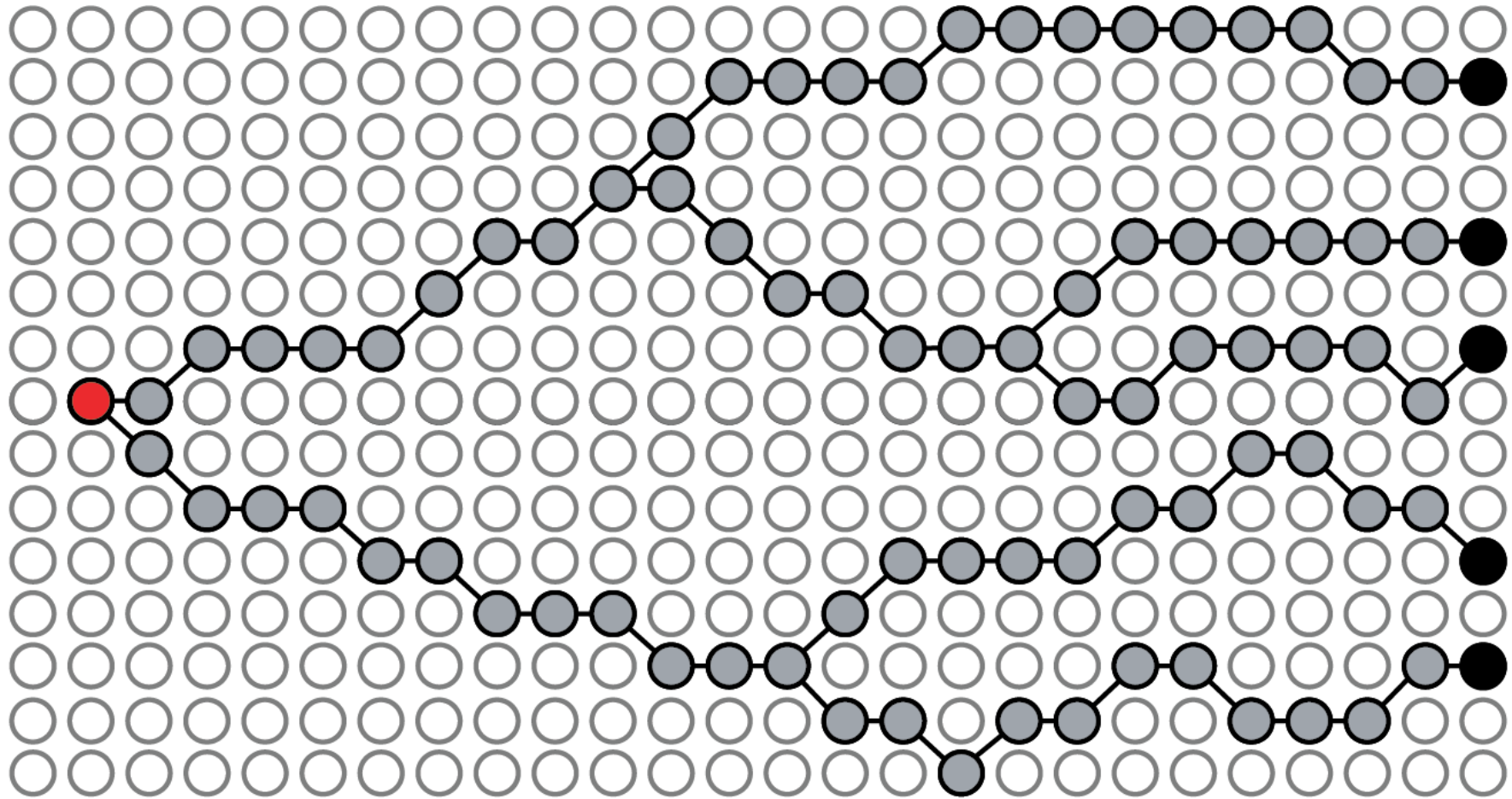


---

Lecture 1.5

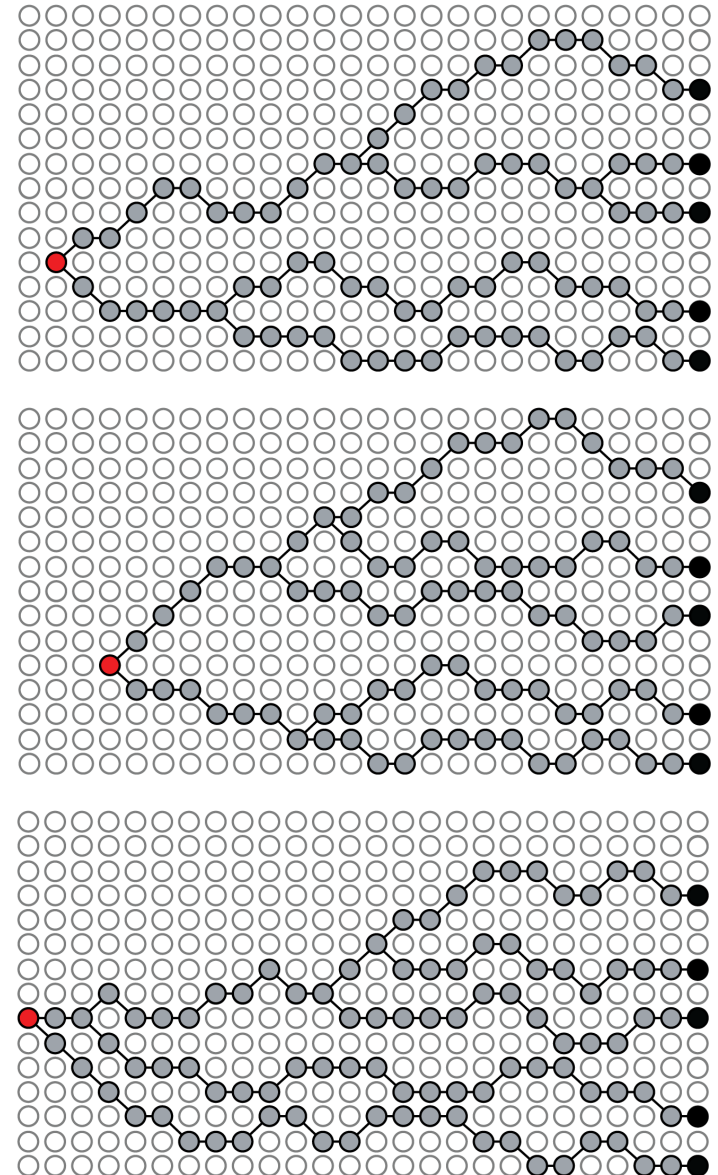
**Gene Trees**

---



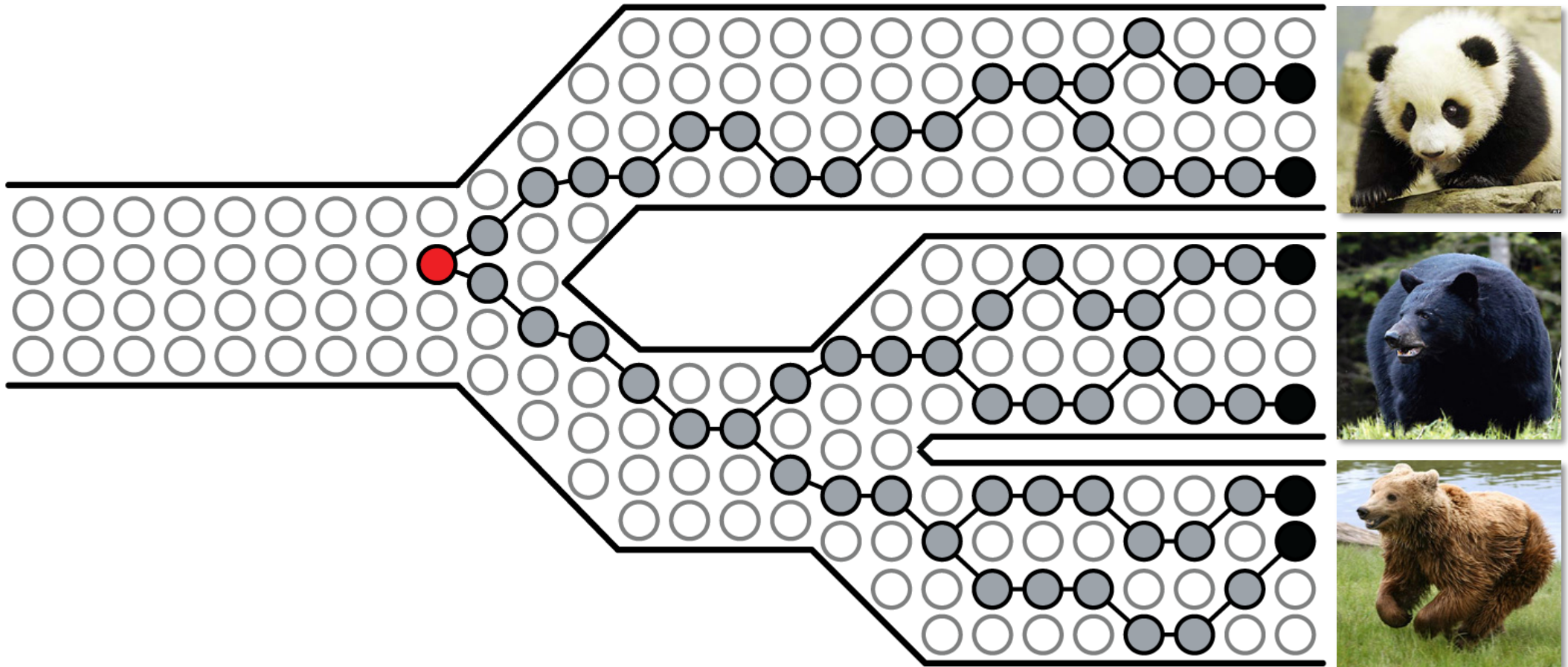
# Gene trees in a species

- Genealogies vary stochastically among unlinked loci
- Should not concatenate independent loci in a phylogenetic analysis of intraspecific data
  - Different trees
  - Different coalescence times



