

Phylogenetic networks to study **reticulate** evolution

Claudia Solís-Lemus, PhD

University of Wisconsin-Madison
Wisconsin Institute for Discovery
Department of Plant Pathology



May 31, 2025



<https://solislemuslab.github.io/>



mstdn.social/@solislemuslab



crsl4



@thestatistician

I want to find signal
of hybridization in
my data

I want to find signal of hybridization in my data

Hybrid detection
methods

HyDe

ABBA-BABA

MSCQuartets

Network inference
methods

PhyloNet
SNaQ
BEAST2
NANUQ
RF-Net
PhyNEST

I want to find signal of hybridization in my data

Hybrid detection
methods

HyDe

ABBA-BABA

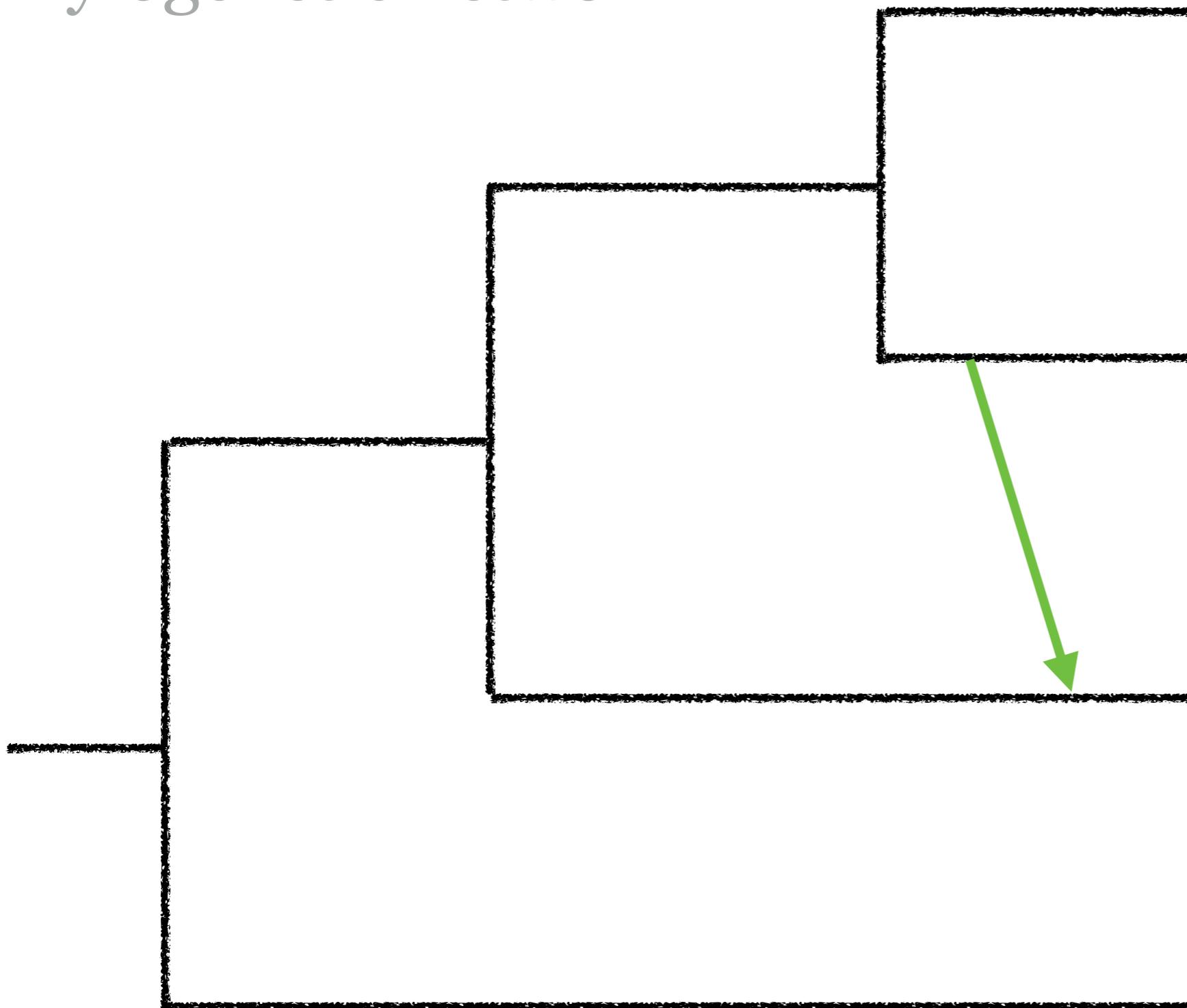
MSCQuartets

Network inference
methods

PhyloNet
SNaQ
BEAST2
NANUQ
RF-Net
PhyNEST

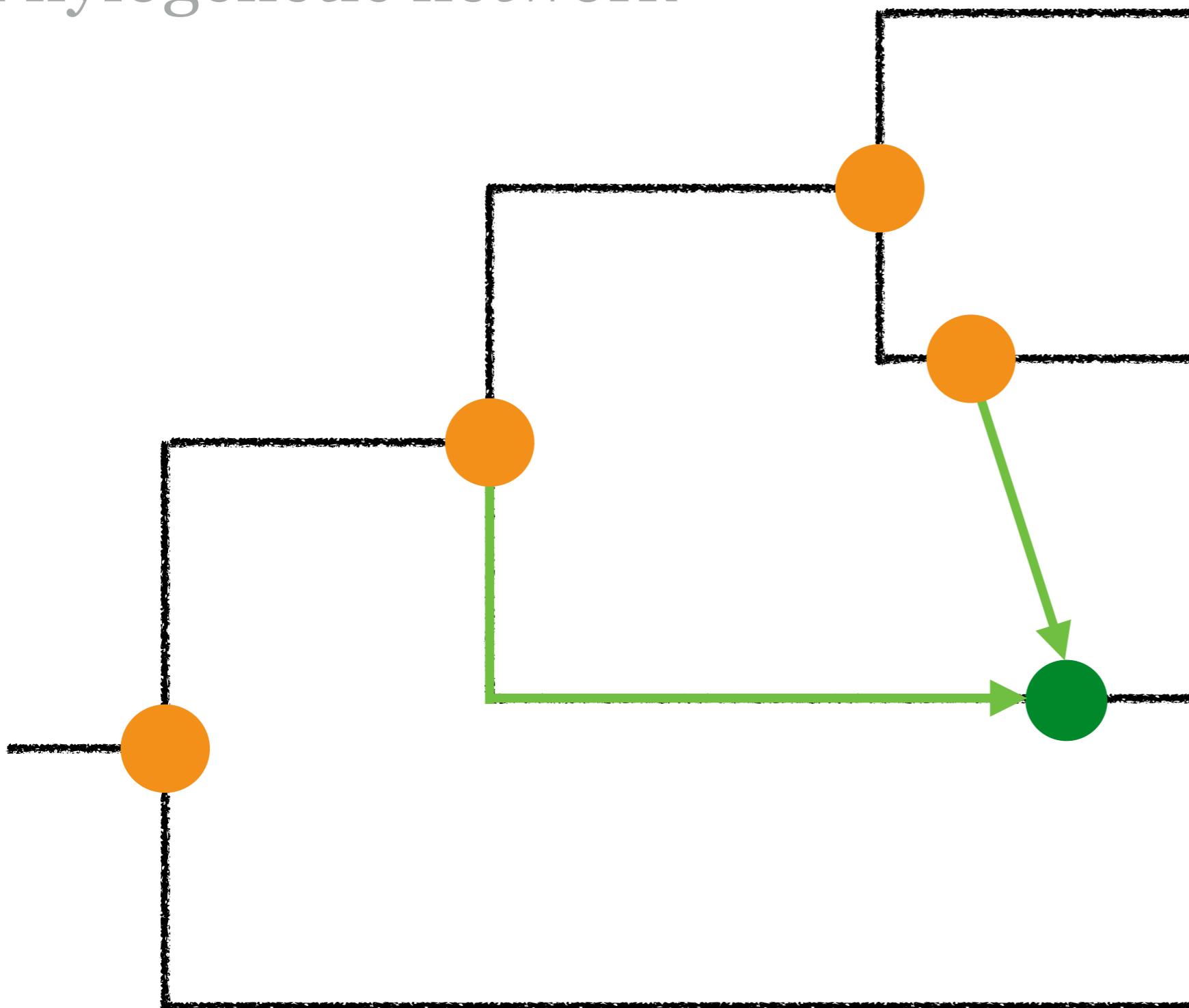
Network thinking

Phylogenetic network



Network thinking

Phylogenetic network

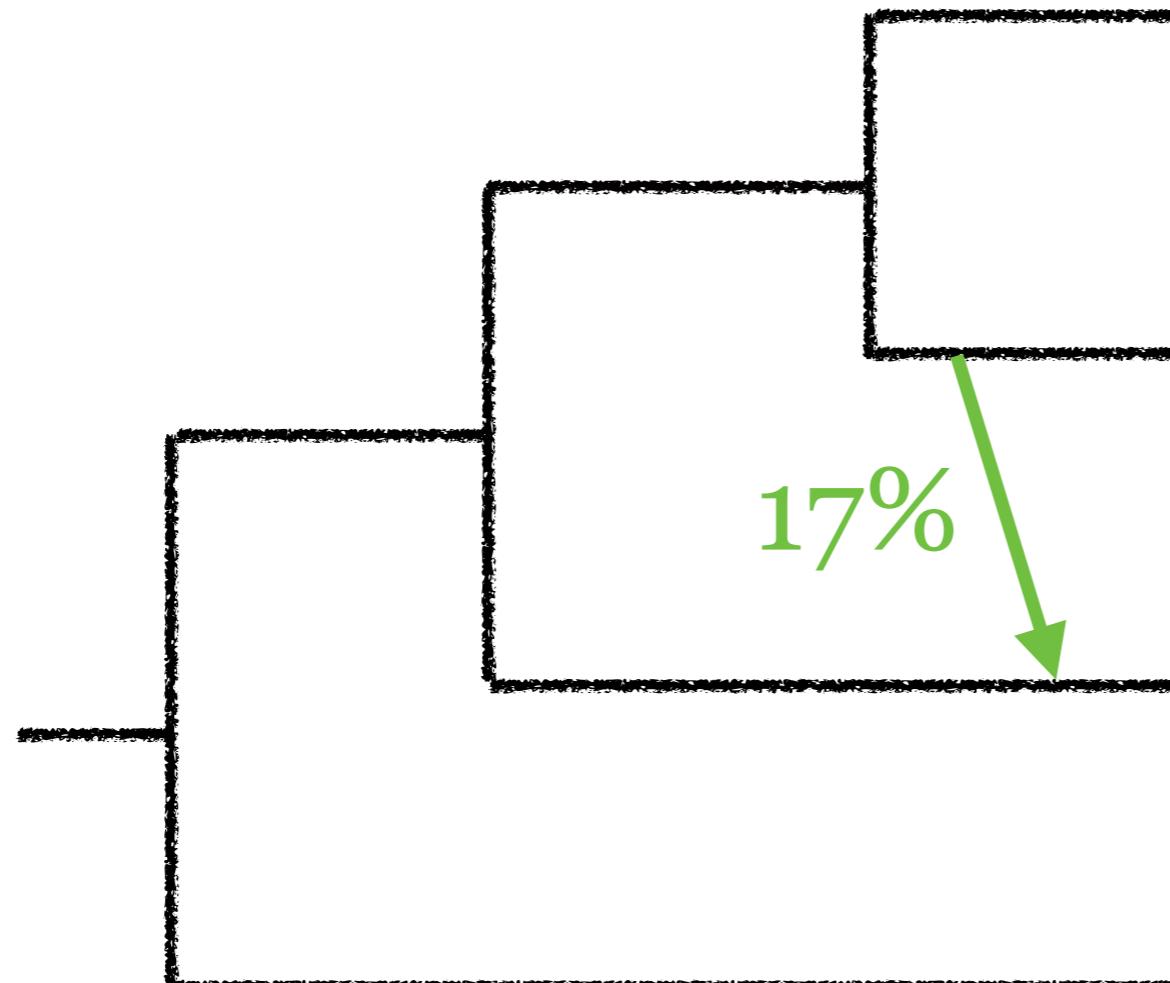


speciation
hybridization



Network thinking

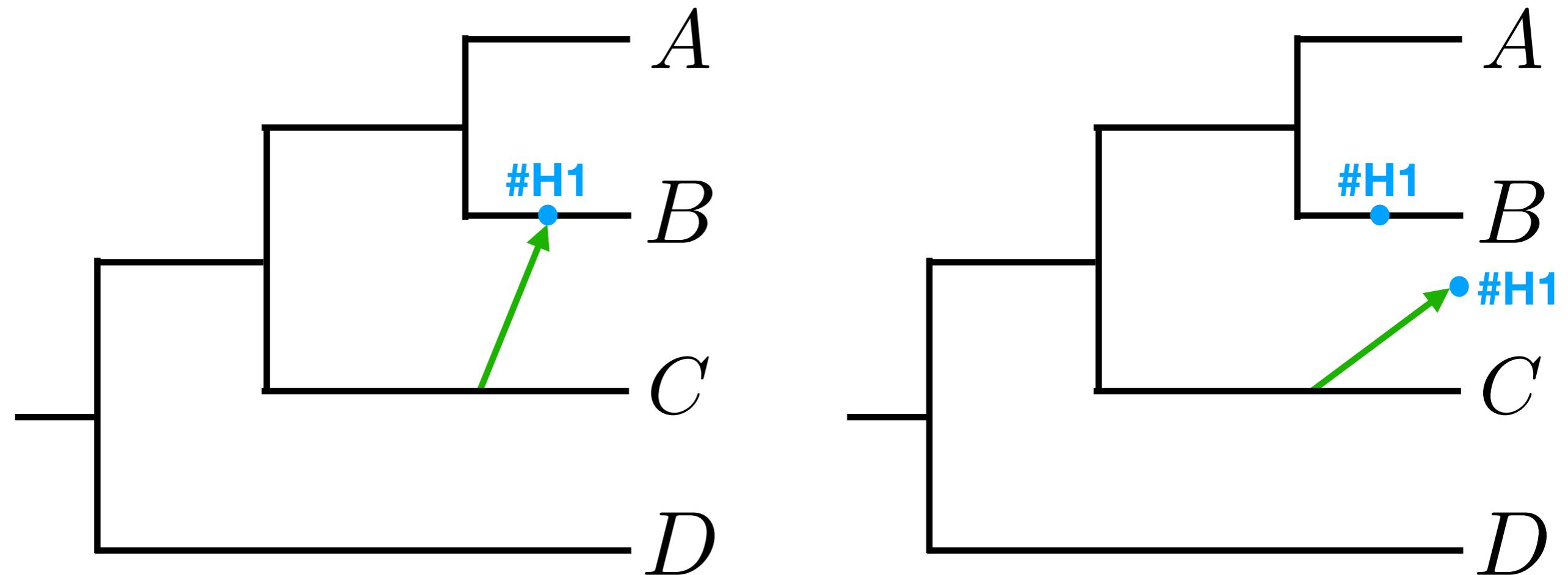
Phylogenetic network



inheritance probability

Network thinking

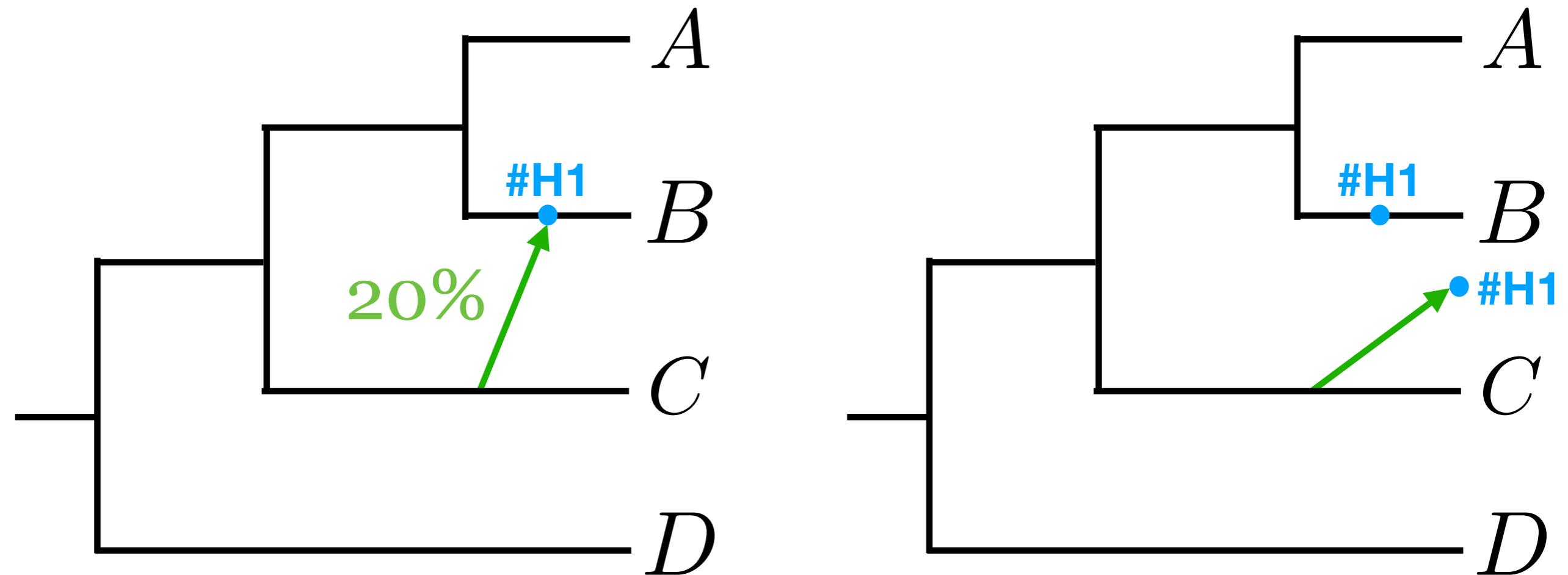
Phylogenetic network: extended Newick format



$((((A,(B)\#H1),(C,\#H1)),D);$

Network thinking

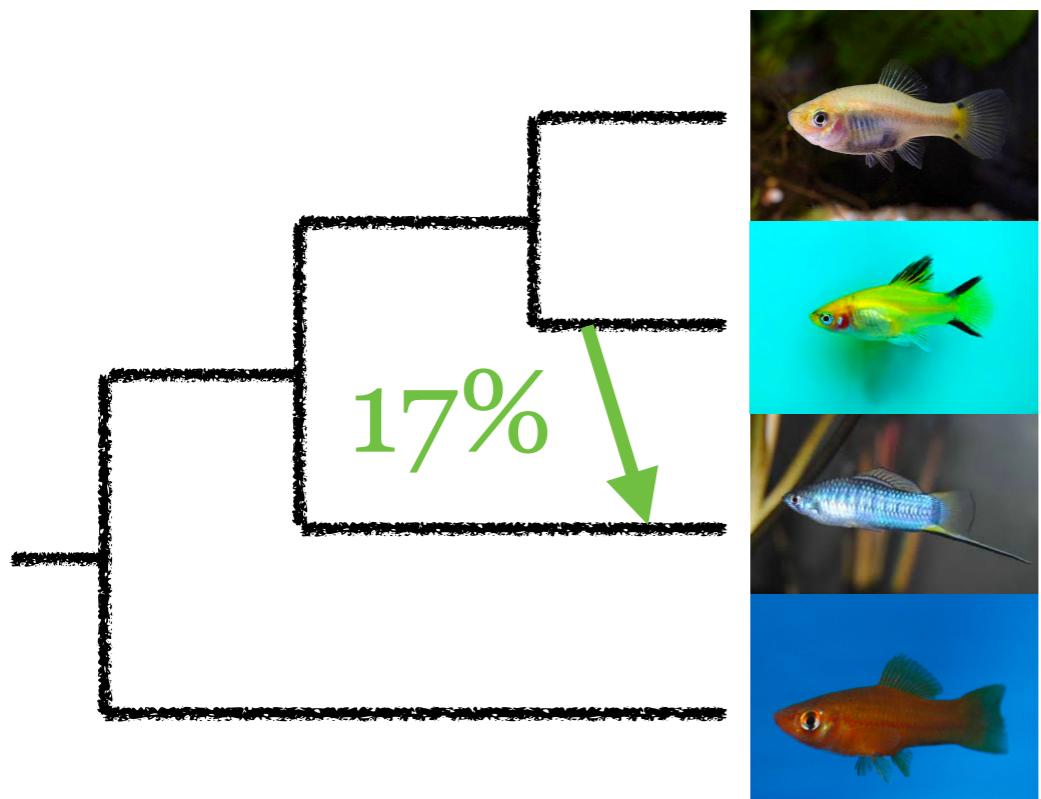
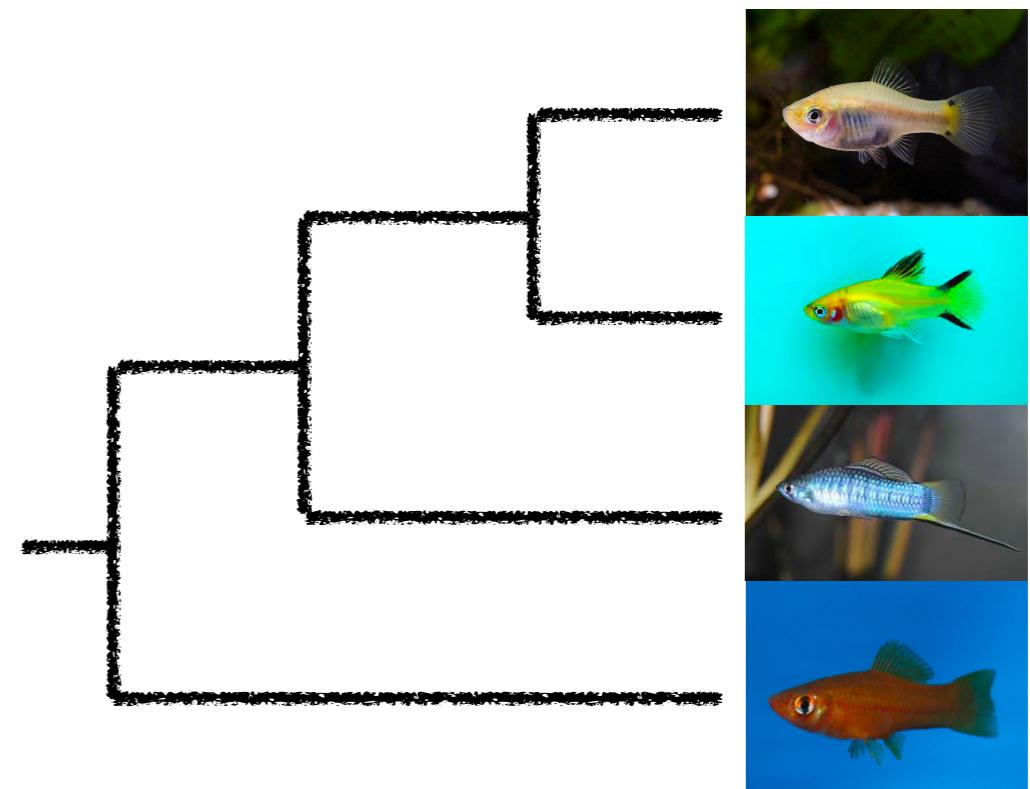
Phylogenetic network: extended Newick format



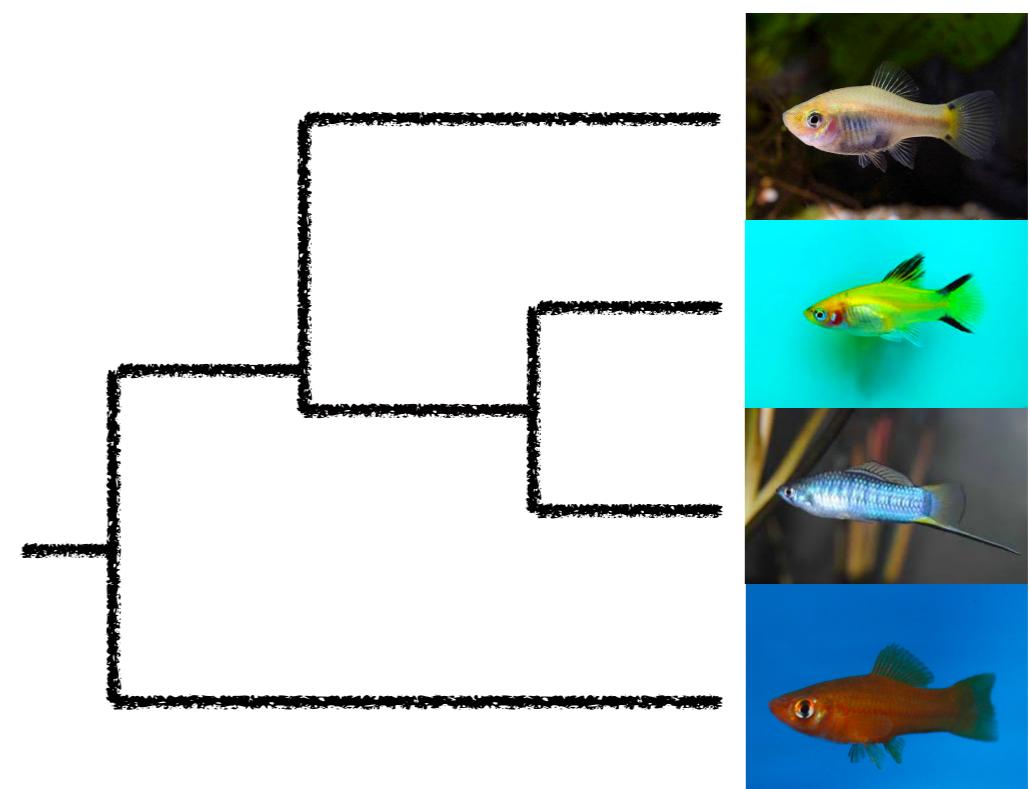
```
((((A,(B)#H1::::0.8),(C,#H1::::0.2)),D);
```

Network thinking

Phylogenetic network



major tree



minor tree

Network thinking

Phylogenetic network

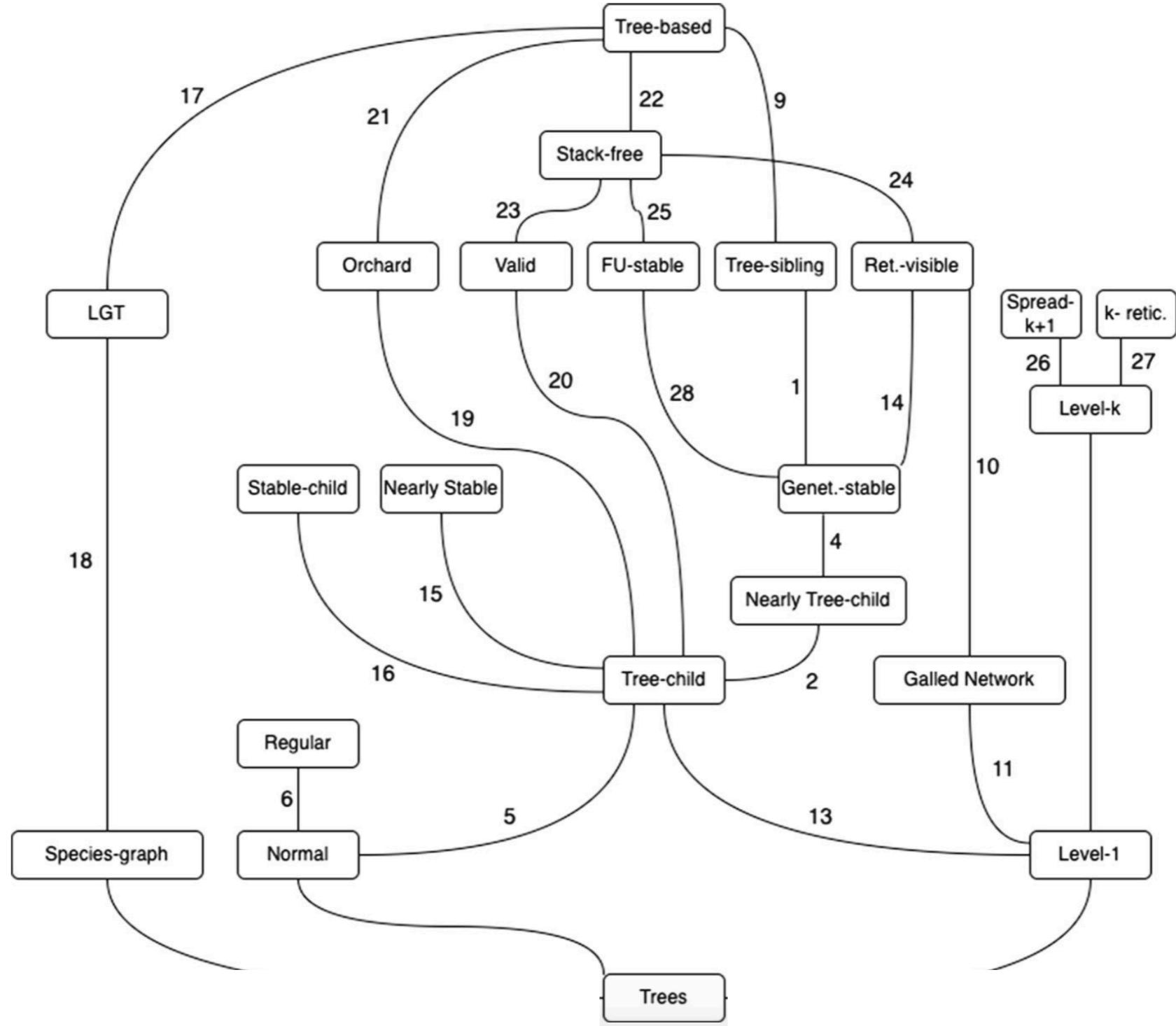


Classes of explicit phylogenetic networks and their biological and mathematical significance

Sungsik Kong¹ · Joan Carles Pons² · Laura Kubatko^{1,3} · Kristina Wicke⁴

Received: 21 September 2021 / Revised: 18 January 2022 / Accepted: 31 March 2022 /

Sungsik (Kevin)
Kong



<https://solislemuslab.github.io/>



<mstdn.social/@solislemuslab>



[crsl4](#)



@thestatistician

Network thinking

Phylogenetic network

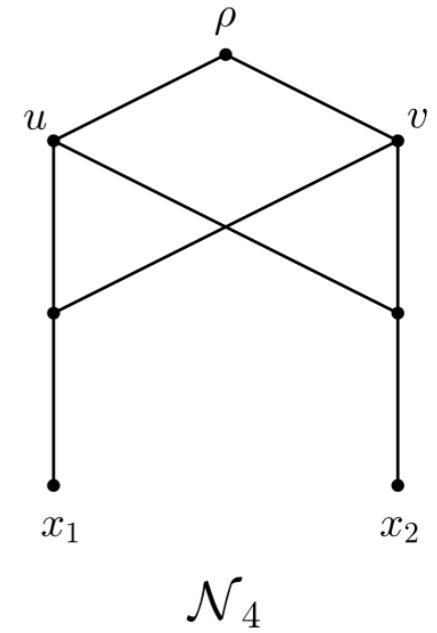
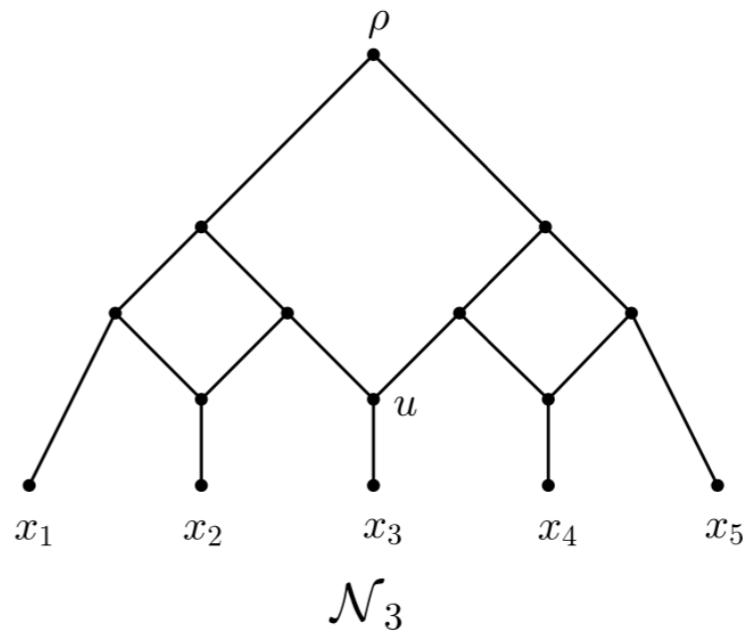
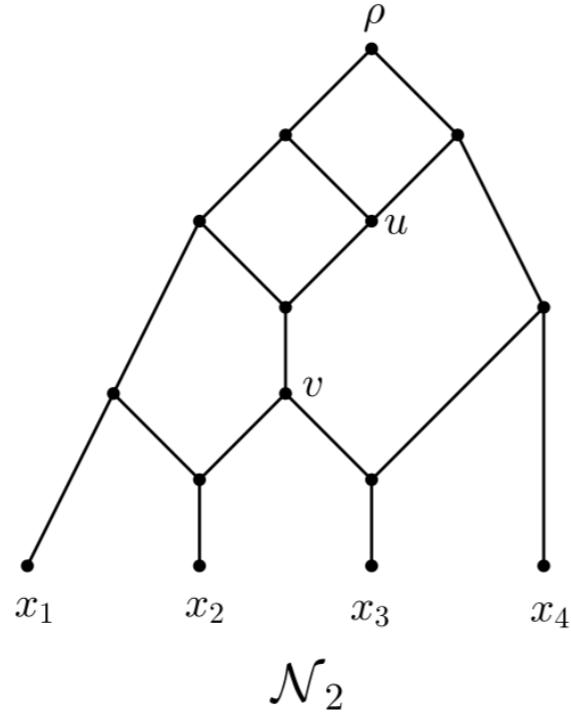
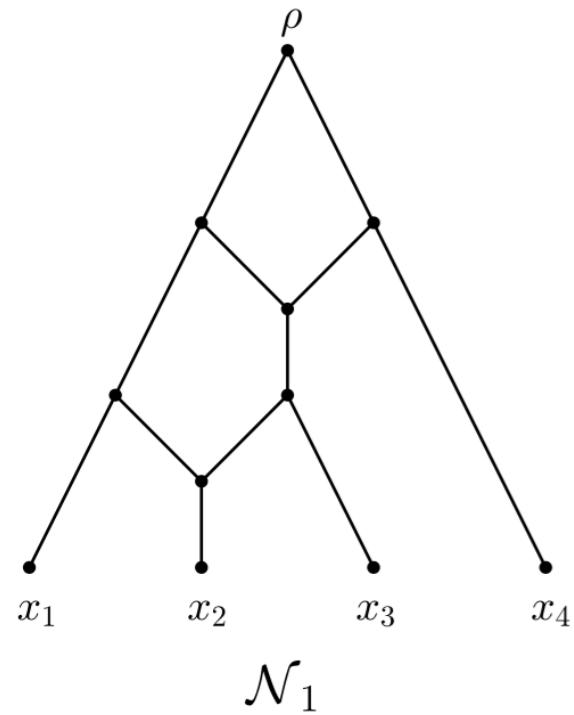


