

# Methodological advances in reticulate evolution

---

# Acknowledgements

---

Botanical Society of America

American Society of Plant Taxonomists

Society of Herbarium Curators



# Conclusions

---

1. Introgression is a process generating gene tree variation
2. Some networks are identifiable from sequence data and gene trees
3. Networks are a valuable tool for investigating plant evolution

# Gene tree variation is normal

---

Now well-accepted that many, sometimes a majority of gene trees, may conflict with a true species history

Gene Duplication and Loss

Deep Coalescence (ILS)

Horizontal Inheritance

Sequence Error

Lack of information

Model misspecification

# Gene tree variation is normal

---

Now well-accepted that many, sometimes a majority of gene trees, may conflict with a true species history

*Syst. Biol.* 46(3):523–536, 1997

## Gene Duplication and Loss

## Deep Coalescence (ILS)

## Horizontal Inheritance

## Sequence Error

## Lack of information

## Model misspecification

## GENE TREES IN SPECIES TREES

WAYNE P. MADDISON

*Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721, USA*

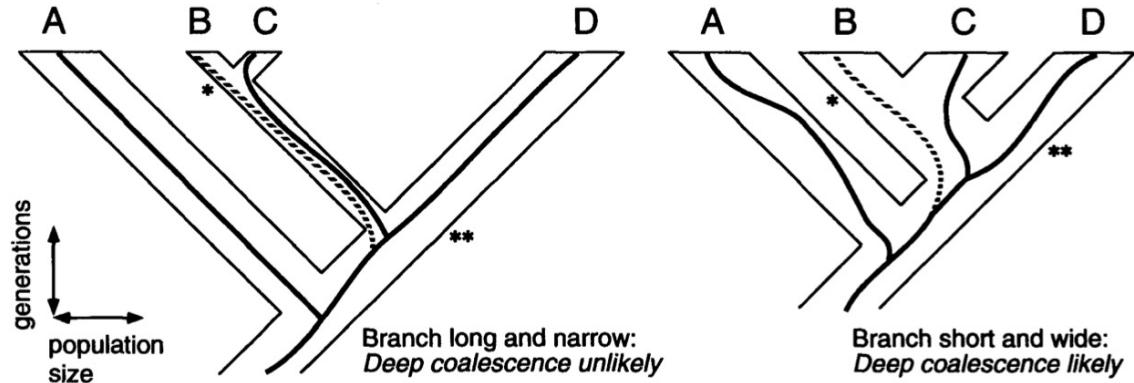
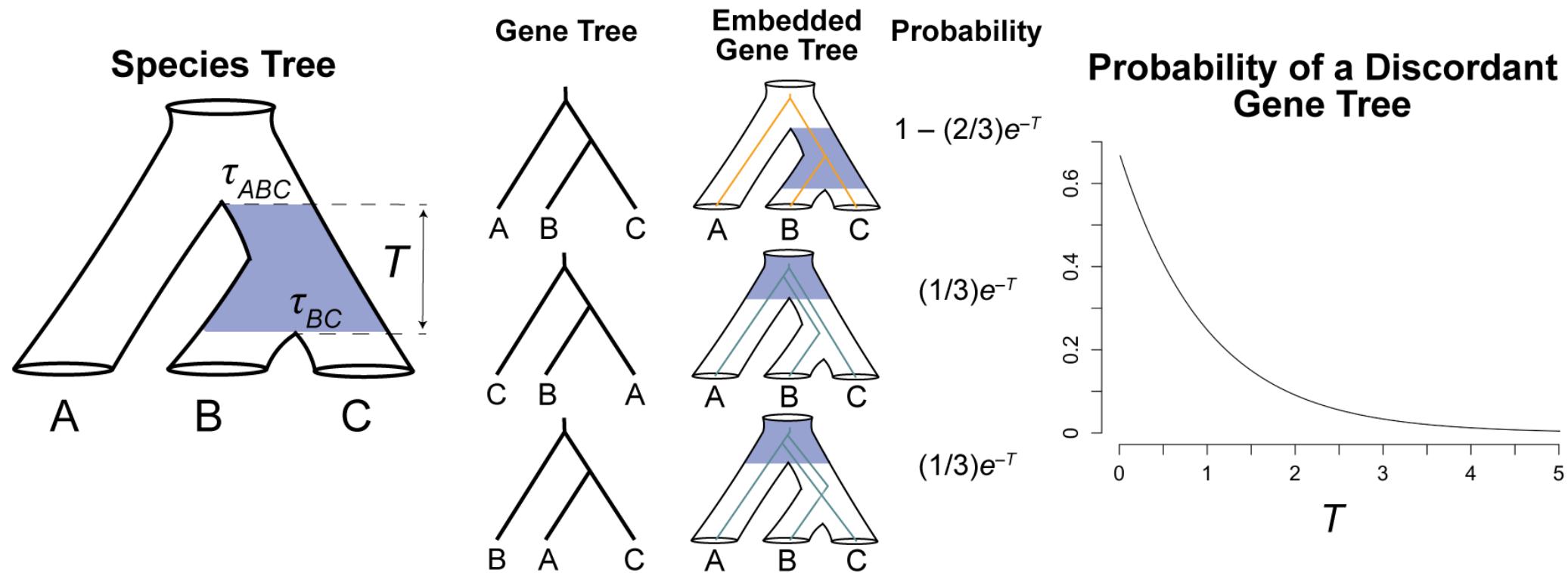


FIGURE 4. Lineage sorting (deep coalescence). Described in a time-forward sense as lineage sorting, an ancestral polymorphism at \*\*\* is retained through a lineage to the next speciation event at \*, where different forms are sampled in different descendant species. Described in a time-backward sense as deep coalescence, two gene copies from species B and C meet at \* but fail to coalesce until deeper than the speciation event at \*\*, at which point the gene from C coalesces first with the gene from D. Failure to coalesce is more likely the shorter (in generations) and wider (in effective population size) the branch is between \*\* and \*.

# Gene tree variation is normal

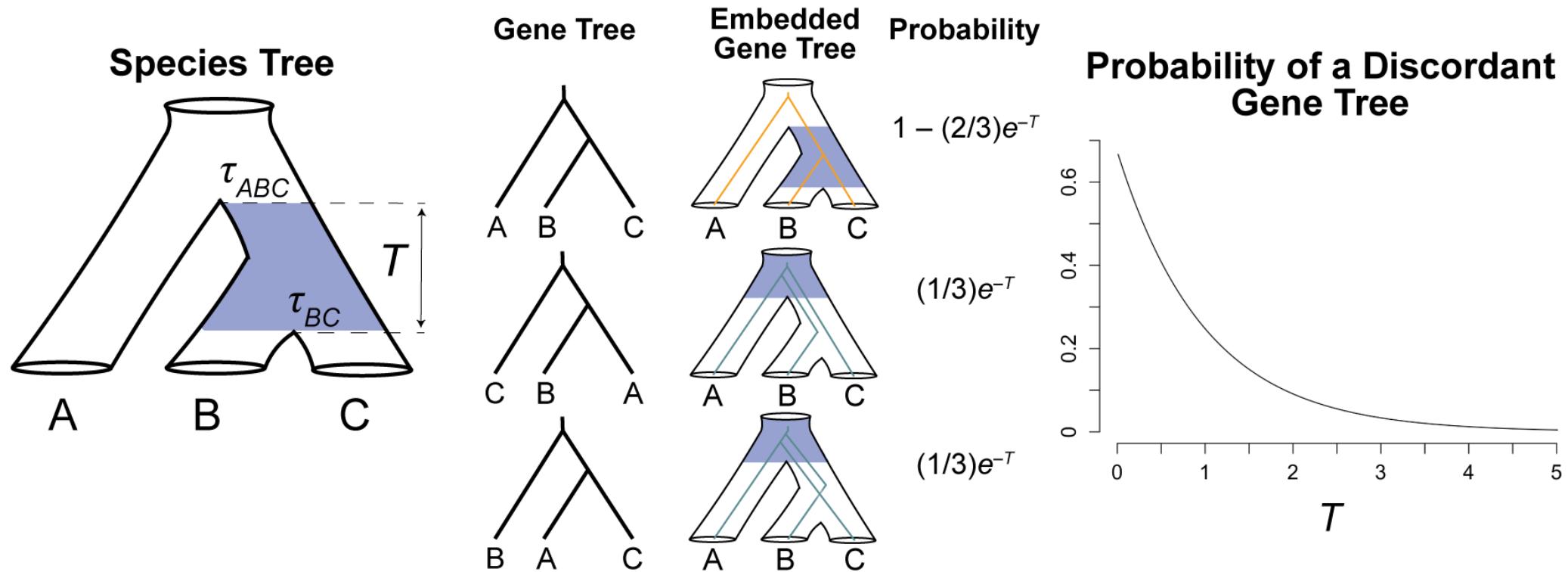
## Deep Coalescence – Incomplete Lineage Sorting (ILS)



$$\tau = \frac{t}{2N_e} = \frac{\mu t}{\theta} \quad \begin{matrix} \text{coalescent branch lengths} \\ \text{mutation-scaled branch lengths} \end{matrix}$$

# Gene tree variation is normal

## Deep Coalescence – Incomplete Lineage Sorting (ILS)



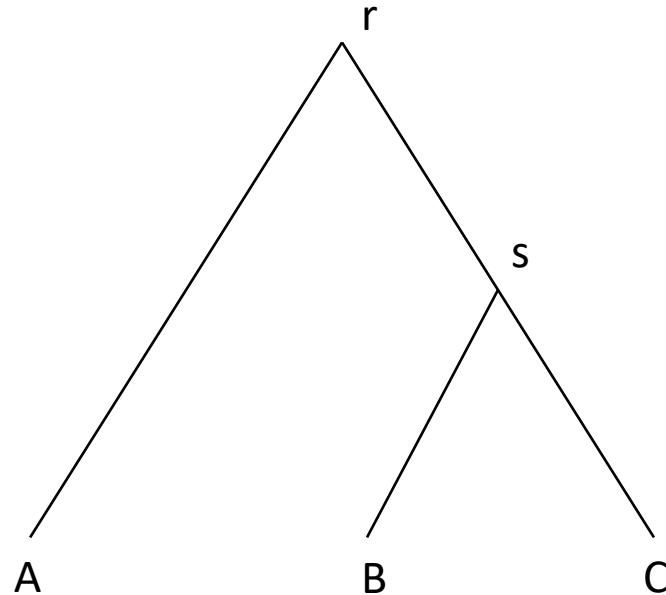
Rannala B and Yang Z. 2003. Genetics. 164:1645-1656.  
Pamilo P and Nei M. 1988. Mol Biol Evol. 5:568-583.  
Hudson RR. 1983. Evolution. 37:203-217.  
Kingman JFC. 1982. J Appl Probab. 19A:27-43.  
Ewans WJ. 1972. Theor Popul Biol. 3:87-112.