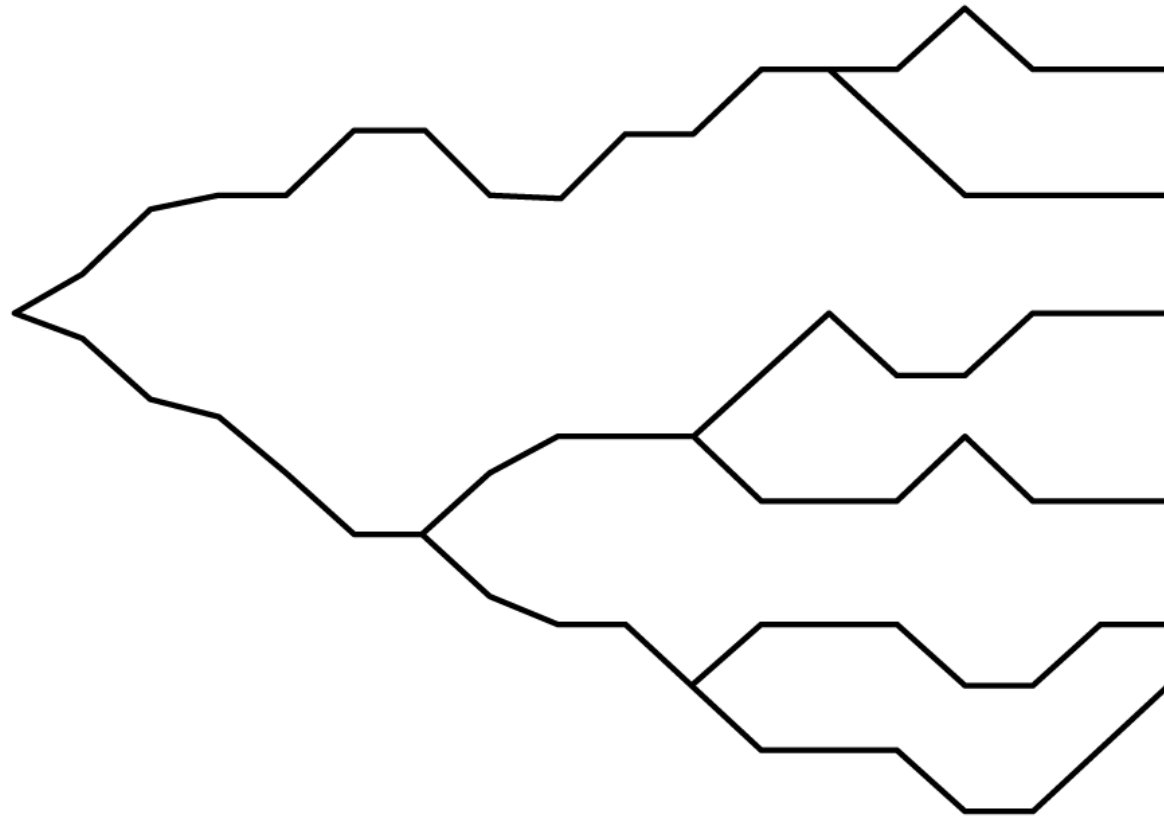
# Gene trees in multiple species

- Gene trees are embedded in the species tree