Classes of explicit phylogenetic networks and their biological and mathematical significance

Sungsik Kong¹ · Joan Carles Pons² · Laura Kubatko^{1,3} · Kristina Wicke⁴

Received: 21 September 2021 / Revised: 18 January 2022 / Accepted: 31 March 2022 /

Sungsik (Kevin)
Kong



tree child

not tree child

not tree child

not tree child

tree sibling

tree sibling

not tree sibling

not tree sibling

stack free

not stack free

stack free

stack free

FU stable

not FU stable

FU stable

not FU stable



<https://solislemuslab.github.io/>



mstdn.social/@solislemuslab



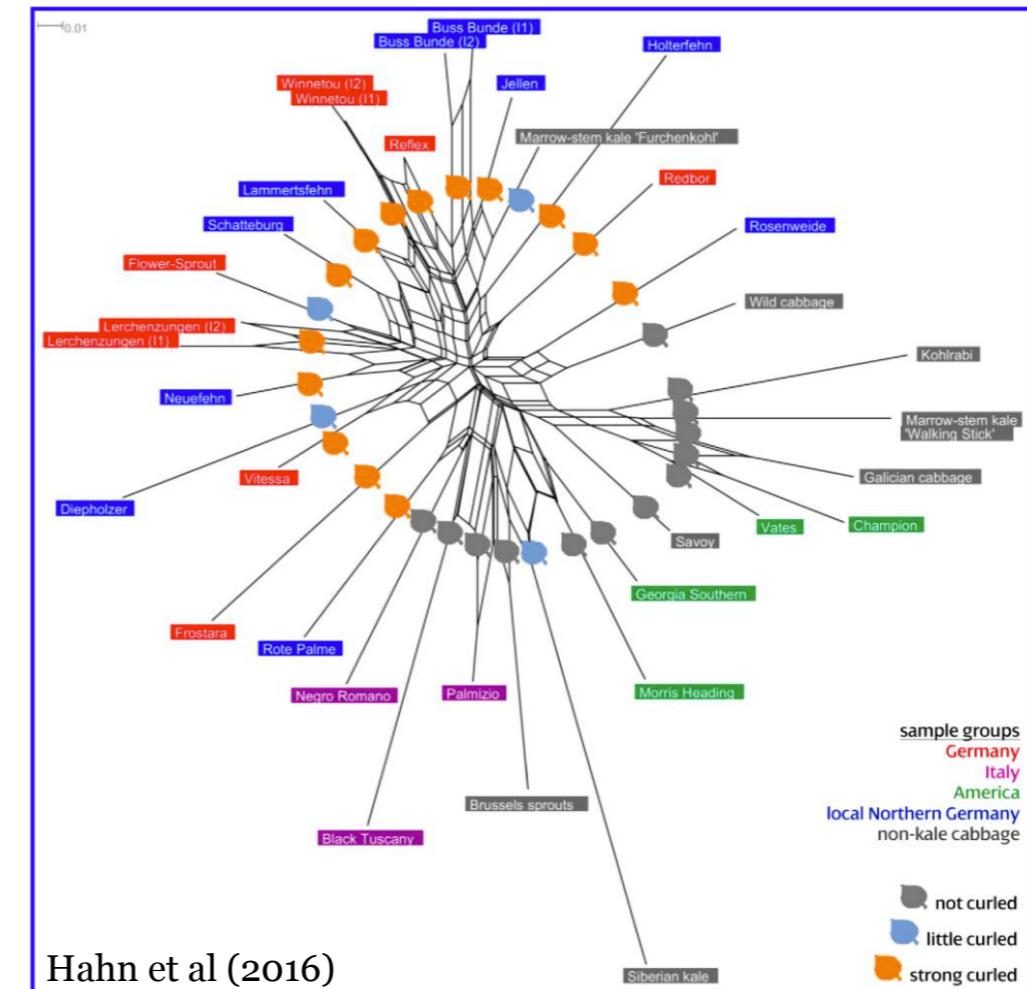
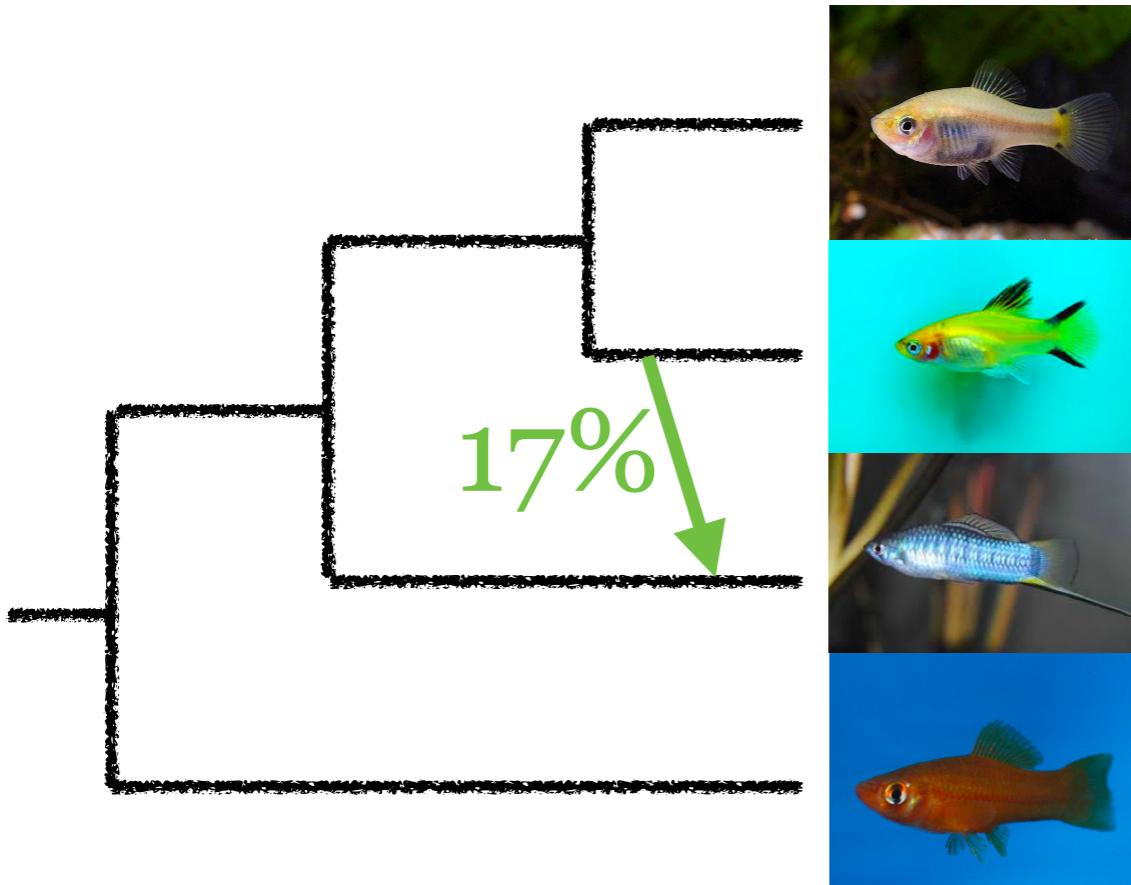
[crsl4](#)



@thestatistician

Network thinking

Types of phylogenetic networks

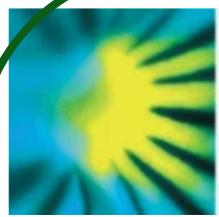


Explicit

Implicit

Why?

Phylogenetic network

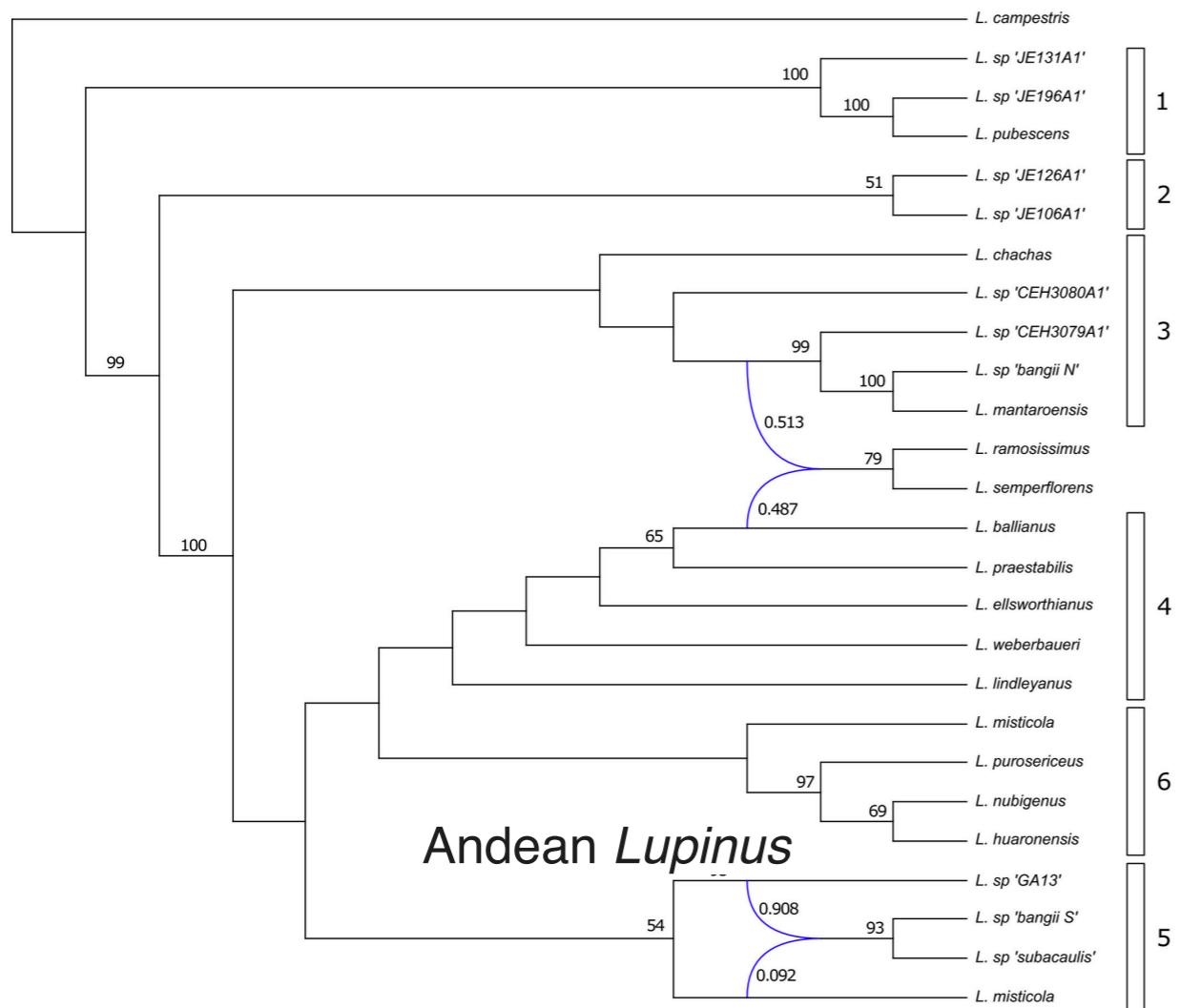


New Phytologist

Full Paper | Free Access

Pleistocene glacial cycles drive isolation, gene flow and speciation in the high-elevation Andes

Bruno Nevado , Natalia Contreras-Ortiz, Colin Hughes, Dmitry A. Filatov



Current Biology

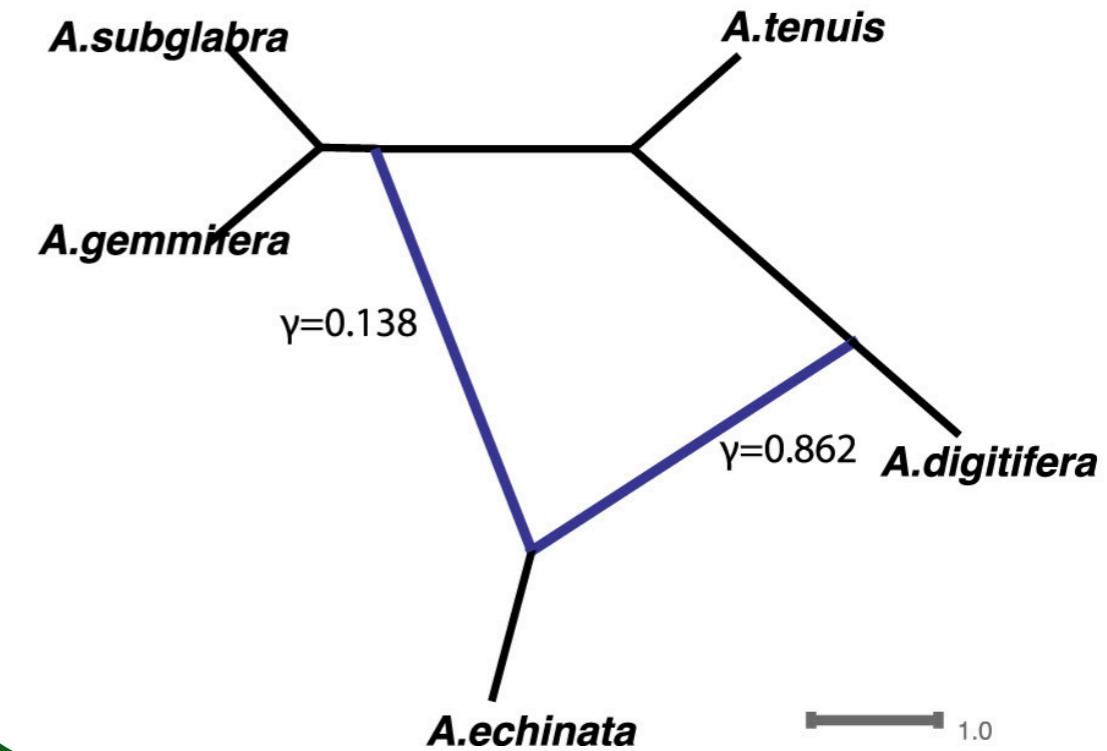
Volume 28, Issue 21, 5 November 2018, Pages 3373-3382.e5

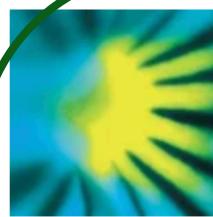


Article

The Roles of Introgression and Climate Change in the Rise to Dominance of *Acropora* Corals

Yafei Mao ^{1, 2, 3} , Evan P. Economo ² , Noriyuki Satoh ¹



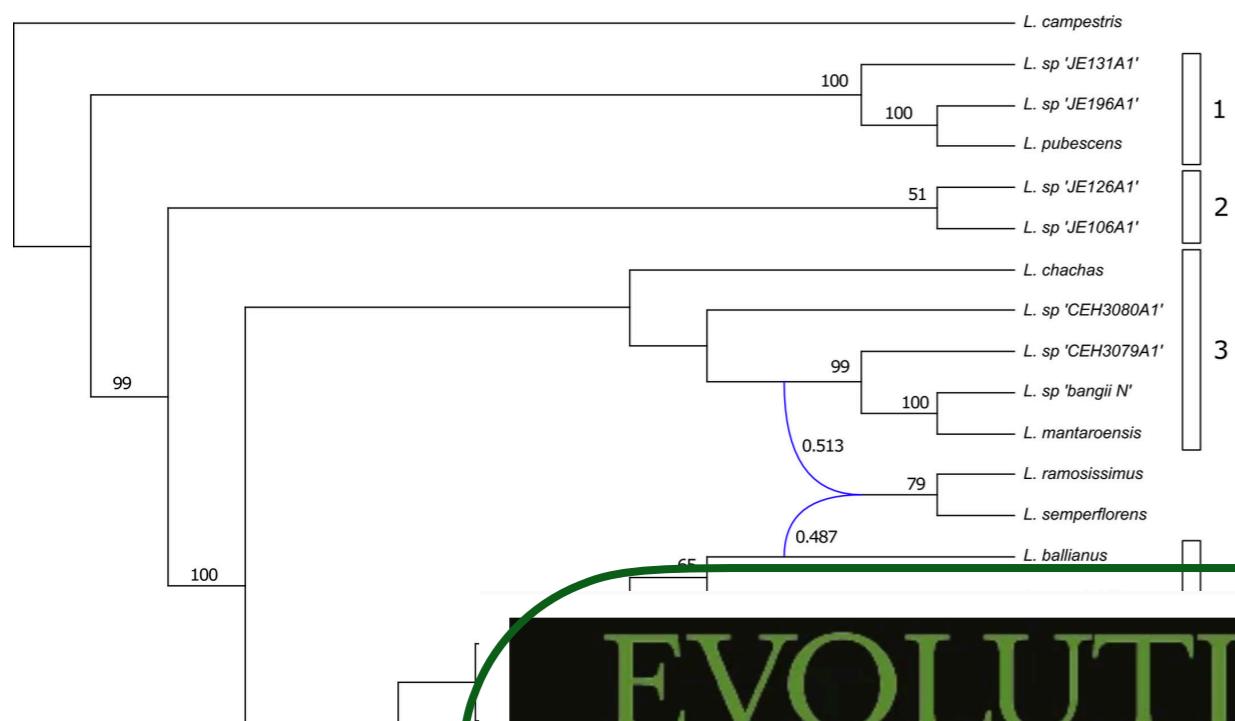


New Phytologist

Full Paper | Free Access

Pleistocene glacial cycles drive isolation, gene flow and speciation in the high-elevation Andes

Bruno Nevado , Natalia Contreras-Ortiz, Colin Hughes, Dmitry A. Filatov



EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

ORIGINAL ARTICLE | Open Access |

Embracing discordance: Phylogenomic analyses provide evidence for allopolyploidy leading to cryptic diversity in a Mediterranean *Campanula* (Campanulaceae) clade

Andrew A. Crowl , Cody Myers, Nico Cellinese

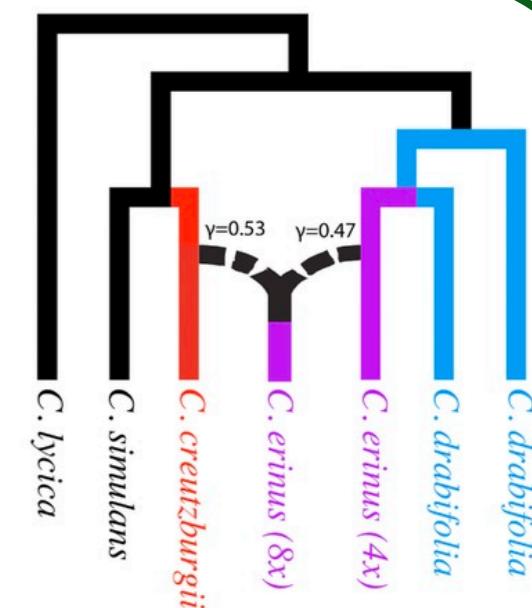
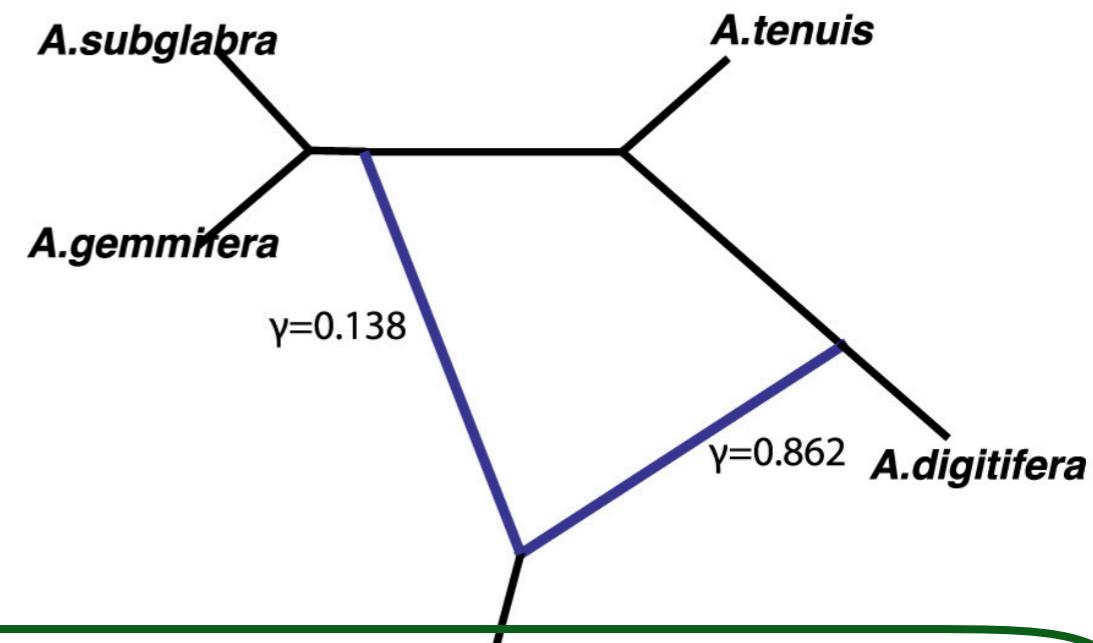
Current Biology

Volume 28, Issue 21, 5 November 2018, Pages 3373-3382.e5

Article

The Roles of Introgression and Climate Change in the Rise to Dominance of *Acropora* Corals

Yafei Mao ^{1, 2, 3} , Evan P. Economo ² , Noriyuki Satoh ¹



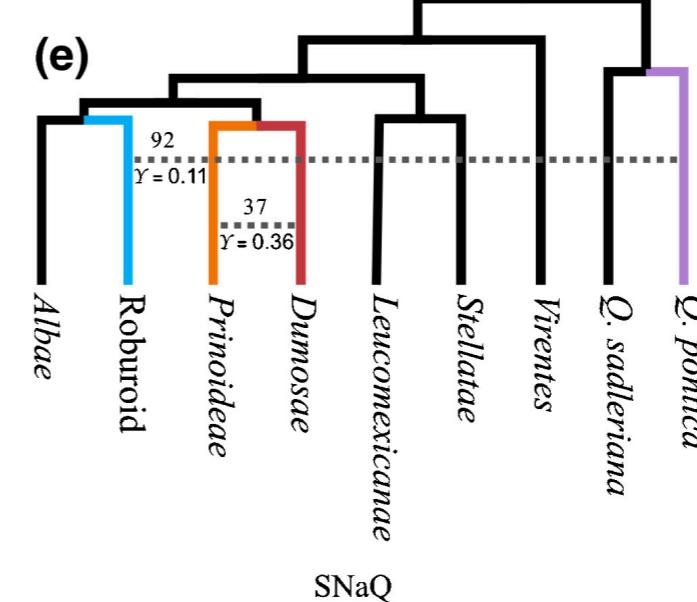


Full paper | Free Access

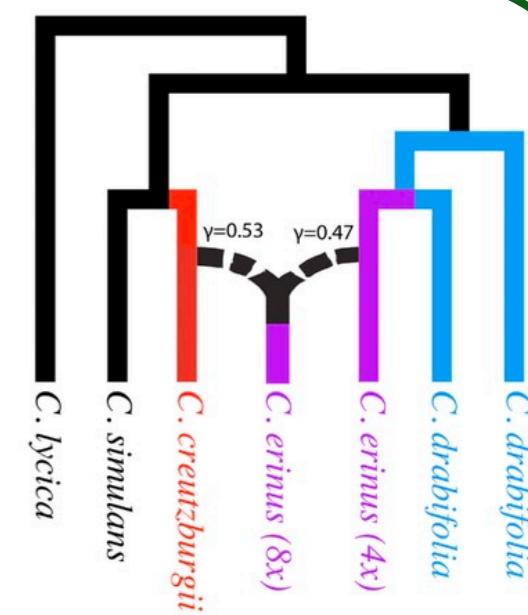
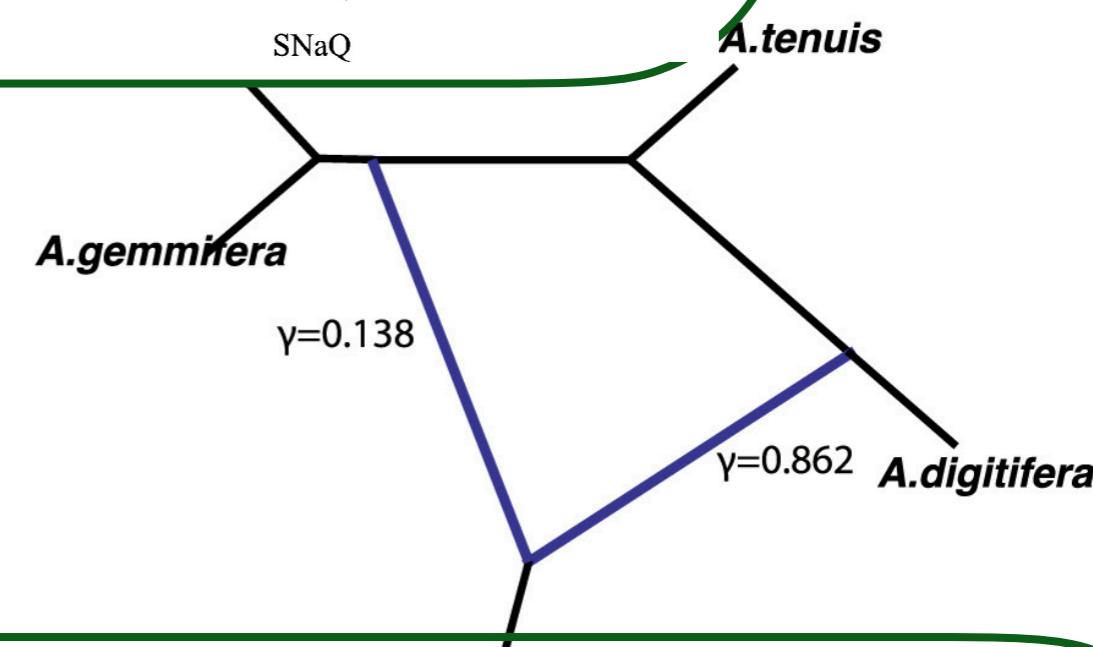
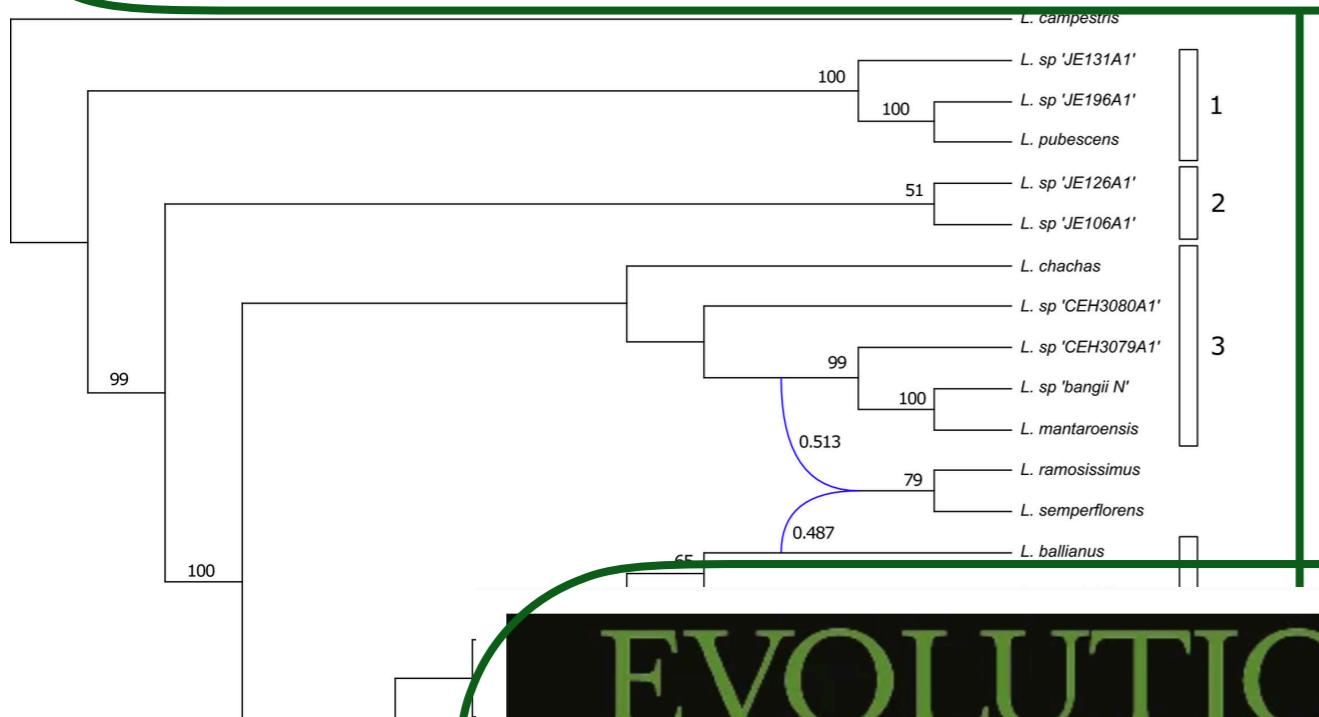
Uncovering the genomic signature of ancient introgression between white oak lineages (*Quercus*)

Andrew A. Crowl Paul S. Manos, John D. McVay, Alan R. Lemmon, Emily Moriarty Lemmon, Andrew L. Hipp ... See fewer authors

First published: 08 April 2019 | <https://doi.org/10.1111/nph.15842> | Citations: 32



imate Change in
Corals



EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

SSE
SOCIETY for the STUDY of EVOLUTION

ORIGINAL ARTICLE | Open Access |

Embracing discordance: Phylogenomic analyses provide evidence for allopolyploidy leading to cryptic diversity in a Mediterranean *Campanula* (Campanulaceae) clade

Andrew A. Crowl Cody Myers, Nico Cellinese



Full pa

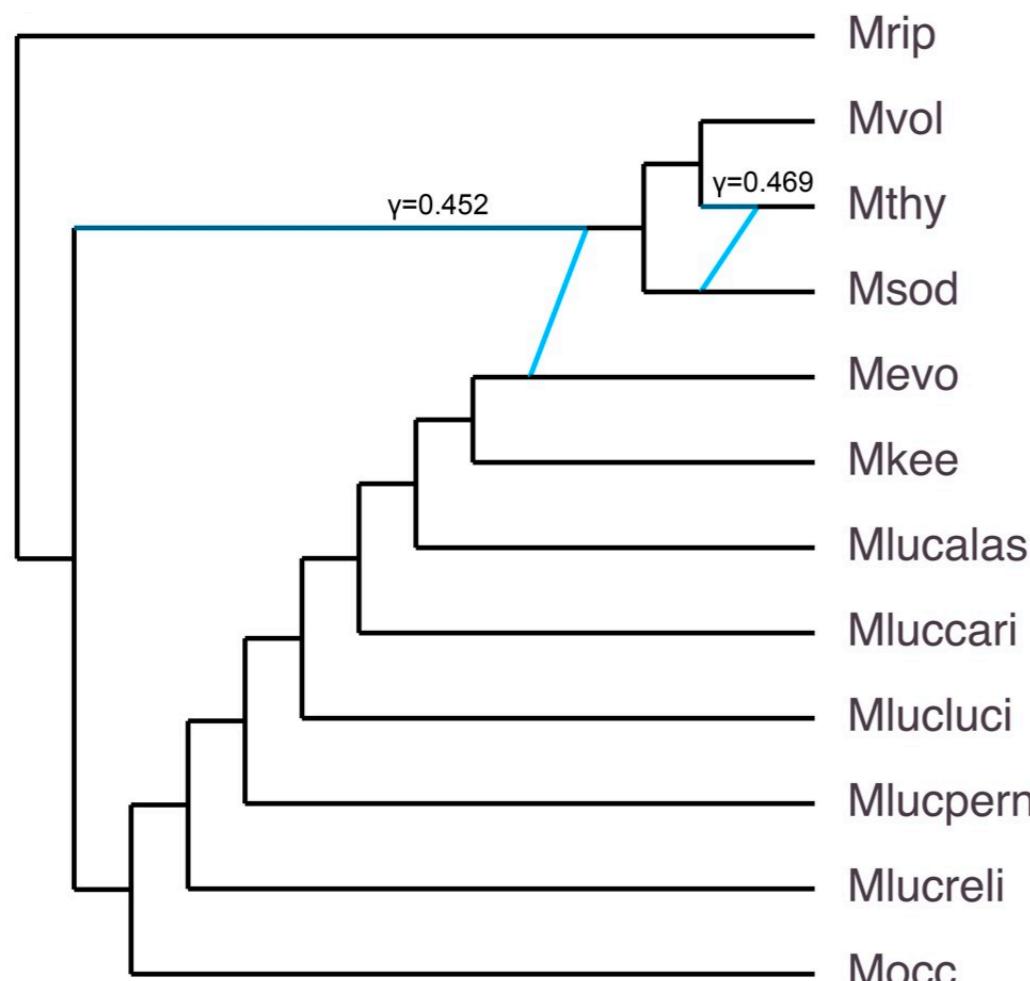
For

Evidence that *Myotis lucifugus* “Subspecies” are Five Nonsister Species, Despite Gene Flow