Fig. 2 — Tiley et al. 2020 Trends Genet. 36:845-856

# Gene tree variation is normal

---

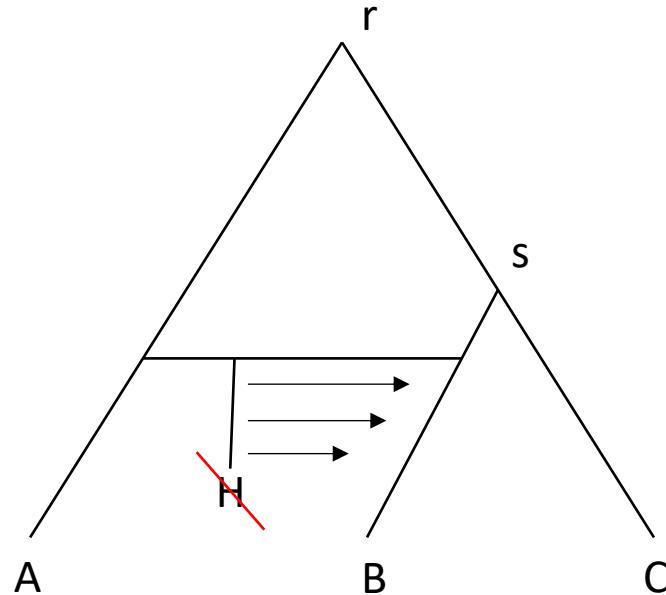
Different mechanisms of gene flow might affect our models



# Gene tree variation is normal

---

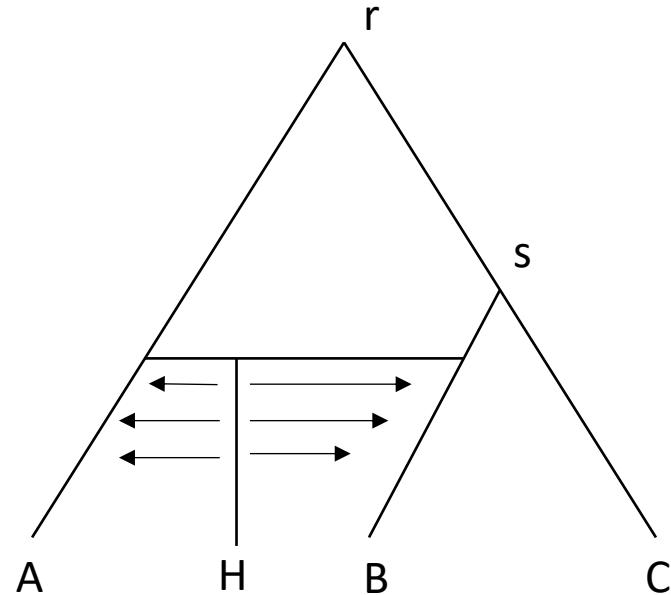
## Introgressive Hybridization (Introgression)



# Gene tree variation is normal

---

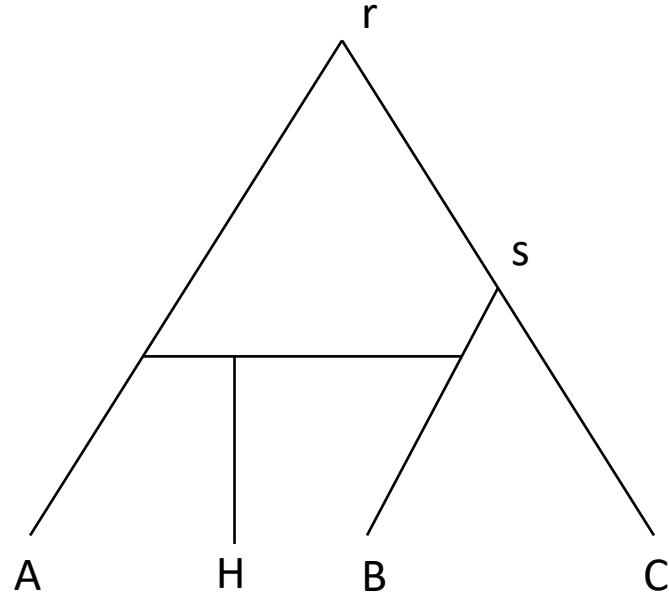
## Homoploid Hybrid Speciation



# Gene tree variation is normal

---

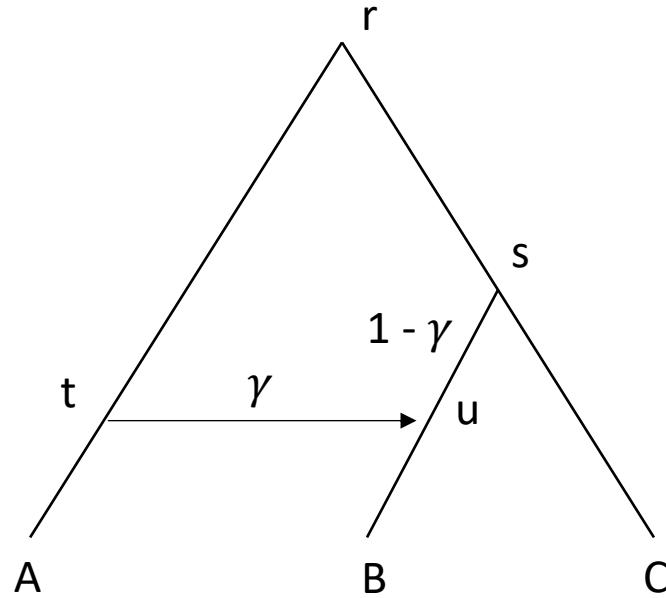
## Allopolyploid Hybrid Speciation



# Gene tree variation is normal

---

## Introgression



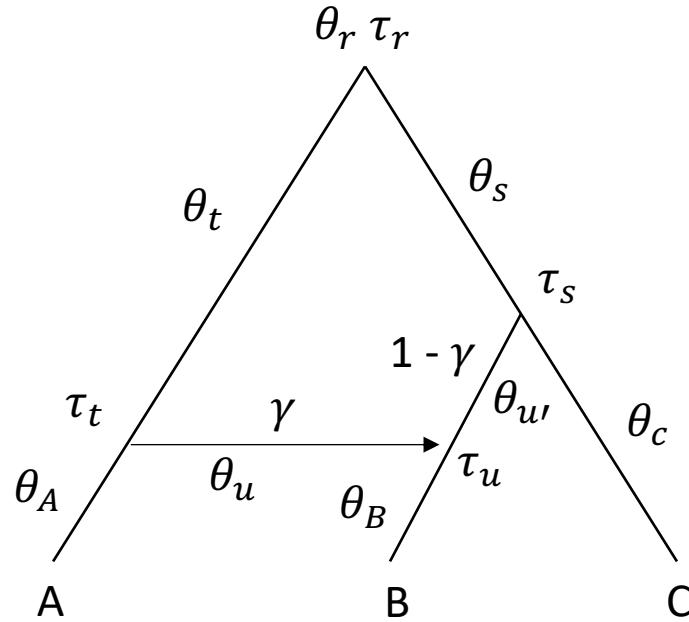
This is modeled as an episodic process  
Meng and Kubatko (2009)  
Yu et al. (2012)

$\gamma$  is the inheritance probability  
The distribution of gene trees is not  
a simple mixture with proportions  $\gamma$   
and  $1-\gamma$ , it is a polynomial function of  $\gamma$   
Solís-Lemus and Ané (2016)

# Gene tree variation is normal

---

## Introgression



Optimizing the likelihood function for this MSC model with introgression (MSNC, NMSC, MSci) is very difficult except for very small examples

Estimation of the likelihood through Bayesian MCMC can incur a heavy time and computing burden

gene tree density from Yu et al. 2014

$$f(\tau, \theta, \gamma | X) \propto f(\tau, \theta, \gamma) \prod_{i=1}^L \int_{G_i} f(G_i | \tau, \theta, \gamma) \times f(X_i | G_i) \partial G_i$$

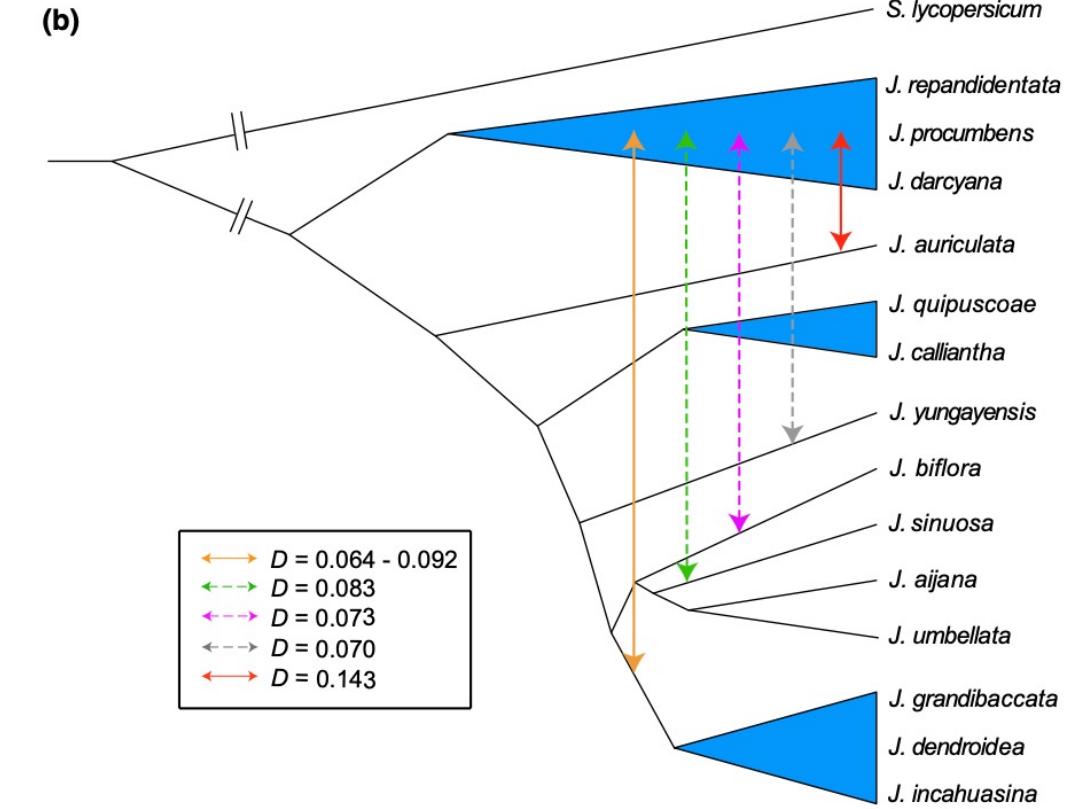
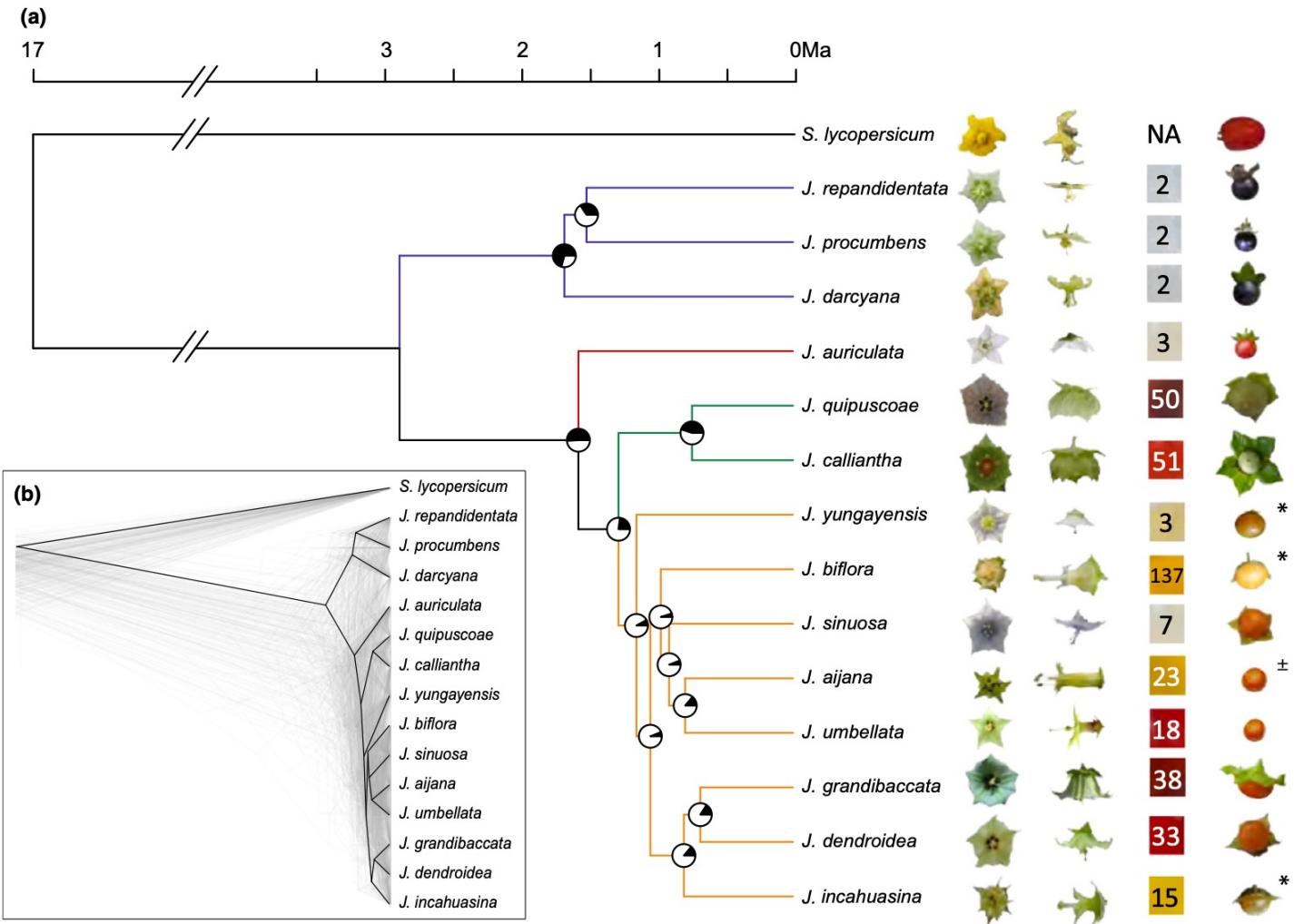
Felsenstein Likelihood

