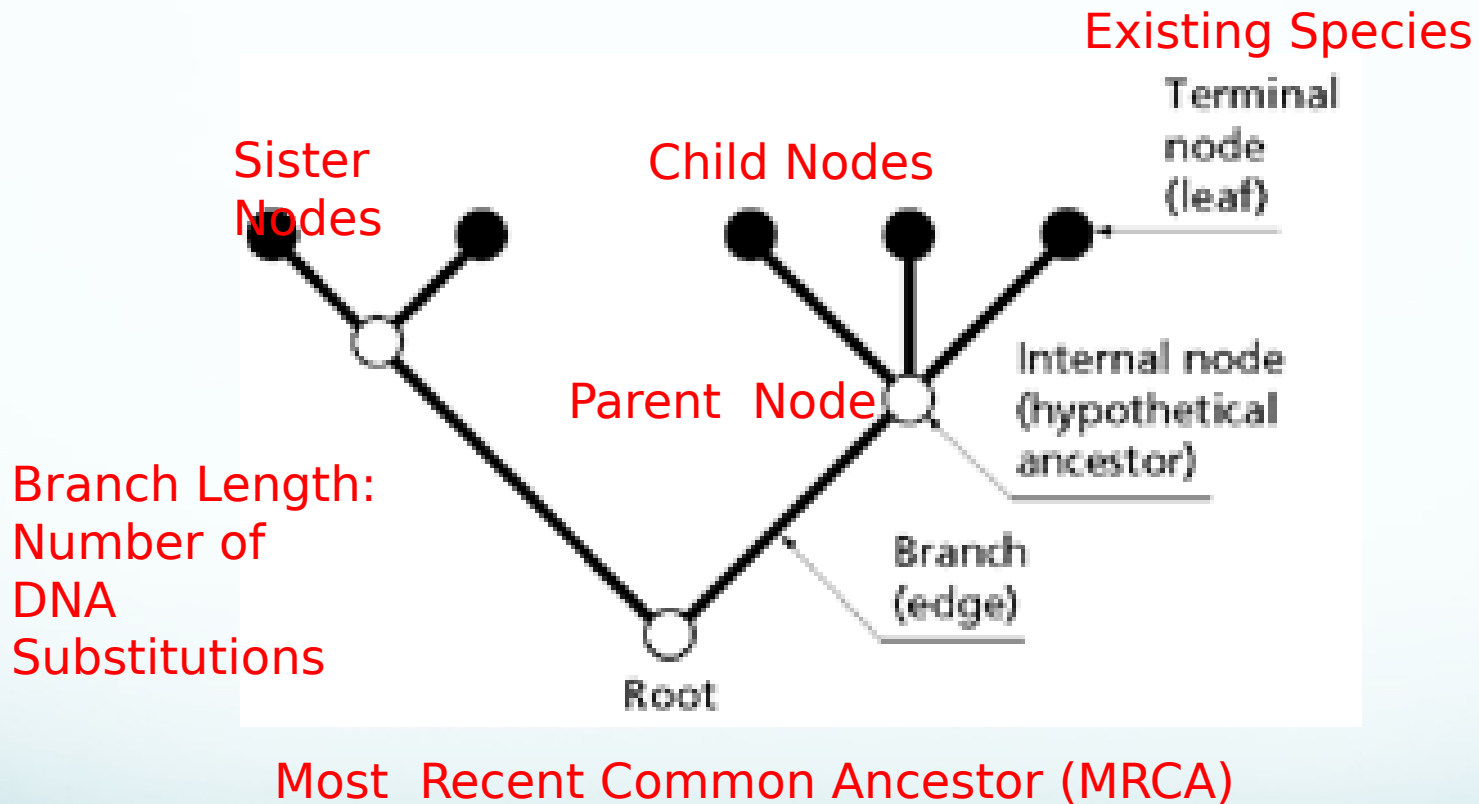
and **label** LCA (last common ancestor) or MRCA (most recent common ancestor)