Ariadna E Morales , Bryan C Carstens

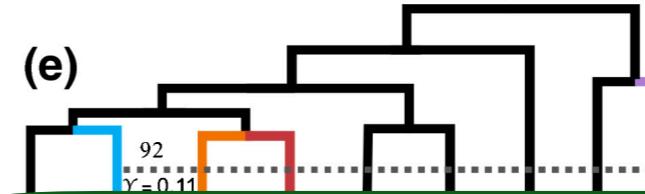
Systematic Biology, Volume 67, Issue 5, September 2018, Pages 756–769,
<https://doi.org/10.1093/sysbio/syy010>

Published: 15 February 2018 Article history ▾



Mediterranean Campanula (Campanulaceae)

Andrew A. Crowl , Cody Myers, Nico Cellinese



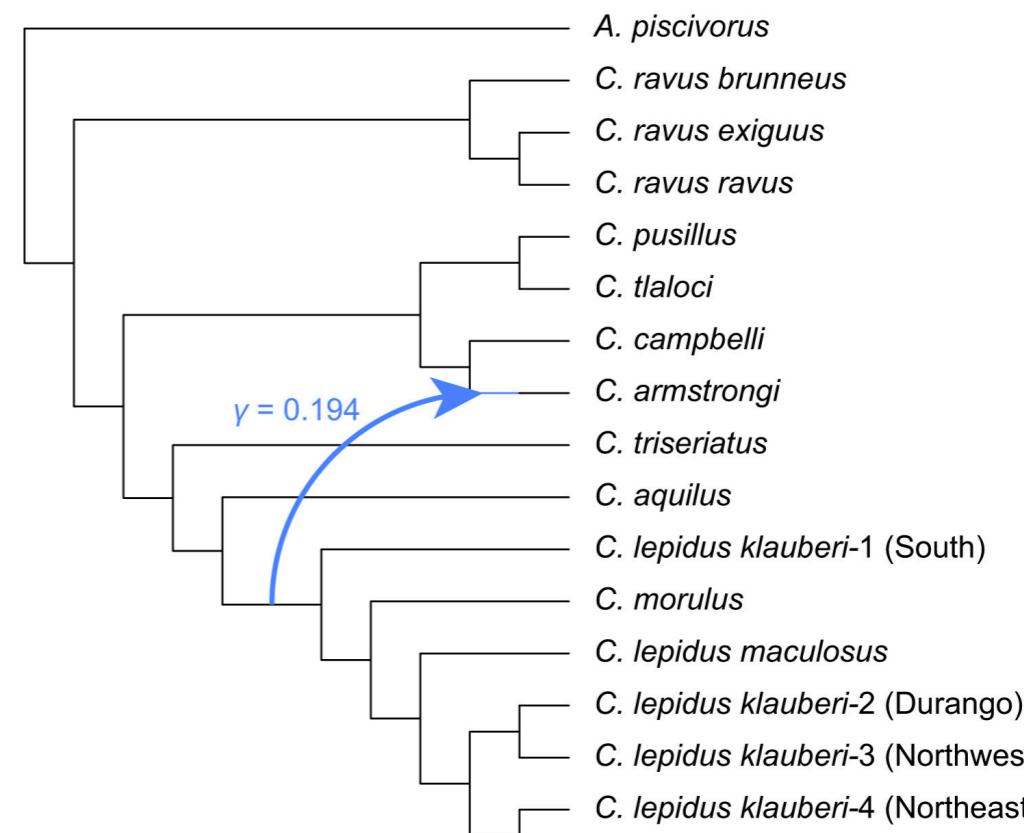
MOLECULAR ECOLOGY RESOURCES

RESOURCE ARTICLE | Full Access

Cryptic diversity in the Mexican highlands: Thousands of UCE loci help illuminate phylogenetic relationships, species limits and divergence times of montane rattlesnakes (Viperidae: *Crotalus*)

Christopher Blair , Robert W. Bryson Jr, Charles W. Linkem, David Lazcano, John Klicka, John E. McCormack ... See fewer authors ^

First published: 22 November 2018 | <https://doi.org/10.1111/1755-0998.12970> | Citations: 20



in

drabifolia

burgii
(8x)
(4x)

bifolia

ms

s

sp



Full p_c

E

Evidence that *Myotis lucifugus* “Subspecies” are Five Nonsister Species, Despite Gene Flow

Ariadna E Morales ✉, Bryan C Carstens

~~Systemati~~

<https://>

Bulk

Phylogenomic Signatures of Ancient Introgression in a Rogue Lineage of Darters (Teleostei: Percidae)

FREE

Daniel J MacGuigan ✉, Thomas J Near

Systematic Biology, Volume 68, Issue 2, March 2019, Pages 329–346,
<https://doi.org/10.1093/sysbio/svy074>

Published: 03 December 2018 Article history

Mediterranean Campanula (Campanula)

Andrew A. Crowl, Cody Myers, Nico Cellinese

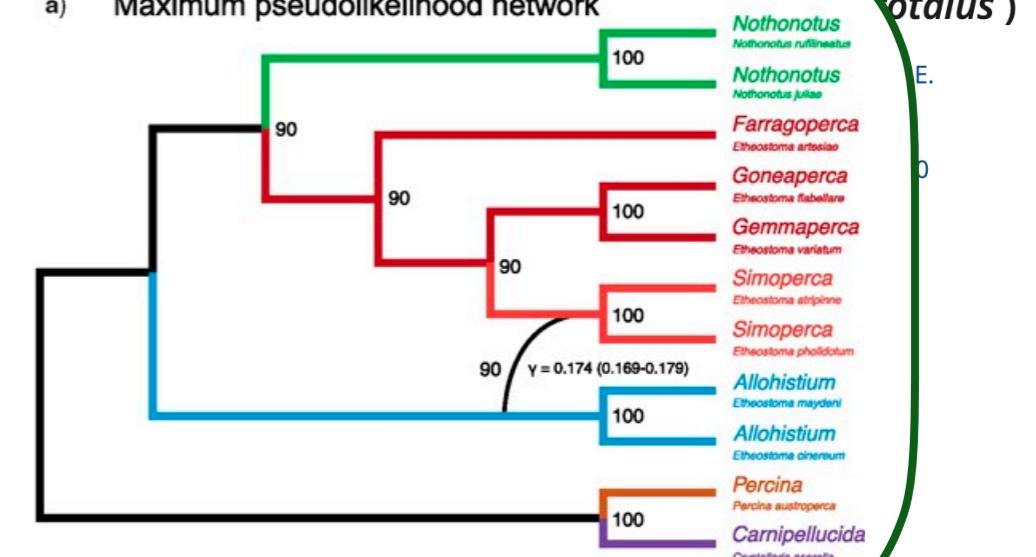
MOLECULAR ECOLOGY
RESOURCES

RESOURCE ARTICLE | Full Access

 Full Acces

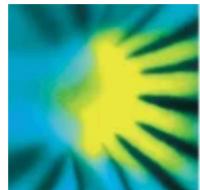
Cryptic diversity in the Mexican highlands: Thousands of UCE loci

a) Maximum pseudolikelihood network



- *C. triseriatus*
- *C. aquilus*
- *C. lepidus klauberi*-1 (South)
- *C. morulus*
- *C. lepidus maculosus*
- *C. lepidus klauberi*-2 (Durango)
- *C. lepidus klauberi*-3 (Northwest)
- *C. lepidus klauberi*-4 (Northeast)
- *C. lepidus lepidus*

ifolia
(4x)
(8x)



Full

Full

Evidence that *Myotis lucifugus* "Subspecies" are Five Nonsister Species, Despite Gene Flow

Ariadna E Morales ✉, Bryan C Carstens

Systematic

https://doi.org/10.1111/nph.14900

Publ

Phylogenomic Signatures of Ancient Introgression in a Rogue Lineage of Darters (Teleostei: Percidae)

FREE

Daniel J MacGuigan ✉, Thomas J Near

Systematic Biology, Volume 68, Issue 2, March 2019, Pages 329–346,
<https://doi.org/10.1093/sysbio/syy074>

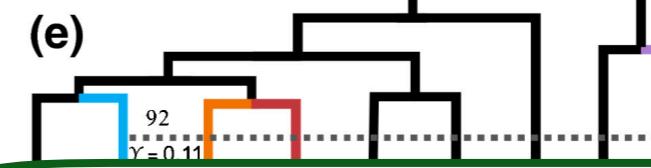
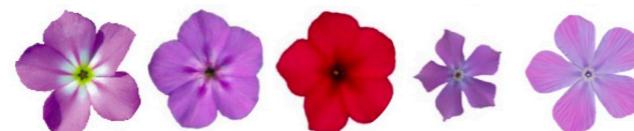
Published: 03 December 2018 Article history ▾

MOLECULAR ECOLOGY

Original Article |

Genomic evidence of gene flow during reinforcement in Texas *Phlox*

Federico Roda ✉, Fábio K. Mendes, Matthew W. Hahn, Robin Hopkins

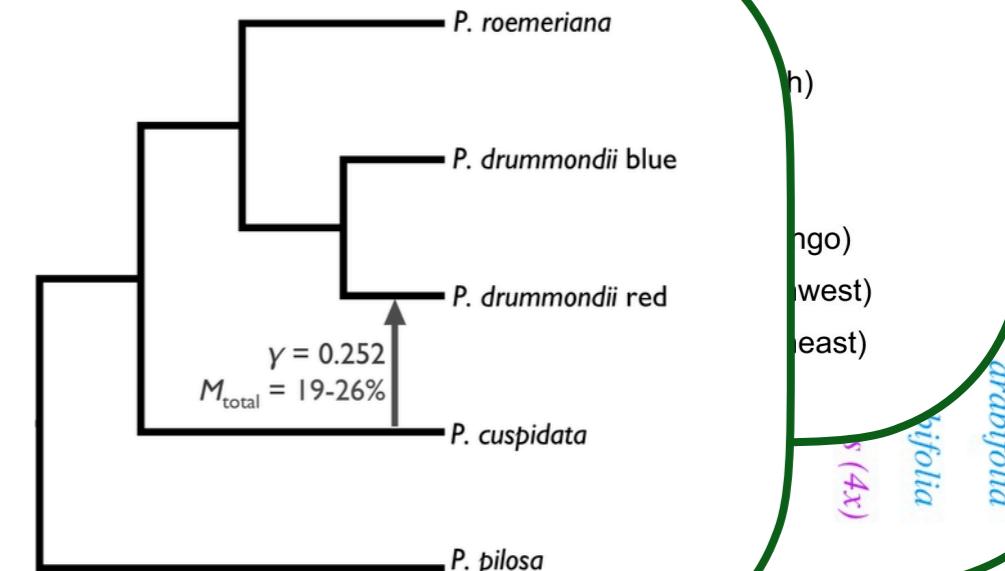
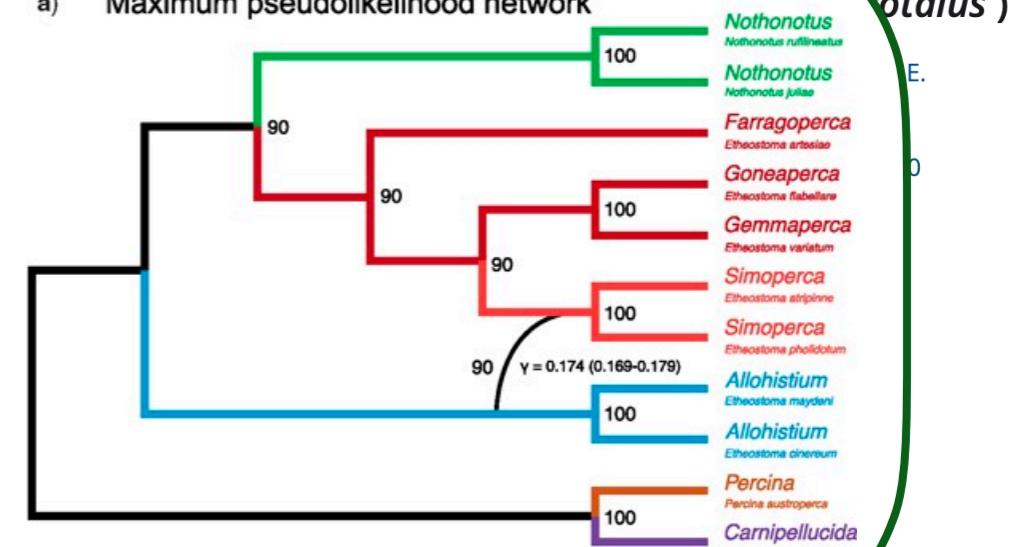
First published: 31 January 2017 | <https://doi.org/10.1111/mec.14041> | Citations: 15

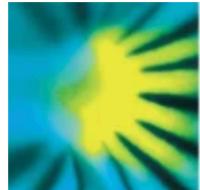
MOLECULAR ECOLOGY RESOURCES

RESOURCE ARTICLE |

Cryptic diversity in the Mexican highlands: Thousands of UCE loci

a) Maximum pseudolikelihood network





Full paper

Free

Evidence that *Myotis lucifugus* "Subspecies" Are Nonsister Species, Despite Gene Flow

Ariadna E Morales ✉, Bryan C Carstens

Systematic Biology

https://doi.org/10.1093/sysbio/syz073

Published

Phylogenomic Signature in a Rogue Lineage of

FREE

Daniel J MacGuigan ✉, Thomas J Near

Systematic Biology, Volume 68, Issue 3, May 2019, Pages 462–478,

https://doi.org/10.1093/sysbio/syy073

Published: 03 December 2018 Article history ▾

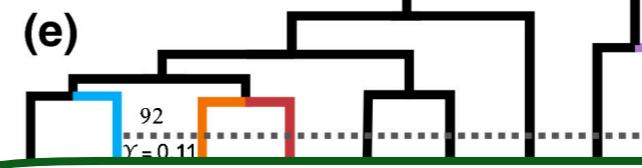
MOLECULAR

Original Article | Full Access

Genomic evidence of gene flow between sister species in *Phlox*

Federico Roda ✉, Fábio K. Mendes, Matthew W. Hahn, Robin Hopkins

First published: 31 January 2017 | https://doi.org/10.1111/mec.14041 | Citations: 15



MOLECULAR ECOLOGY RESOURCES

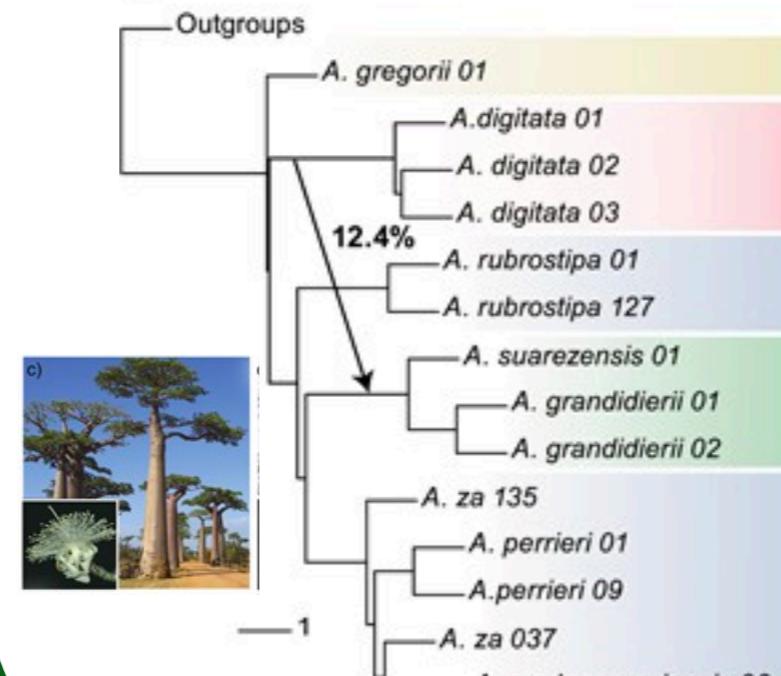
Reticulate Evolution Helps Explain Apparent Homoplasy in Floral Biology and Pollination in Baobabs (*Adansonia*; Bombacoideae; Malvaceae)

Nisa Karimi ✉, Corinne E Grover, Joseph P Gallagher, Jonathan F Wendel, Cécile Ané, David A Baum ✉

Systematic Biology, Volume 69, Issue 3, May 2020, Pages 462–478,

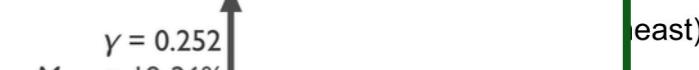
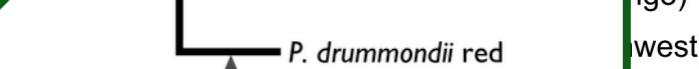
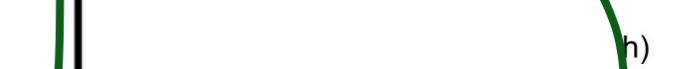
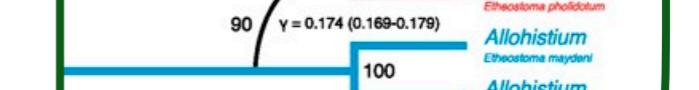
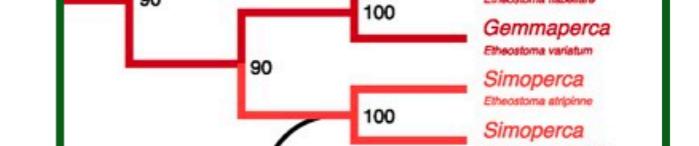
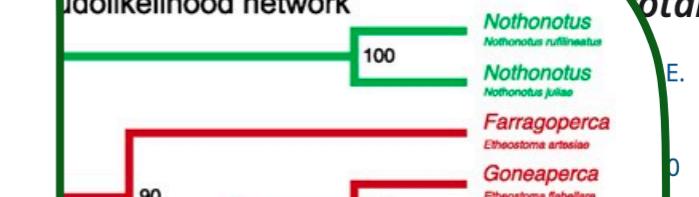
https://doi.org/10.1093/sysbio/syz073

Published: 06 November 2019 Article history ▾



Mexican Highlands: Thousands of UCE Loci Reveal Reticulate Evolution in *Phlox* (4x)

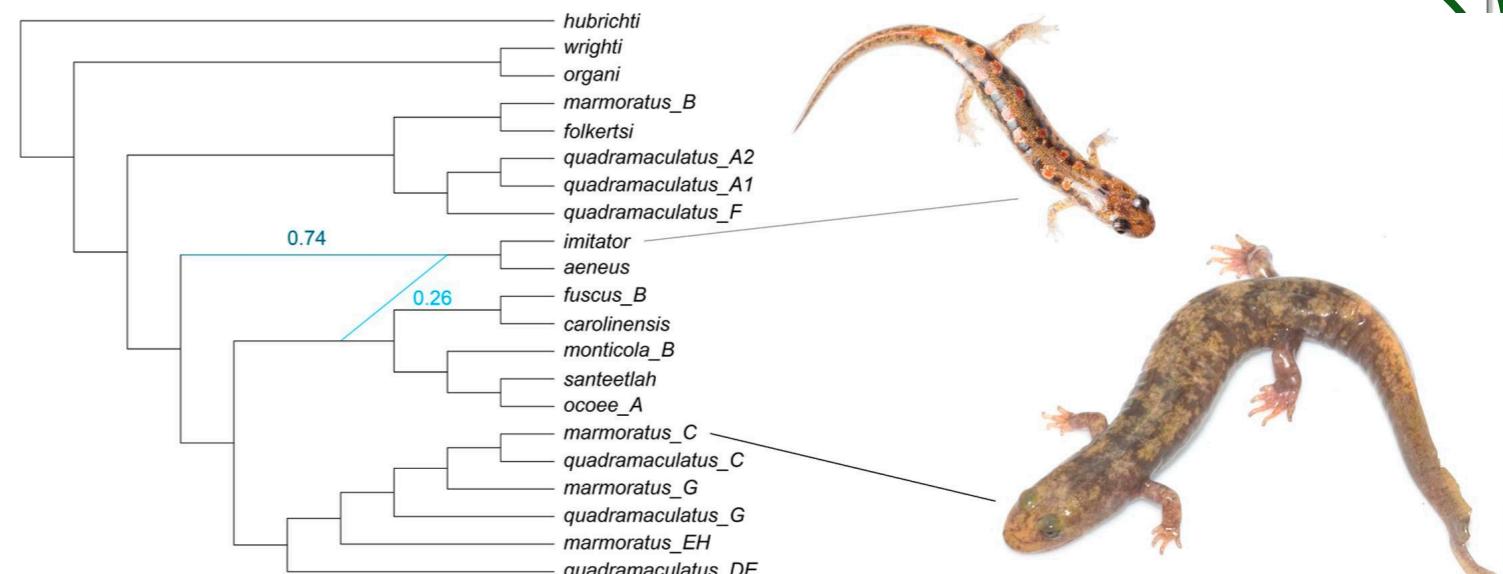
Maximum likelihood network





Phylogenomic data reveal reticulation and incongruence among mitochondrial candidate species in Dusky Salamanders (*Desmognathus*)

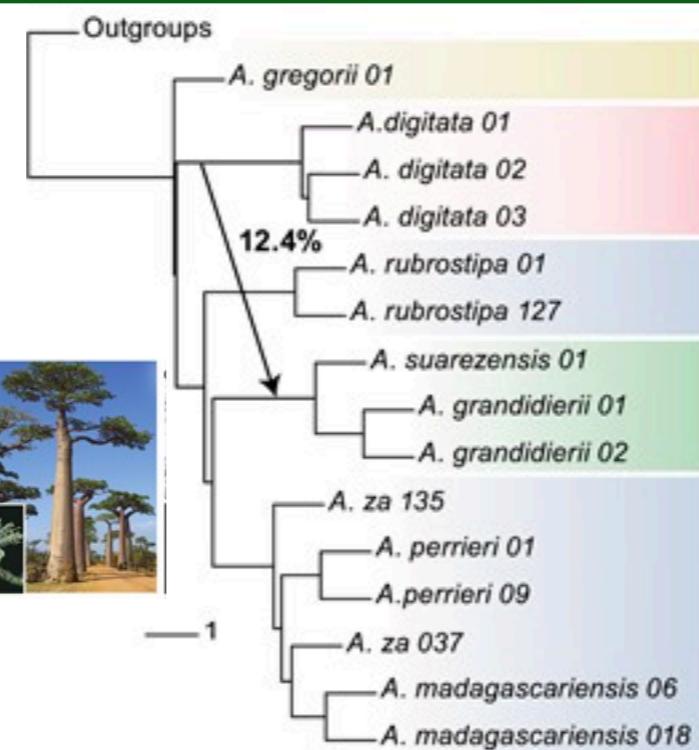
R. Alexander Pyron ^{a, b} , Kyle A. O'Connell ^{a, b, c}, Emily Moriarty Lemmon ^d, Alan R. Lemmon ^e, David A. Beamer ^f



Daniel J MacGuigan , Thomas J Ne

Systematic Biology, Volume 68, Issue
<https://doi.org/10.1093/sysbio/syy0>

Published: 03 December 2018

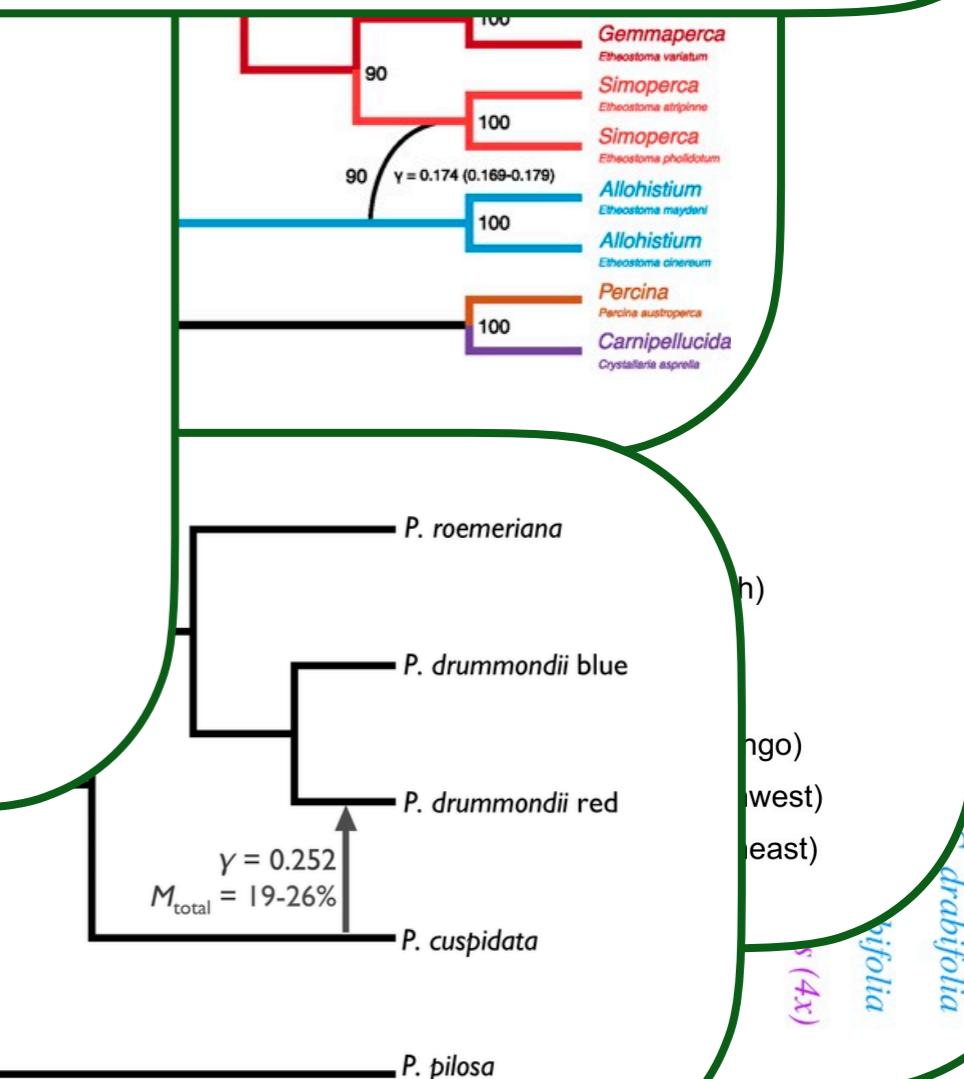


Original Article |  Full Access

Genomic evidence of gene conversion in *Phlox*

Federico Roda, Fábio K. Mendes, Matthew W. Hahn, Robin Hopkins

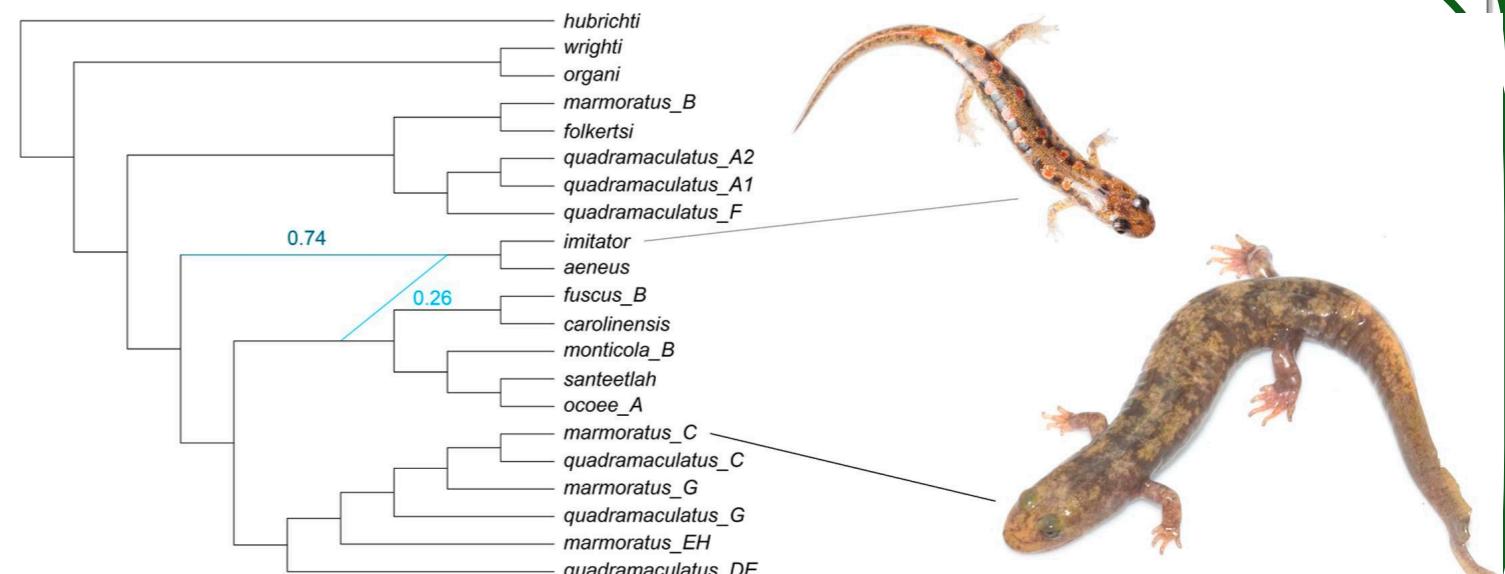
First published: 31 January 2017 | <https://doi.org/10.1111/mec.14041> | Citations: 15





Phylogenomic data reveal reticulation and incongruence among mitochondrial candidate species in Dusky Salamanders (*Desmognathus*)

R. Alexander Pyron ^{a, b} Kyle A. O'Connell ^{a, b, c}, Emily Moriarty Lemmon ^d, Alan R. Lemmon ^e, David A. Beamer ^f



Daniel J MacGuigan , Thomas J Neff

Systematic Biology, Volume 68, Issue 5, September 2019, Pages 1000–1014

<https://doi.org/10.1093/sysbio/syy031>

Outgroups

A. gregorii 01
A. digitata 01

Gemmoperca
Etheostoma variatum
Simoperca
Etheostoma atripinne
Simoperca

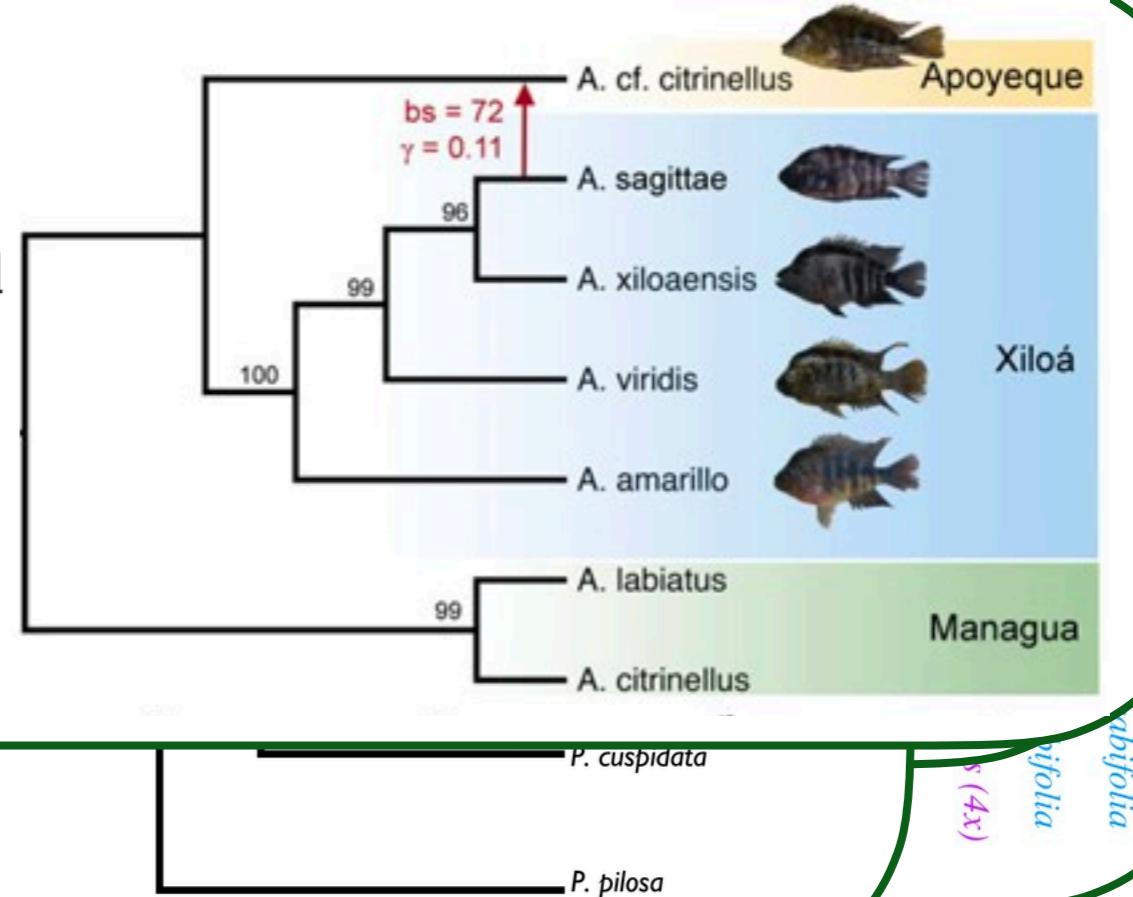
Implementing Large Genomic Single Nucleotide Polymorphism Data Sets in Phylogenetic Network Reconstructions: A Case Study of Particularly Rapid Radiations of Cichlid Fish