Eq. 1 from Flouri et al. (2020)

Most strategies employed will fix some aspects of the model and use smaller pieces of information  
**(unrooted quartets or rooted triples)** for searches

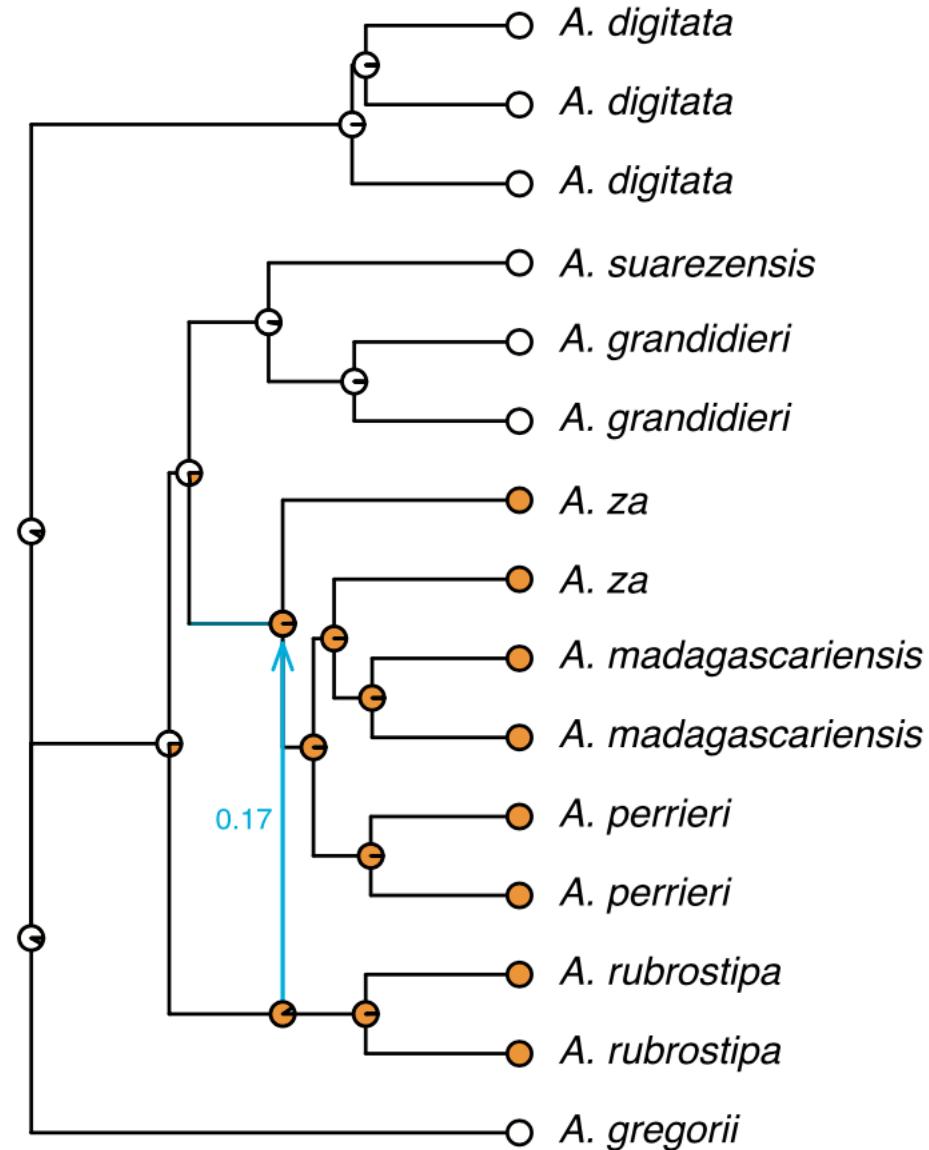
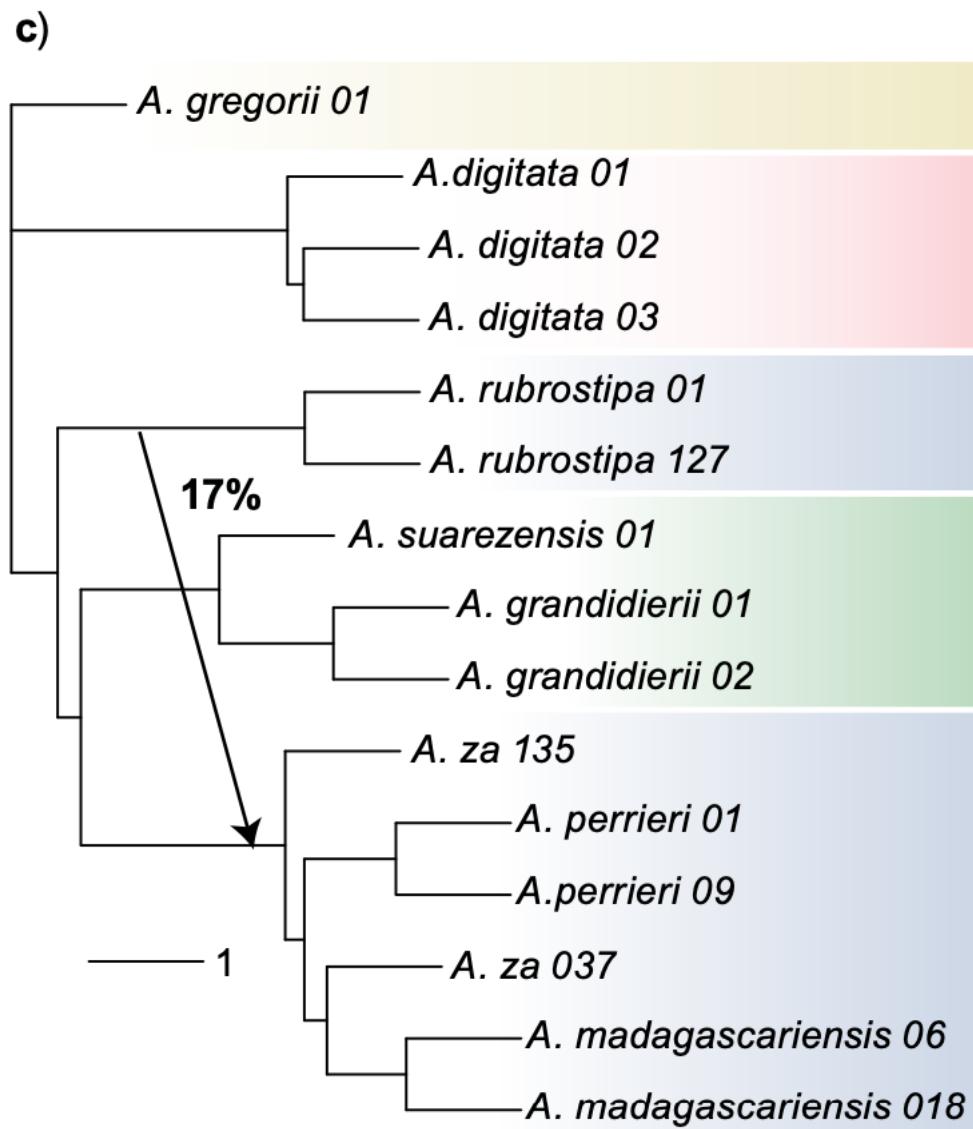
# Plants!

Within the genus *Jaltomata*, introgressive hybridization could explain gene tree discordance and fruit color evolution.



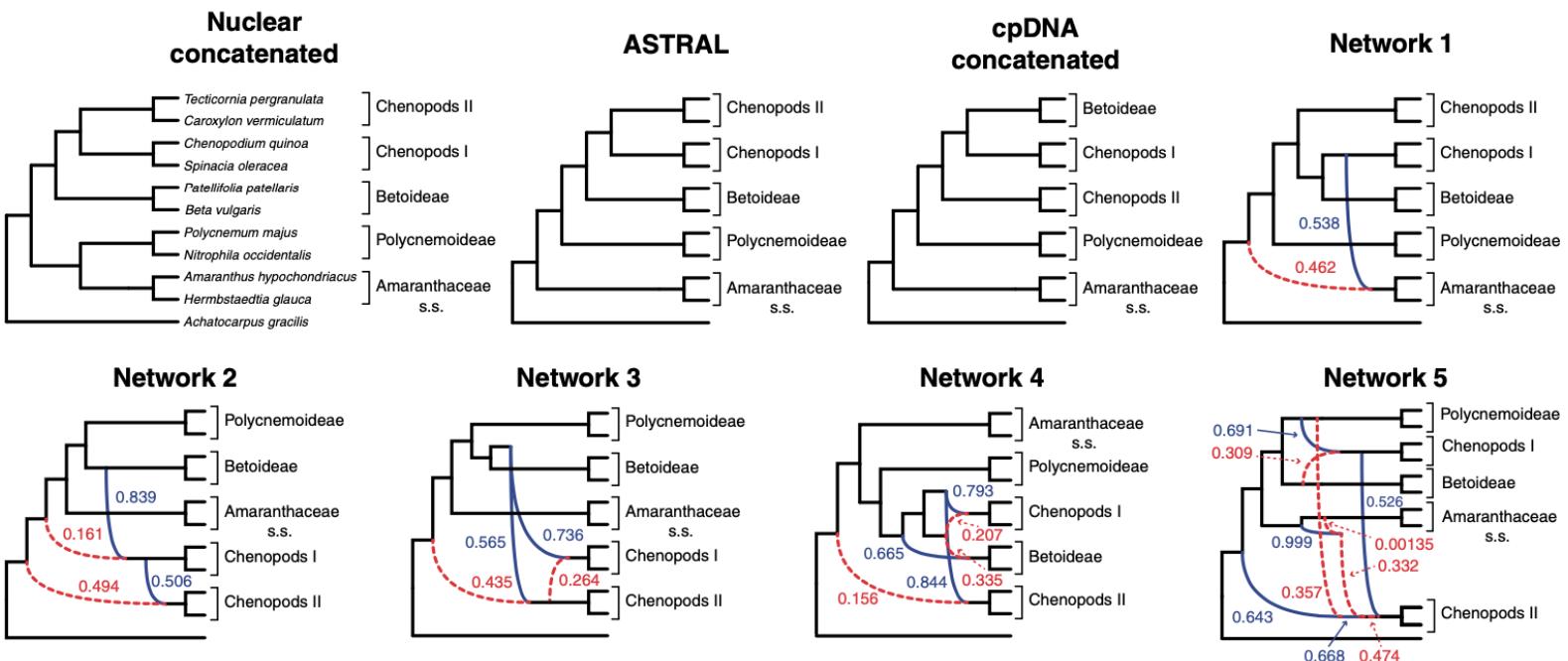
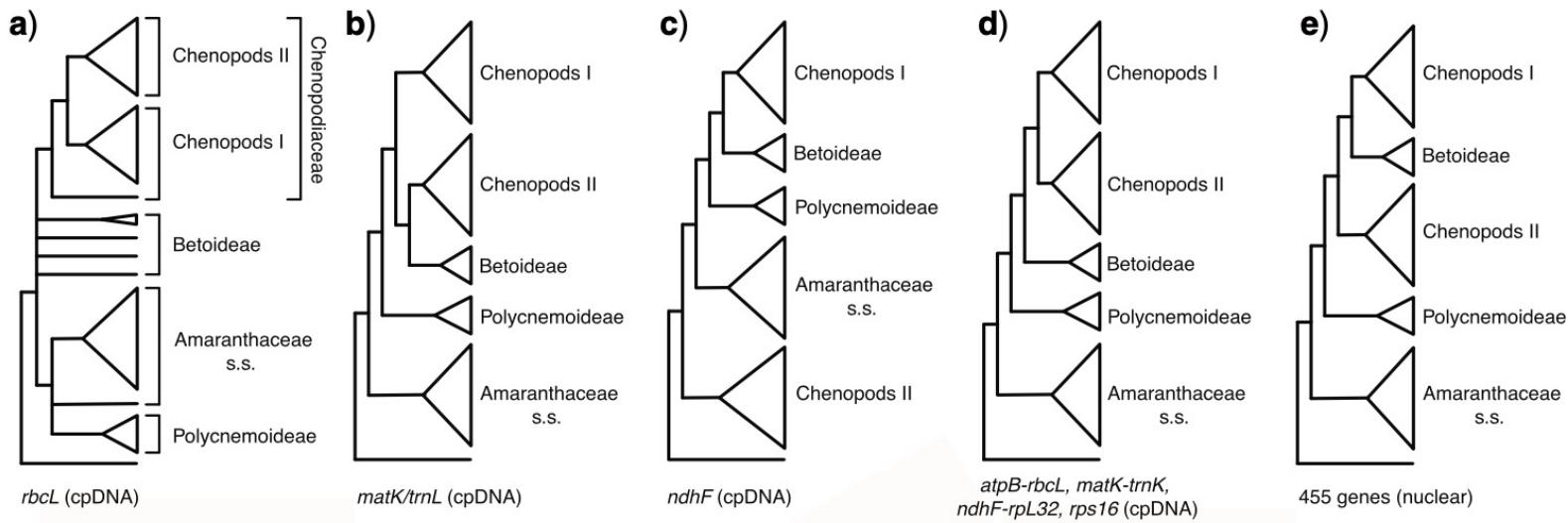
# Plants!