# Gene trees in multiple species

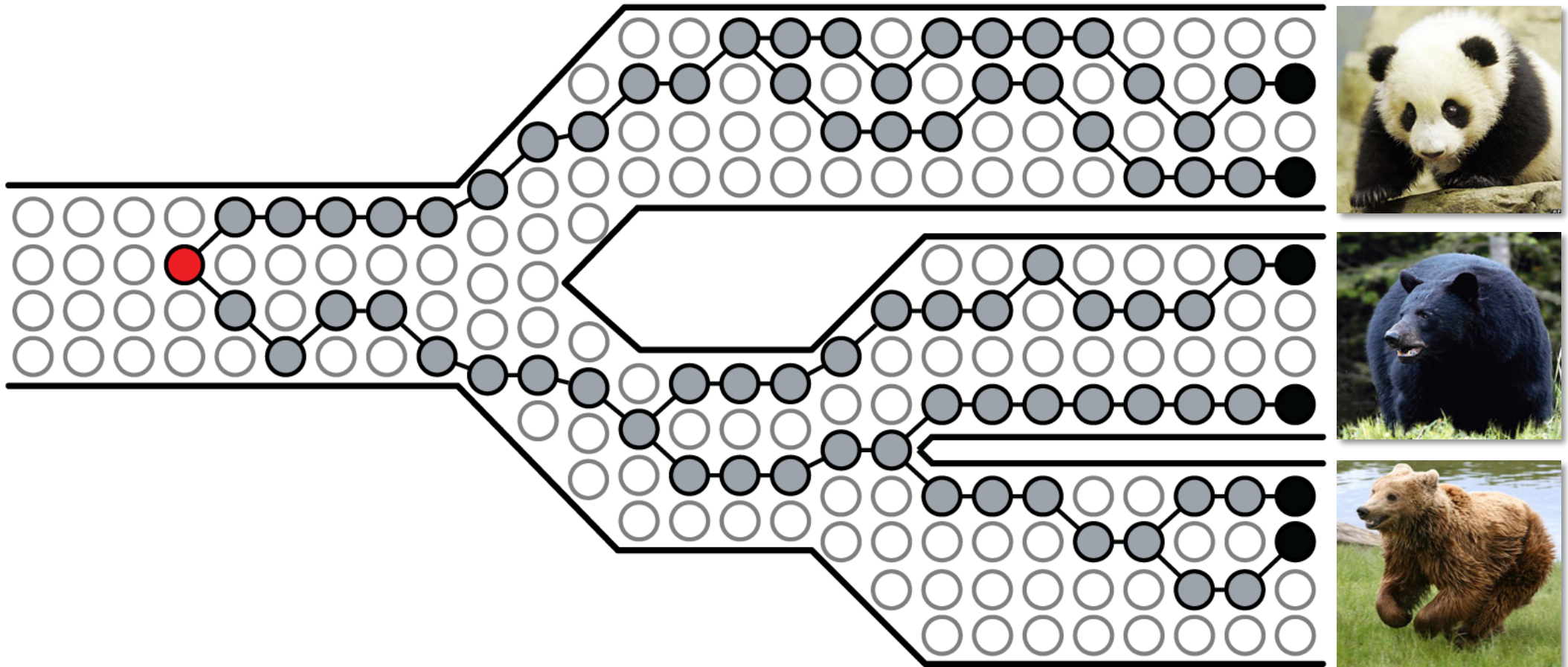
- Gene trees are embedded in the species tree