Tree of Life



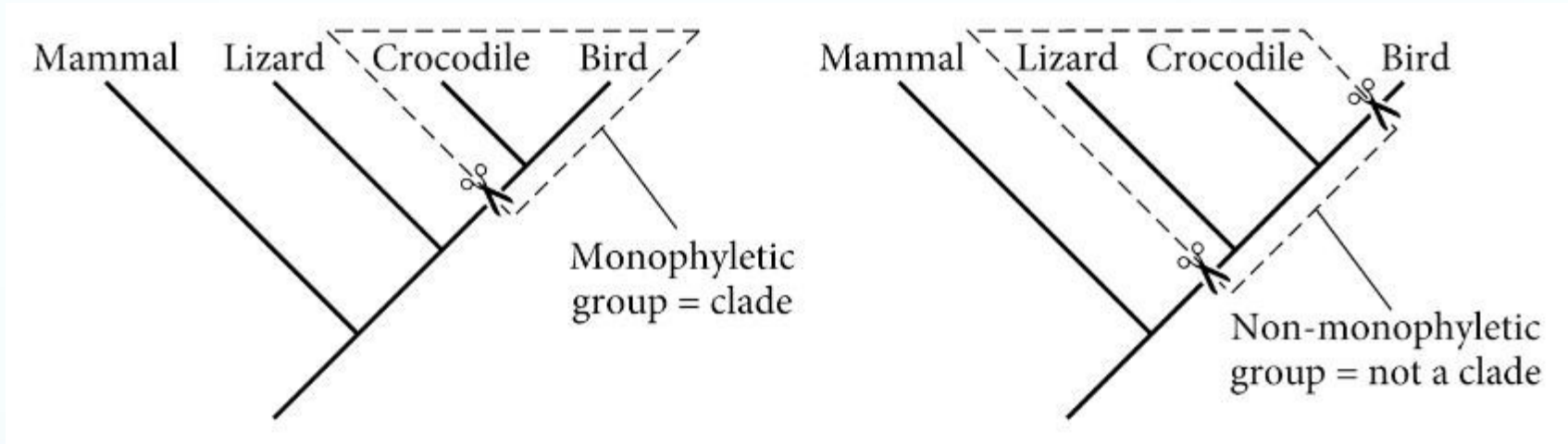
Homology (common ancestry) \neq Similarity (e.g., fish & whale)