Introgression  
explains  
paraphyly of a  
section of  
*Adansonia* and  
the evolution of  
pollinator  
syndrome



# Plants!

Reticulate evolution can explain previous phylogenetic disagreements and discrepancies among methods in Amaranthaceae.

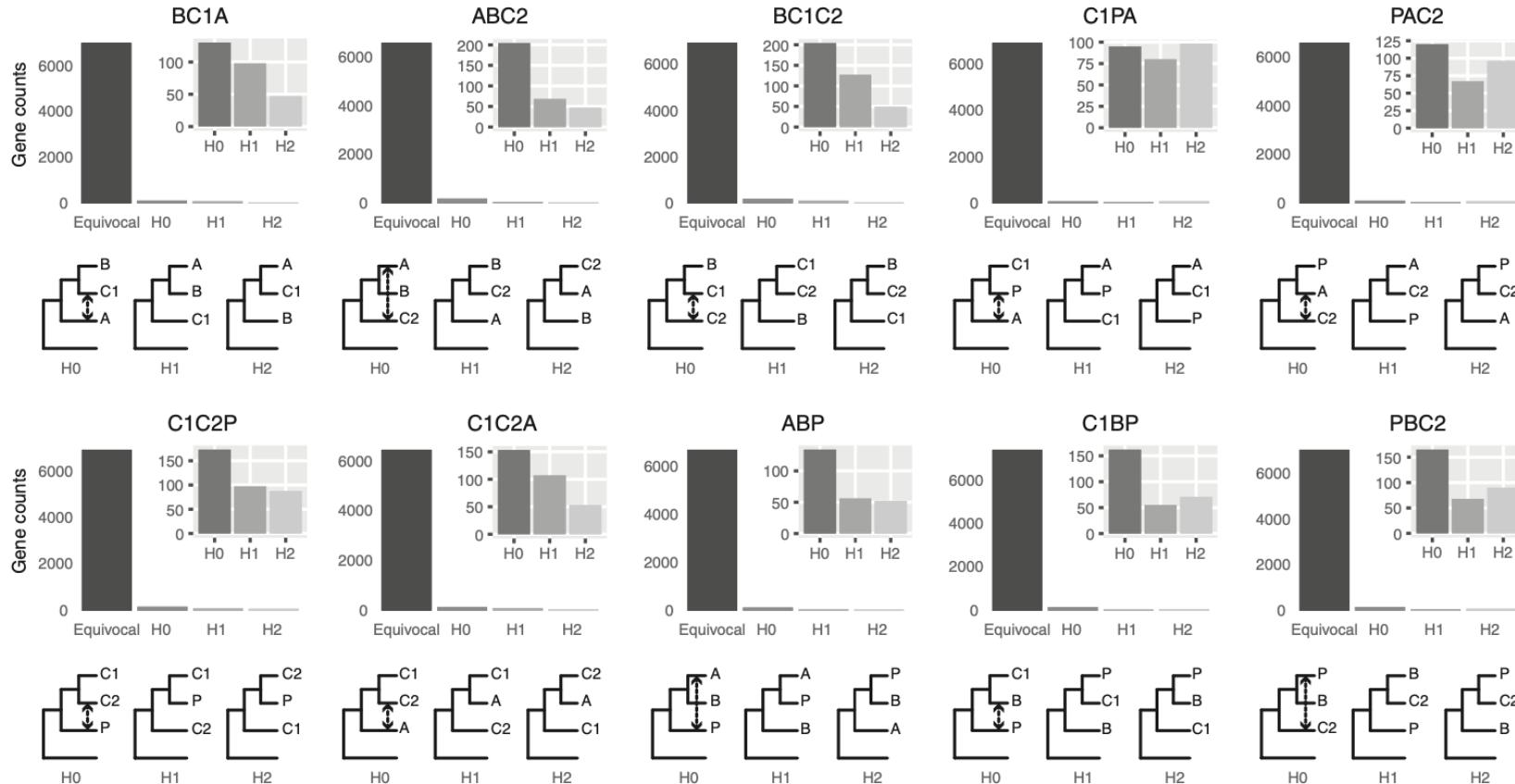


# Plants!

---

Reticulate evolution  
can explain previous  
phylogenetic  
disagreements and  
discrepancies among  
methods in  
Amaranthaceae.

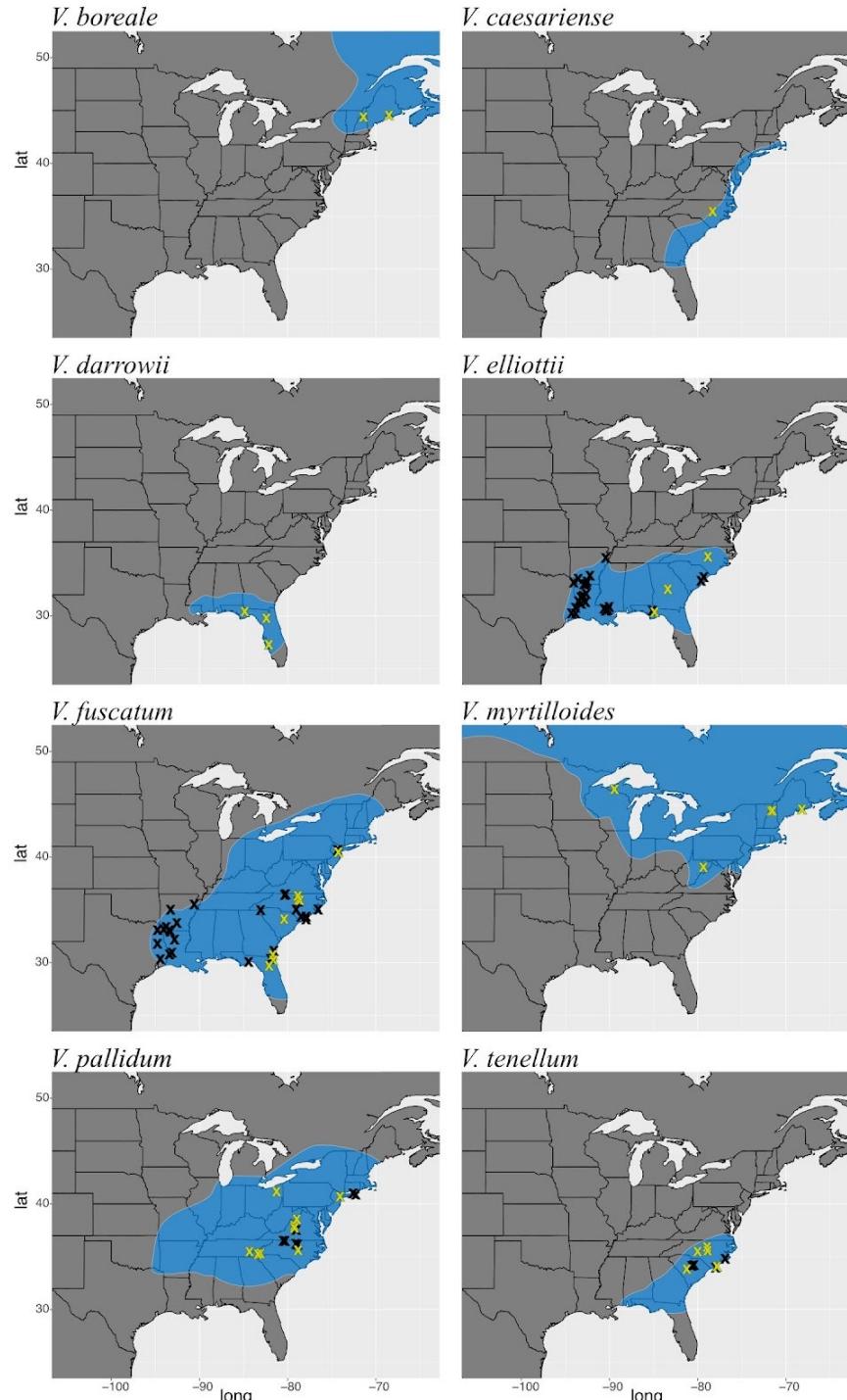
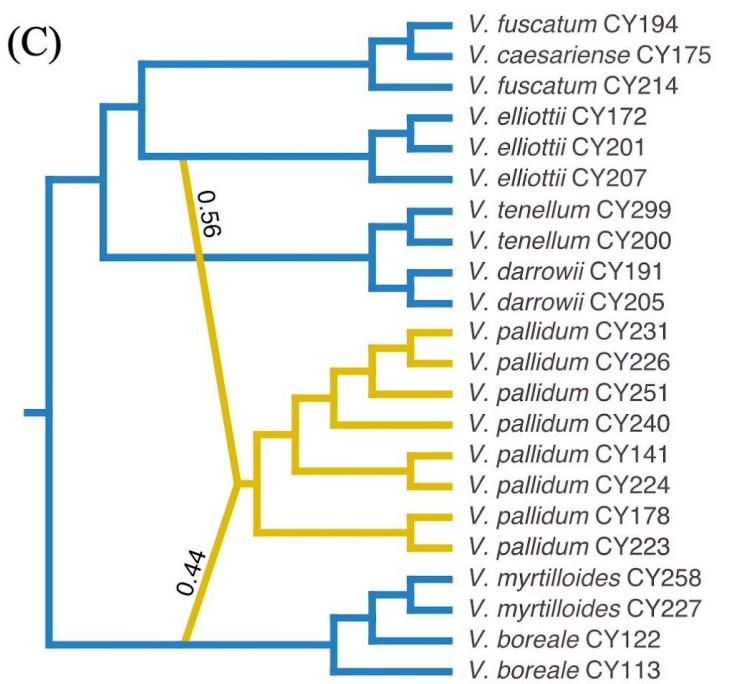
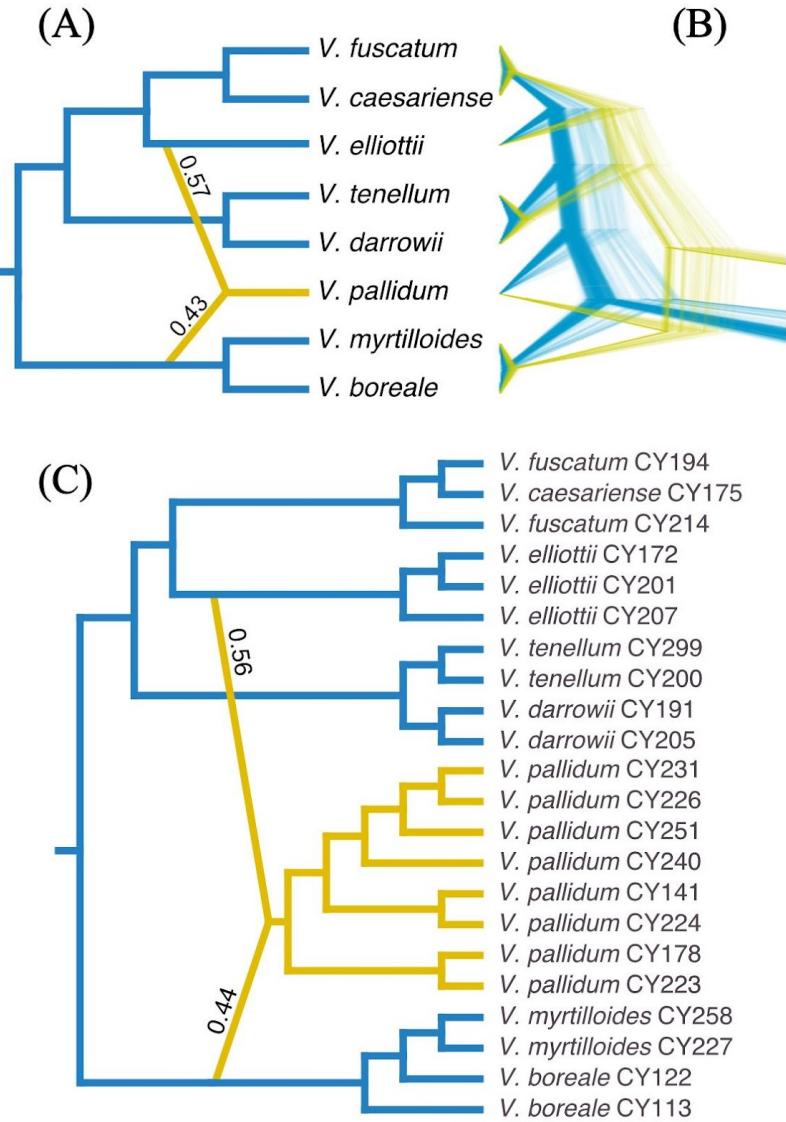
We can use networks  
to generate testable  
hypotheses.



