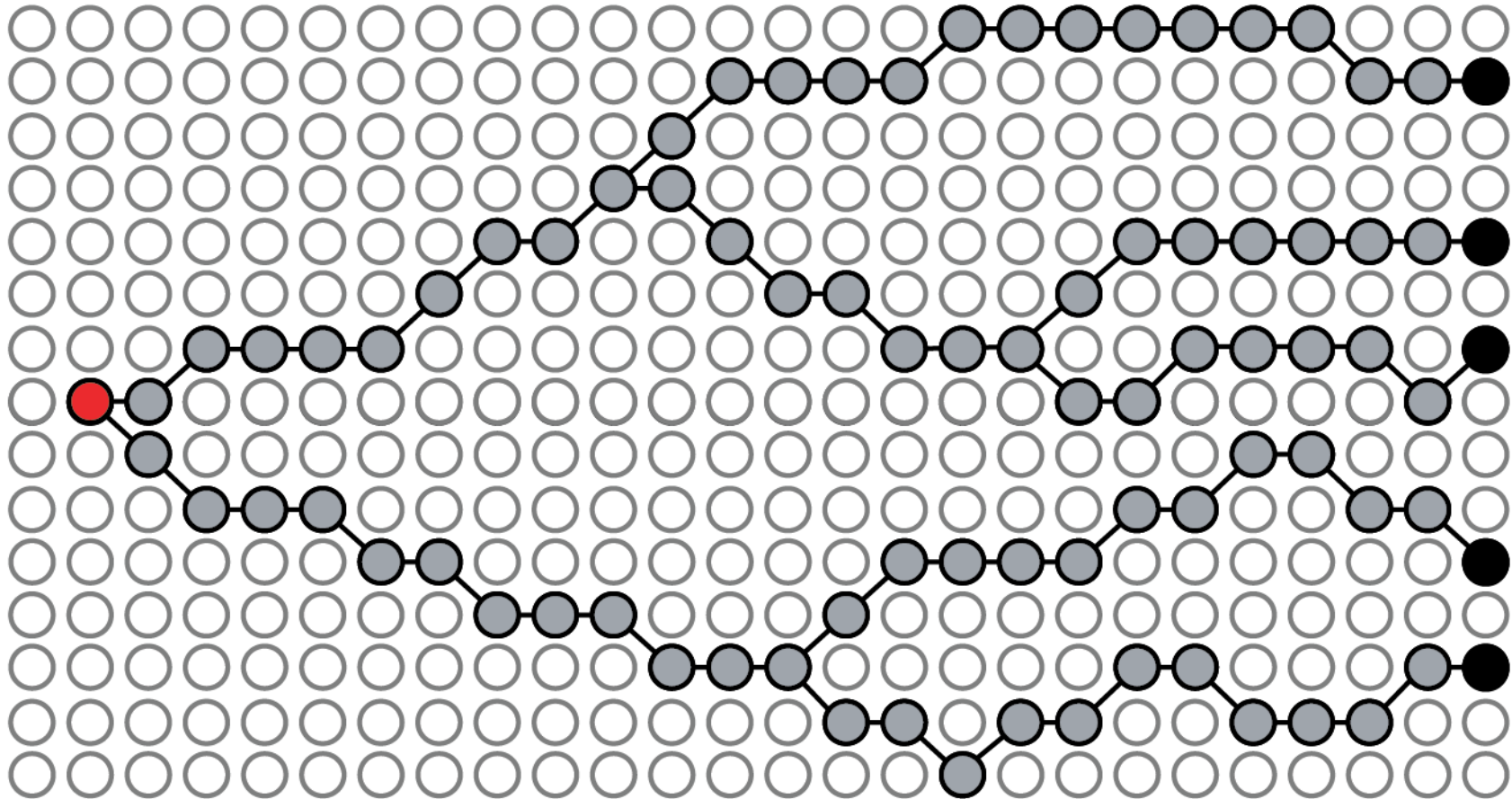

Lecture 4

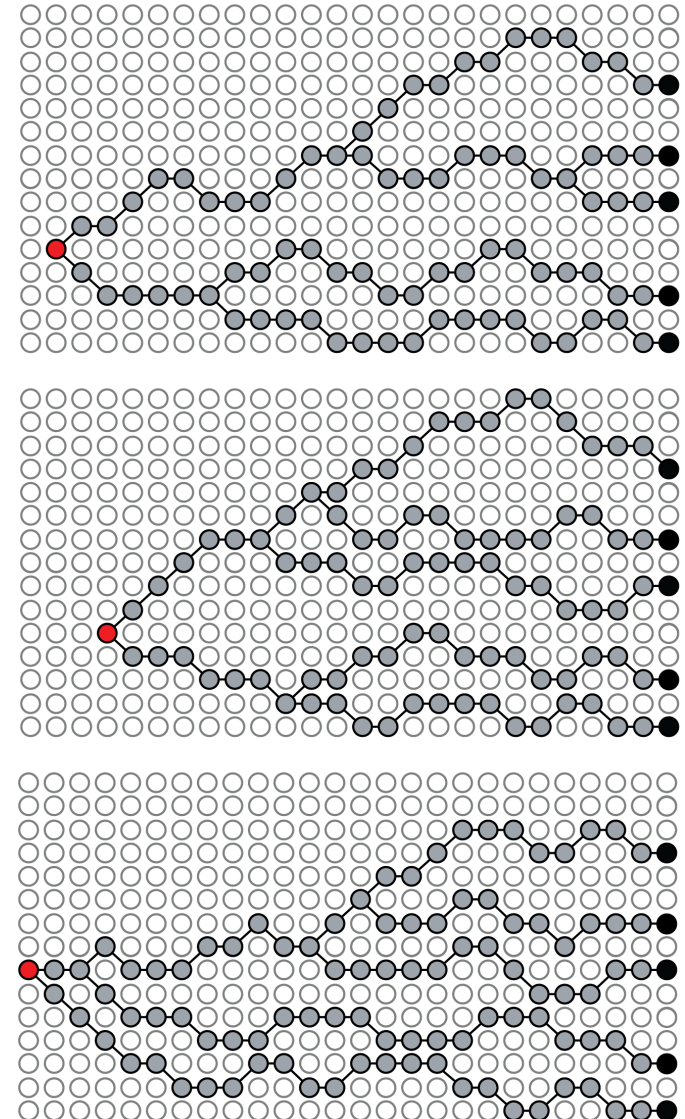
Gene trees

Coalescence theory



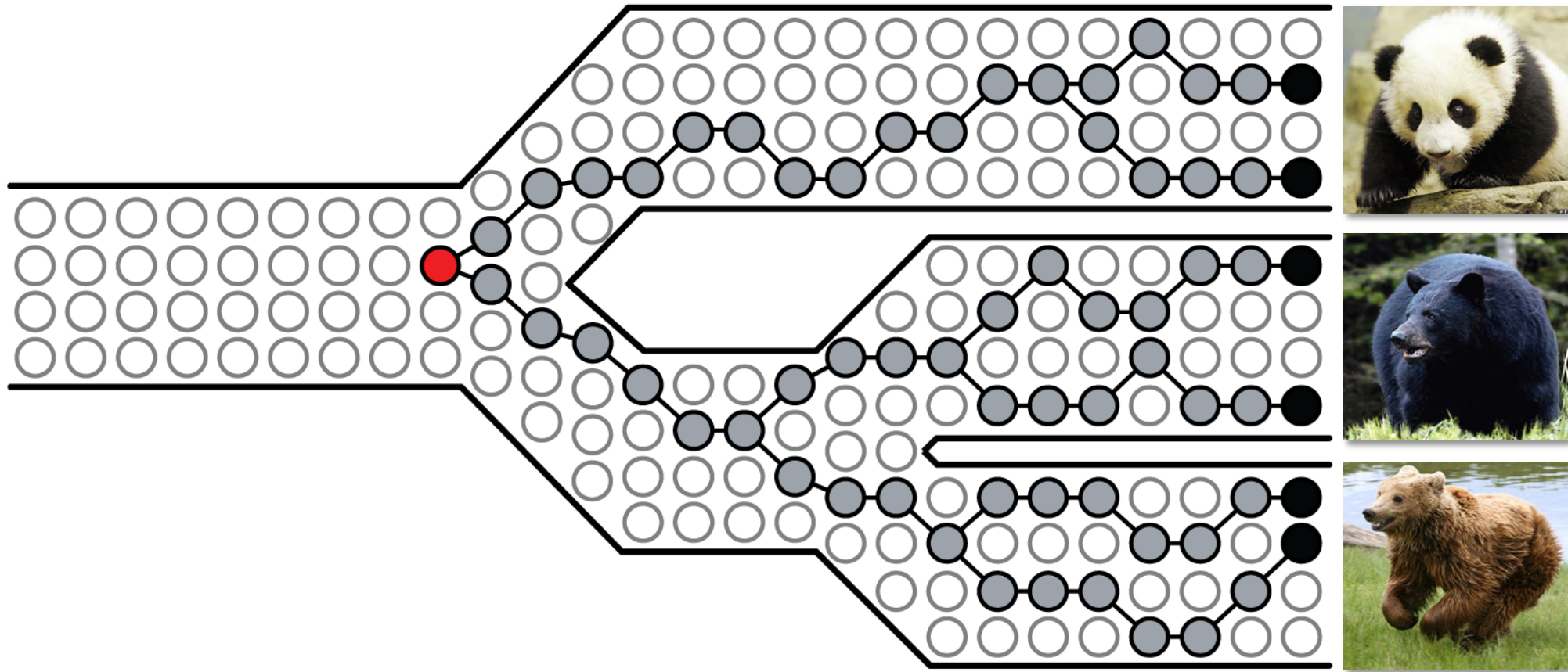