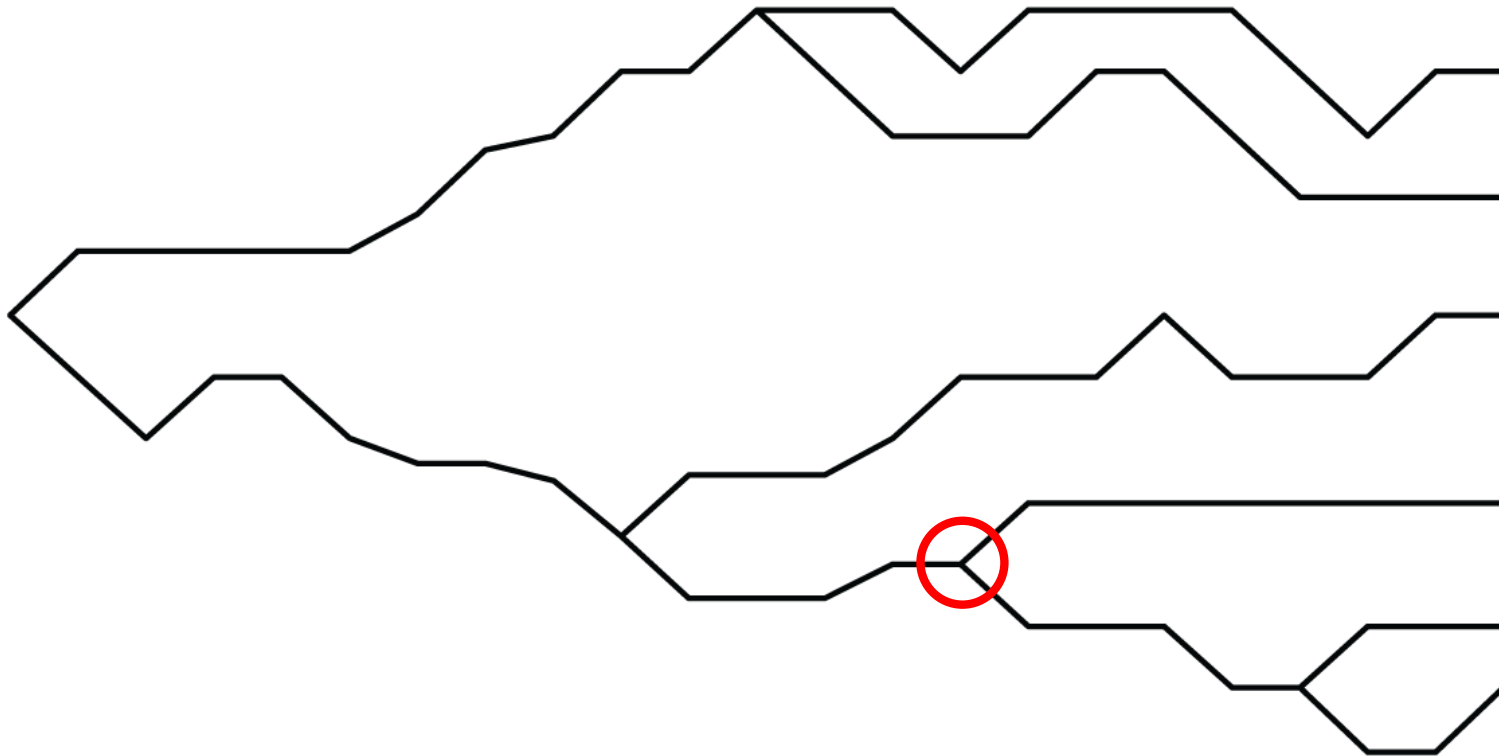
# Gene trees in multiple species

- Incomplete lineage sorting



# Gene trees in multiple species

- Incomplete lineage sorting



# Species tree

---

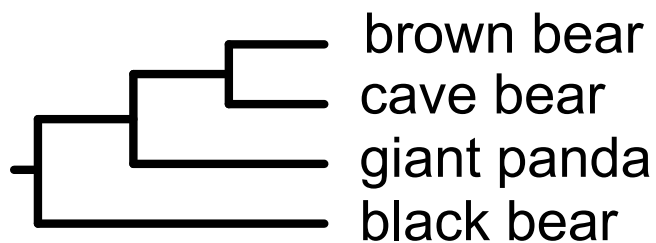
- **Incomplete lineage sorting** can lead to gene trees that do not match the species tree
- We can infer the species tree from multiple gene trees
- Three approaches
  1. Consensus
  2. Concatenation
  3. Species-tree methods



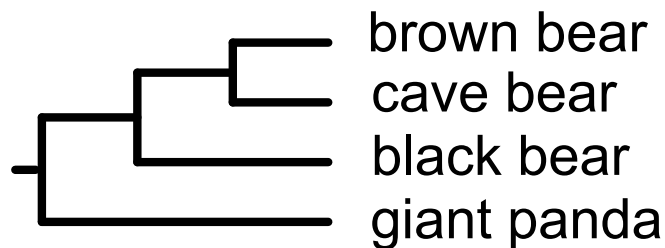
# Species tree

## 1. Consensus

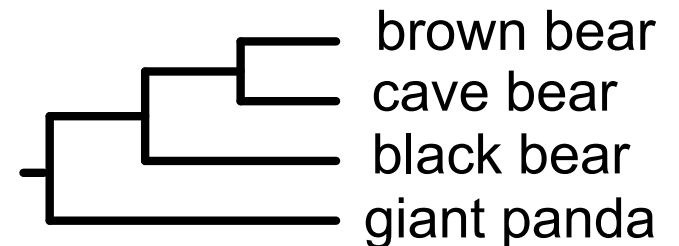
Estimate genealogy from each locus and find the consensus