# Plants!

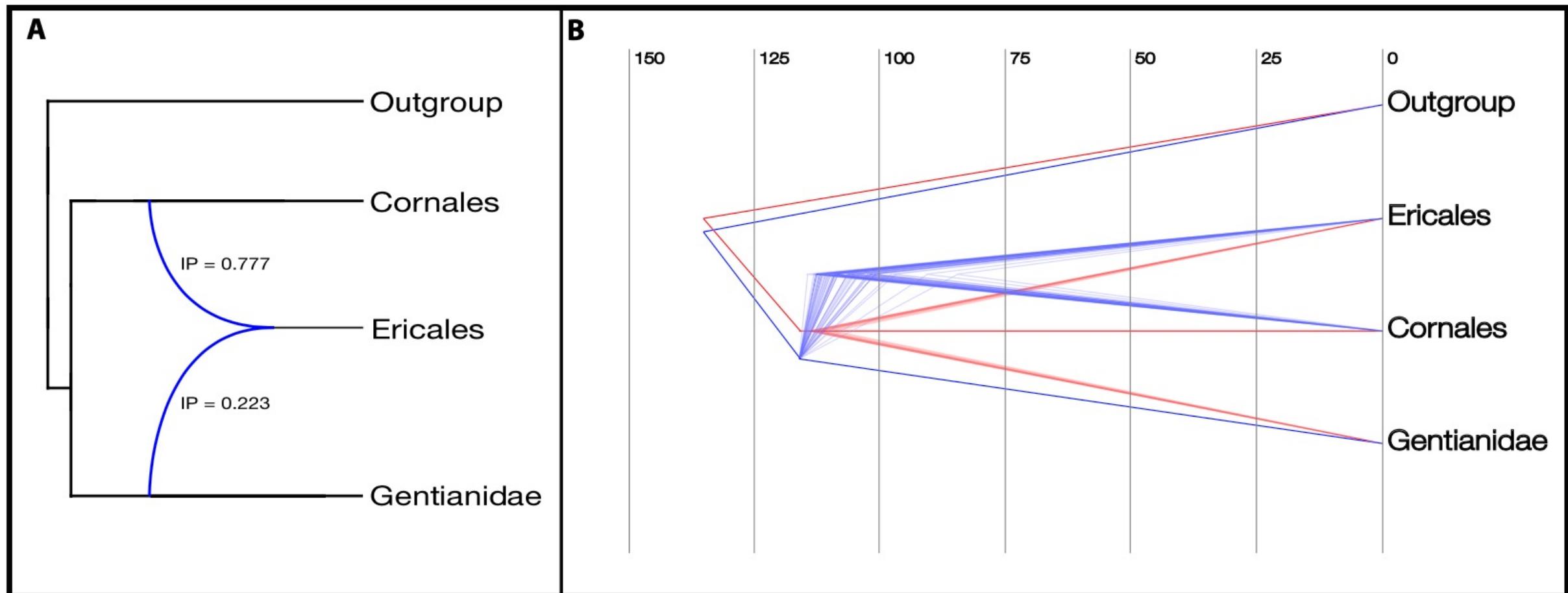
We can place networks in the context of species distributions and their ploidy levels.

Potential homoploid hybrid origins of  
*Vaccinium pallidum*



# Plants!

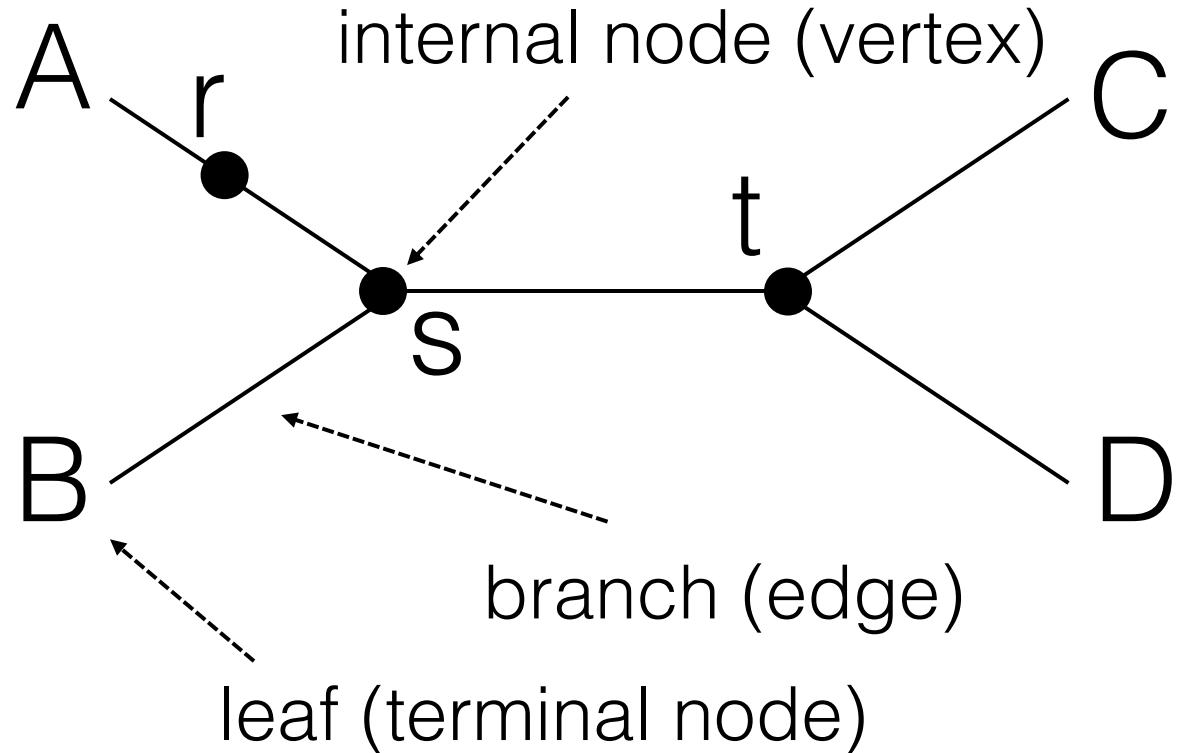
The processes of reticulate evolution can be found at all levels in the Tree of Life. It is not just restricted to closely related species.



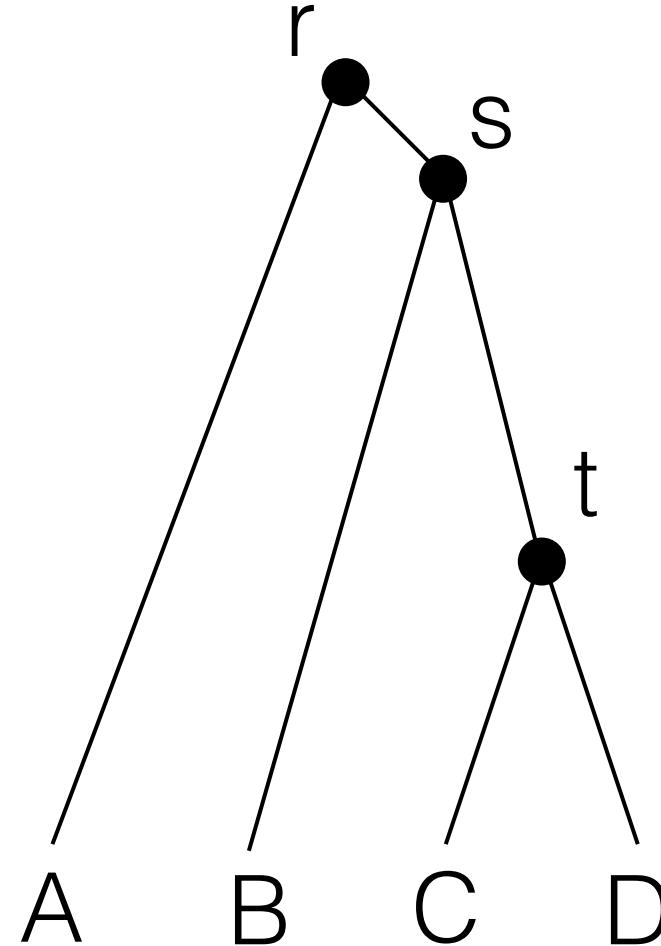
# How to write a network

---

**Unrooted Tree**



**Rooted Tree**



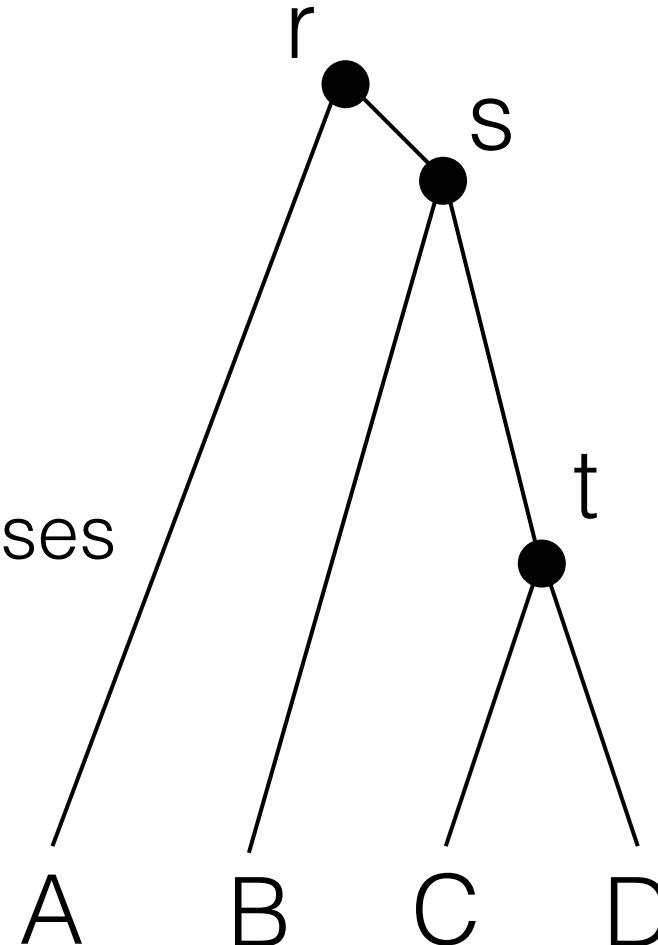
# How to write a network

---

## The Newick Format

(A,(B,(C,D)t)s)r;