Tree Terms



Clades

ALL descendants of a node

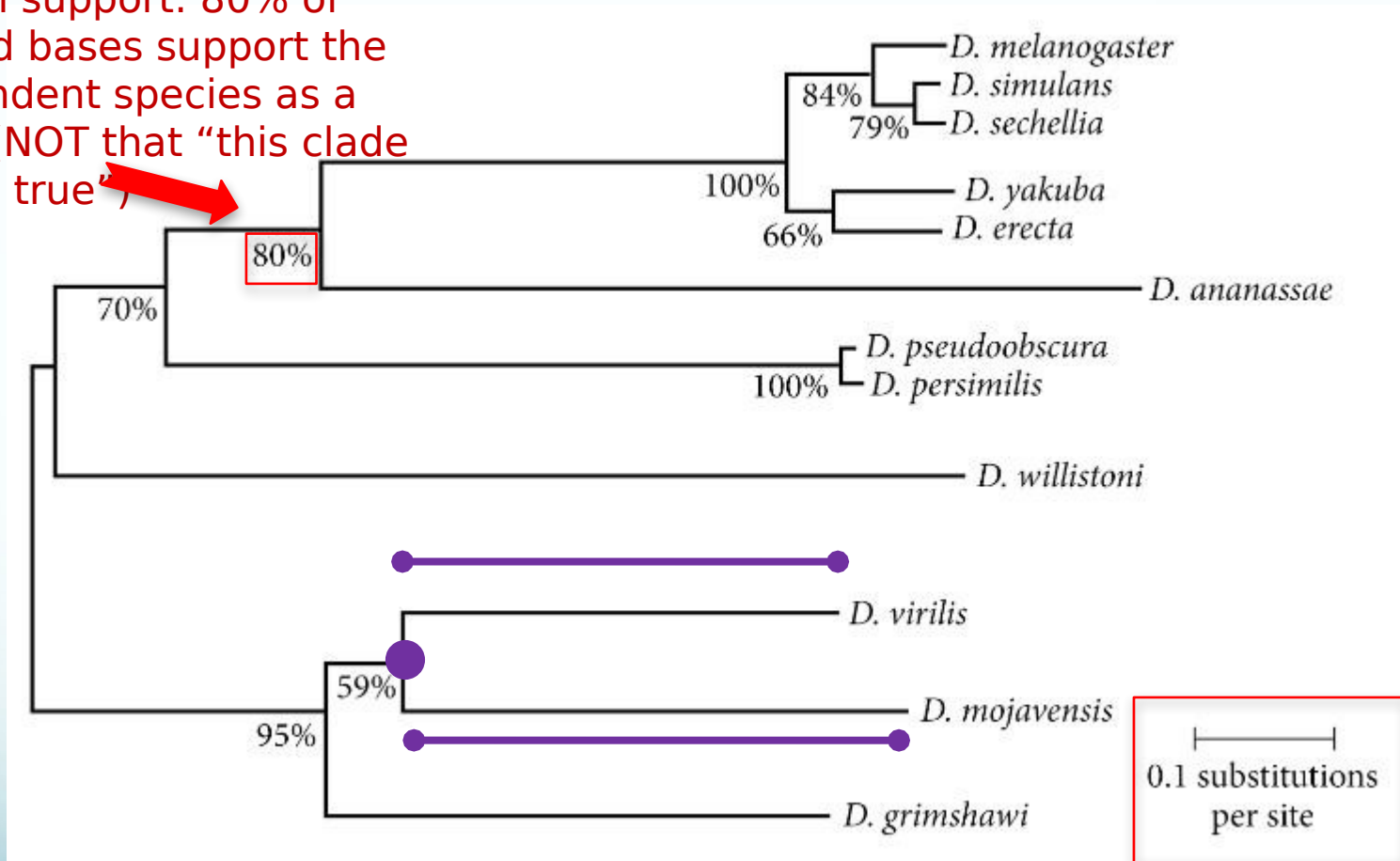


Groups that are not a clade:

- Apes excluding humans
- Prokaryotes
- Dinosaur
- Reptile
- Gram-Negatives
- Animals excluding humans

Tree Distances & Branch Support

Branch support: 80% of aligned bases support the descendent species as a clade (NOT that "this clade is 80% true")