Melisa Olave , Axel Meyer

Systematic Biology, Volume 69, Issue 5, September 2020, Pages 848–862,

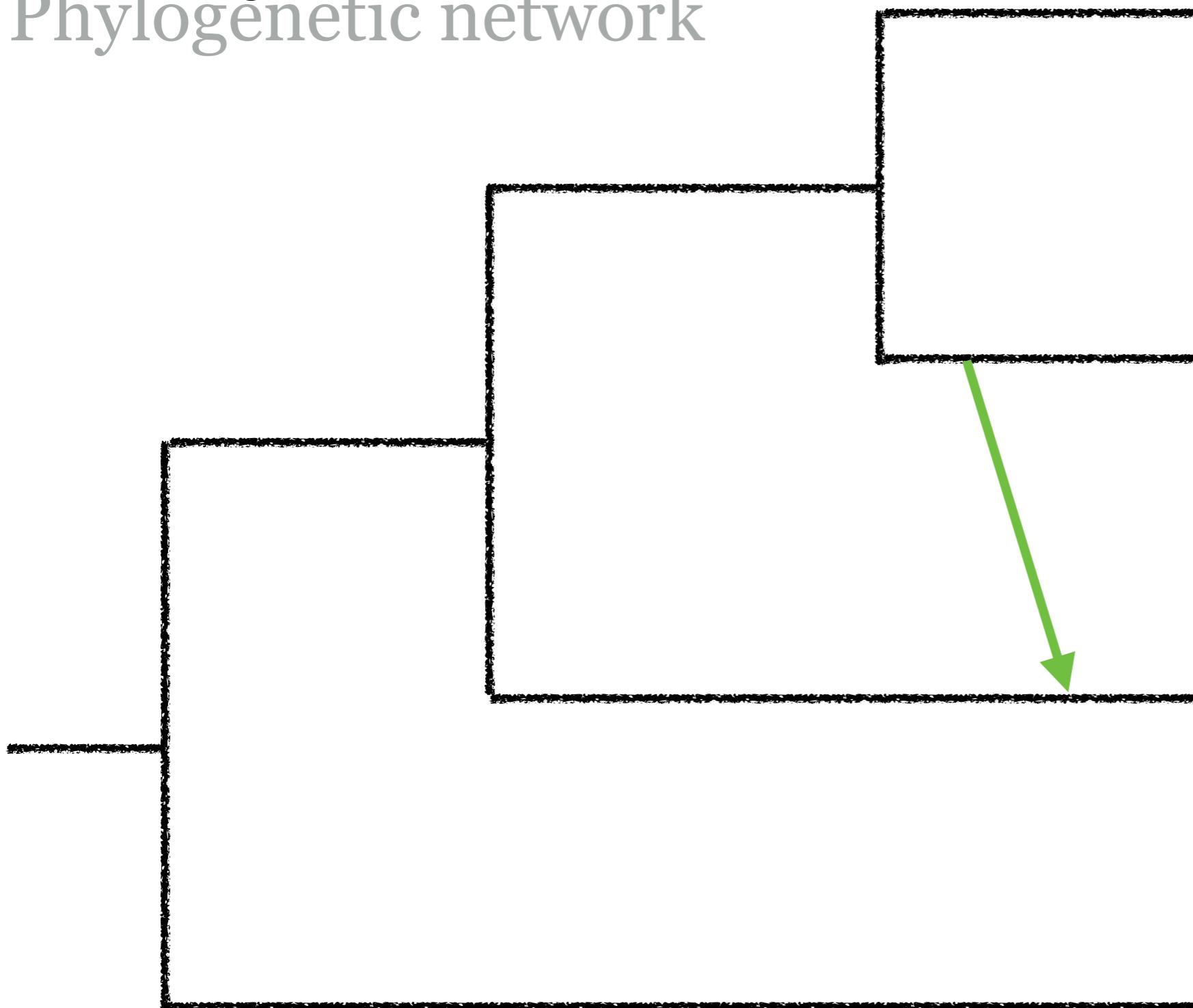
<https://doi.org/10.1093/sysbio/syaa005>

Published: 03 February 2020 Article history ▾



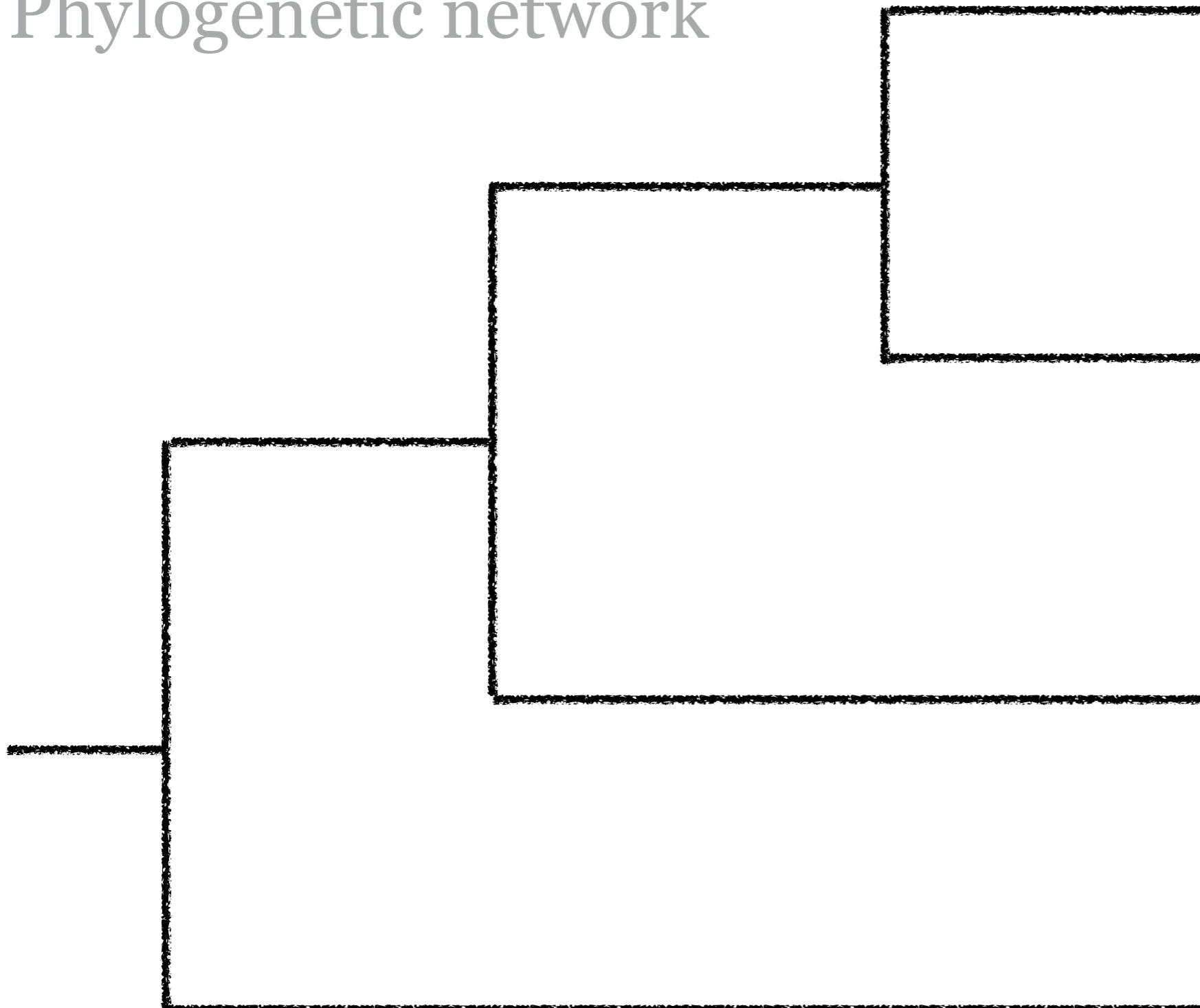
Why?

Phylogenetic network



Why?

Phylogenetic network



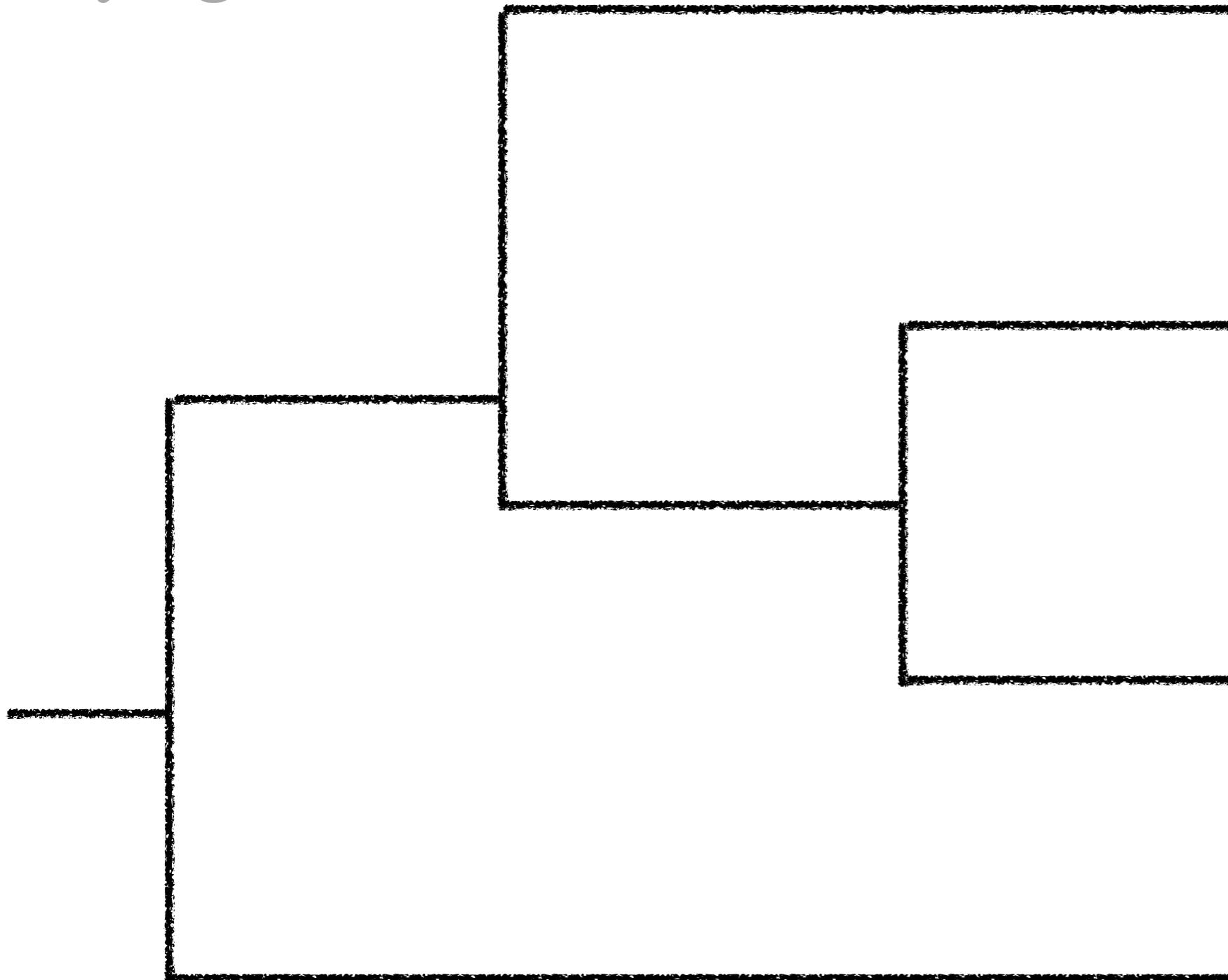
Main tree



Why?

Phylogenetic network

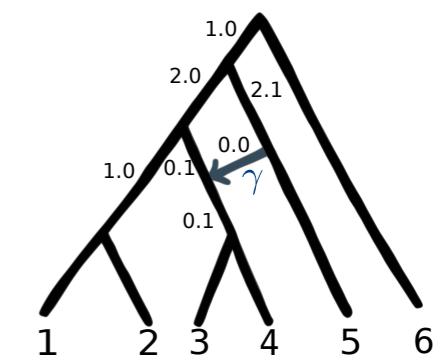
Ignore gene flow
=>Wrong tree!



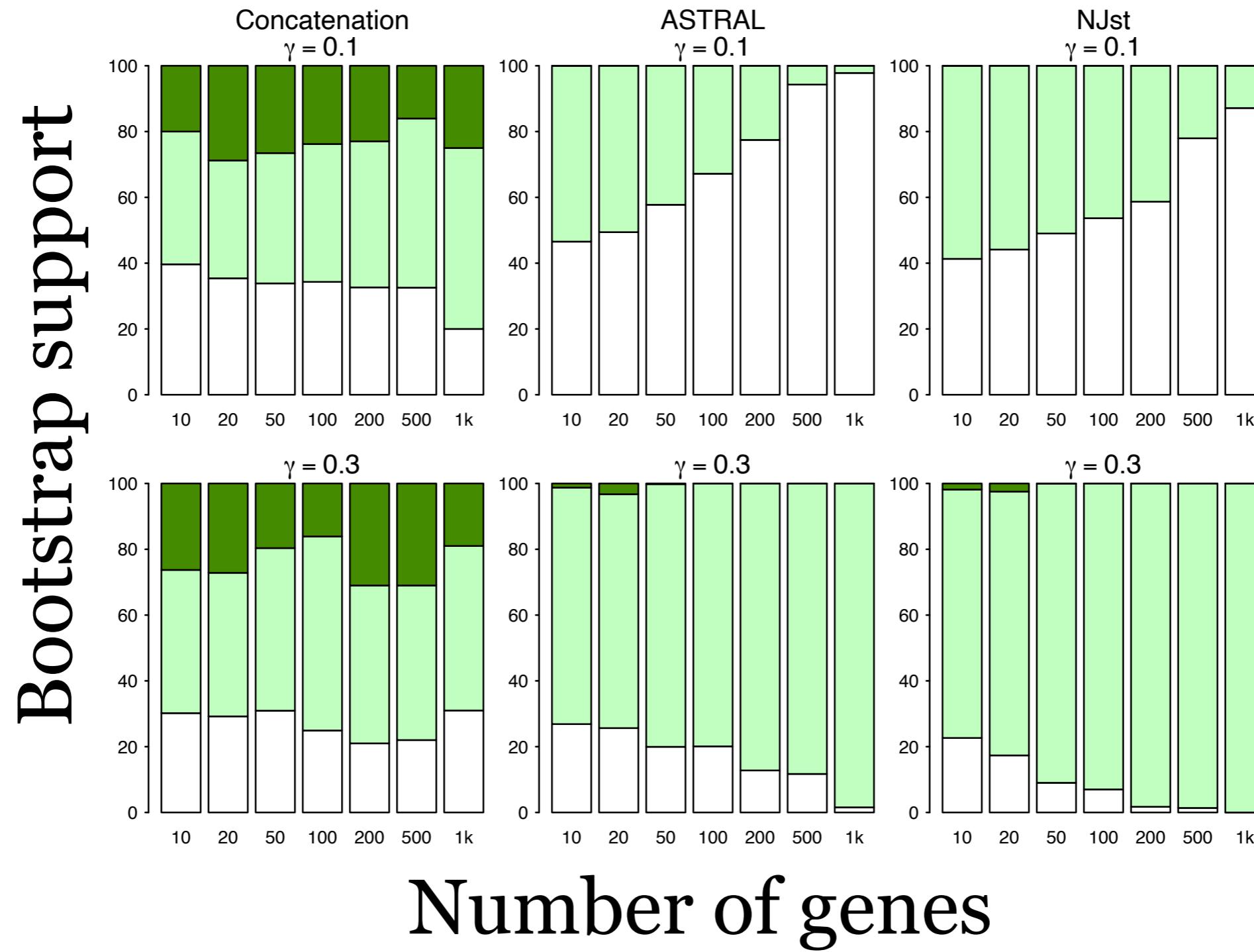
Why?

Phylogenetic network

Coalescent tree methods
not robust to gene flow



White:
true tree

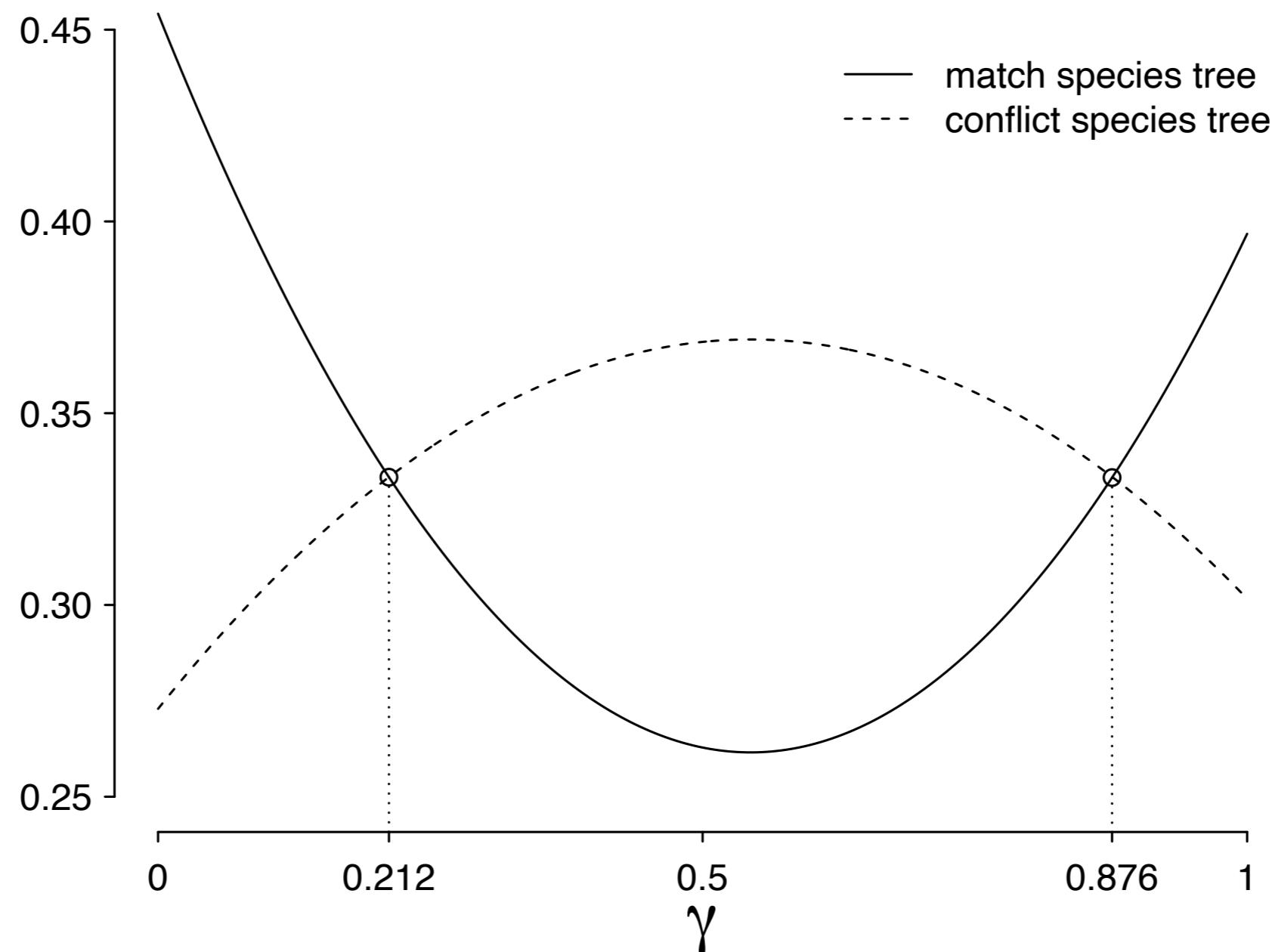


(S.-L., Yang, Ané, 2016, Syst Bio)

ASTRAL (Mirarab et al, 2014)
NJst (Liu&Yu, 2011)

Why? Phylogenetic network

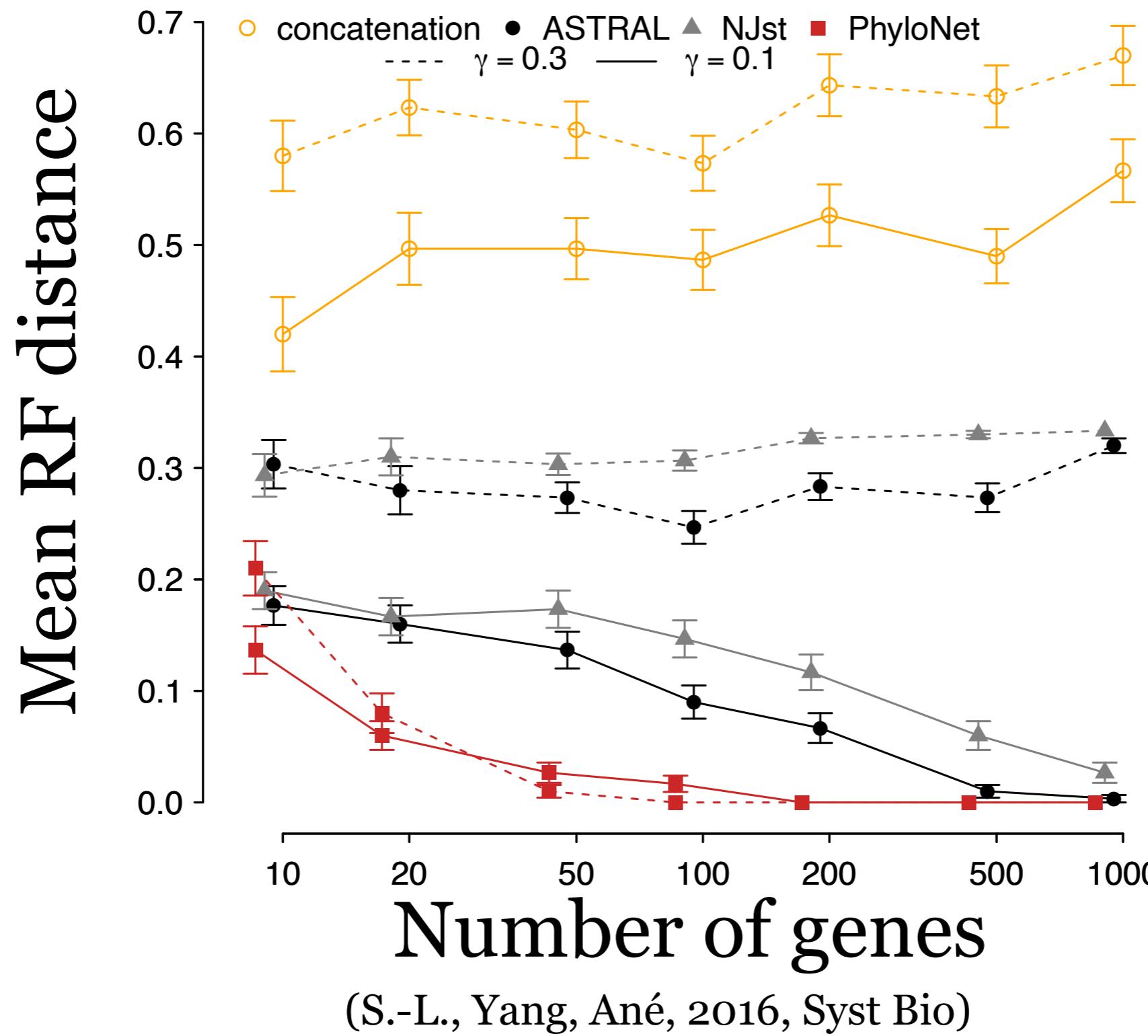
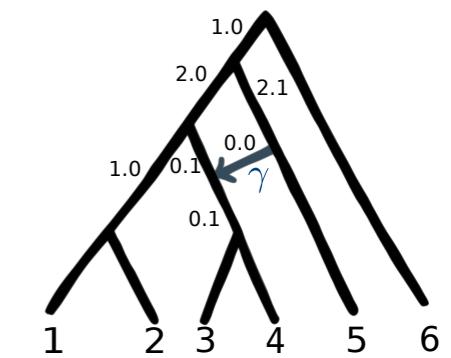
Anomaly zone with
gene flow



Why?

Phylogenetic network

Coalescent tree methods
not robust to gene flow

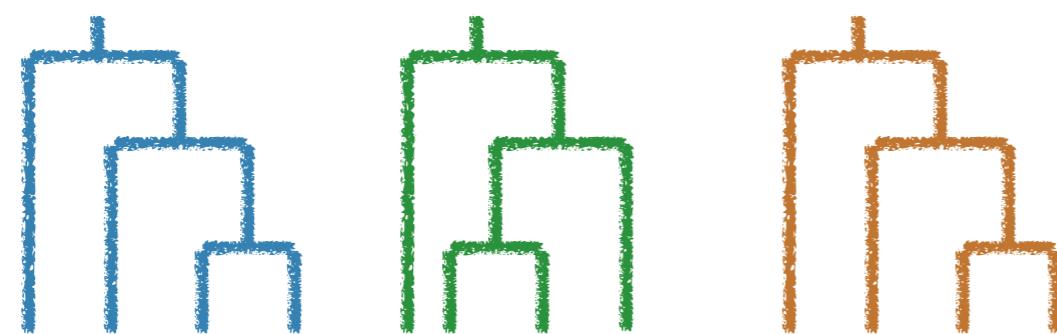


How?

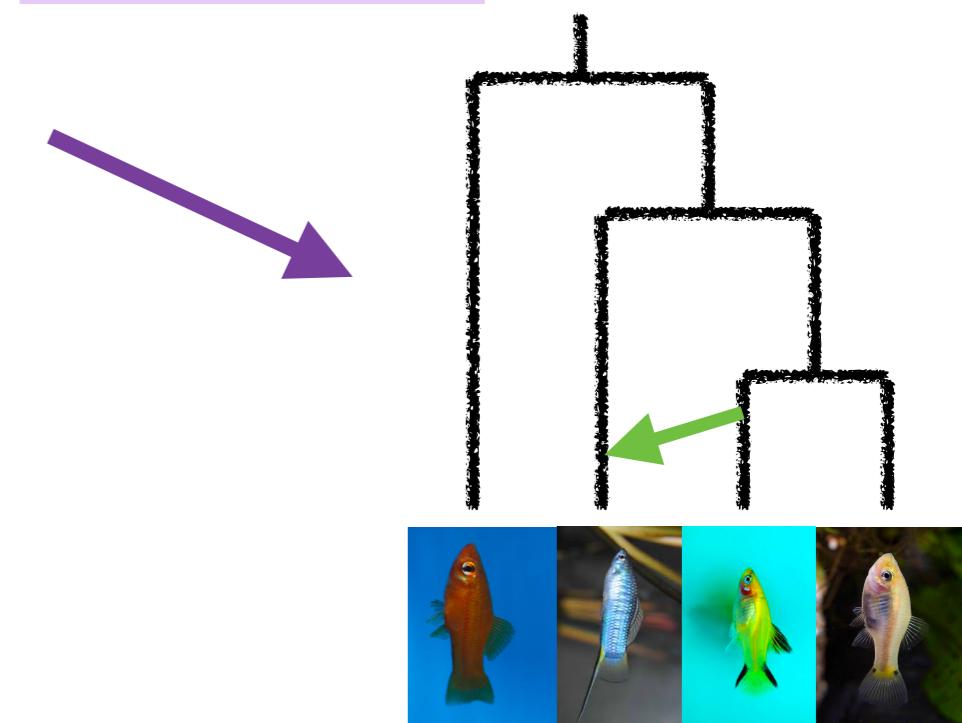
Phylogenetic network



MrBayes
(Huelsenbeck, Ronquist, 2001)
RAxML
(Stamatakis, 2014)
PhyML
(Guindon et al, 2010)
RevBayes
(Hoehna et al, 2016)
IQ-TREE
Nguyen et al. (2015)

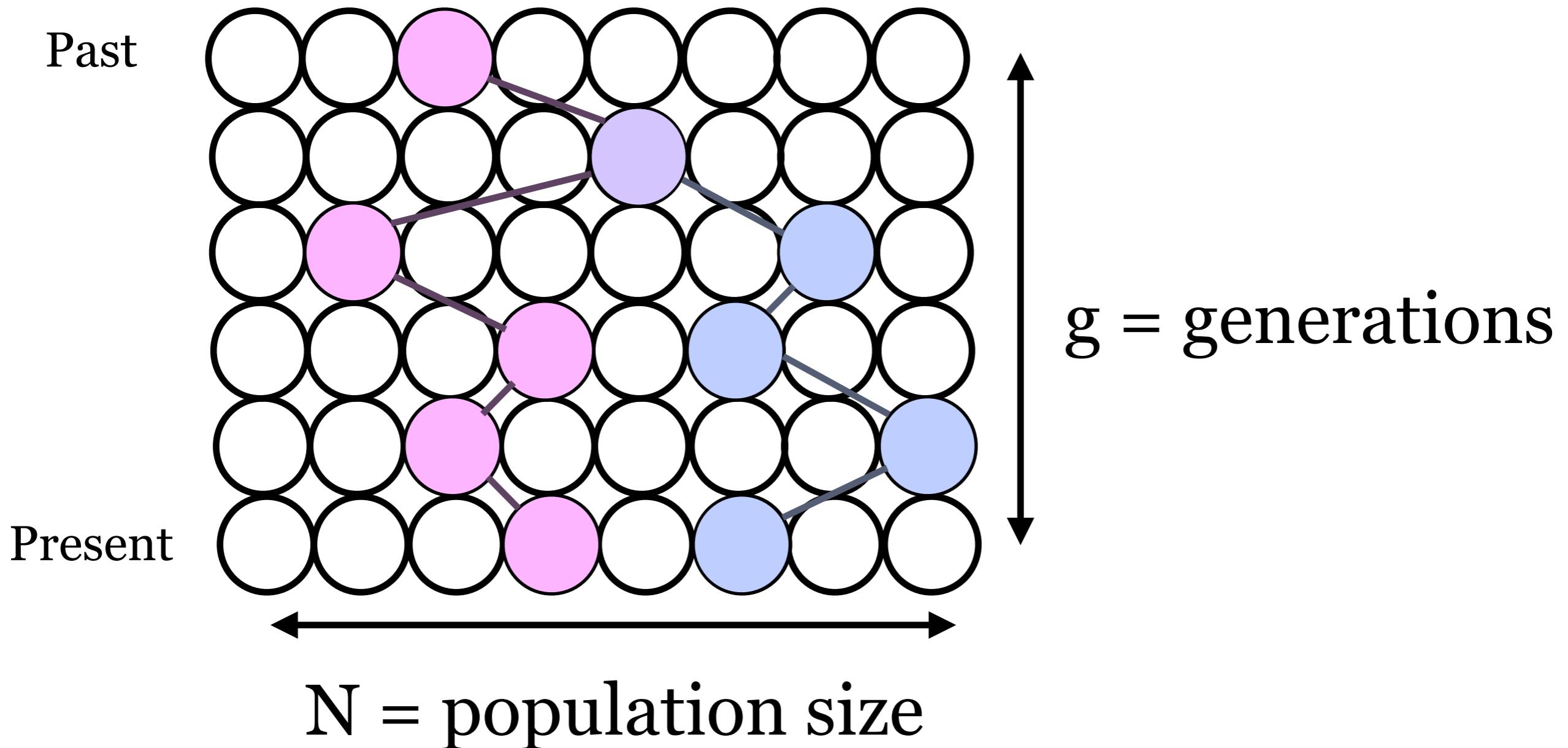


BEAST2
(Zhang et al, 2017)
PhyloNet
(Wen et al, 2016)



SNaQ
(S.-L., Ane, 2016)
PhyloNet
(Yu et al, 2014)

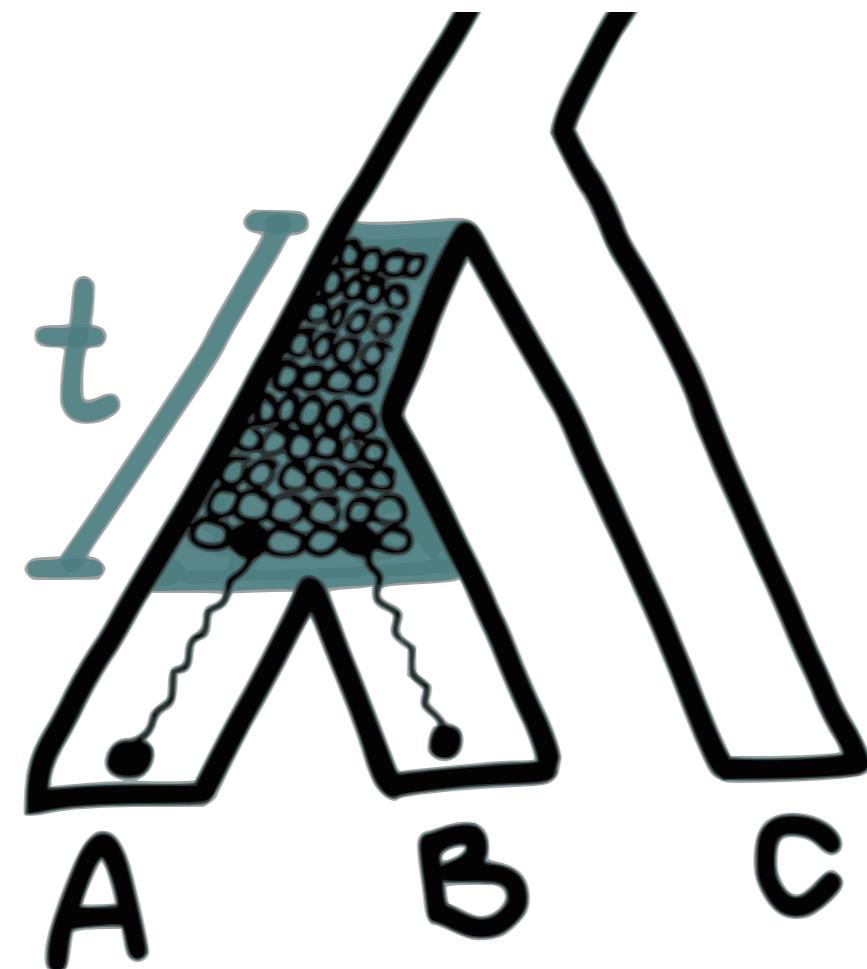
Coalescent model within 1 population



Probability of no coalescence in g generations:

$$\left(1 - \frac{1}{N}\right)^g$$
$$t = g/N \Rightarrow \left(1 - \frac{t}{Nt}\right)^{Nt} \xrightarrow[N \rightarrow \infty]{} e^{-t}$$