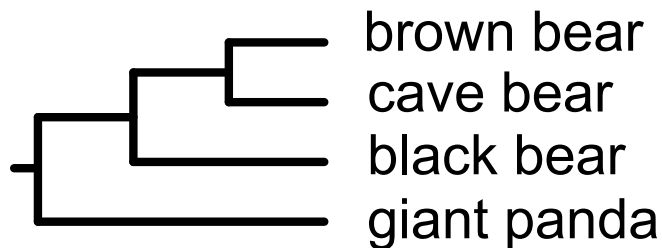
**Locus 1**



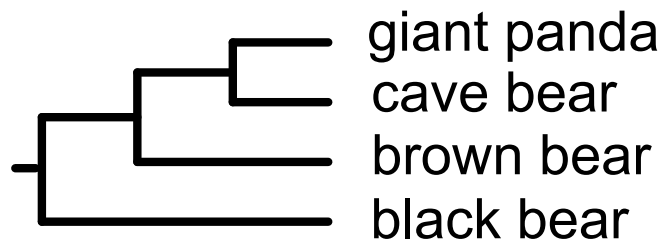
**Locus 2**



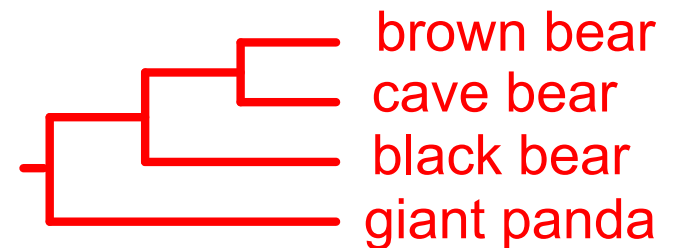
**Locus 5**



**Locus 4**



**Locus 5**



**Consensus**

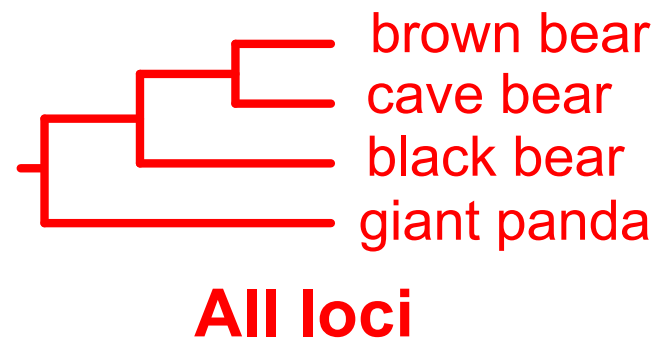
But the most frequent gene tree does not always match the true species tree (anomaly zone)

# Analysing multiple loci

---

## 2. Concatenation

Assume that all loci share the same evolutionary history



But this ignores the occurrence of different gene trees

# Species tree

---

## 3. Species-tree methods

Estimate the species tree based on gene trees

- Gene trees are independent realisations of a stochastic process (the coalescent) on the same species tree
- Various methods
  - \**BEAST* co-estimates gene trees and the species tree

**BIOINFORMATICS**

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

---

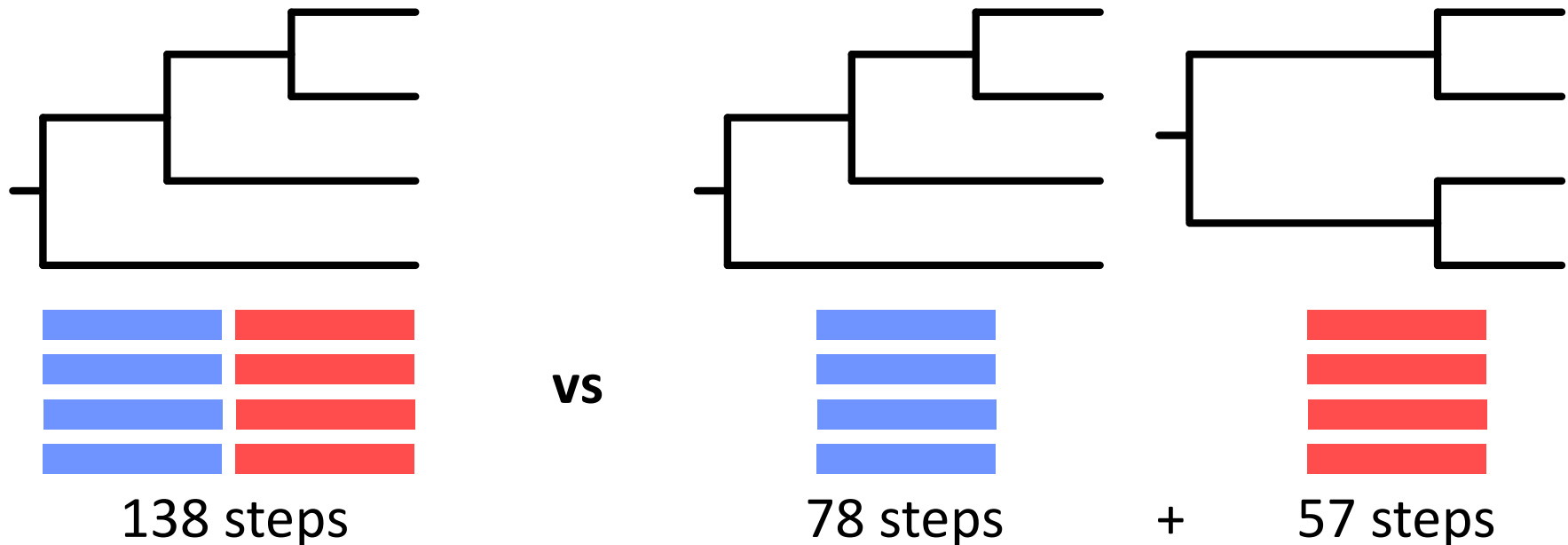
## **ASTRAL: genome-scale coalescent-based species tree estimation**