Gene trees within a species

- Gene genealogies vary stochastically between unlinked genes
- Independent genomic regions should not be grouped in analyses of a single species
 - They have different tree signals
 - They have different coalescent times



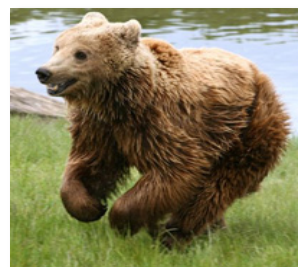
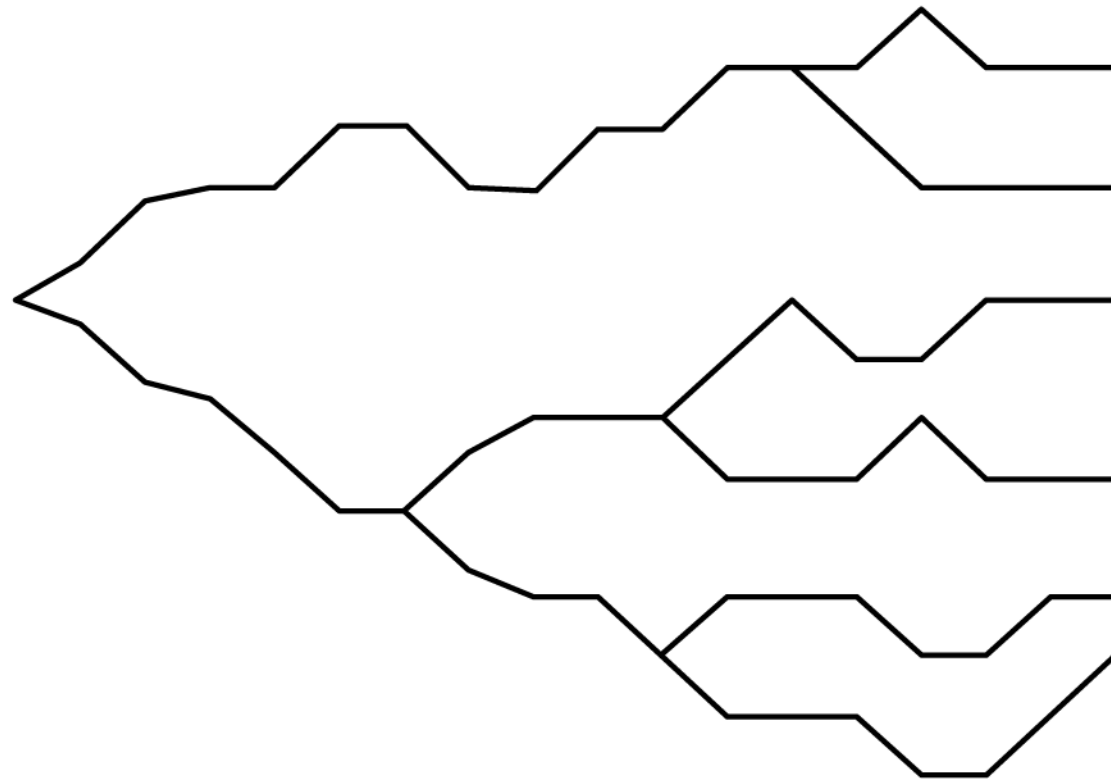
Gene trees across multiple species