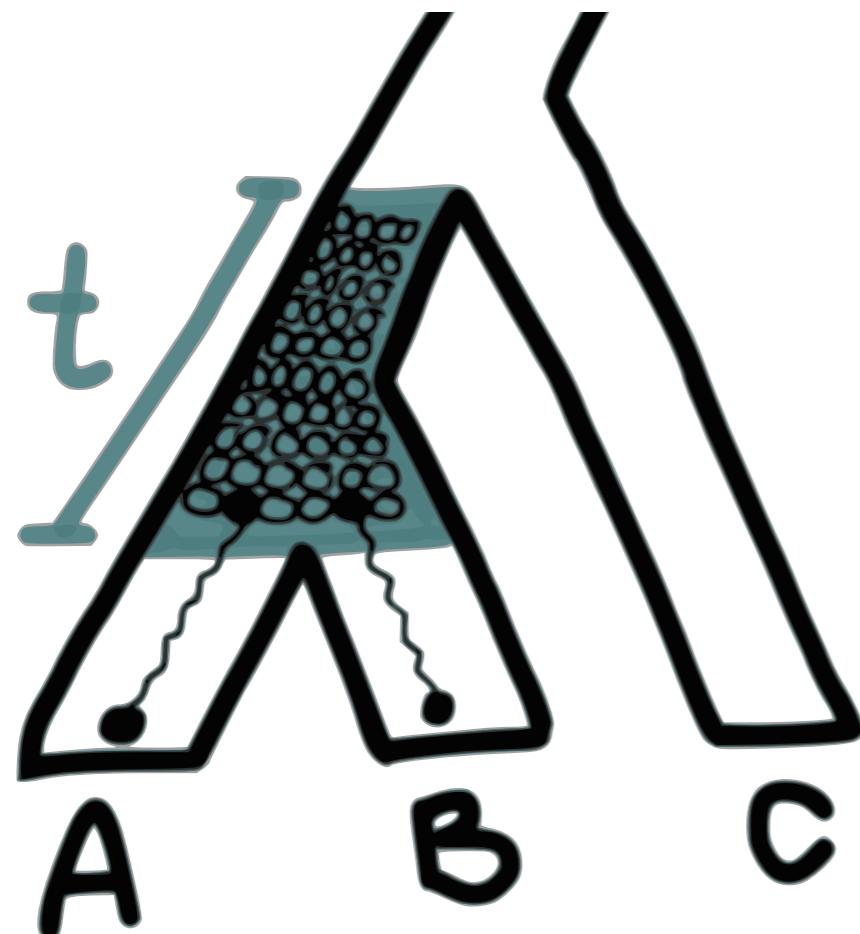
Multispecies coalescent on a tree



$$P(T > t) = e^{-t}$$

$$T = \frac{g}{N} \text{ coalescent units} \sim Exp(1)$$

Multispecies coalescent on a tree

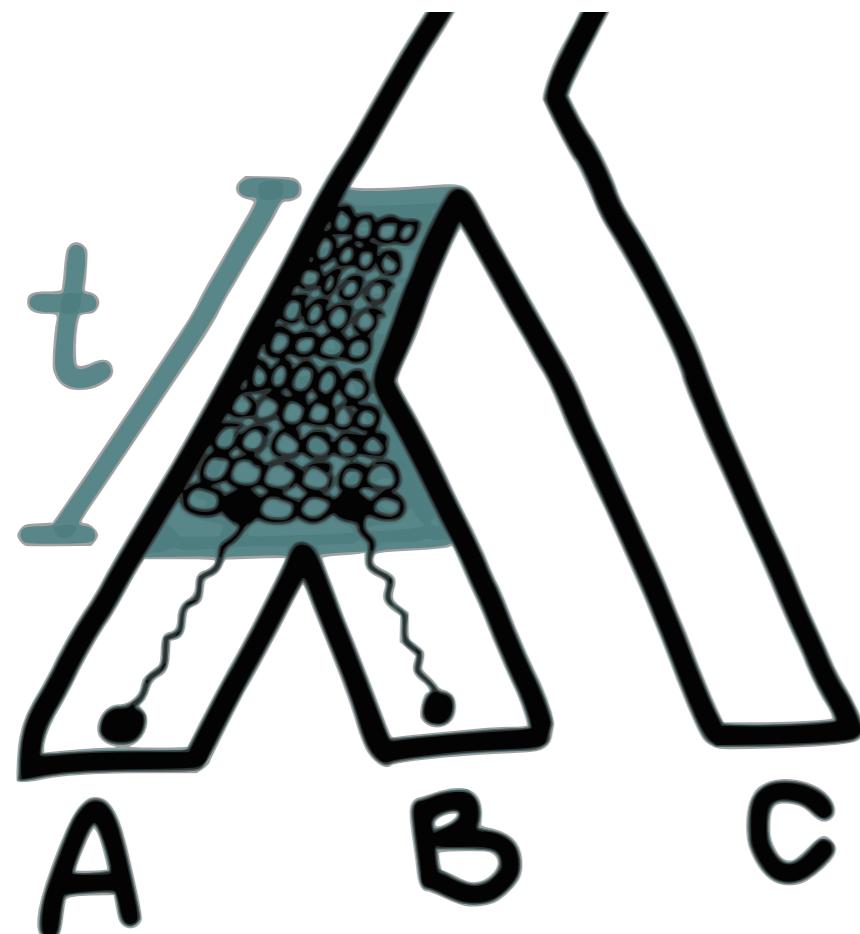


$$P(\text{ } \text{ } \text{ } \text{ } \text{ } \text{ }) =$$

A large black letter P followed by a large black parenthesis containing a phylogenetic tree with three tips labeled A, B, and C. The entire expression is followed by an equals sign.

$$P(T > t) = e^{-t}$$

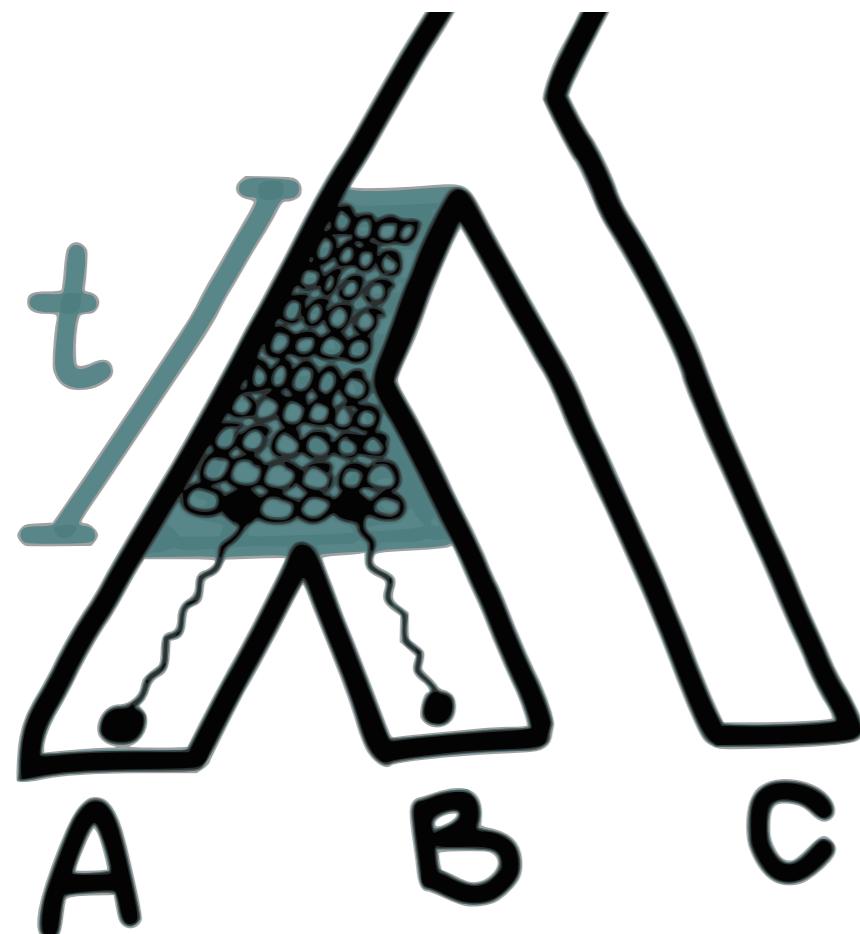
Multispecies coalescent on a tree



$$P(\text{ } \text{ } \text{ } \text{ } \text{ } \text{ }) = \\ 1 - e^{-t}$$

$$P(T > t) = e^{-t}$$

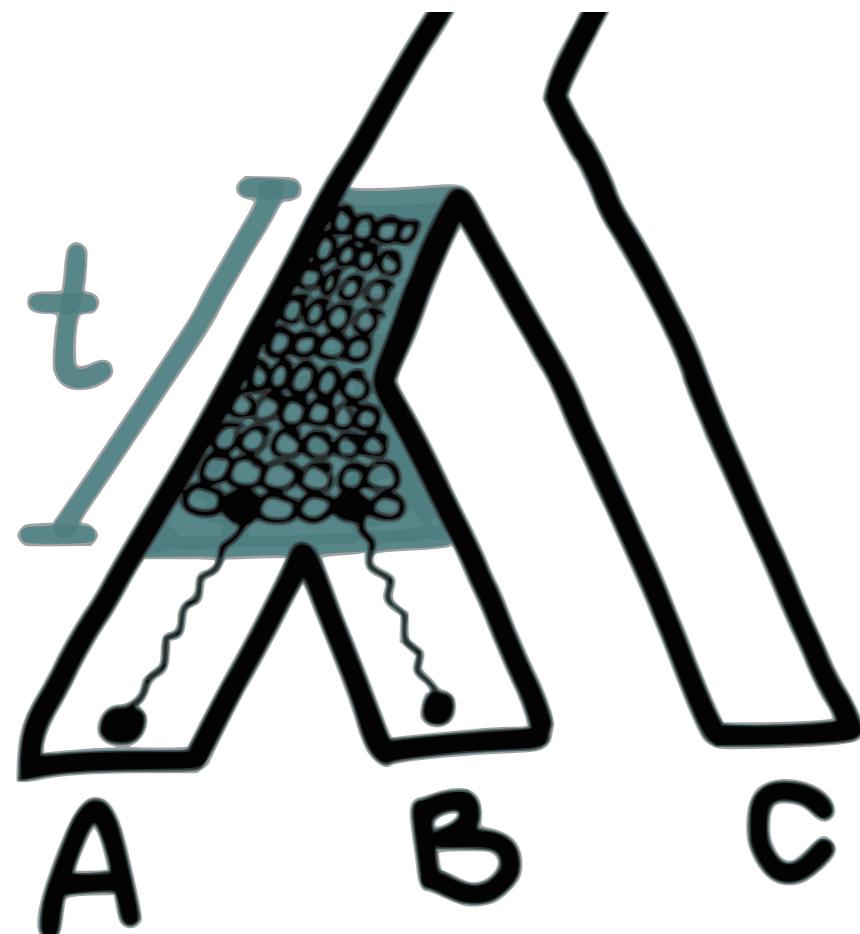
Multispecies coalescent on a tree



$$P(\text{ } \text{ } \text{ } \text{ } \text{ } \text{ }) = \\ 1 - e^{-t} \\ +$$

$$P(T > t) = e^{-t}$$

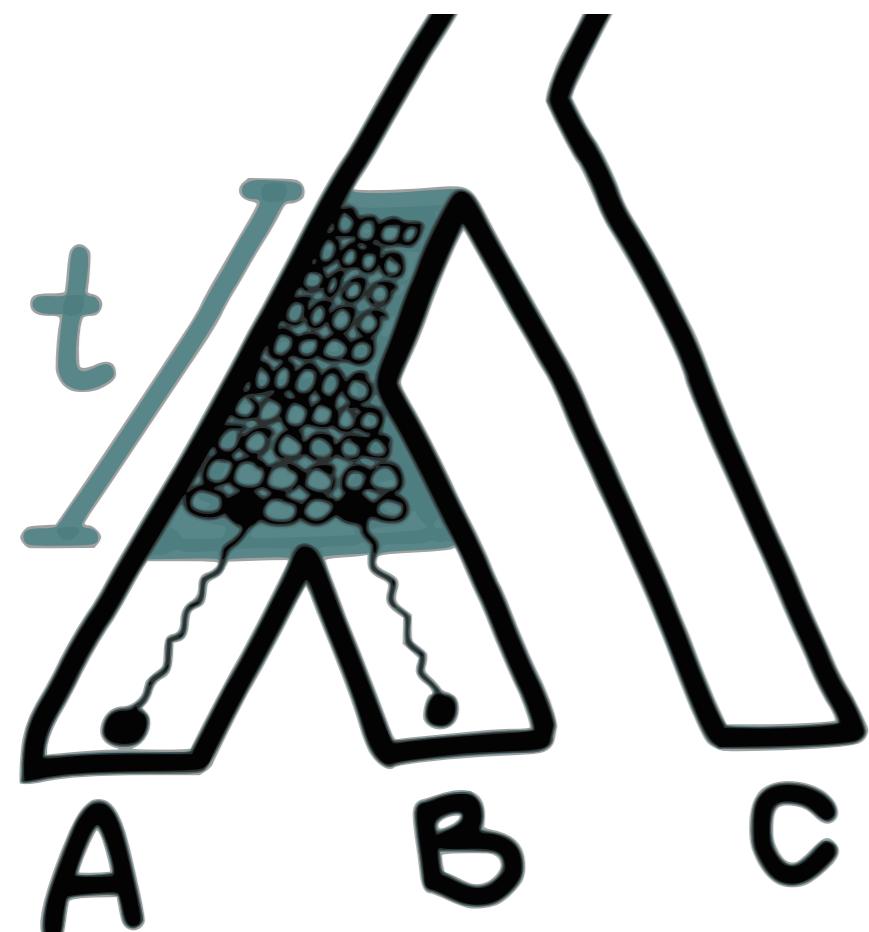
Multispecies coalescent on a tree



$$P(\text{ } \text{ } \text{ } \text{ } \text{ } \text{ }) = \\ 1 - e^{-t} \\ + \\ e^{-t} \times 1/3$$

$$P(T > t) = e^{-t}$$

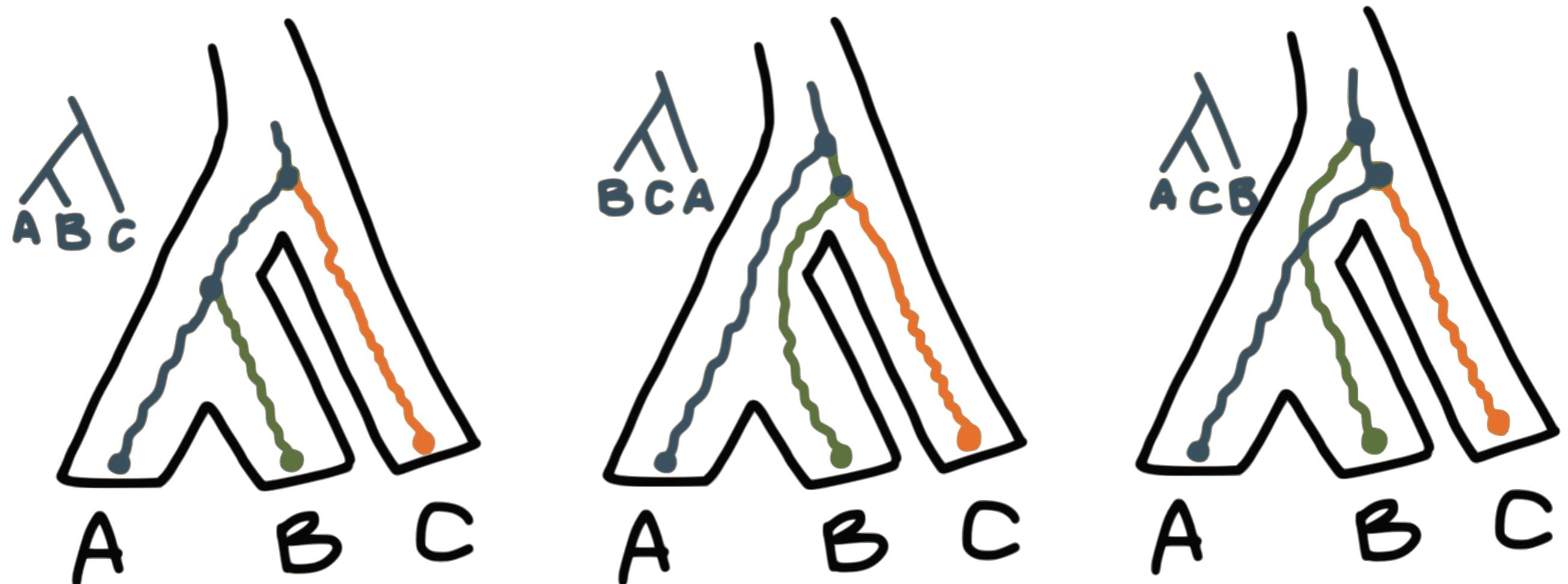
Multispecies coalescent on a tree



$$P(T > t) = e^{-t}$$

$$\begin{aligned}
& P(\bigwedge_{A \in \mathcal{B}} A) = \\
& 1 - e^{-t} \\
& + \\
& e^{-t} \times 1/3 \\
& = 1 - \frac{2}{3}e^{-t}
\end{aligned}$$

Multispecies coalescent on a tree



$$1 - \frac{2}{3}e^{-t}$$

$$\frac{1}{3}e^{-t}$$

$$\frac{1}{3}e^{-t}$$

Multispecies coalescent on a network



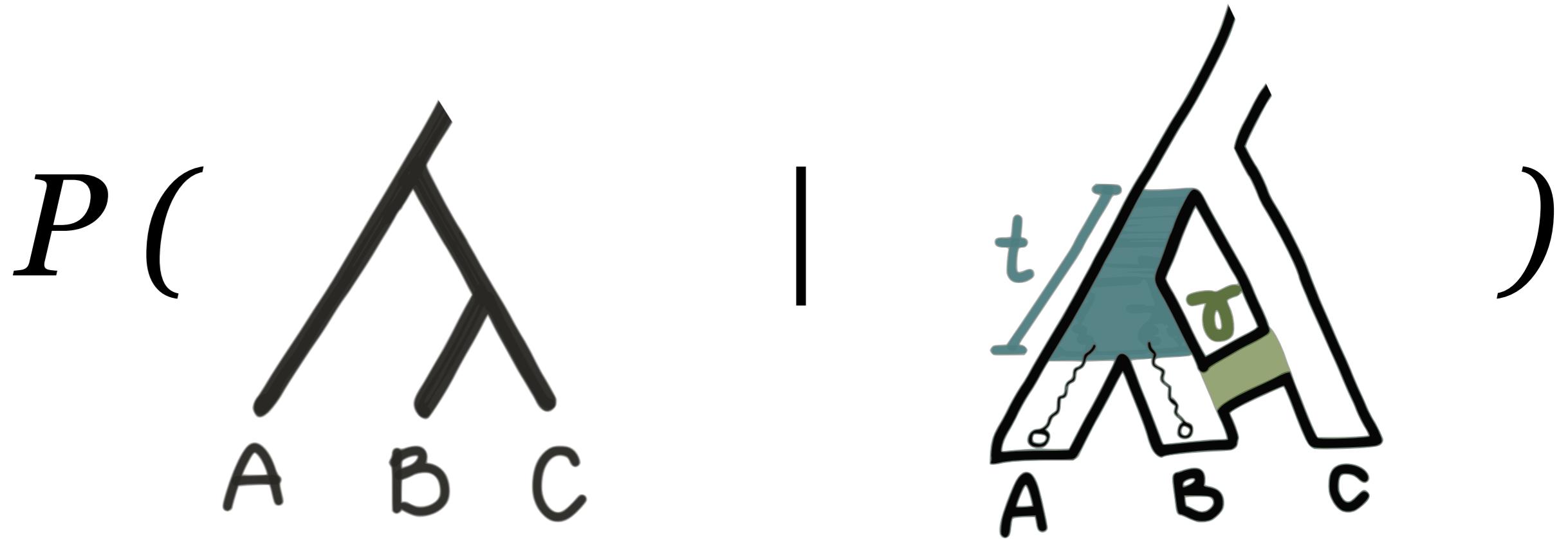
(Meng, Kubatko, 2009)
(Yu, Degnan, Nakhleh, 2012)

Multispecies coalescent on a network



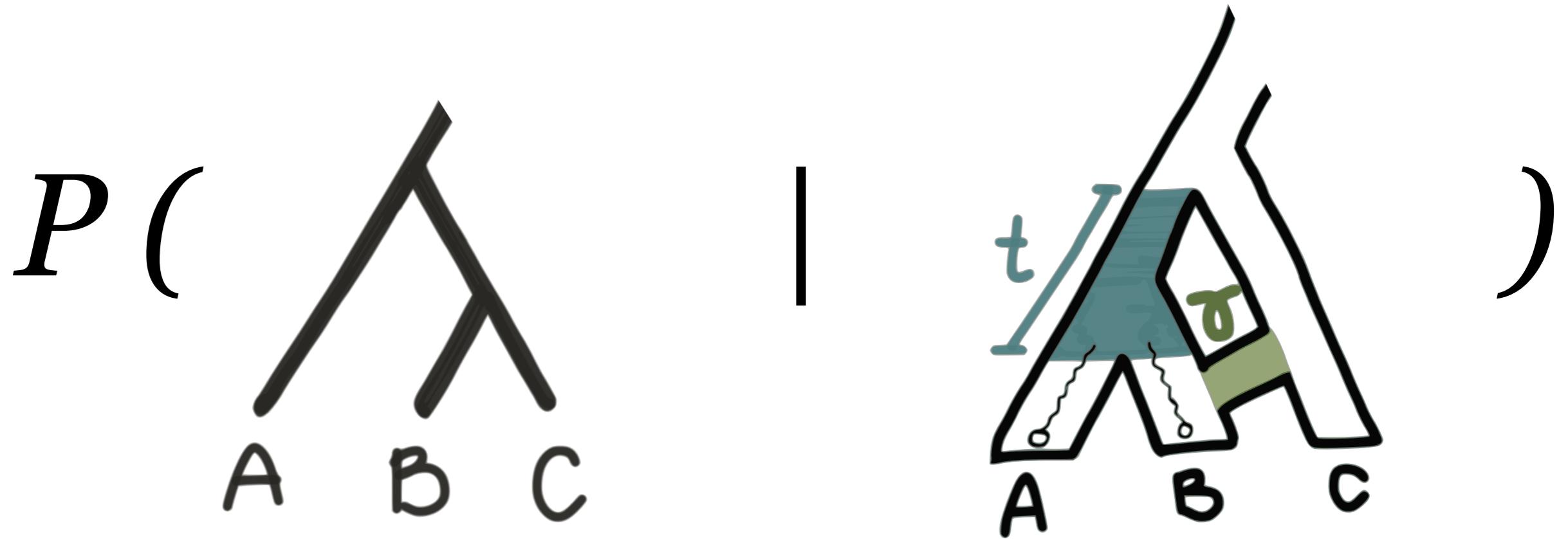
(Meng, Kubatko, 2009)
(Yu, Degnan, Nakhleh, 2012)

Multispecies coalescent on a network



(Meng, Kubatko, 2009)
(Yu, Degnan, Nakhleh, 2012)

Multispecies coalescent on a network



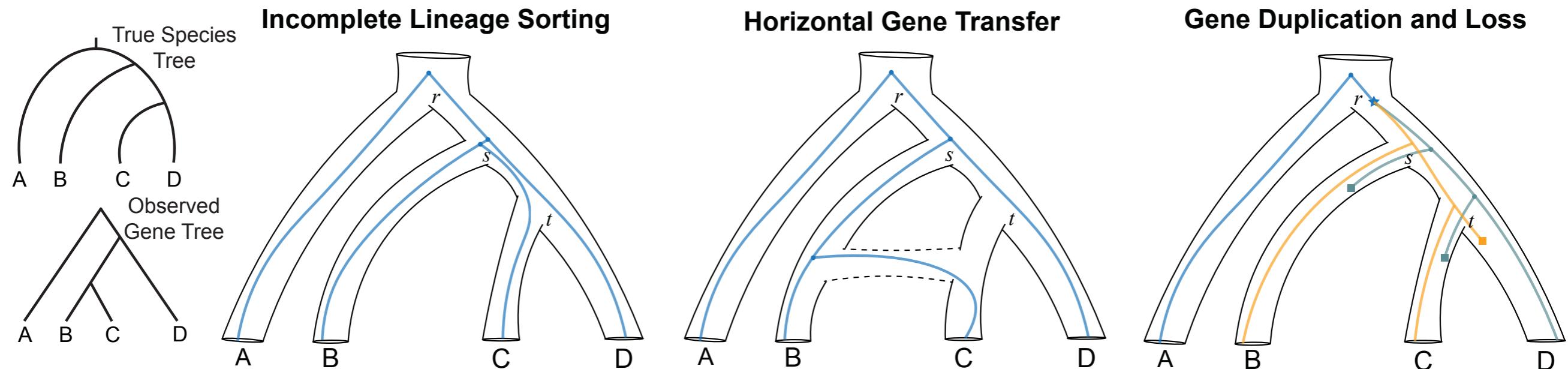
$$p_{BC|AD}(t, t_2, \gamma) = (1 - \gamma) \frac{1}{3} e^{-t} + \gamma (1 - \frac{2}{3} e^{-t_2})$$

(Meng, Kubatko, 2009)
(Yu, Degnan, Nakhleh, 2012)



George Tiley

Causes of Gene Tree Variation



Models of Processes

Multispecies Coalescent

Multispecies Network
Coalescent

Isolation-with-Migration

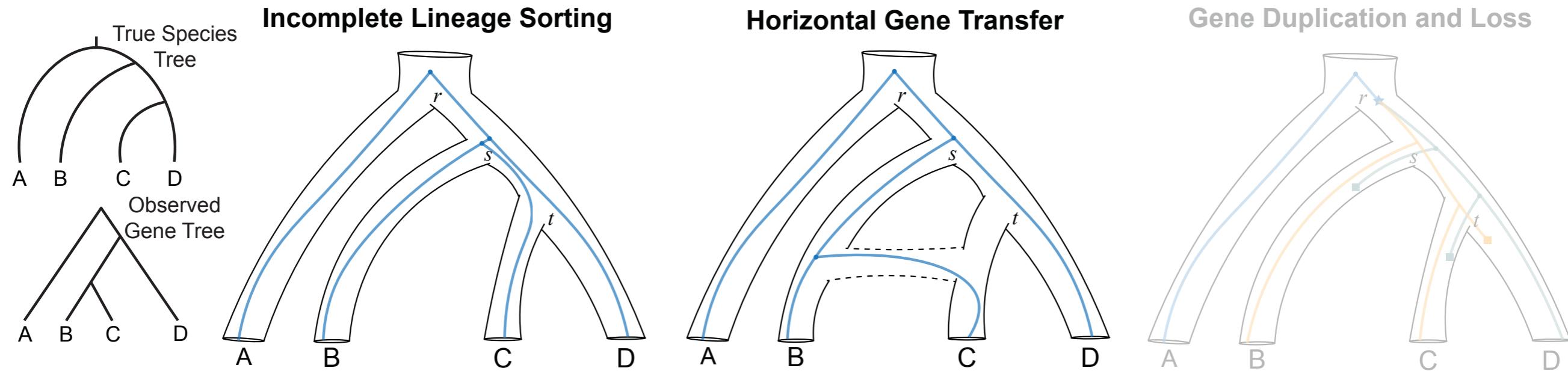
Duplication, Loss, and
Coalescence

Multispecies coalescent on a network



George Tiley

Causes of Gene Tree Variation



Models of Processes

Multispecies Coalescent

Multispecies Network Coalescent

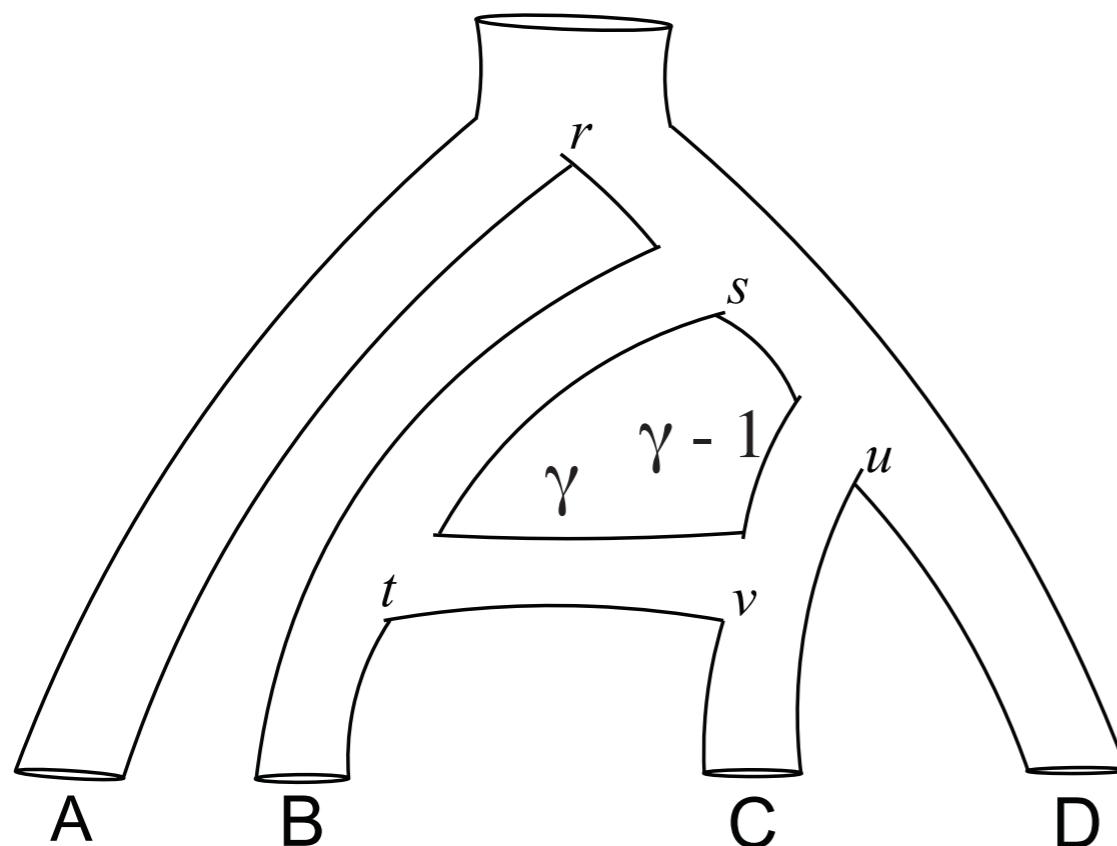
Isolation-with-Migration

Duplication, Loss, and Coalescence



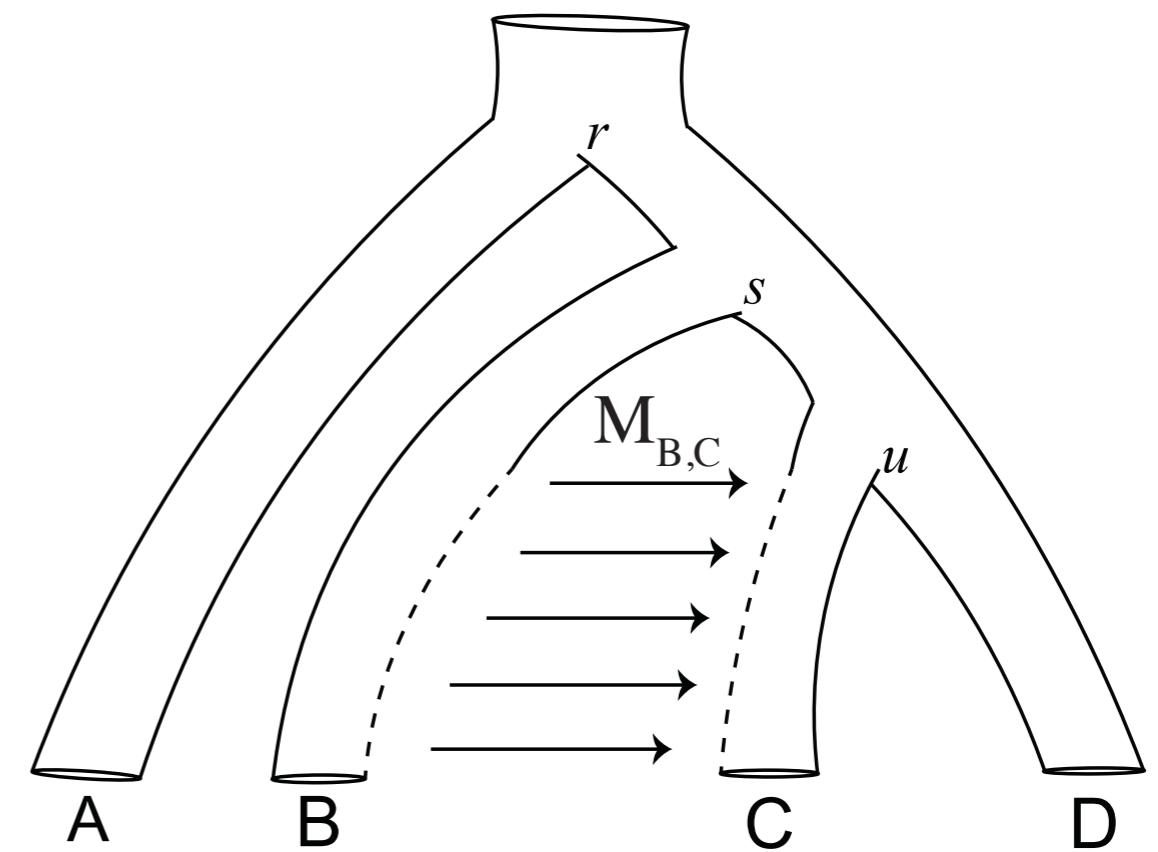
George Tiley

Episodic Gene Flow



Genes trace back from C into B
at time $\tau_t = \tau_v$ with probability γ .