S. Mirarab<sup>1</sup>, R. Reaz<sup>1</sup>, Md. S. Bayzid<sup>1</sup>, T. Zimmermann<sup>1,2</sup>, M. S. Swenson<sup>3</sup> and  
T. Warnow<sup>1,\*</sup>

# Partition-homogeneity test

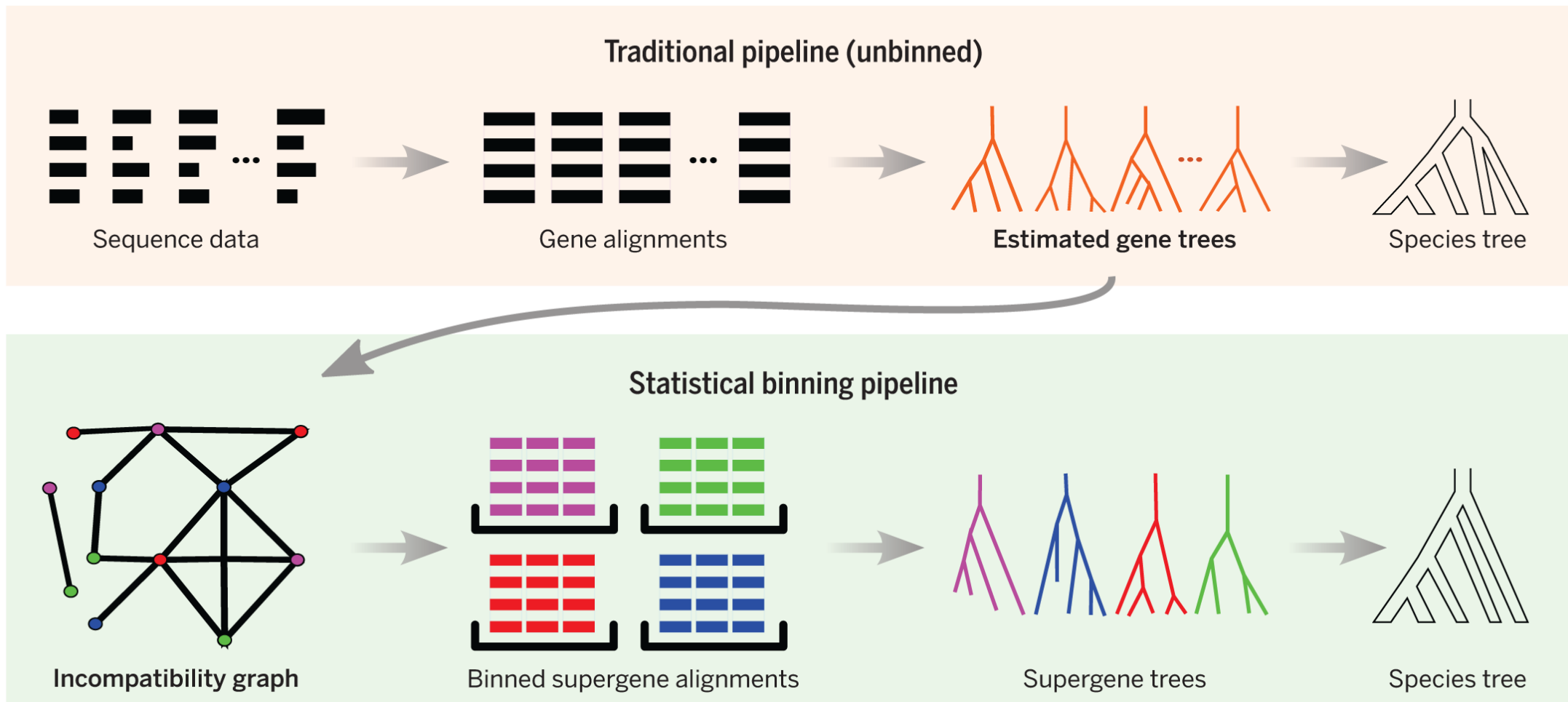
---

- Test for phylogenetic congruence across markers
- Partition-homogeneity (incongruence length difference) test