- Gene trees within a tree of multiple species



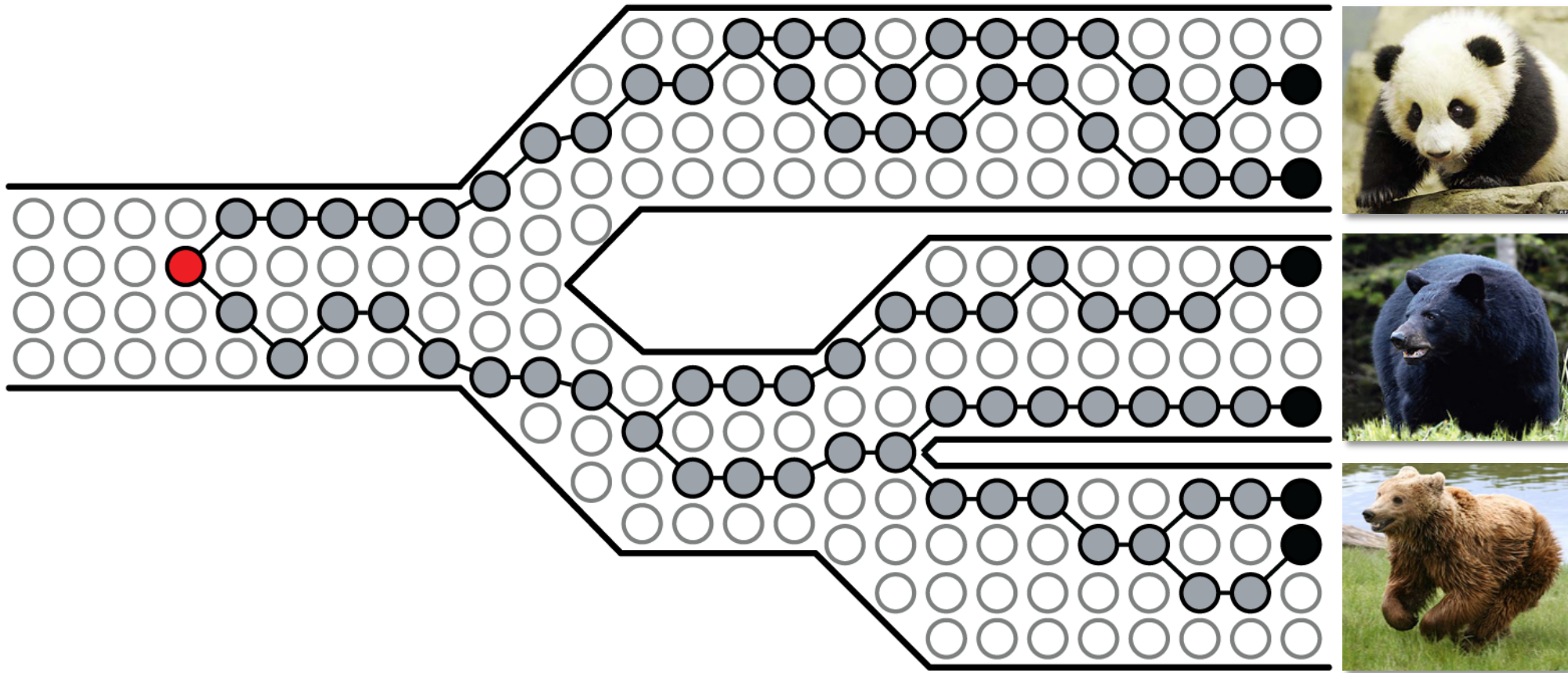
Gene trees across multiple species

- Gene trees within a tree of multiple species



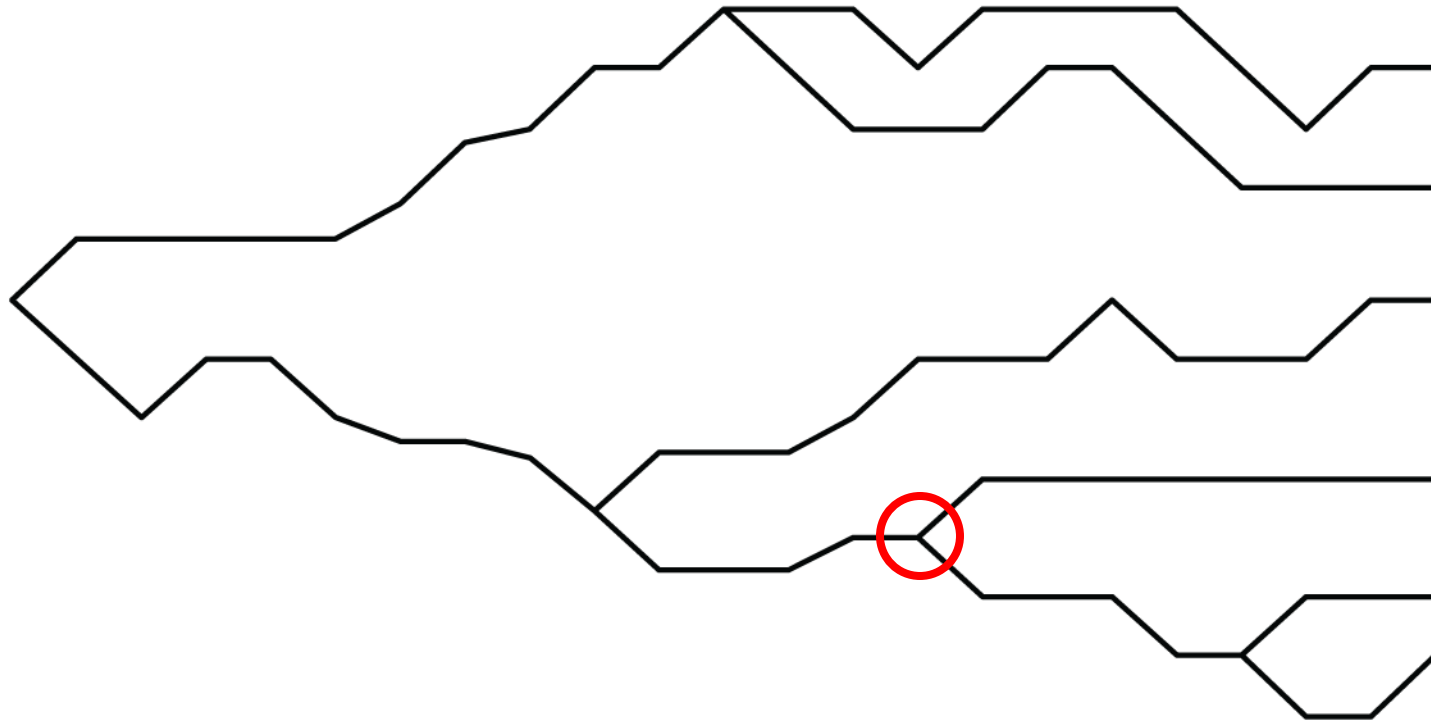
Gene trees across multiple species

- Incomplete lineage sorting



Gene trees across multiple species

- Incomplete lineage sorting



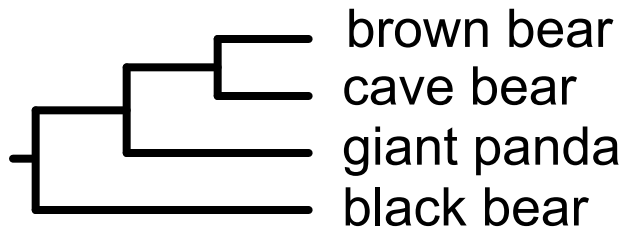
The species tree