Continuous Gene Flow

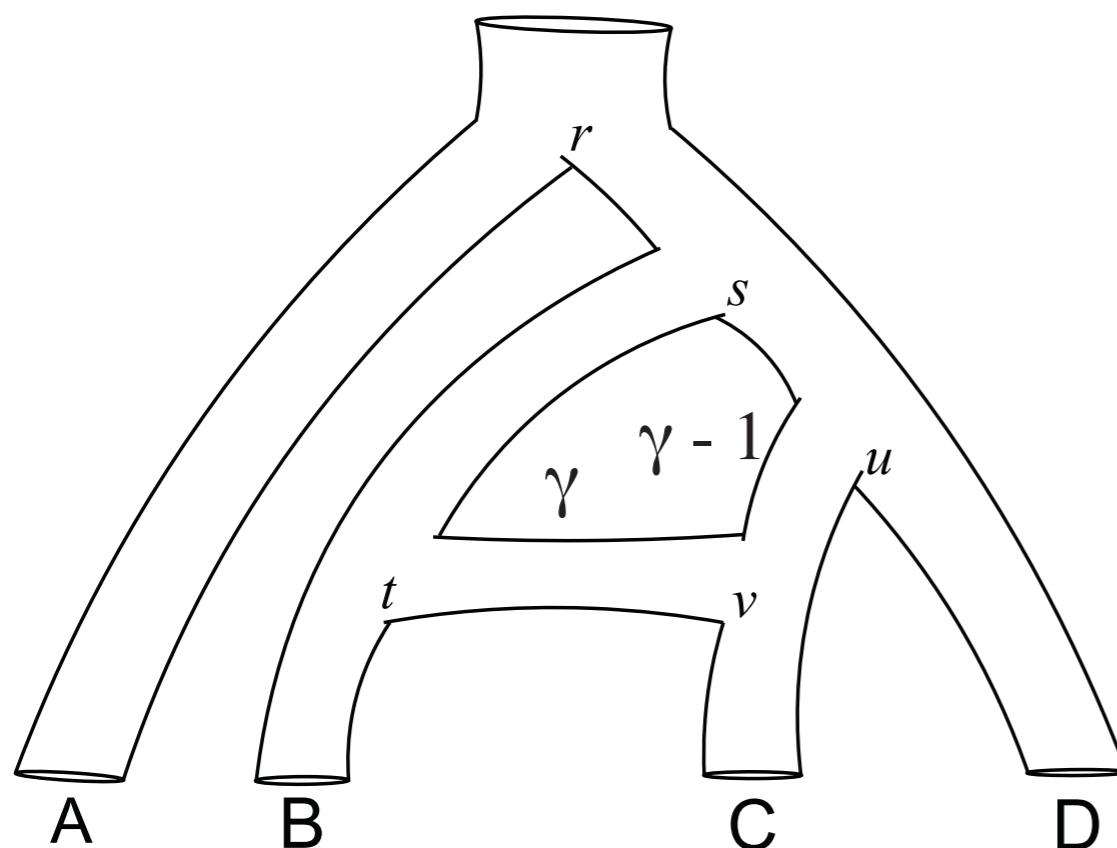


Genes trace back from C into B
at a rate of $M_{B,C}$ migrants per
generation over the interval τ_u .



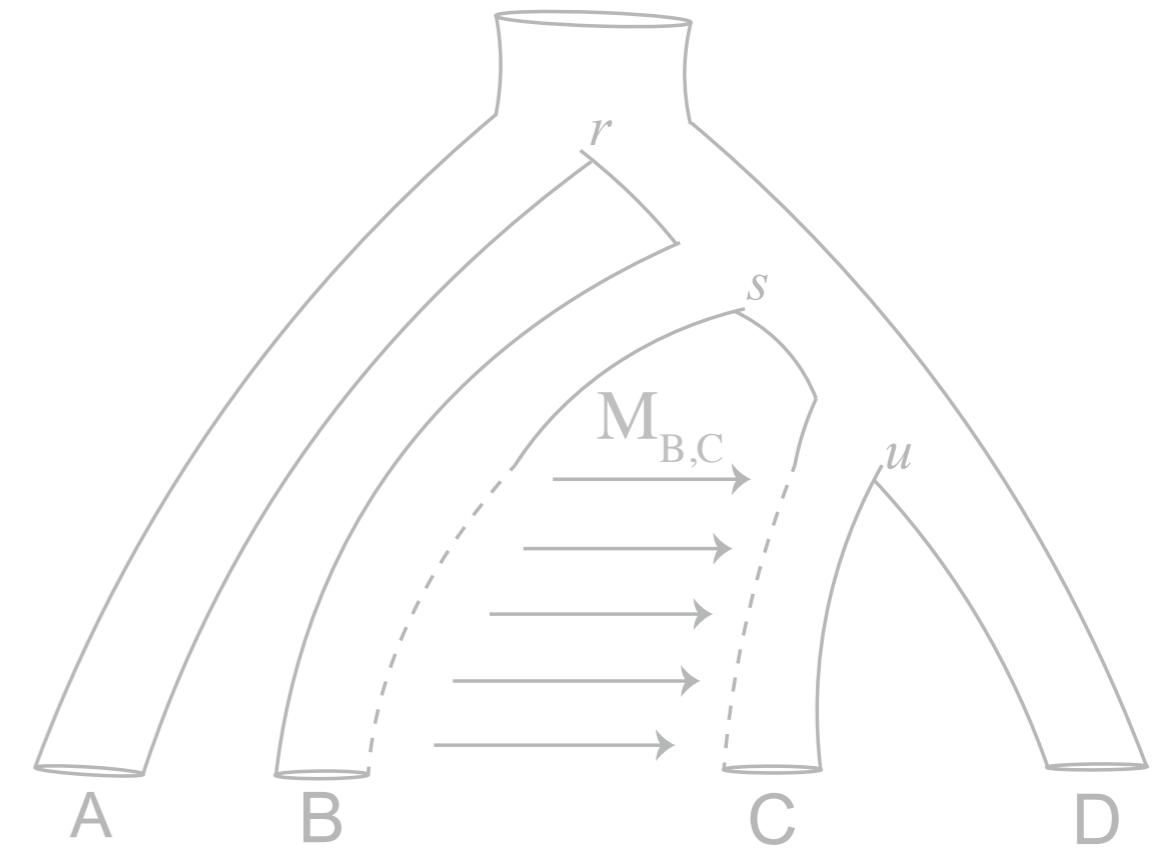
George Tiley

Episodic Gene Flow

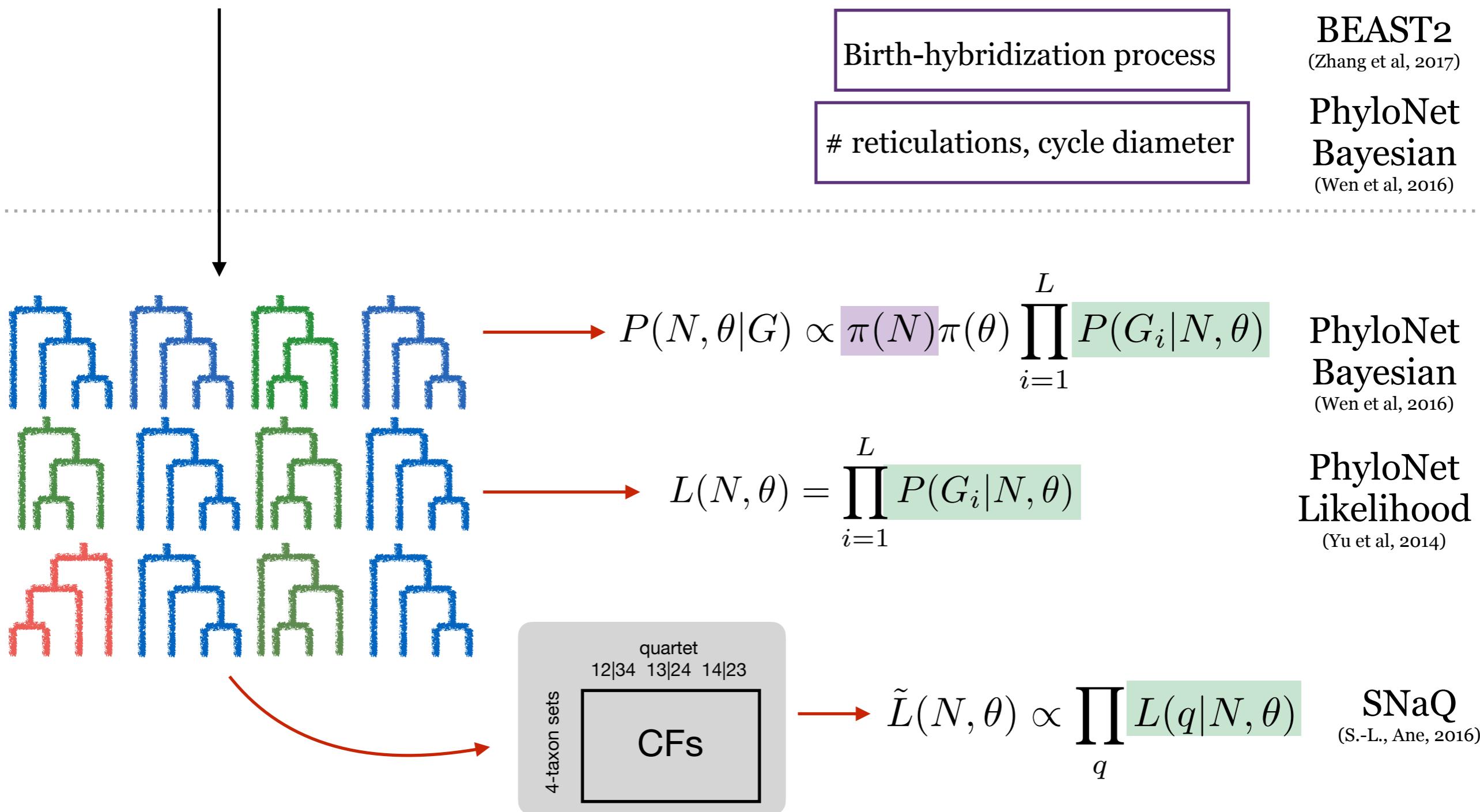
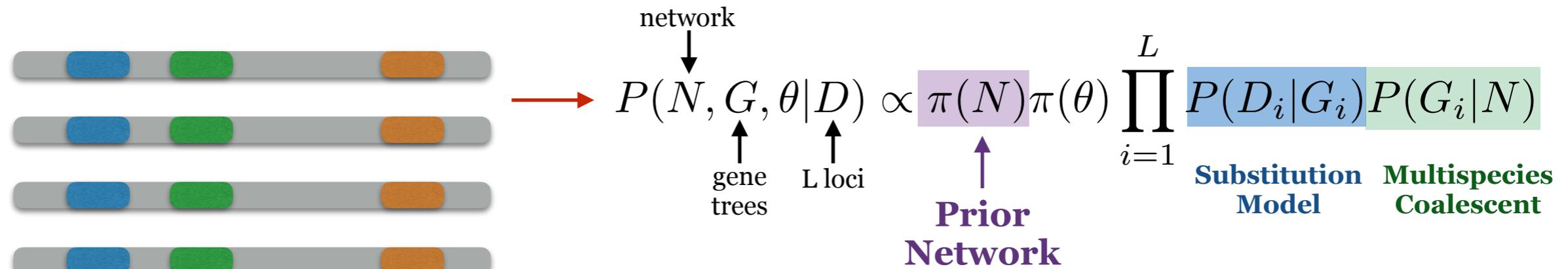


Genes trace back from C into B
at time $\tau_t = \tau_v$ with probability γ .

Continuous Gene Flow



Genes trace back from C into B
at a rate of $M_{B,C}$ migrants per
generation over the interval τ_u .



STEM-hy	gene trees rooted, BL	likelihood	hybridization b/w sister lineages
PhyloNet InferNetwork_ML	gene trees rooted	likelihood	
PhyloNet InferNetwork_MPL	gene trees rooted	triplet likelihood	
PhyloNetworks SNaQ	gene trees or quartet CFs	quartet likelihood	level-1 network
PhyloNet MCMC_GT	gene trees rooted	Bayesian	compound prior
PhyloNet MCMC_SEQ	alignments	Bayesian	compound prior no rate variation
BEAST2 SpeciesNetwork	alignments	Bayesian	birth-hyb prior
PhyloNet MLE_BiMarkers	biallelic sites	likelihood	compound prior
PhyloNet MCMC_BiMarkers	biallelic sites	Bayesian	compound prior
HyDe	sites	invariants	4 taxa, 1 hyb.

STEM-hy	gene trees rooted, BL	likelihood	hybridization b/w sister lineages
PhyloNet InferNetwork_ML	gene trees rooted	likelihood	
PhyloNet InferNetwork_MPL	gene trees rooted	triplet likelihood	
PhyloNetworks SNaQ	gene trees or quartet CFs	quartet likelihood	level-1 network
PhyloNet MCMC_GT	gene trees rooted	Bayesian	compound prior
PhyloNet MCMC_SEQ	alignments	Bayesian	compound prior no rate variation
BEAST2 SpeciesNetwork	alignments	Bayesian	birth-hyb prior
PhyloNet MLE_BiMarkers	biallelic sites	likelihood	compound prior
PhyloNet MCMC_BiMarkers	biallelic sites	Bayesian	compound prior
HyDe	sites	invariants	4 taxa, 1 hyb.
D-statistics	sites	freq. patterns	4 taxa, 1 hyb.
MSCQuartets	gene trees or quartet CF	quartet likelihood	4 taxa, 1 hyb.
NANUQ	gene trees or quartet CF	quartet likelihood	split (level-1) network
RF-Net	gene trees	min RF distance	No ILS
PhyNEST.jl	sites	quartet likelihood	level-1 network
PhyloDiamond.jl	gene trees or quartet CF	quartet likelihood	level-1 4-cycle network

STEM-hy	gene trees rooted, BL	likelihood	hybridization b/w sister lineages
PhyloNet InferNetwork_ML	gene trees rooted	likelihood	
PhyloNet InferNetwork_MPL	gene trees rooted	triplet likelihood	
PhyloNetworks SNaQ	gene trees or quartet CFs	quartet likelihood	level-1 network
PhyloNet MCMC_GT	gene trees rooted	Bayesian	compound prior
PhyloNet MCMC_SEQ	alignments	Bayesian	compound prior no rate variation
BEAST2 SpeciesNetwork	alignments	Bayesian	birth-hyb prior
PhyloNet MLE_BiMarkers	biallelic sites	likelihood	compound prior
PhyloNet MCMC_BiMarkers	biallelic sites	Bayesian	compound prior
HyDe	sites	invariants	4 taxa, 1 hyb.
D-statistics	sites	freq. patterns	4 taxa, 1 hyb.
MSCQuartets	gene trees or quartet CF	quartet likelihood	4 taxa, 1 hyb.
NANUQ	gene trees or quartet CF	quartet likelihood	split (level-1) network
RF-Net	gene trees	min RF distance	No ILS
PhyNEST.jl	sites	quartet likelihood	level-1 network
PhyloDiamond.jl	gene trees or quartet CF	quartet likelihood	level-1 4-cycle network

Hybrid detection methods

Estimate explicit network under the coalescent

Hybrid detection methods

STEM-hy	gene trees rooted, BL	likelihood	hybridization b/w sister lineages
PhyloNet InferNetwork_ML	gene trees rooted	likelihood	
PhyloNet InferNetwork_MPL	gene trees rooted	triplet likelihood	
Phylogenetworks SNaQ	gene trees or quartet CFs	quartet likelihood	level-1 network
PhyloNet MCMC_GT	gene trees rooted	Bayesian	compound prior
PhyloNet MCMC_SEQ	alignments	Bayesian	compound prior no rate variation
BEAST2 SpeciesNetwork	alignments	Bayesian	birth-hyb prior
PhyloNet MLE_BiMarkers	biallelic sites	likelihood	compound prior
PhyloNet MCMC_BiMarkers	biallelic sites	Bayesian	compound prior
HyDe	sites	invariants	4 taxa, 1 hyb.
D-statistics	sites	freq. patterns	4 taxa, 1 hyb.
MSCQuartets	gene trees or quartet CF	quartet likelihood	4 taxa, 1 hyb.
NANUQ	gene trees or quartet CF	quartet likelihood	split (level-1) network
RF-Net	gene trees	min RF distance	No ILS
PhyNEST.jl	sites	quartet likelihood	level-1 network
PhyloDiamond.jl	gene trees or quartet CF	quartet likelihood	level-1 4-cycle network

Estimate explicit network under the coalescent

Hybrid detection methods

Estimate split network

STEM-hy	gene trees rooted, BL	likelihood	hybridization b/w sister lineages
PhyloNet InferNetwork_ML	gene trees rooted	likelihood	
PhyloNet InferNetwork_MPL	gene trees rooted	triplet likelihood	
Phylogenetworks SNaQ	gene trees or quartet CFs	quartet likelihood	level-1 network
PhyloNet MCMC_GT	gene trees rooted	Bayesian	compound prior
PhyloNet MCMC_SEQ	alignments	Bayesian	compound prior no rate variation
BEAST2 SpeciesNetwork	alignments	Bayesian	birth-hyb prior
PhyloNet MLE_BiMarkers	biallelic sites	likelihood	compound prior
PhyloNet MCMC_BiMarkers	biallelic sites	Bayesian	compound prior
HyDe	sites	invariants	4 taxa, 1 hyb.
D-statistics	sites	freq. patterns	4 taxa, 1 hyb.
MSCQuartets	gene trees or quartet CF	quartet likelihood	4 taxa, 1 hyb.
NANUQ	gene trees or quartet CF	quartet likelihood	split (level-1) network
RF-Net	gene trees	min RF distance	No ILS
PhyNEST.jl	sites	quartet likelihood	level-1 network
PhyloDiamond.jl	gene trees or quartet CF	quartet likelihood	level-1 4-cycle network



Cécile Ané

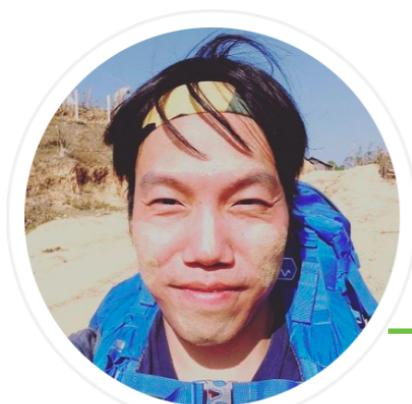
STEM-hy	gene trees rooted, BL	likelihood	hybridization b/w sister lineages
PhyloNet InferNetwork_ML	gene trees rooted	likelihood	
PhyloNet InferNetwork_MPL	gene trees rooted	triplet likelihood	
PhyloNetworks SNaQ	gene trees or quartet CFs	quartet likelihood	level-1 network
PhyloNet MCMC_GT	gene trees rooted	Bayesian	compound prior
PhyloNet MCMC_SEQ	alignments	Bayesian	compound prior no rate variation
BEAST2 SpeciesNetwork	alignments	Bayesian	birth-hyb prior
PhyloNet MLE_BiMarkers	biallelic sites	likelihood	compound prior
PhyloNet MCMC_BiMarkers	biallelic sites	Bayesian	compound prior
HyDe	sites	invariants	4 taxa, 1 hyb.
D-statistics	sites	freq. patterns	4 taxa, 1 hyb.
MSCQuartets	gene trees or quartet CF	quartet likelihood	4 taxa, 1 hyb.
NANUQ	gene trees or quartet CF	quartet likelihood	split (level-1) network
RF-Net	gene trees	min RF distance	No ILS
PhyNEST.jl	sites	quartet likelihood	level-1 network
PhyloDiamond.jl	gene trees or quartet CF	quartet likelihood	level-1 4-cycle network

Estimate
explicit
network
under the
coalescent

Hybrid
detection
methods
Estimate
split
network



Cécile Ané



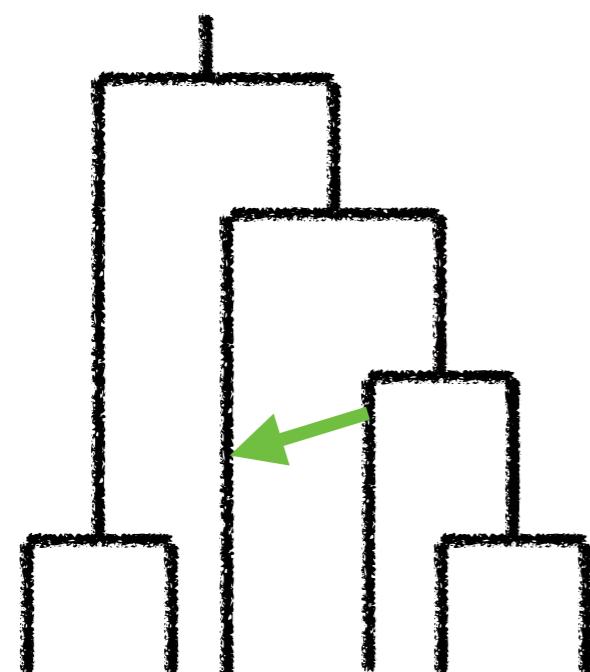
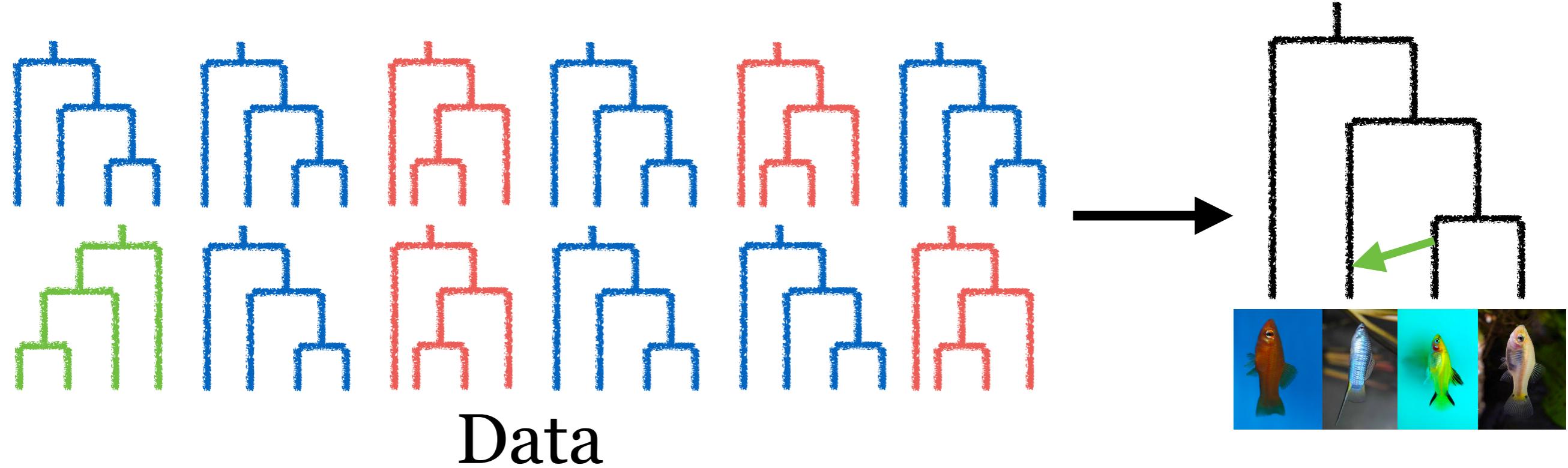
Sungsik (Kevin)
Kong

STEM-hy	gene trees rooted, BL	likelihood	hybridization b/w sister lineages
PhyloNet InferNetwork_ML	gene trees rooted	likelihood	
PhyloNet InferNetwork_MPL	gene trees rooted	triplet likelihood	
PhyloNetworks SNaQ	gene trees or quartet CFs	quartet likelihood	level-1 network
PhyloNet MCMC_GT	gene trees rooted	Bayesian	compound prior
PhyloNet MCMC_SEQ	alignments	Bayesian	compound prior no rate variation
BEAST2 SpeciesNetwork	alignments	Bayesian	birth-hyb prior
PhyloNet MLE_BiMarkers	biallelic sites	likelihood	compound prior
PhyloNet MCMC_BiMarkers	biallelic sites	Bayesian	compound prior
HyDe	sites	invariants	4 taxa, 1 hyb.
D-statistics	sites	freq. patterns	4 taxa, 1 hyb.
MSCQuartets	gene trees or quartet CF	quartet likelihood	4 taxa, 1 hyb.
NANUQ	gene trees or quartet CF	quartet likelihood	split (level-1) network
RF-Net	gene trees	min RF distance	No ILS
PhyNEST.jl	sites	quartet likelihood	level-1 network
PhyloDiamond.jl	gene trees or quartet CF	quartet likelihood	level-1 4-cycle network

Estimate
explicit
network
under the
coalescent

Hybrid
detection
methods
Estimate
split
network

Maximum pseudolikelihood



Quartet-based inference

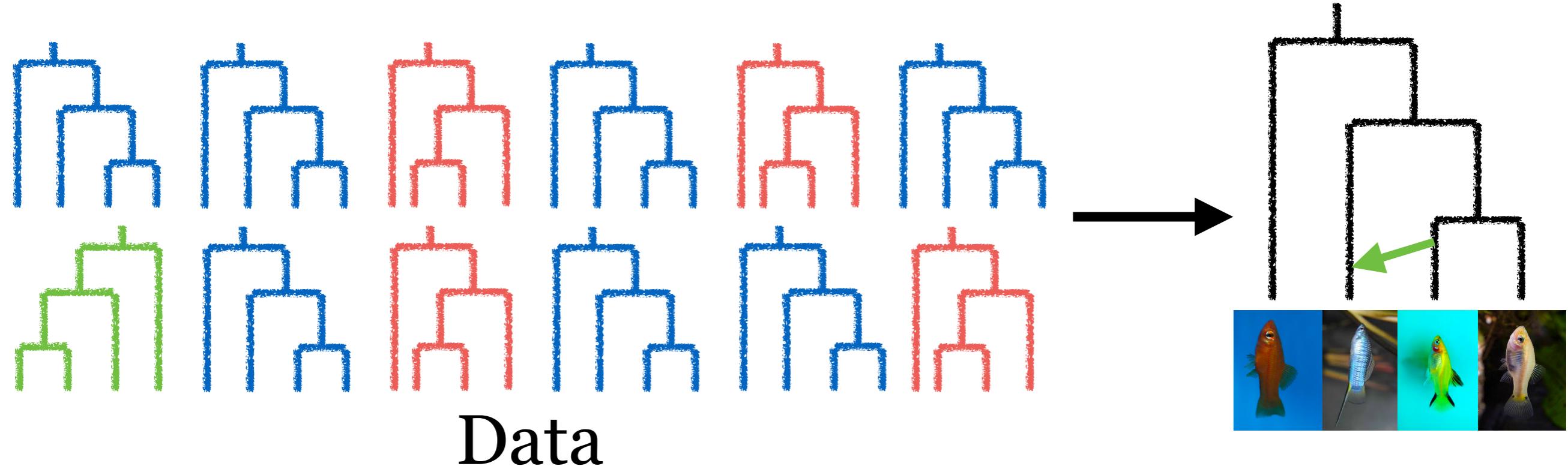
$$\tilde{L}(\textit{network}) = \prod L(\textit{quartet})$$

(S-L, Ané, 2016, PLoS Genetics)

www.github.com/CRSL4/PhyloNetworks

snaQ julia

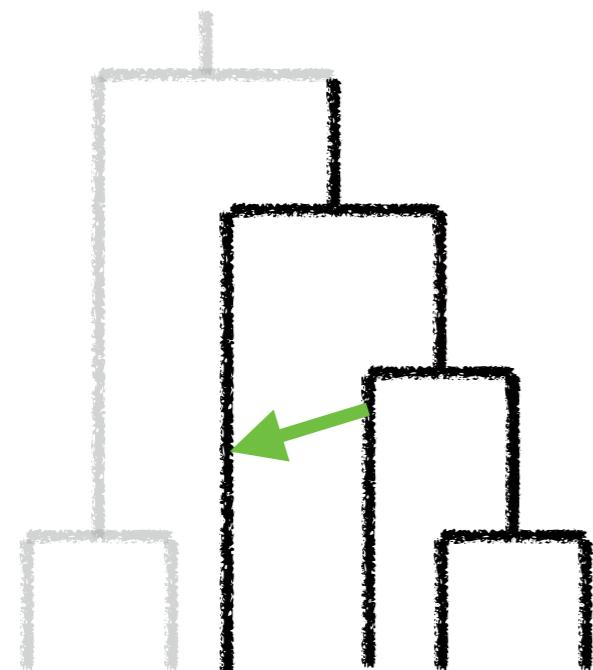
Maximum pseudolikelihood



$$\tilde{L}(\textit{network}) = \prod L(\textit{quartet})$$

(S-L, Ané, 2016, PLoS Genetics)

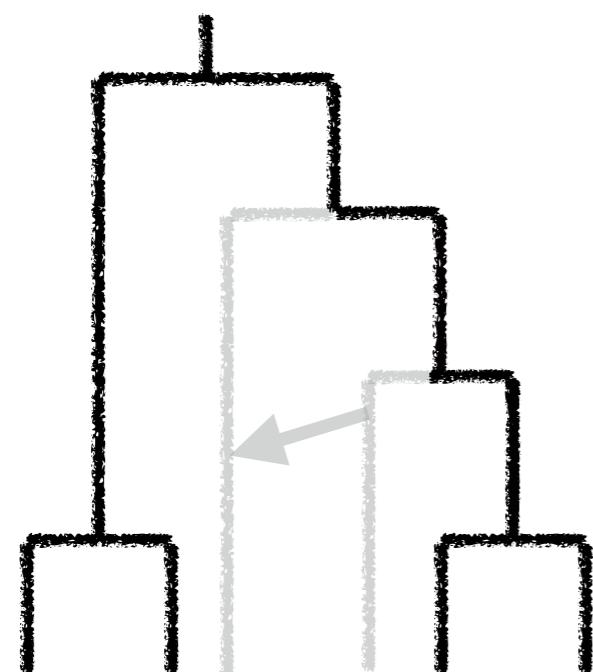
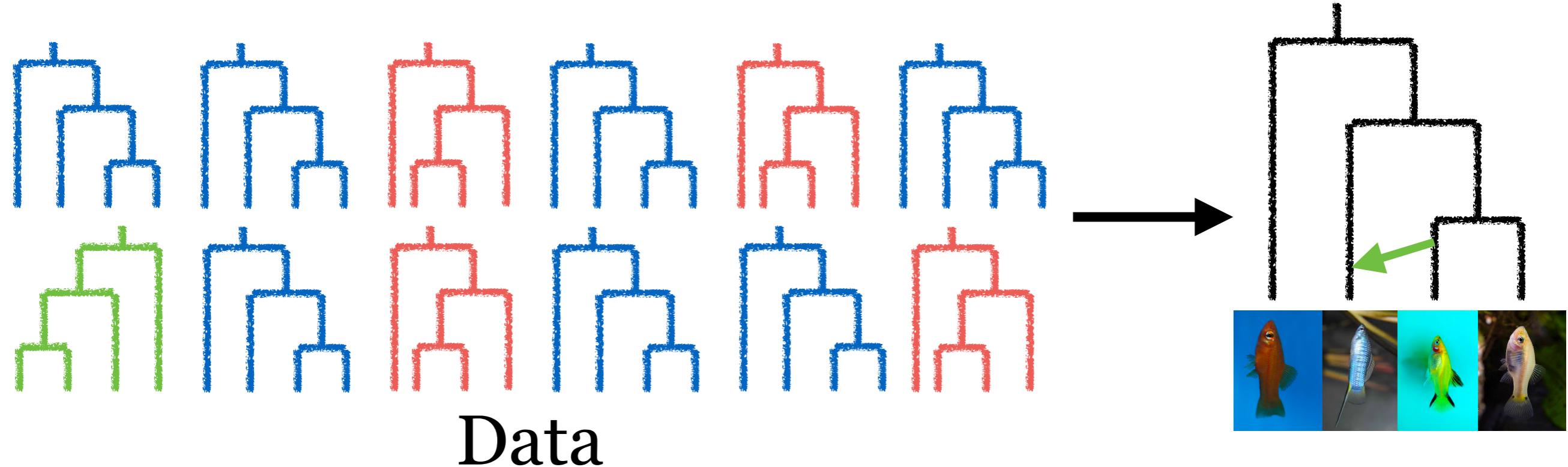
www.github.com/CRSL4/PhyloNetworks