## Rooted Tree



new node = open parentheses and go left

leaf = go back and go right

if you already went right = close the parentheses

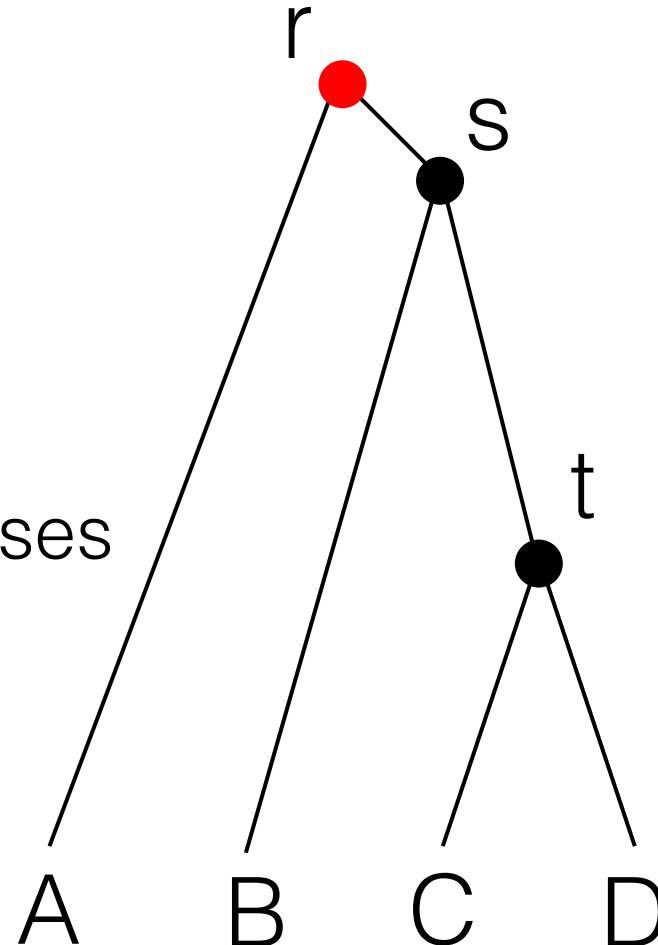
# How to write a network

---

## The Newick Format

(A,(B,(C,D)t)s)r;

## Rooted Tree



new node = open parentheses and go left

leaf = go back and go right

if you already went right = close the parentheses

# How to write a network

---

## The Newick Format

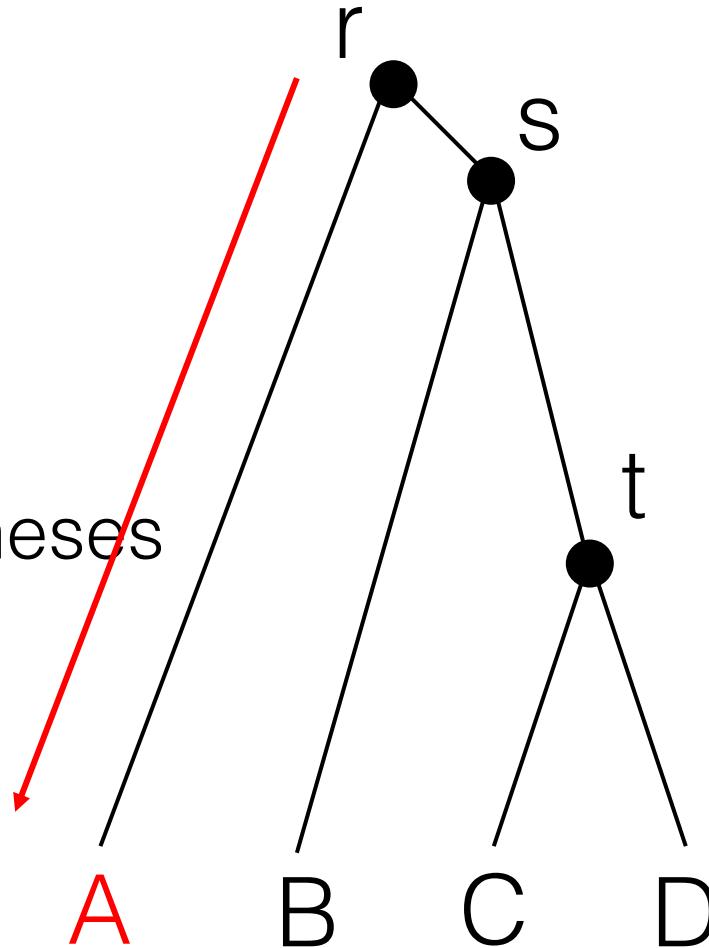
(A,(B,(C,D)t)s)r;

new node = open parentheses and go left

leaf = go back and go right

if you already went right = close the parentheses

## Rooted Tree



# How to write a network

---

## The Newick Format

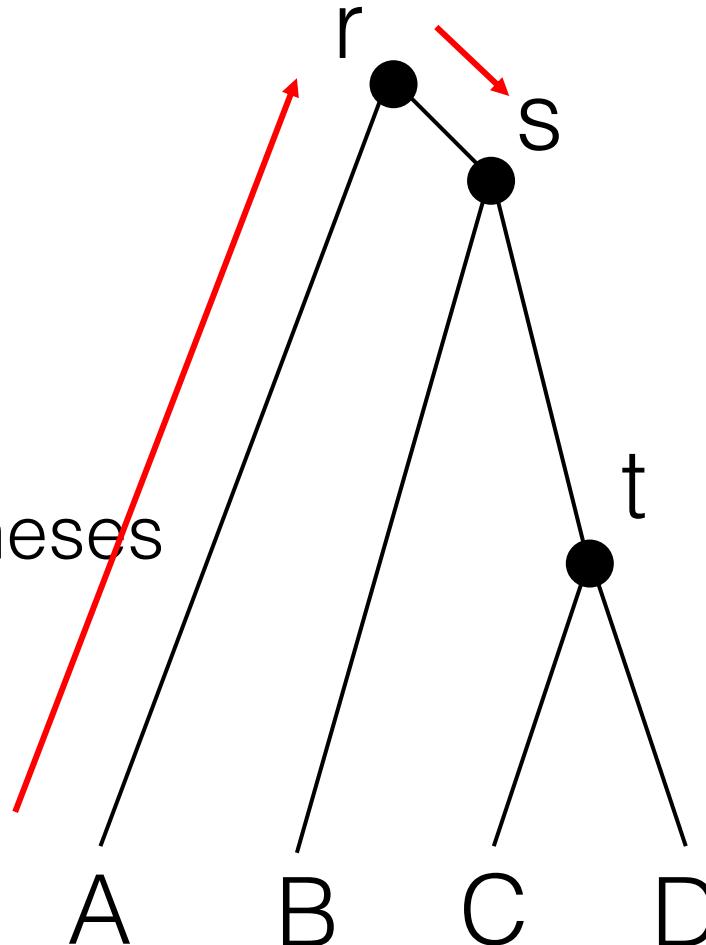
(A,(B,(C,D)t)s)r;

new node = open parentheses and go left

leaf = go back and go right

if you already went right = close the parentheses

## Rooted Tree



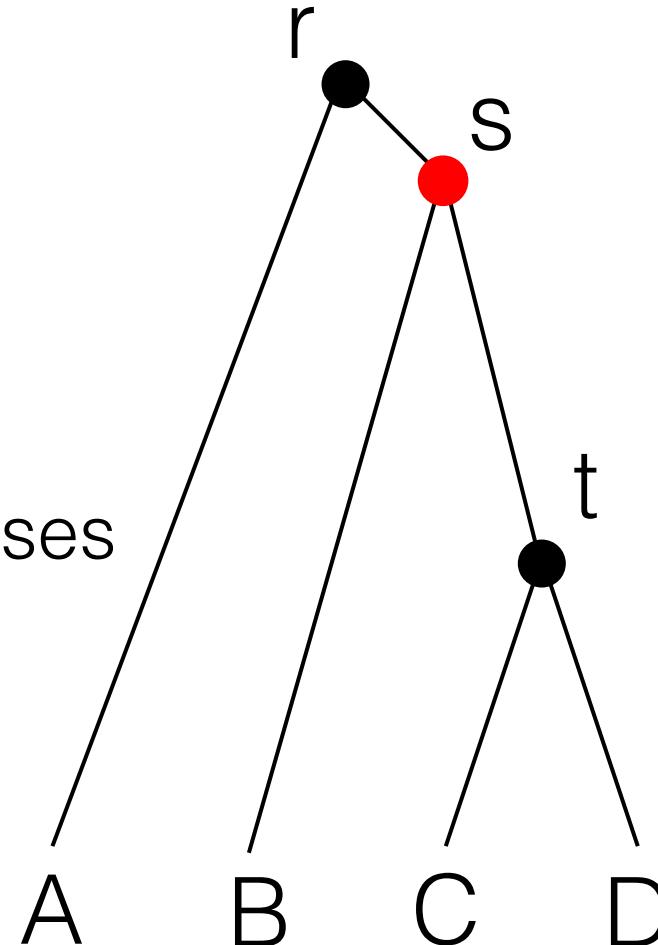
# How to write a network

---

## The Newick Format

(A,(B,(C,D)t)s)r;

## Rooted Tree



new node = open parentheses and go left

leaf = go back and go right

if you already went right = close the parentheses

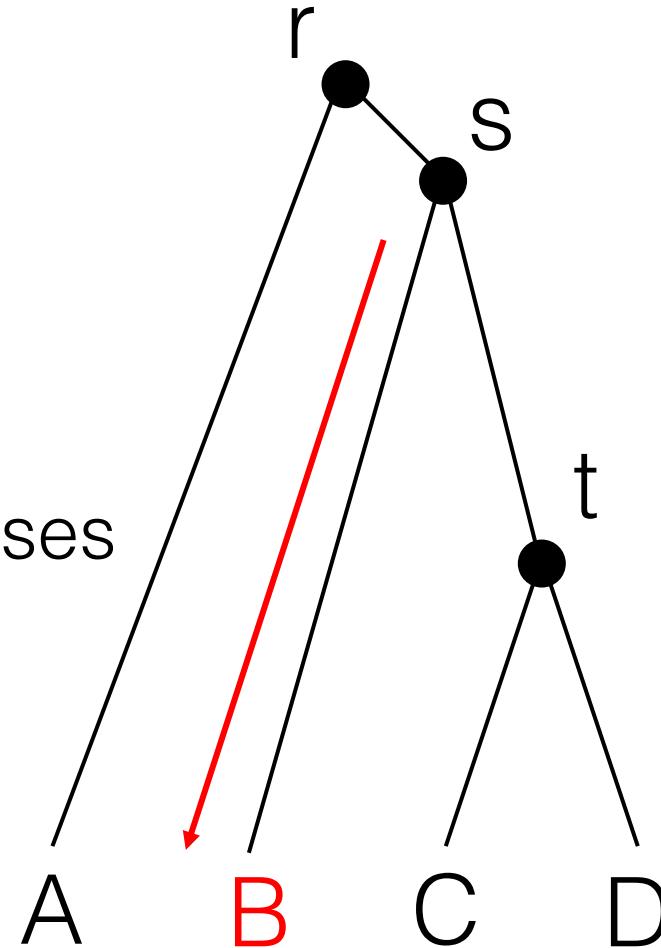
# How to write a network

---

## The Newick Format

(A,(B,(C,D)t)s)r;