- **Incomplete lineage sorting** can lead to gene trees that are different from the species tree
- We can still infer the species tree using multiple genes
- Three approaches
 1. Consensus
 2. Concatenation
 3. Coalescent methods

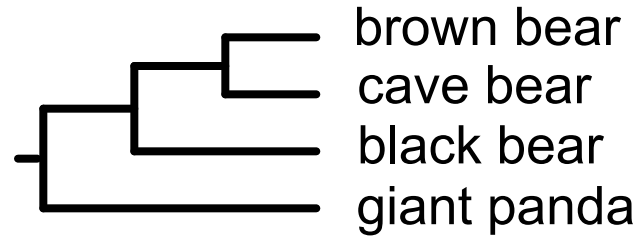
The species tree

1. Consensus

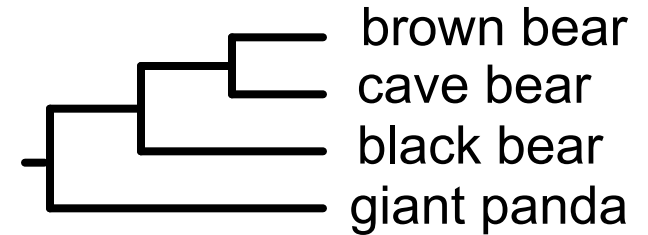
Estimate each gene genealogy and generate a consensus