Quartet-based inference

snaQ julia

Maximum pseudolikelihood



Quartet-based inference

$$\tilde{L}(\textit{network}) = \prod L(\textit{quartet})$$

(S-L, Ané, 2016, PLoS Genetics)

www.github.com/CRSL4/PhyloNetworks

snaQ julia

Maximum pseudolikelihood

Unrooted gene trees

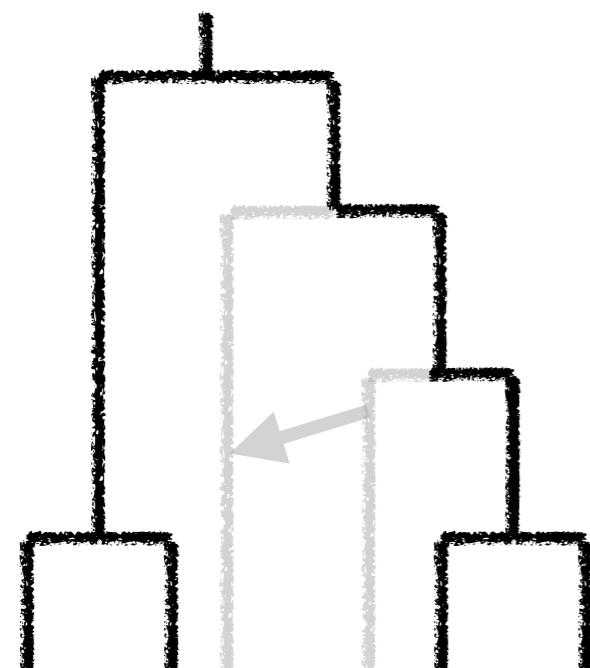
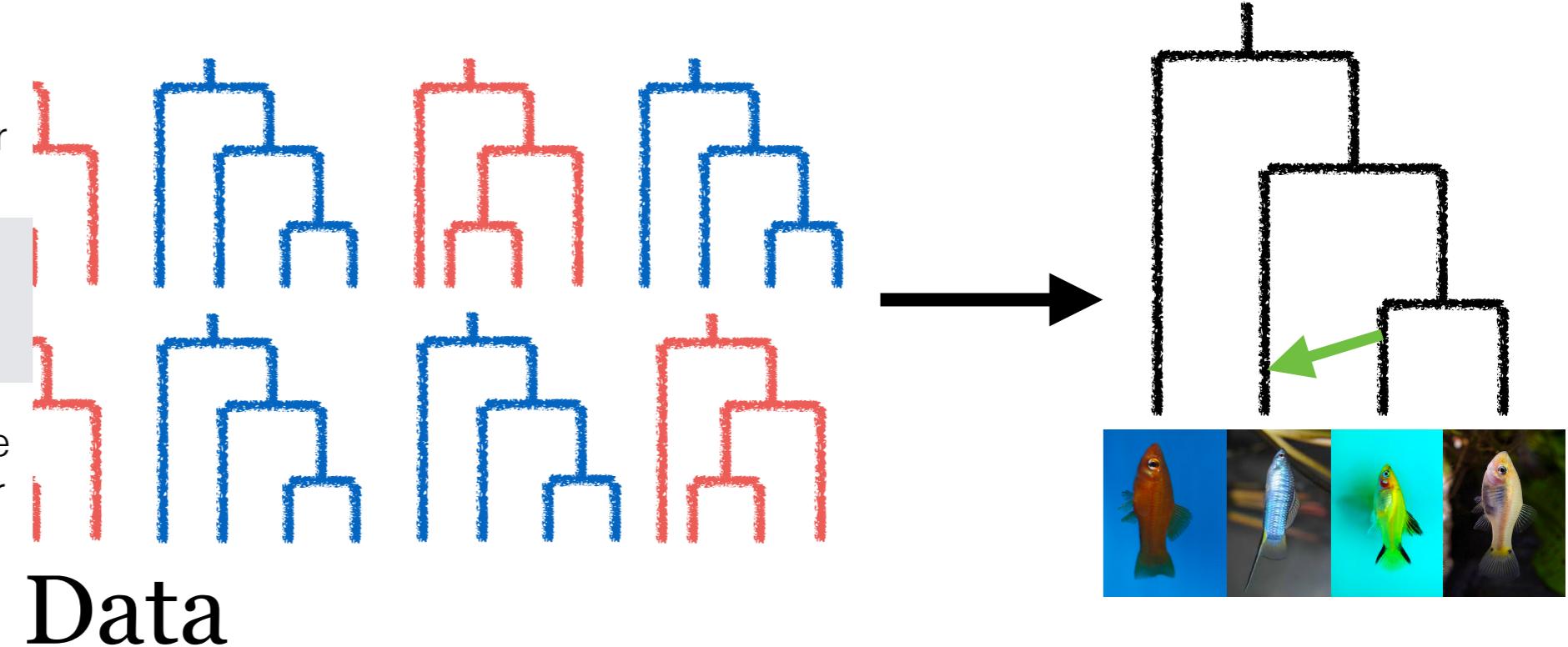
No branch lengths

Concordance factors

No rooting error

No molecular clock assumption

Account for tree estimation error



$$\tilde{L}(\textit{network}) = \prod L(\textit{quartet})$$

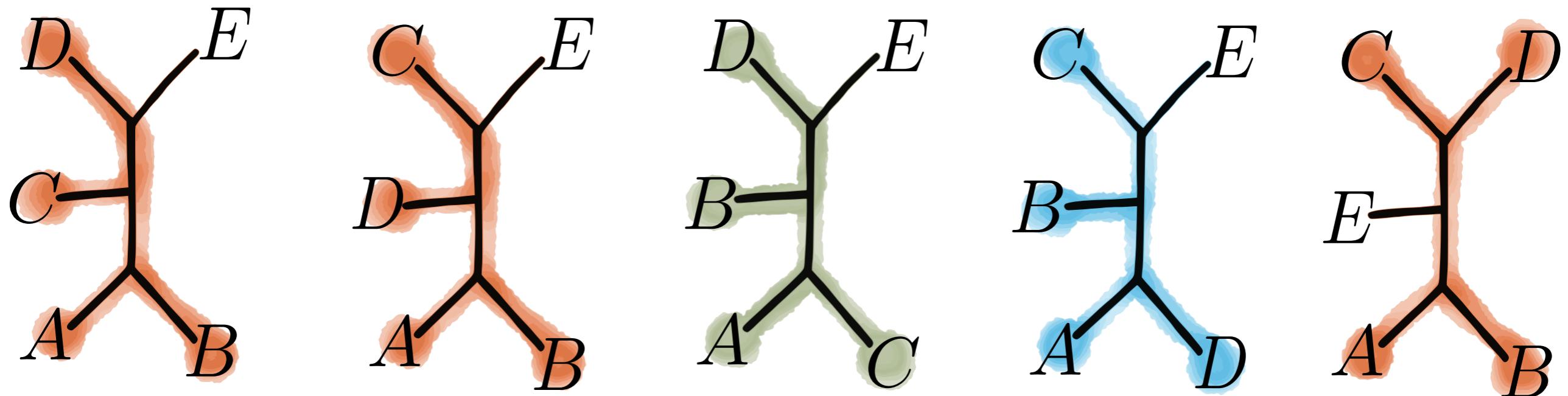
(S-L, Ané, 2016, PLoS Genetics)

www.github.com/CRSL4/PhyloNetworks

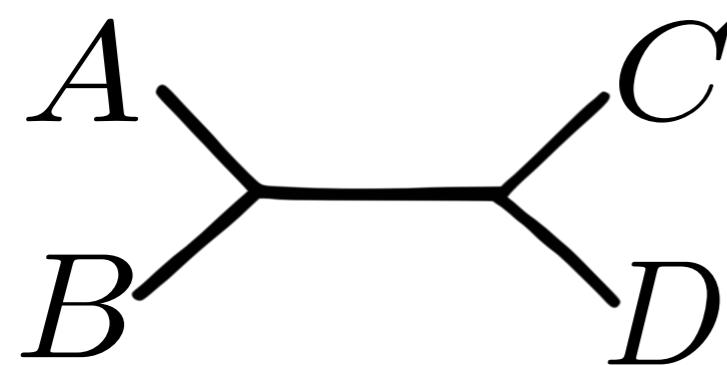
Quartet-based inference

snaQ julia

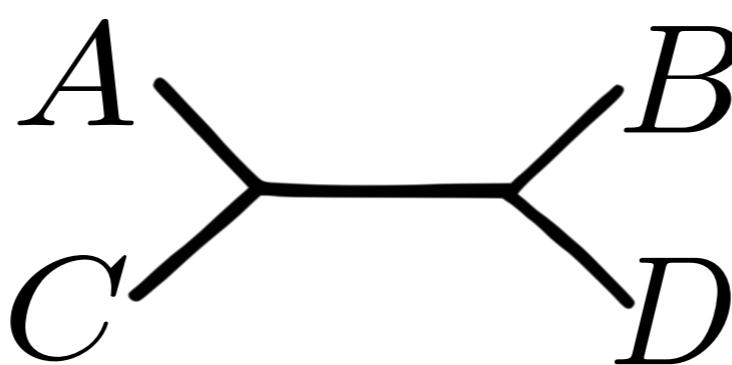
Quartet-based inference



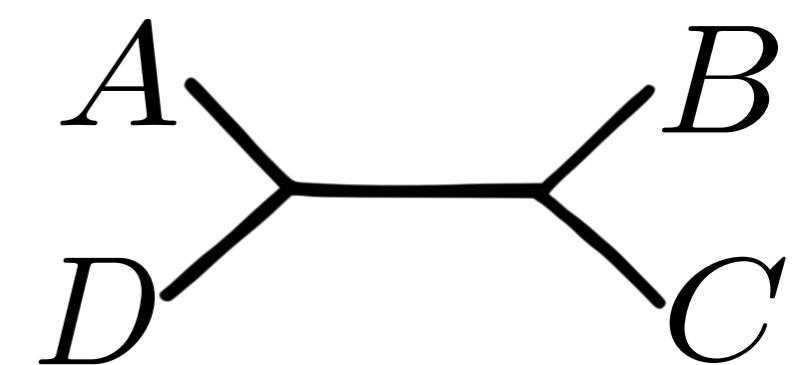
Concordance factors (CF):
% of genes having the quartet in their tree



3/5



1/5



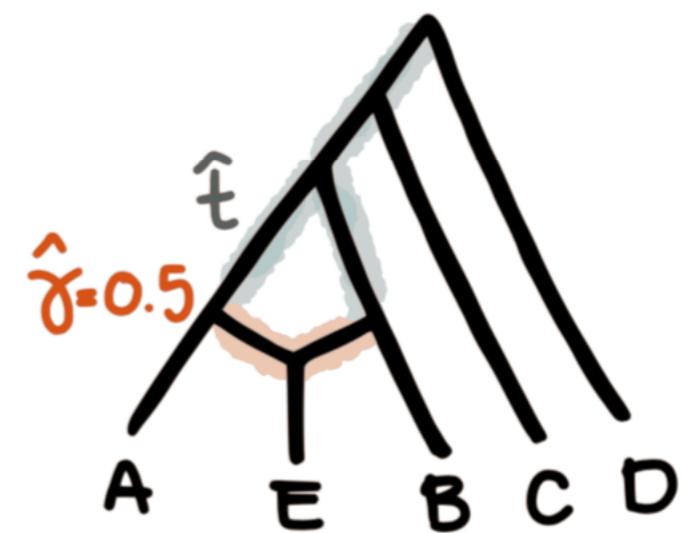
1/5

Quartet-based inference

Observed **quartet** CFs:

4 taxon set	CF_1	CF_2	CF_3
A B C D	.80	.10	.10
A B C E	.40	.40	.20
A B D E	.40	.40	.20
A C D E	.84	.08	.08
B C D E	.82	.10	.08

inferred network:



Maximum Pseudo-Likelihood:

$$\log \tilde{L} = \sum_{q \in Q(N)} CF_{in,1} \log(CF_{net,1}) + CF_{in,2} \log(CF_{net,2}) + CF_{in,3} \log(CF_{net,3})$$

I want to find signal of hybridization in my data

Hybrid detection
methods

HyDe

ABBA-BABA

MSCQuartets

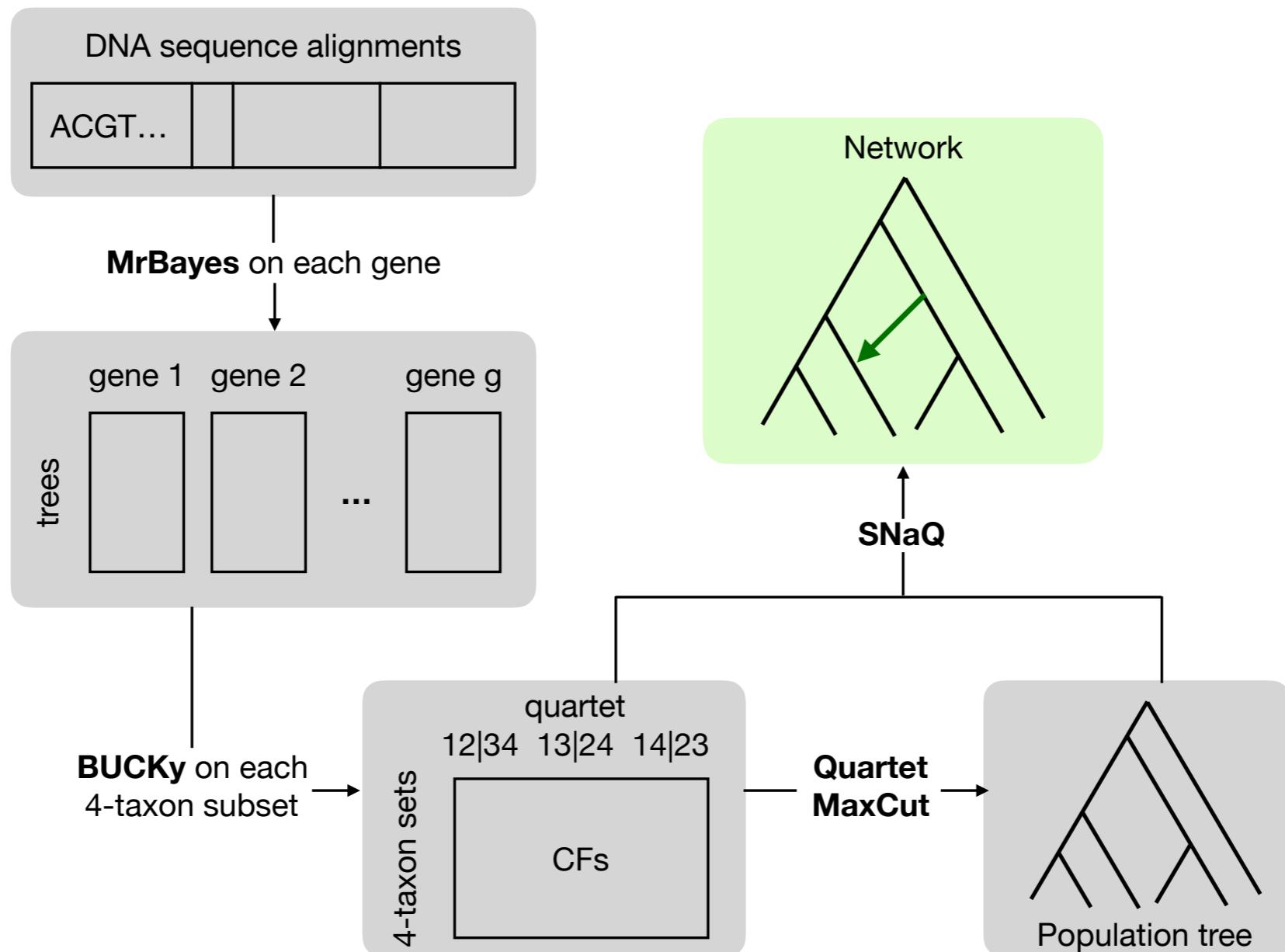
Network inference
methods

PhyloNet
SNaQ
BEAST2
NANUQ
RF-Net
PhyNEST



How?

Phylogenetic network



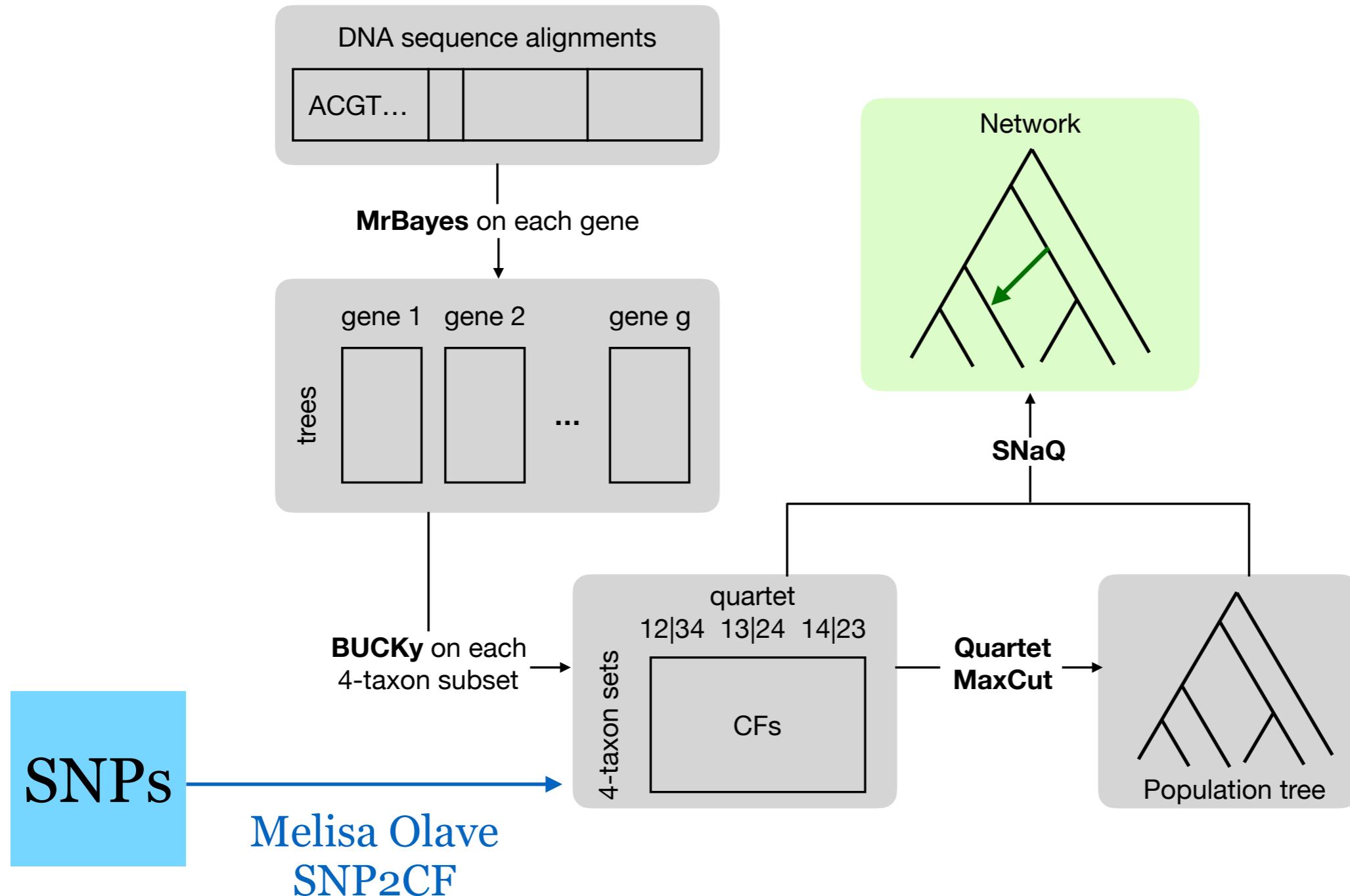
Cécile Ané



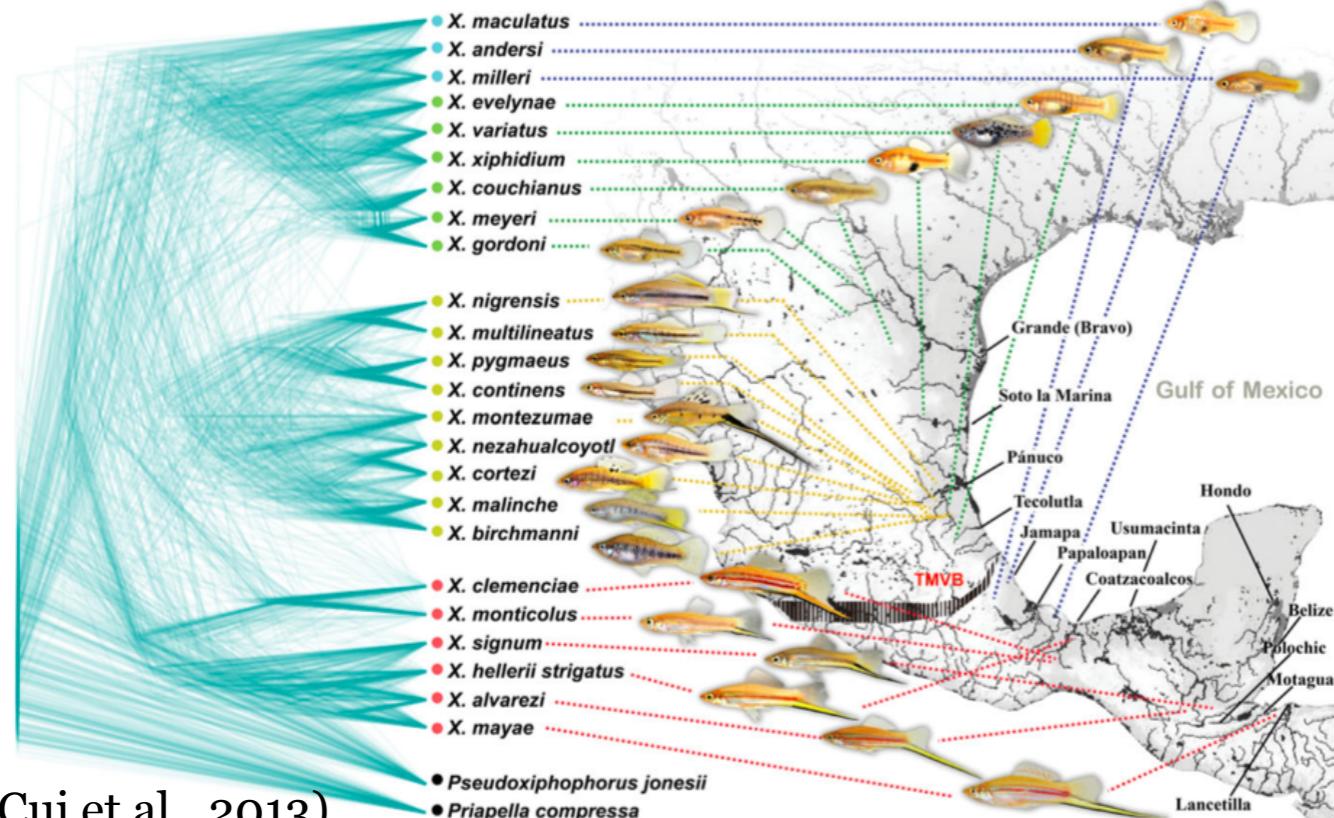
Cécile Ané

How?

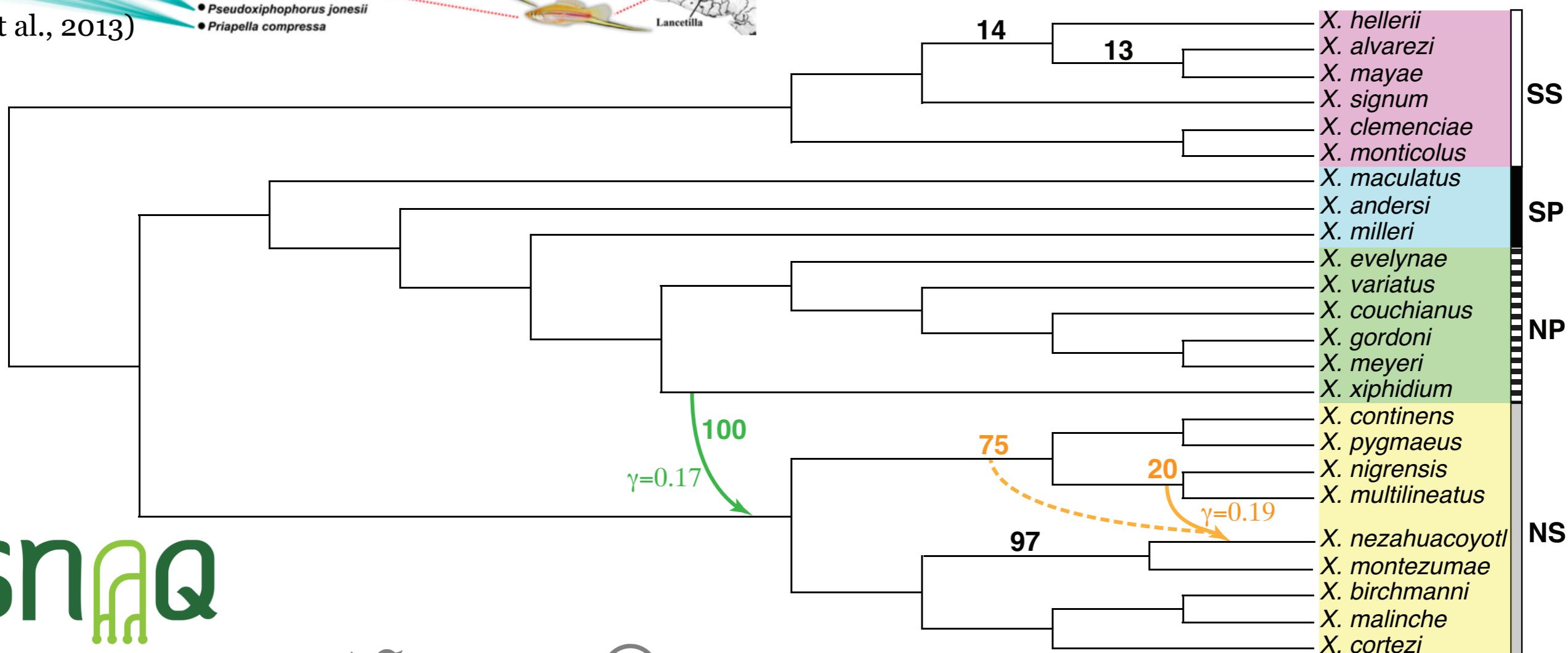
Phylogenetic network



1183 genes, 24 swordtails and platyfish



Xiphophorus fish data



snaQ



<http://solislemuslab.github.io/>



@solislemuslab



crsl4

(Solís-Lemus, Ané, 2016, PLoS Genetics)

Network challenges



Interpretation



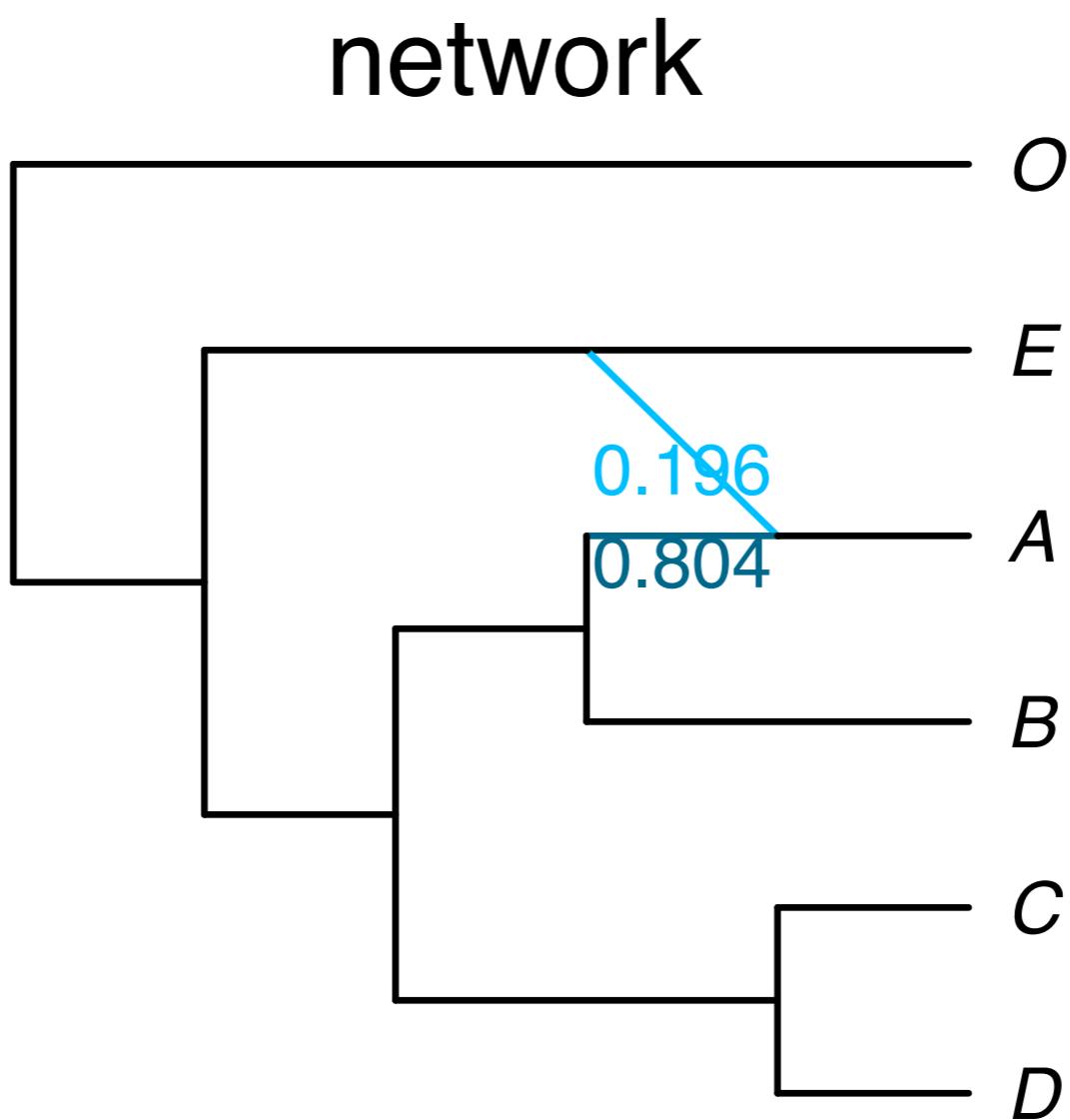
Scalability



Identifiability

Interpretation

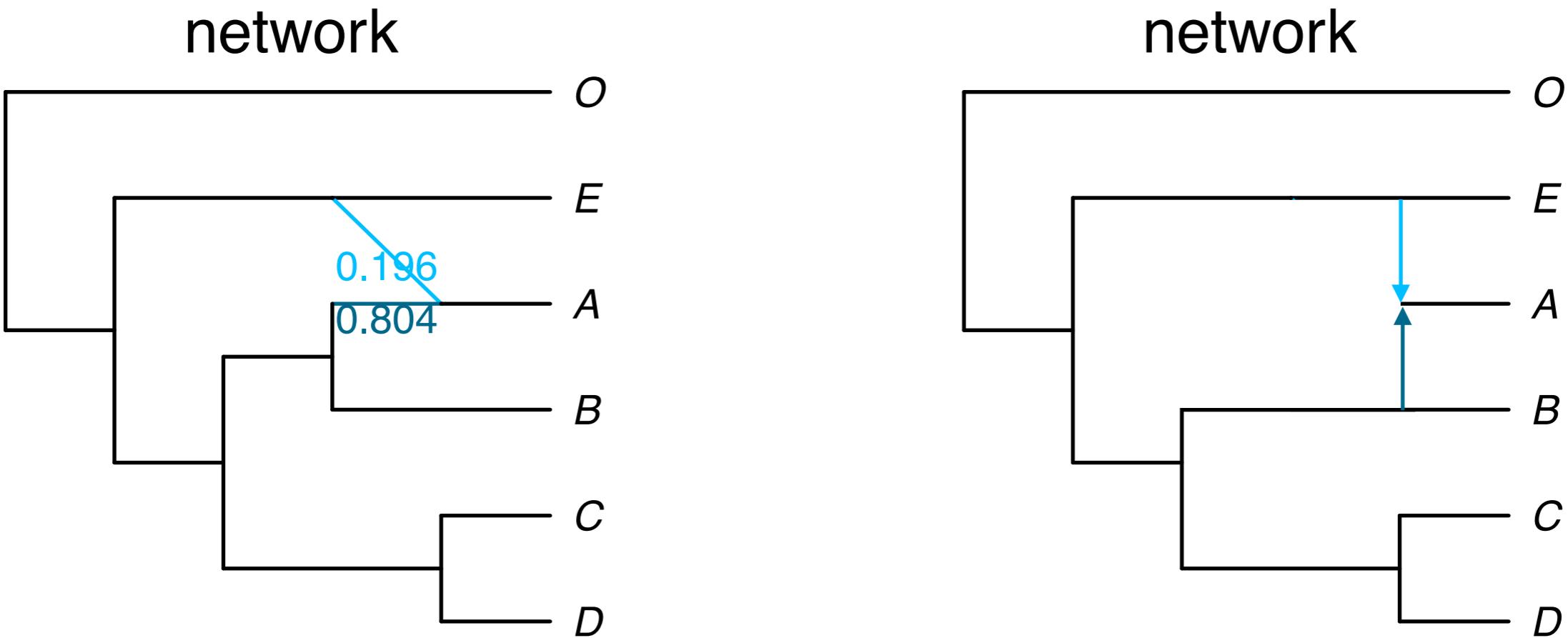
Phylogenetic networks



What does this represent?

Interpretation

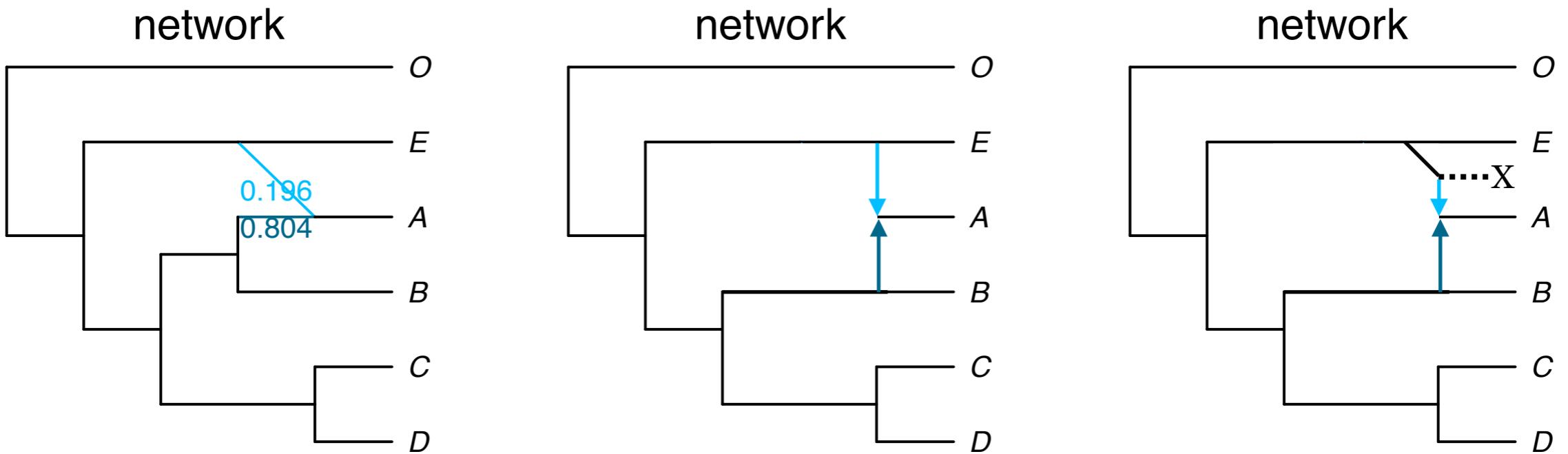
Phylogenetic networks



What does this represent?