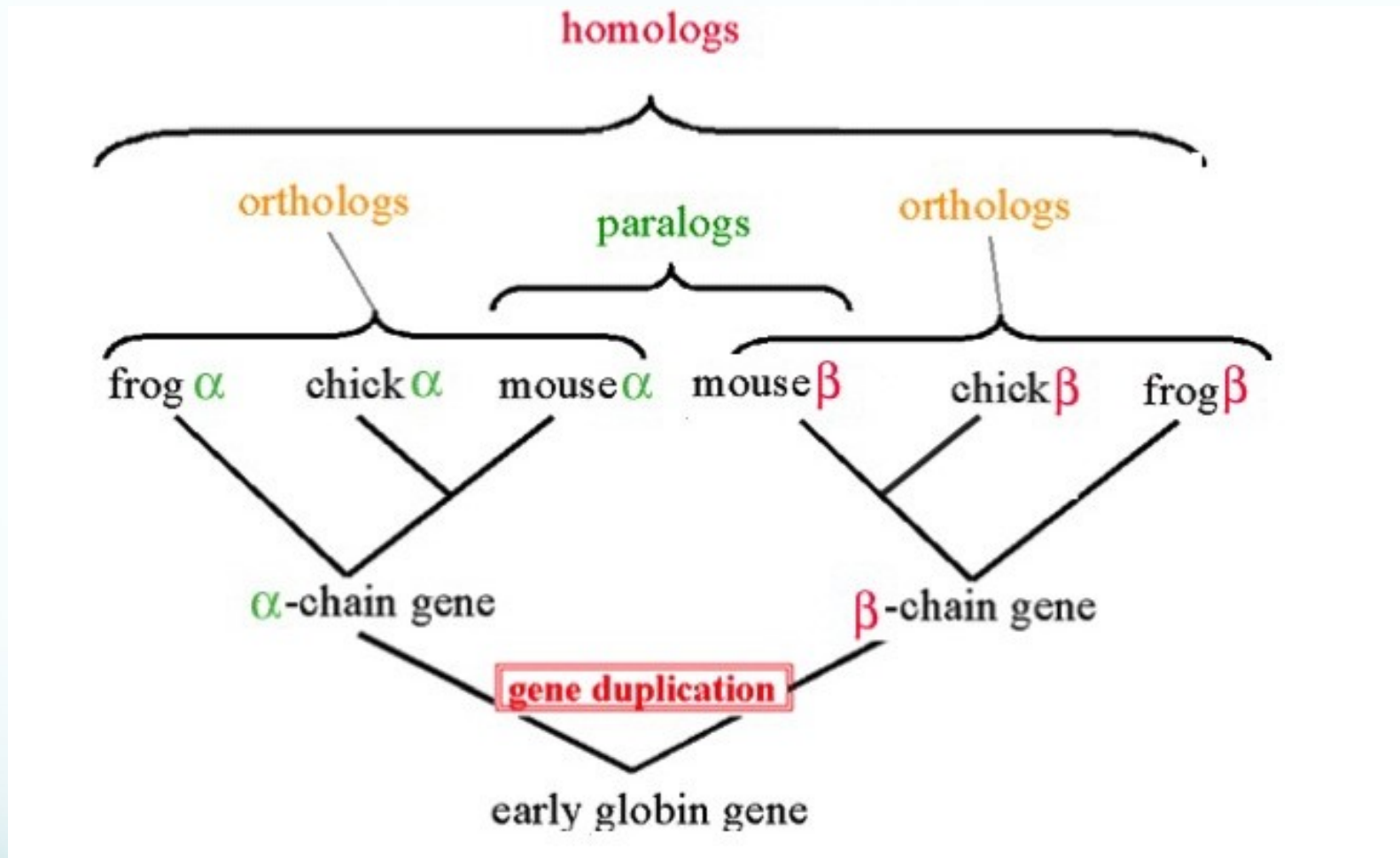


Evolutionary distance between *D. virilis* & *D. mojavensis* is ~ 0.7 sub/site, estimated by adding two distances since their last common ancestor (LCA, dot) & then multiplying it by the scale bar (0.1 per site = 10% sites)

Session II. How to Build a Tree?

- Choice of molecular markers:
 - **Slow-evolving** sequences for **distantly related** species/individuals (e.g., protein sequence for vertebrates, 16S rRNA for bacteria)
 - **Fast-evolving** sequences for **closely related** species/individuals (e.g., mtDNA or non-coding sequences for humans, DNA sequences for primates)
- Identify homologs (BLAST, HMMER)
- Obtain multiple alignment (“muscle”, “clustal”)
- Infer tree (Neighbor-Joining; Parsimony; Likelihood; Bayesian)
- Obtain branch support (Bootstrap; Posterior support)

Gene Tree vs Species Tree



- Orthologs: Descendants of a speciation event; similar mol. function
- Paralogs: Descendants of a gene duplication event; often differ in function
- Paralogs from the same species (genome) are often more divergent than orthologs from different species (genomes) because duplication

Conclusions

- Bio-diversity is **tree-like**, not ladder-like
 - All species share common ancestors (**Homology**)
 - Existing species do not evolve from each other
 - Species diversify over time, not climb a ladder
- How to **read** a tree? Biological interpretations of roots, internal & external nodes, and branch lengths. Identify LCA. Compare relatedness of species based on positions of LCAs
- How to **create** a tree? Four major steps: Identify homologs, align sequences, build tree, obtain branch support