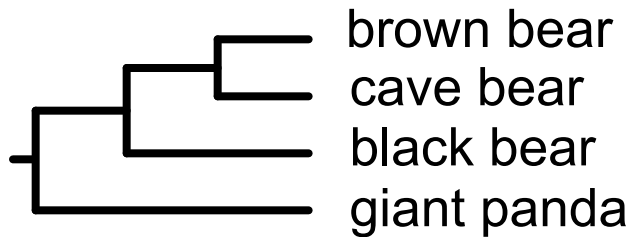
Gene 1



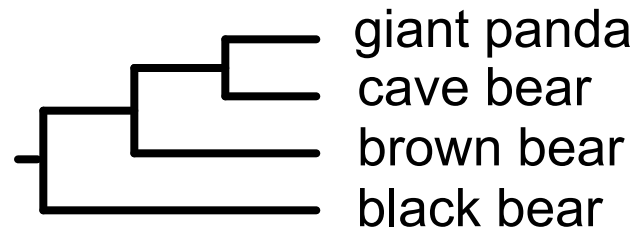
Gene 2



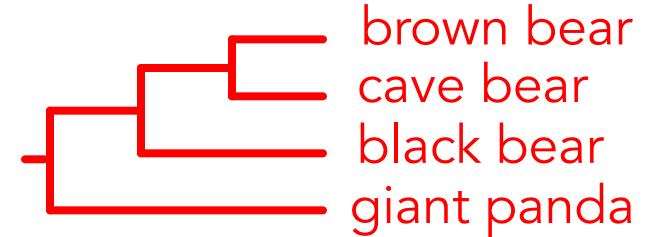
Gene 3



Gene 4



Gene 5



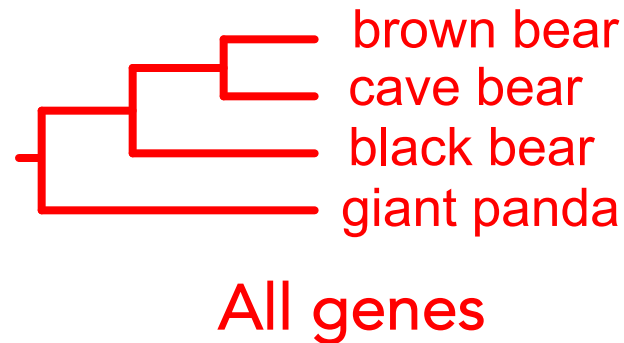
Consensus

But the most common gene tree is not always similar to the species tree ("anomaly zone")

Analysis of multiple genes

2. Concatenation

Assume that all genes share the same evolutionary history



Ignores differences among gene trees

The species tree

3. Coalescent methods

Estimate the species tree from gene trees

- Assume that gene trees are independent samples from a stochastic process withing the species tree (the multi-species coalescent)
- Various methods
 - **BEAST* co-estimates gene- and species-trees

BIOINFORMATICS

Vol. 30 ECCB 2014, pages i541–i548
doi:10.1093/bioinformatics/btu462

ASTRAL: genome-scale coalescent-based species tree estimation

S. Mirarab¹, R. Reaz¹, Md. S. Bayzid¹, T. Zimmermann^{1,2}, M. S. Swenson³ and
T. Warnow^{1,*}

Useful references

- Gene tree discordance, phylogenetic inference and the multispecies coalescent
Degnan & Rosenberg (2009) *Trends Ecol Evol*, 24: 332–340.
- Lineage sorting in apes
Mailund *et al.* (2014) *Annu Rev Genet*, 48: 519–535.
- Estimating phylogenetic trees from genome-scale data
Liu *et al.* (2015) *Annals New York Acad Sci*, 1360: 36–53.
- Analysis of phylogenomic tree space resolves relationships among marsupial families
Duchêne *et al.* (2018) *Syst Biol*, 67: 400–412.