## Rooted Tree



new node = open parentheses and go left

leaf = go back and go right

if you already went right = close the parentheses

# How to write a network

---

## The Newick Format

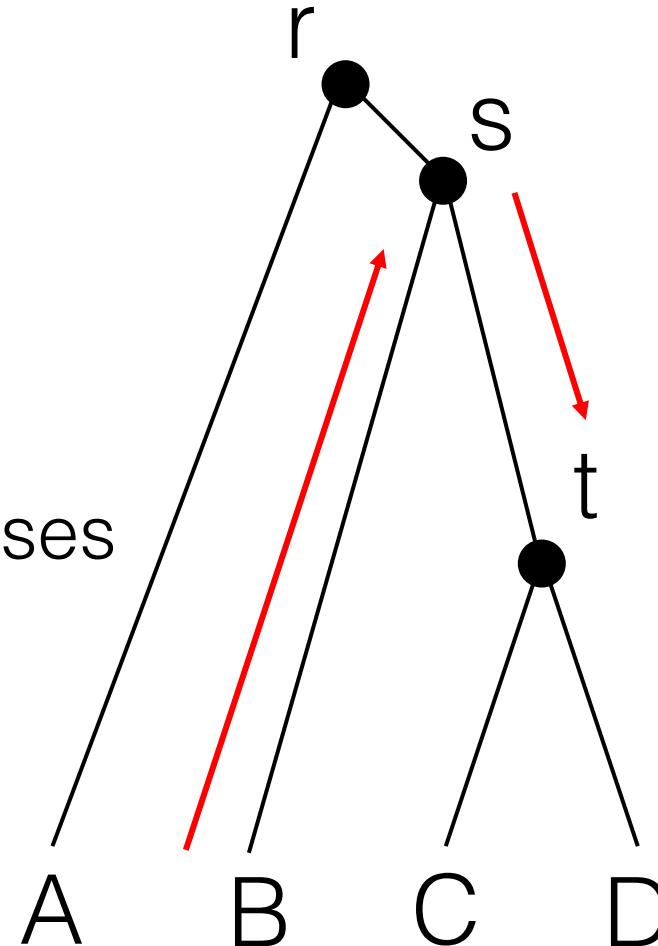
(A,(B,(C,D)t)s)r;

new node = open parentheses and go left

leaf = go back and go right

if you already went right = close the parentheses

## Rooted Tree



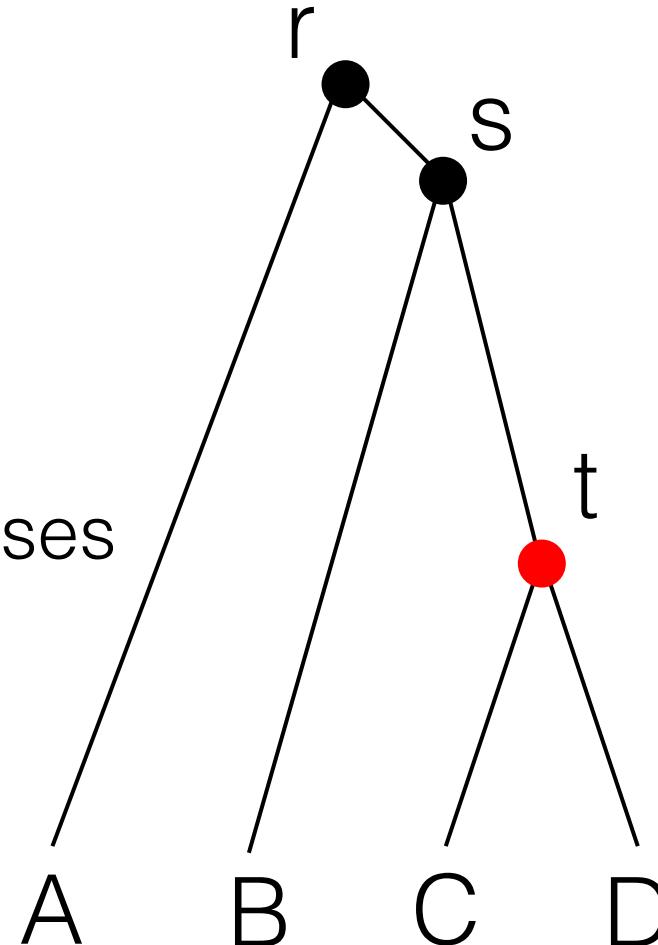
# How to write a network

---

## The Newick Format

(A,(B,(C,D)t)s)r;

## Rooted Tree



new node = open parentheses and go left

leaf = go back and go right

if you already went right = close the parentheses

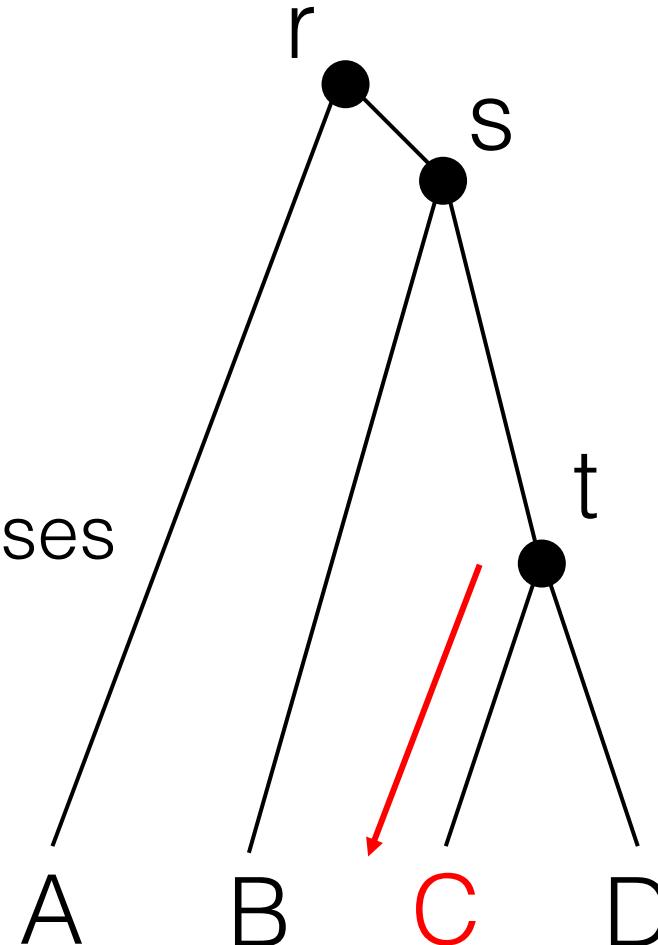
# How to write a network

---

## The Newick Format

(A,(B,(C,D)t)s)r;

## Rooted Tree



new node = open parentheses and go left

leaf = go back and go right

if you already went right = close the parentheses

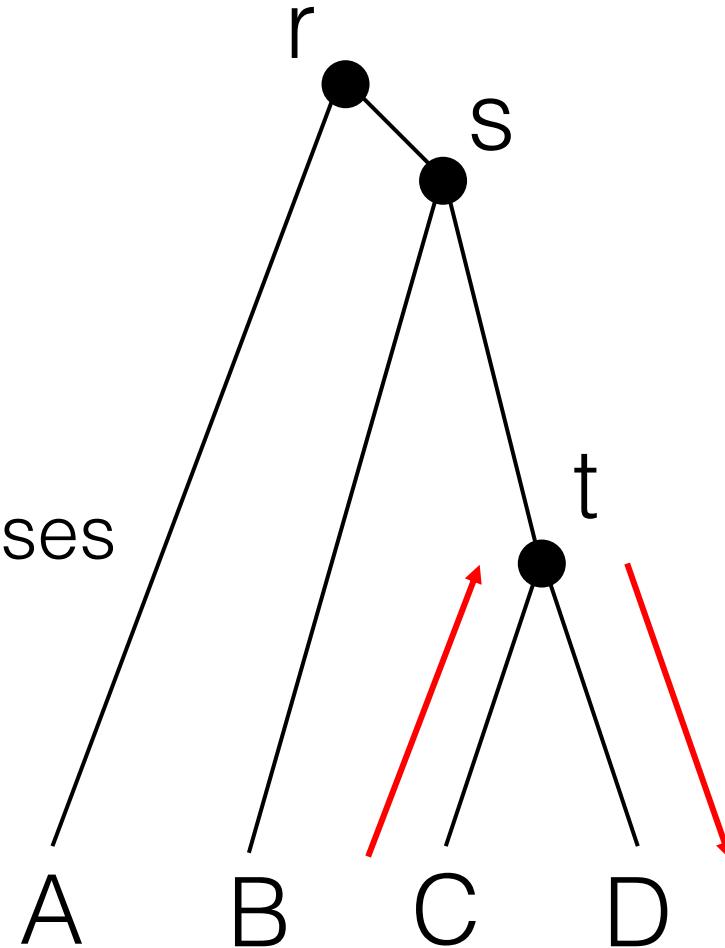
# How to write a network

---

## The Newick Format

(A,(B,(C,D)t)s)r;

## Rooted Tree



new node = open parentheses and go left

leaf = go back and go right

if you already went right = close the parentheses

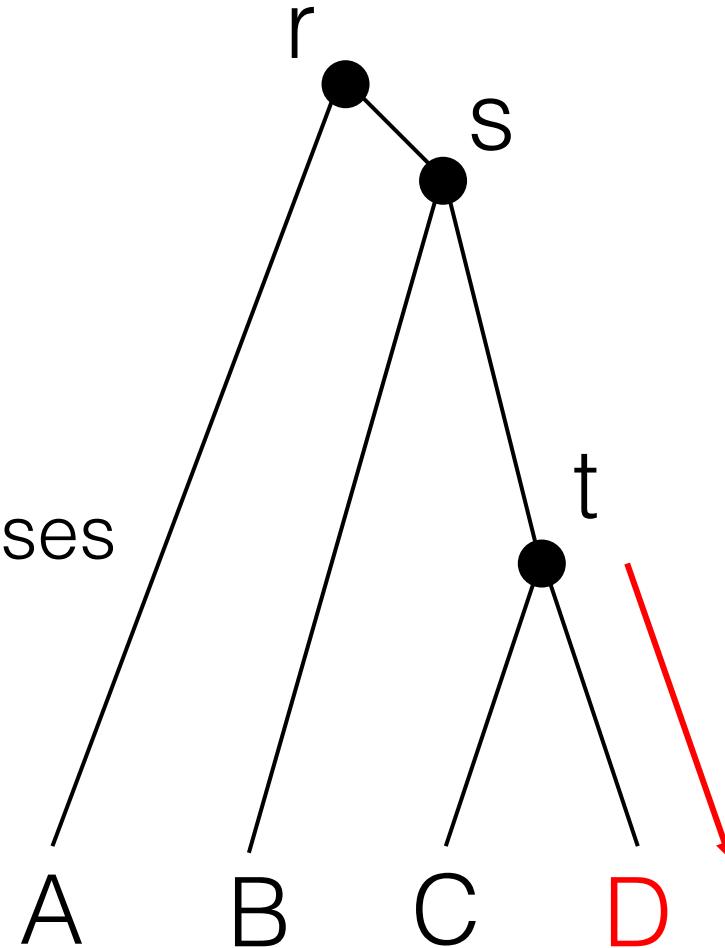
# How to write a network

---

## The Newick Format

(A,(B,(C,D)t)s)r;

## Rooted Tree



new node = open parentheses and go left

leaf = go back and go right

if you already went right = close the parentheses

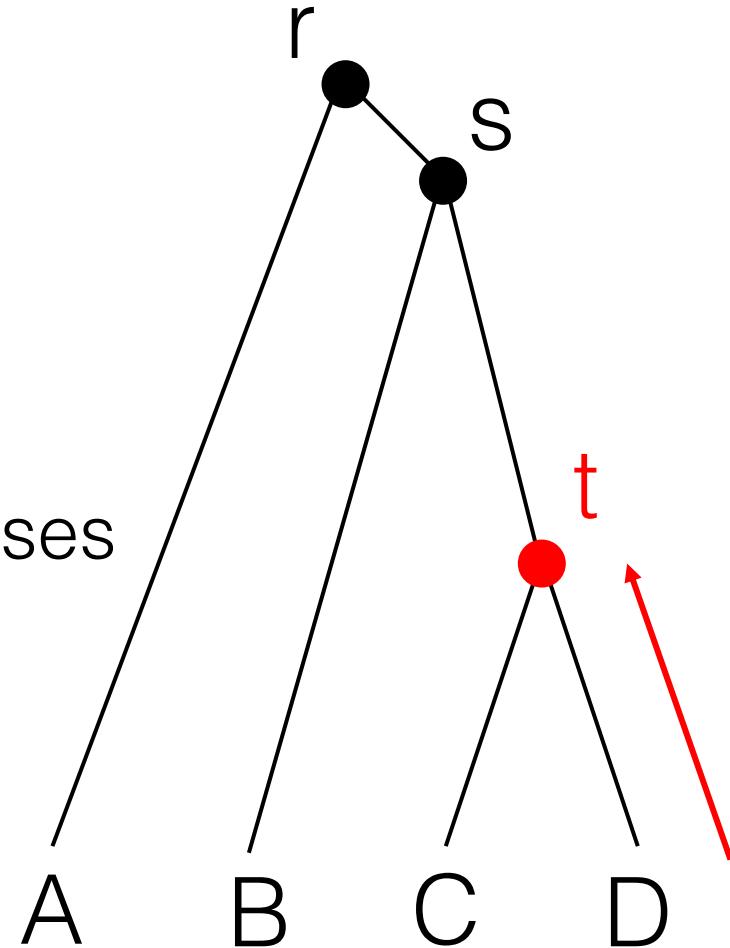
# How to write a network

---

## The Newick Format

(A,(B,(C,D)t)s)r;