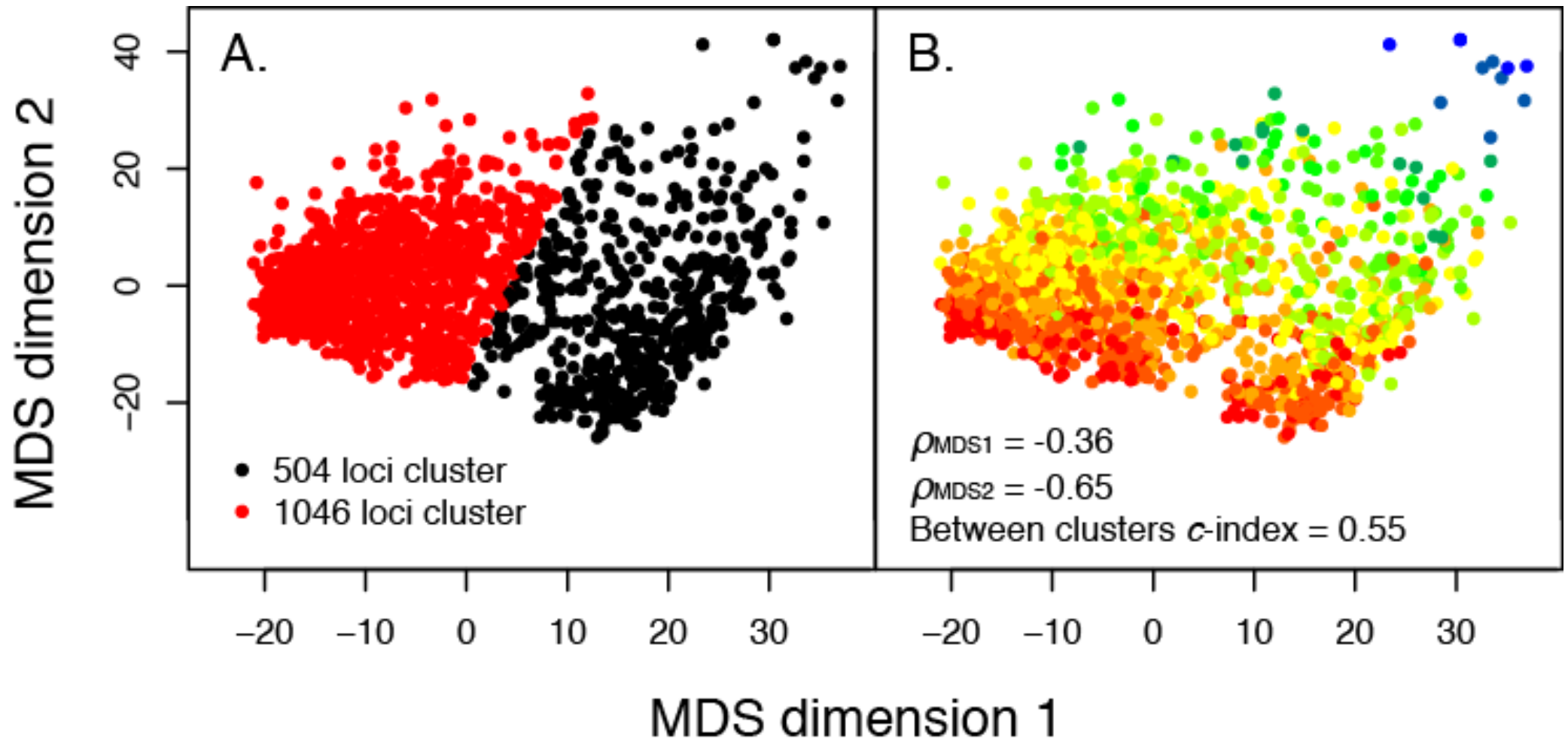
# Species tree

- Statistical binning



# Species tree

- Topology clustering





# Useful references

---

- **Gene tree discordance, phylogenetic inference and the multispecies coalescent**  
Degnan & Rosenberg (2094) *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.