## Rooted Tree



new node = open parentheses and go left

leaf = go back and go right

if you already went right = close the parentheses

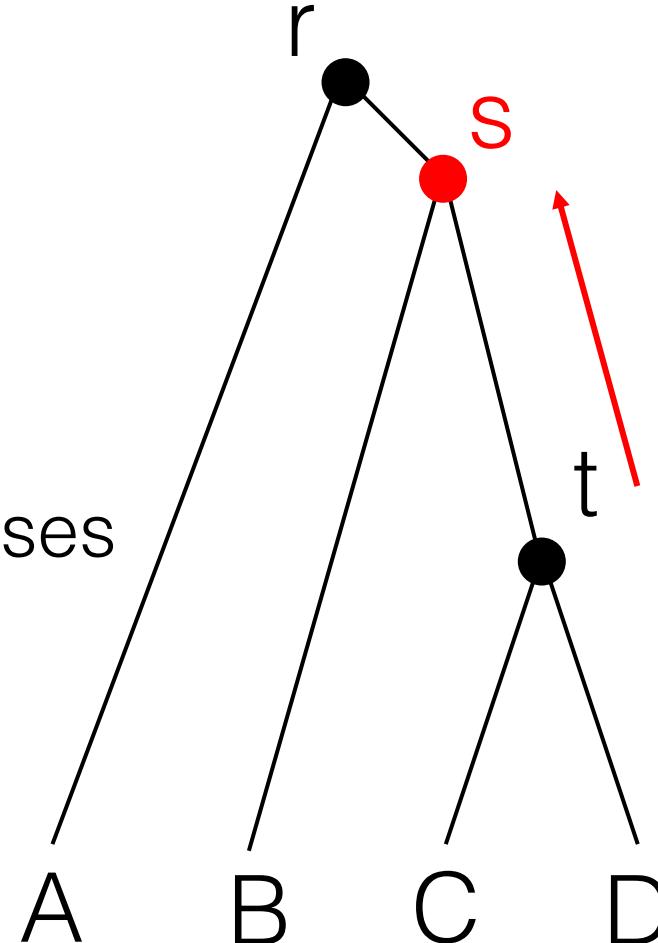
# How to write a network

---

## The Newick Format

(A,(B,(C,D)t)s)r;

## Rooted Tree



new node = open parentheses and go left

leaf = go back and go right

if you already went right = close the parentheses

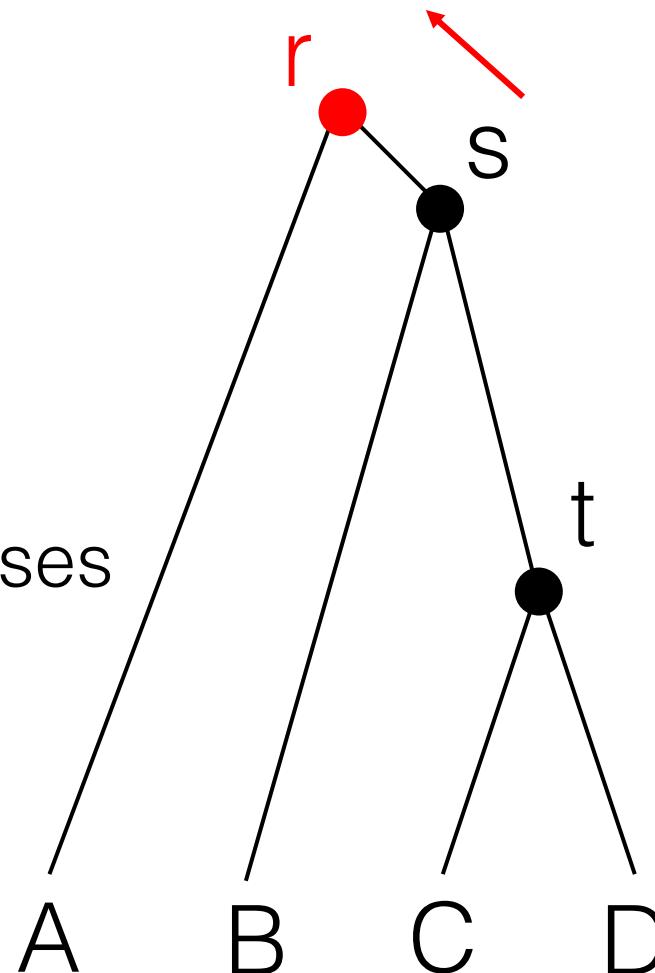
# How to write a network

---

## The Newick Format

(A,(B,(C,D)t)s)r;

## Rooted Tree



new node = open parentheses and go left

leaf = go back and go right

if you already went right = close the parentheses

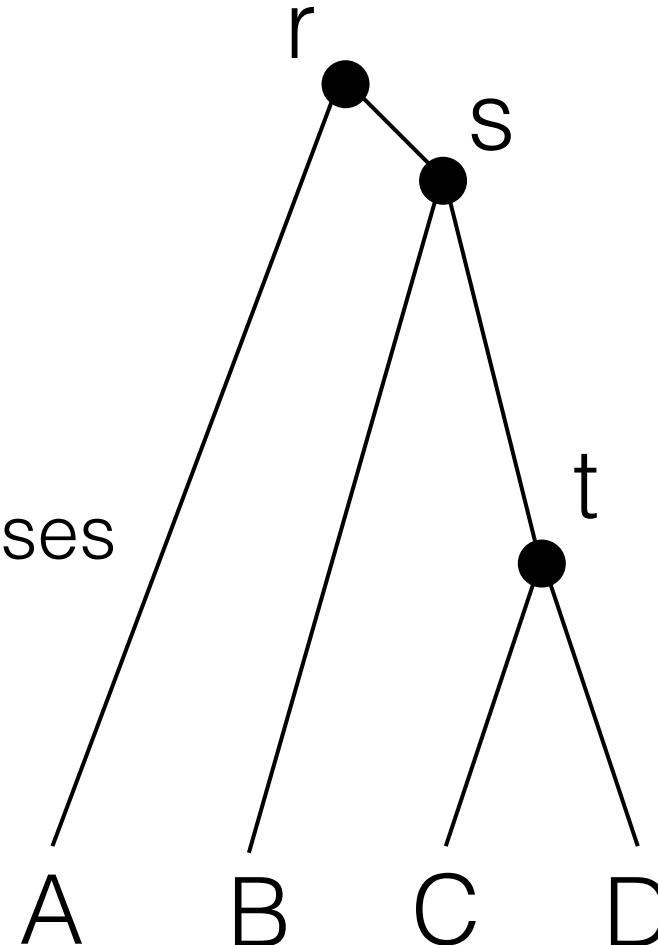
# How to write a network

---

## The Newick Format

(A,(B,(C,D)t)s)r;

## Rooted Tree



new node = open parentheses and go left

leaf = go back and go right

if you already went right = close the parentheses

end with a semicolon (;)

# How to write a network

---

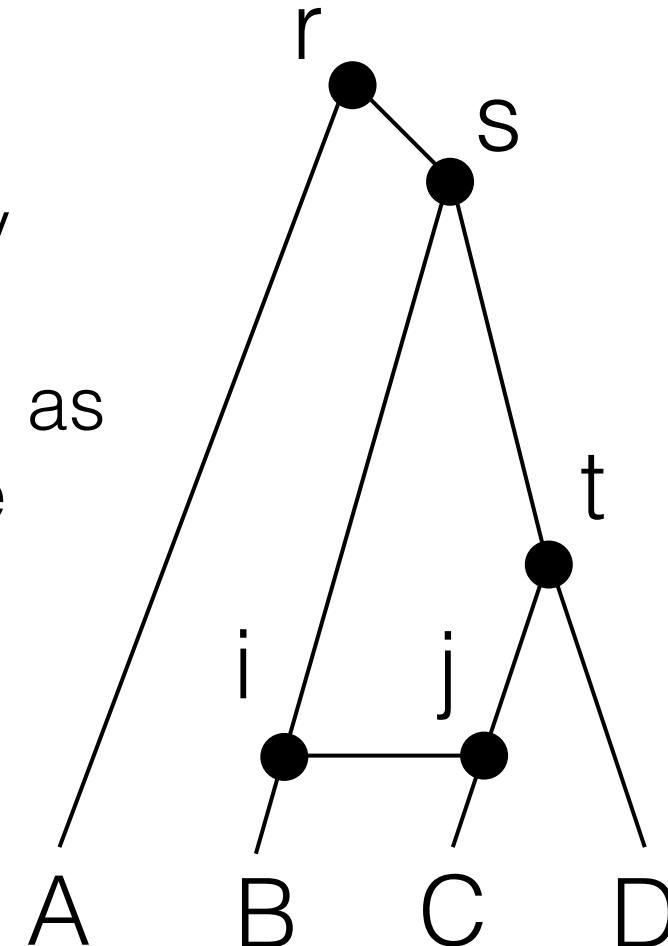
**A Challenge! Can you write the newick string for the network?**

We will work the answer on the board shortly

Add your answer to the google sheet as well as the organism or group that brought you here

Hint: Leaves can only appear once. Once a node is closed, we do not visit it again.

**Rooted Network**



# How to write a network

---

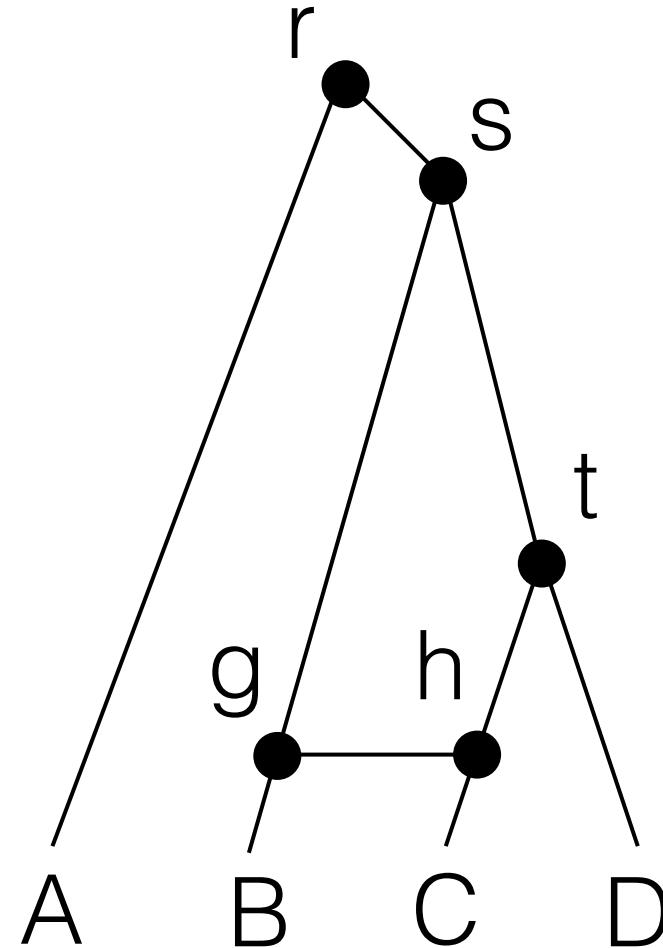
**A Challenge! Can you write the newick string for the network?**

(A,((B,h)g,((C)h,D)t)s)r;

(A,((B)g,((g,C)h,D)t)s)r;

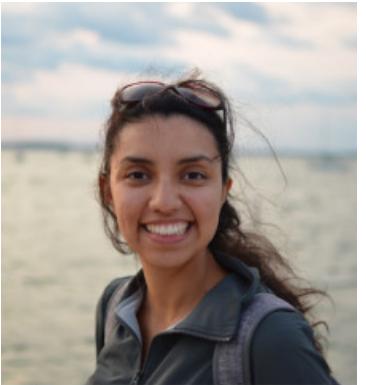
(A,((B,(C)h)g,(h,D)t)s)r;

**Rooted Network**



# Today's instructors

---



Claudia Solís-Lemus  
Assistant Professor  
Dept of Plant Pathology  
UW-Madison



George Tiley  
Marie Curie Fellow  
RBG Kew



Huw Ogilvie  
Assistant Research Professor  
Dept of Comp Sci  
Rice University



Kevin Liu  
Associate Professor  
Dept of Comp Sci & Engineering  
MSU



Mark Kessler  
Software Engineer  
Dept Comp Sci  
Rice University



Kevin Kong  
Postdoctoral Research Associate  
Wisconsin Institute for Discovery  
UW-Madison



Lauren Frankel  
PhD Candidate  
Dept of Botany  
UW-Madison