Interpretation

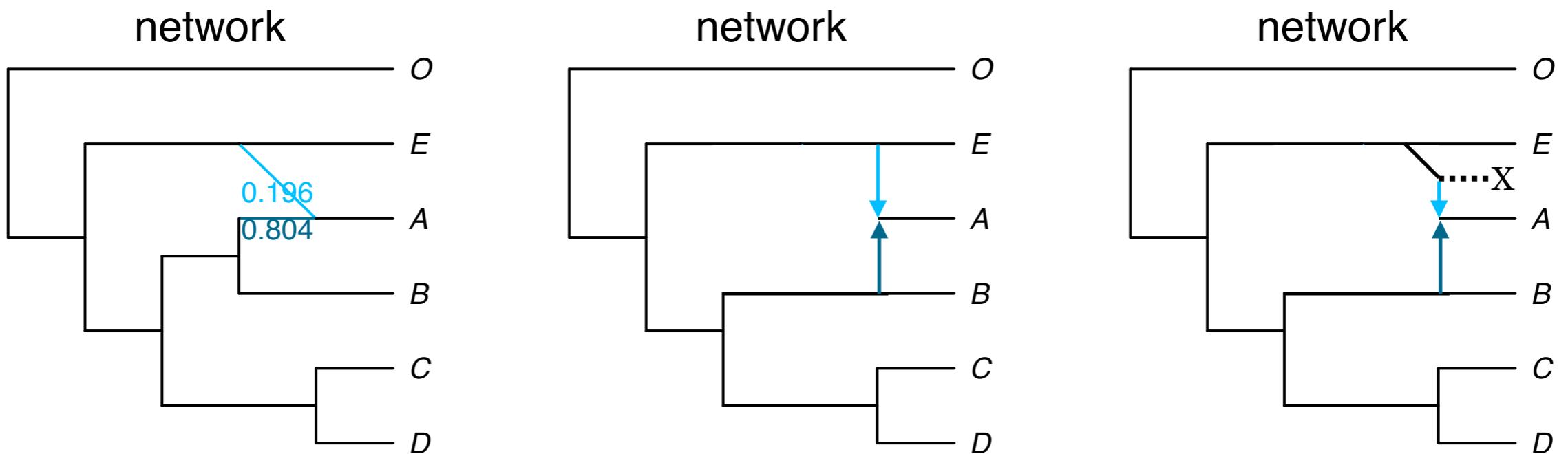
Phylogenetic networks



What does this represent?

Interpretation

Phylogenetic networks



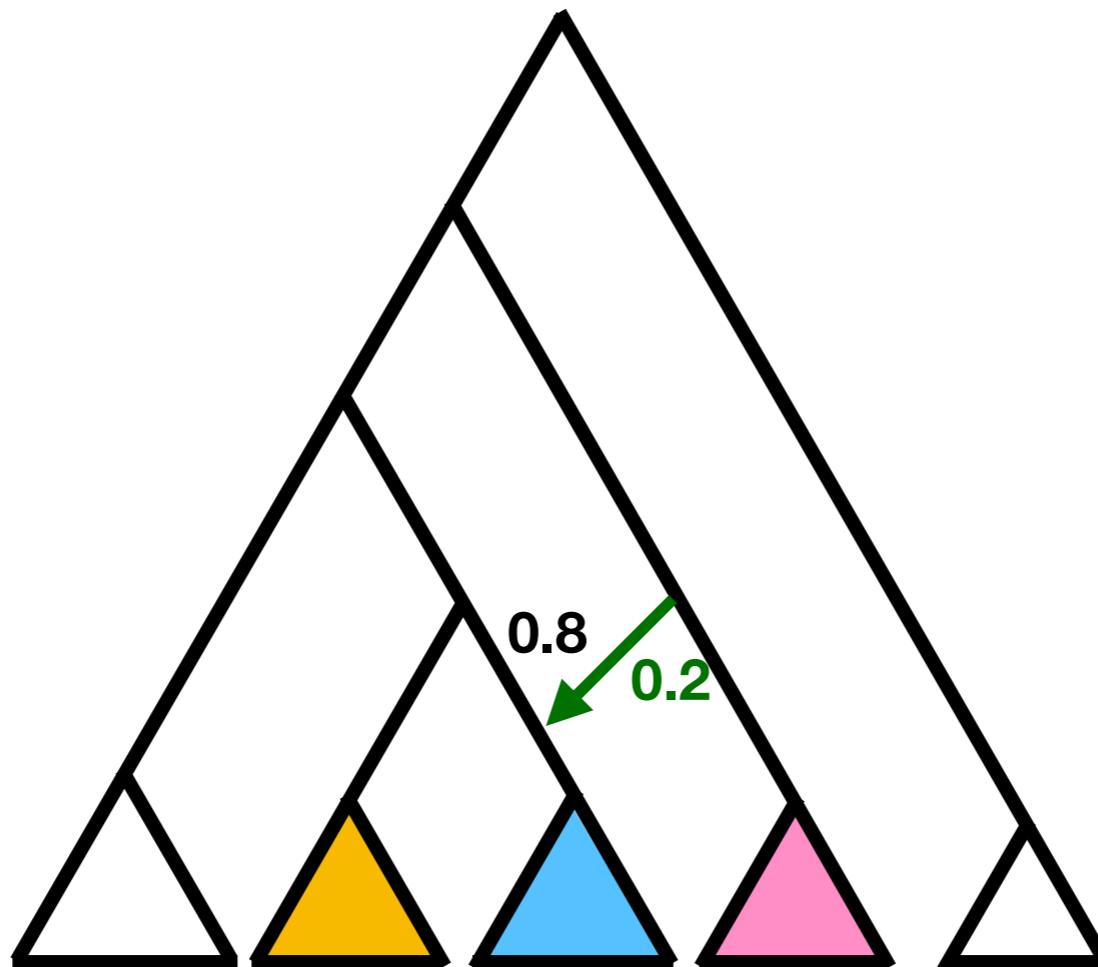
What does this represent?

How can we know which one is it?

We need more information, e.g. dating

Interpretation

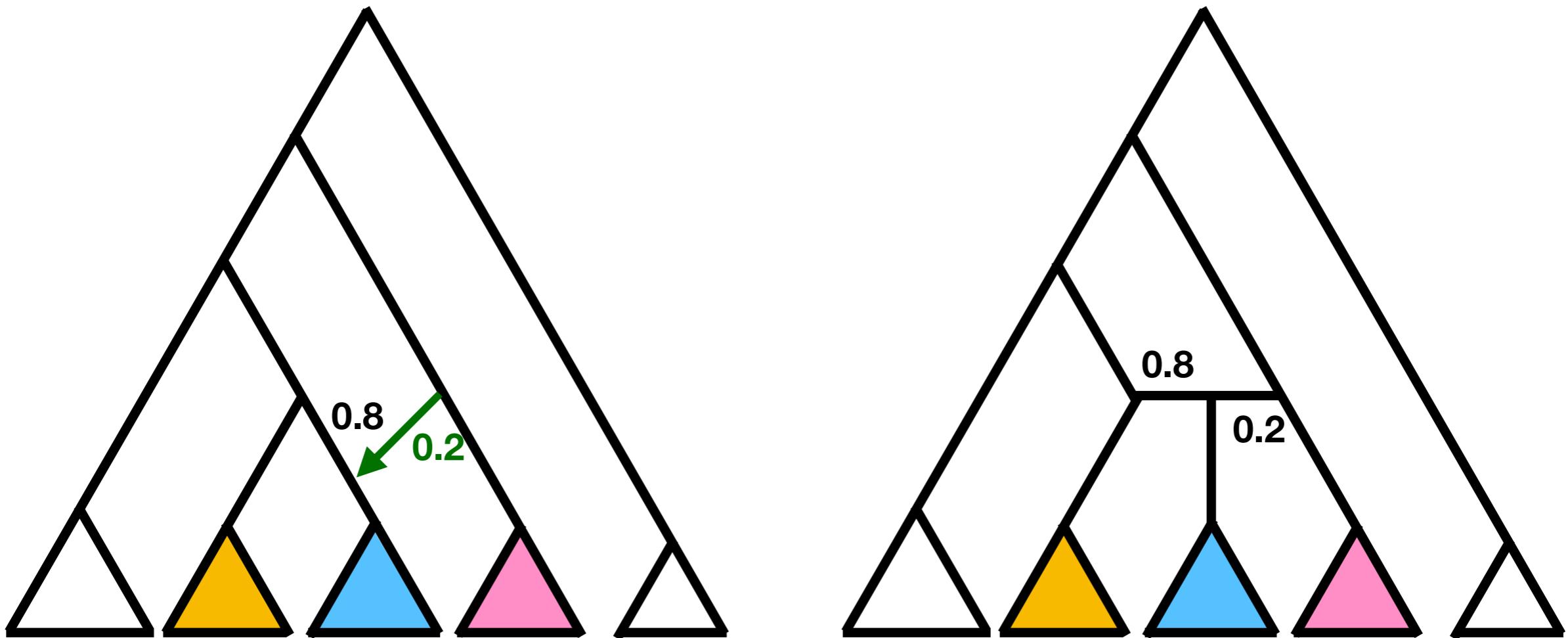
Phylogenetic networks



Example: How else can we draw this?

Interpretation

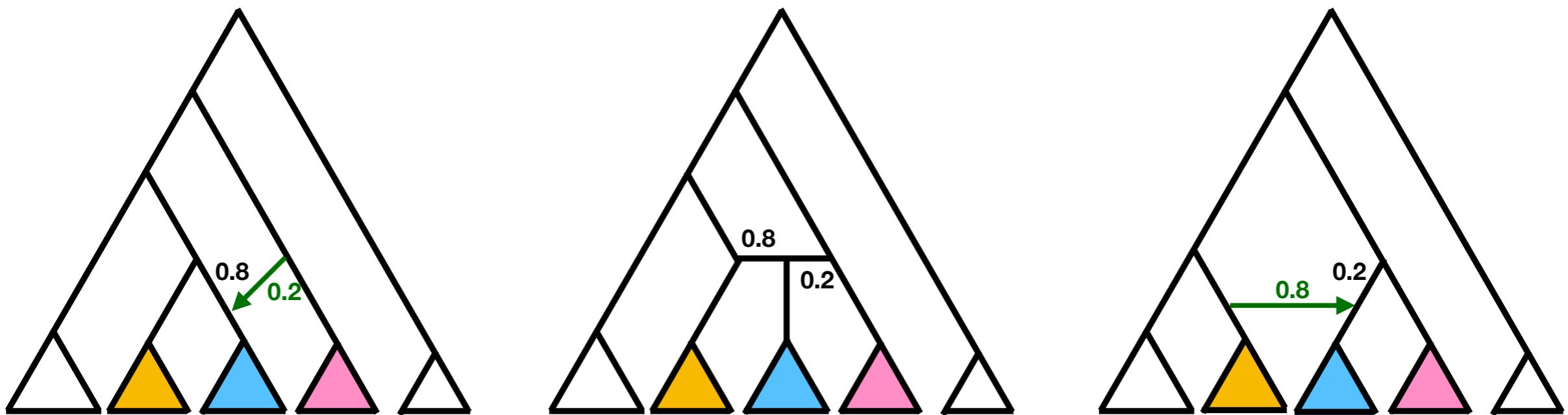
Phylogenetic networks



Example: How else can we draw this?

Interpretation

Phylogenetic networks

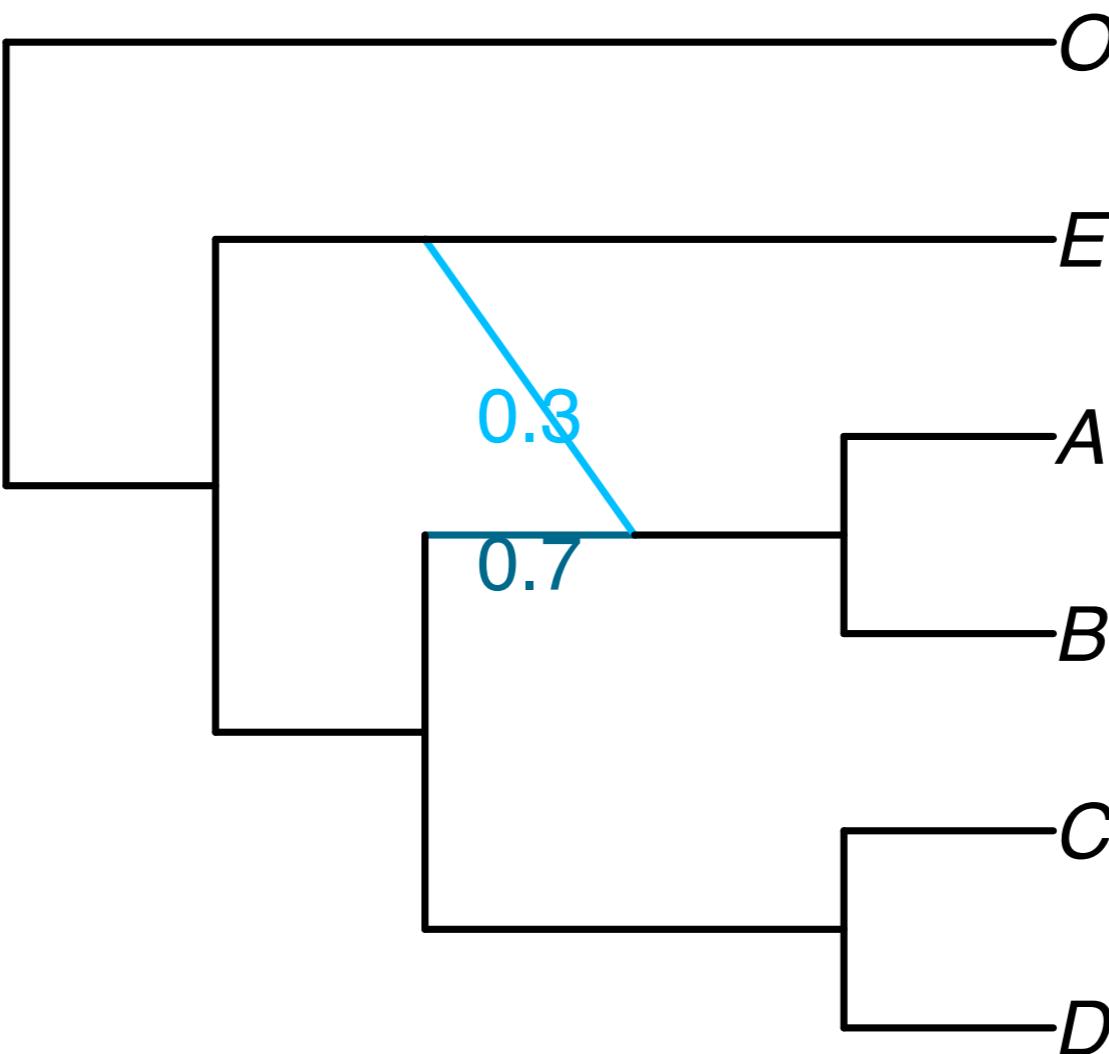


Example: How else can we draw this?

Interpretation

Phylogenetic networks

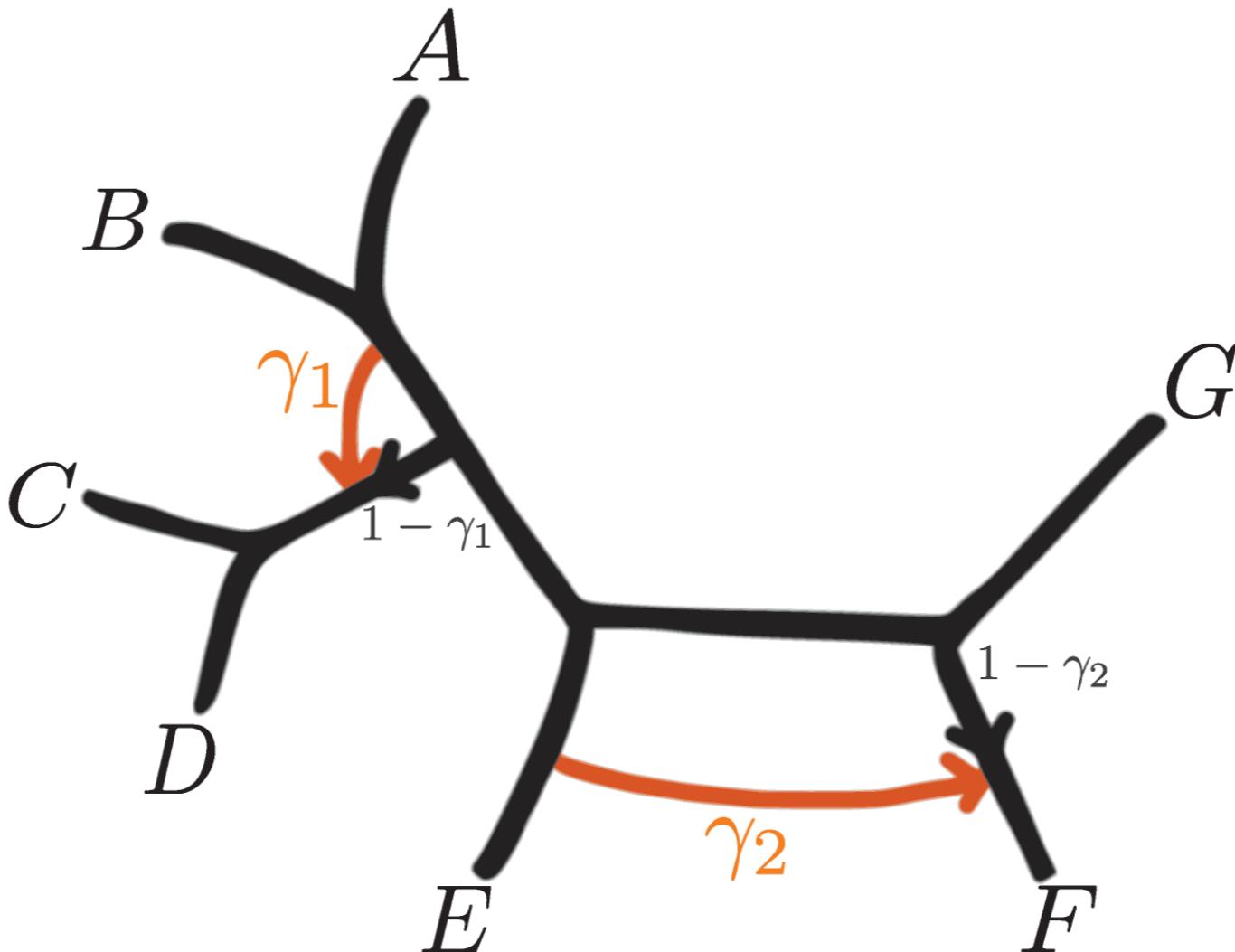
other network



Exercise: How else can we draw this?

Interpretation

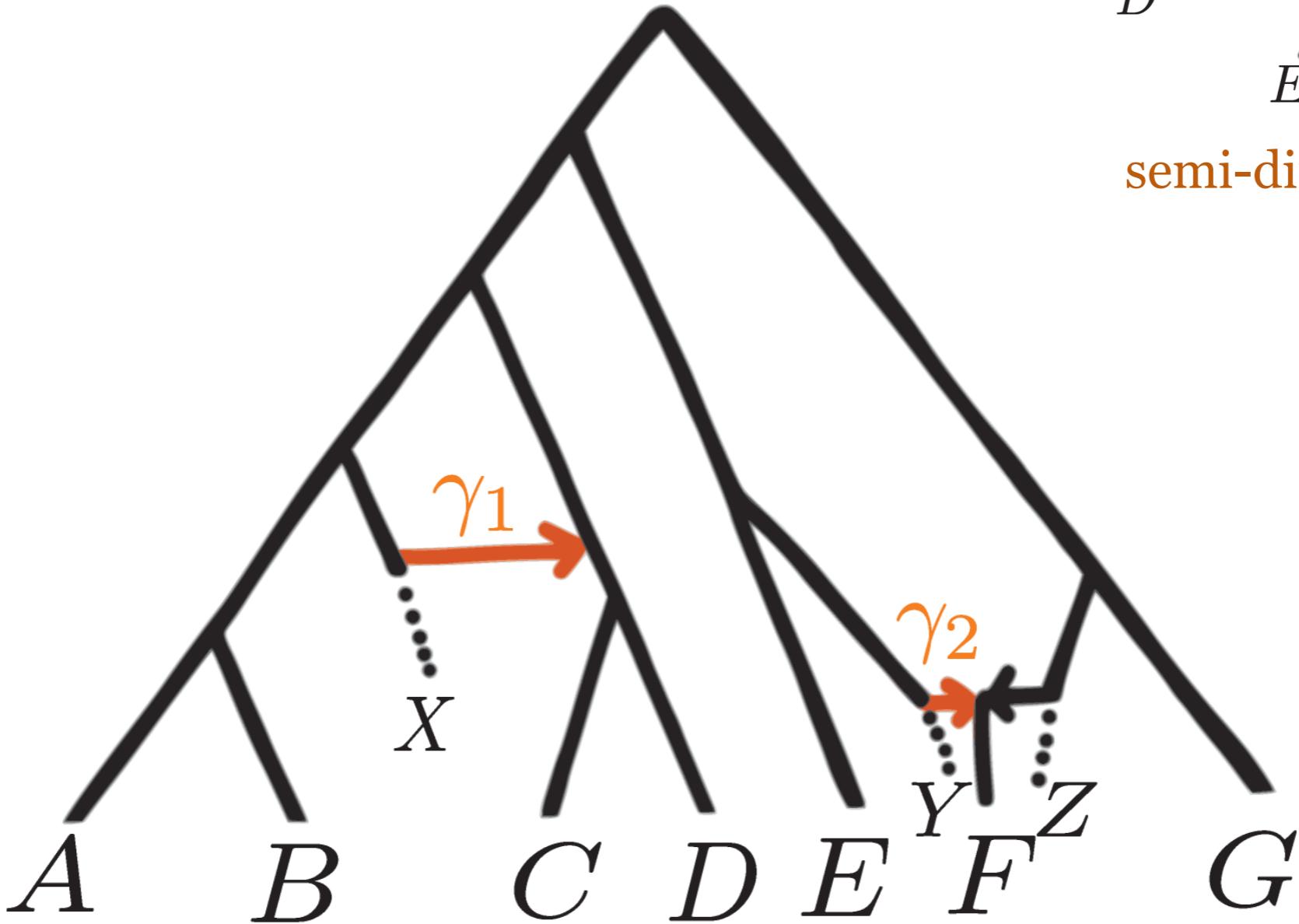
Phylogenetic networks



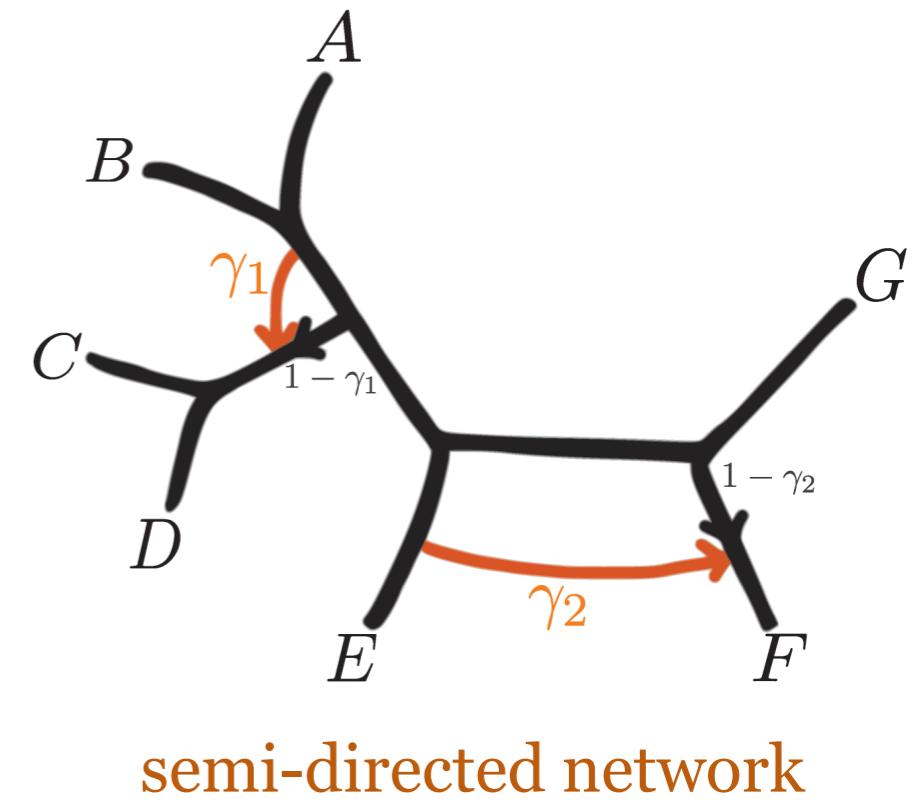
semi-directed network

Interpretation

Phylogenetic networks



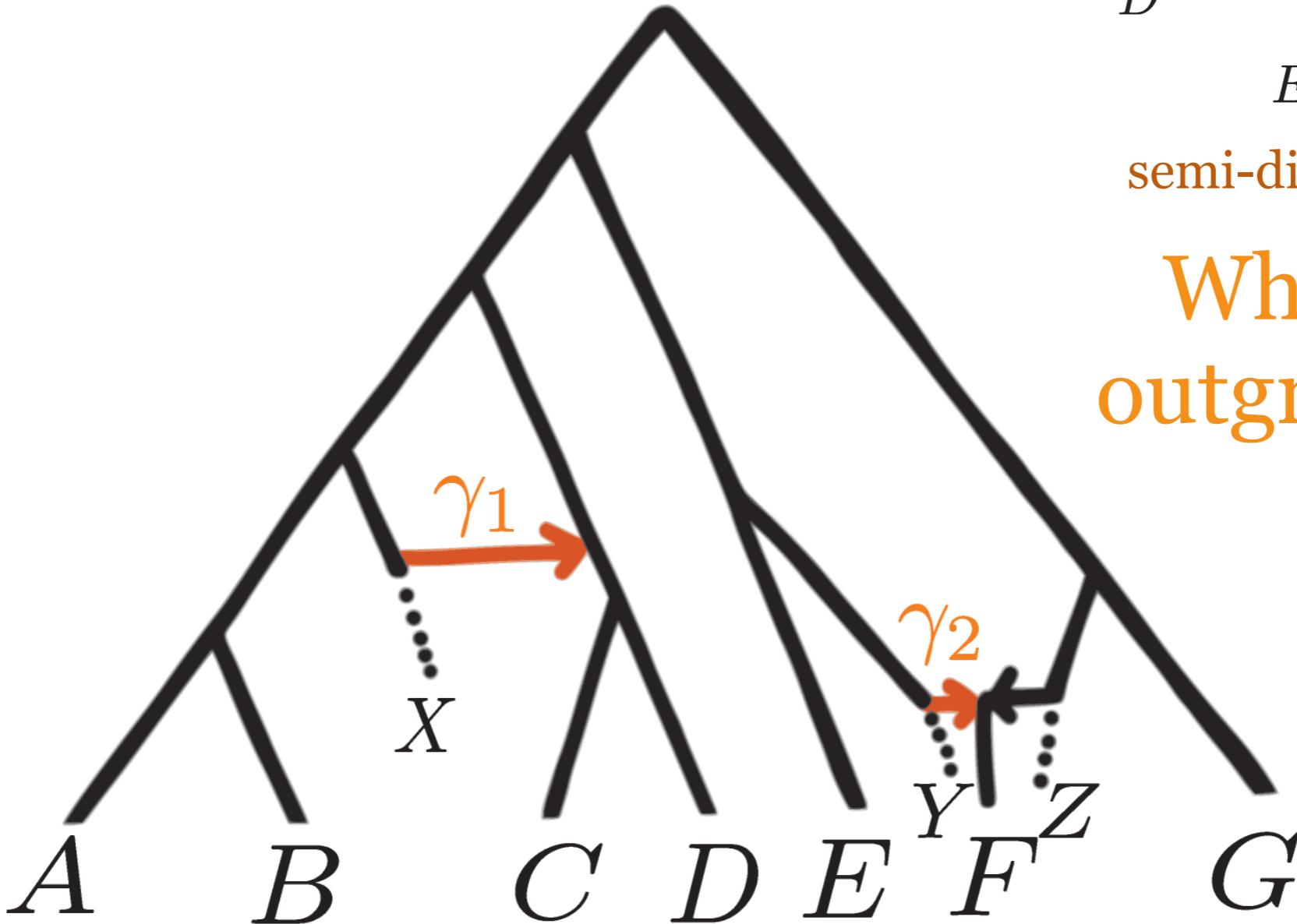
rooted network (outgroup G)



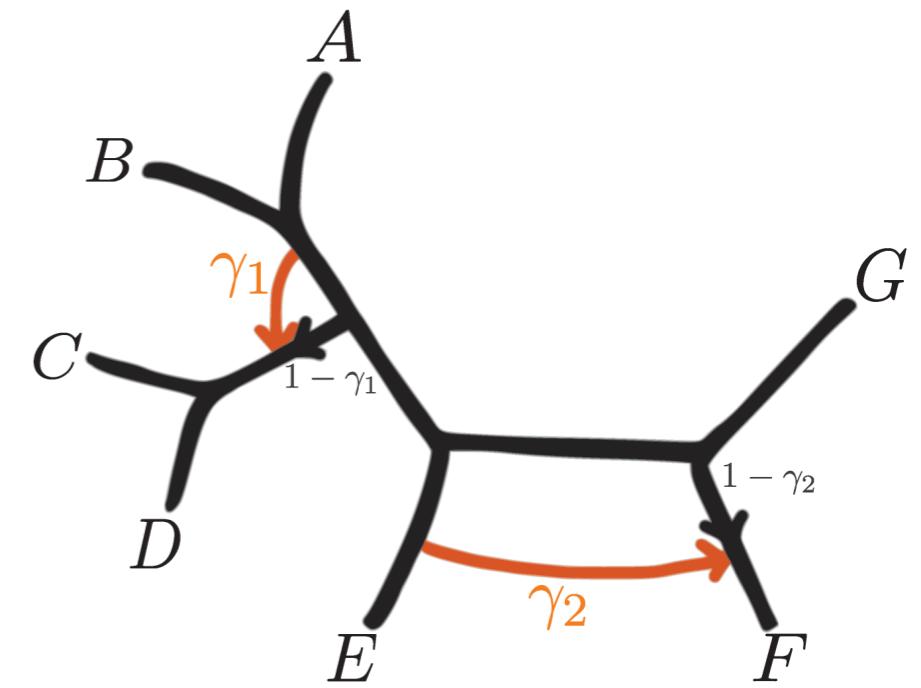
semi-directed network

Interpretation

Phylogenetic networks



rooted network (outgroup G)

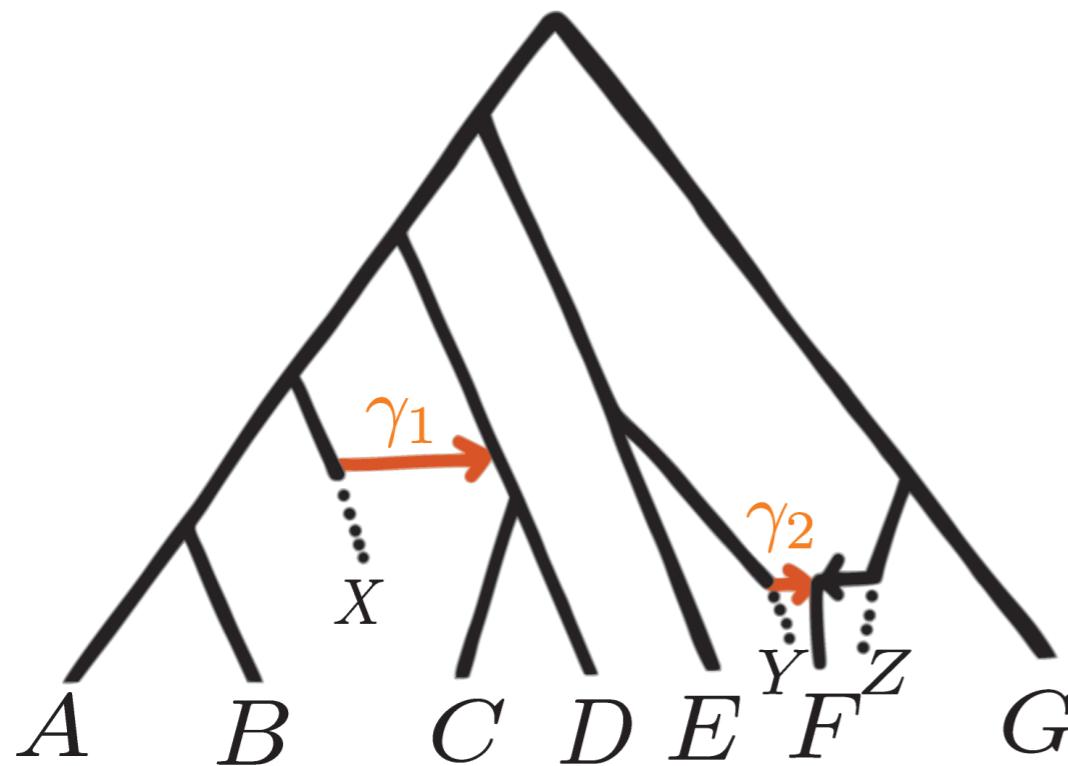


semi-directed network

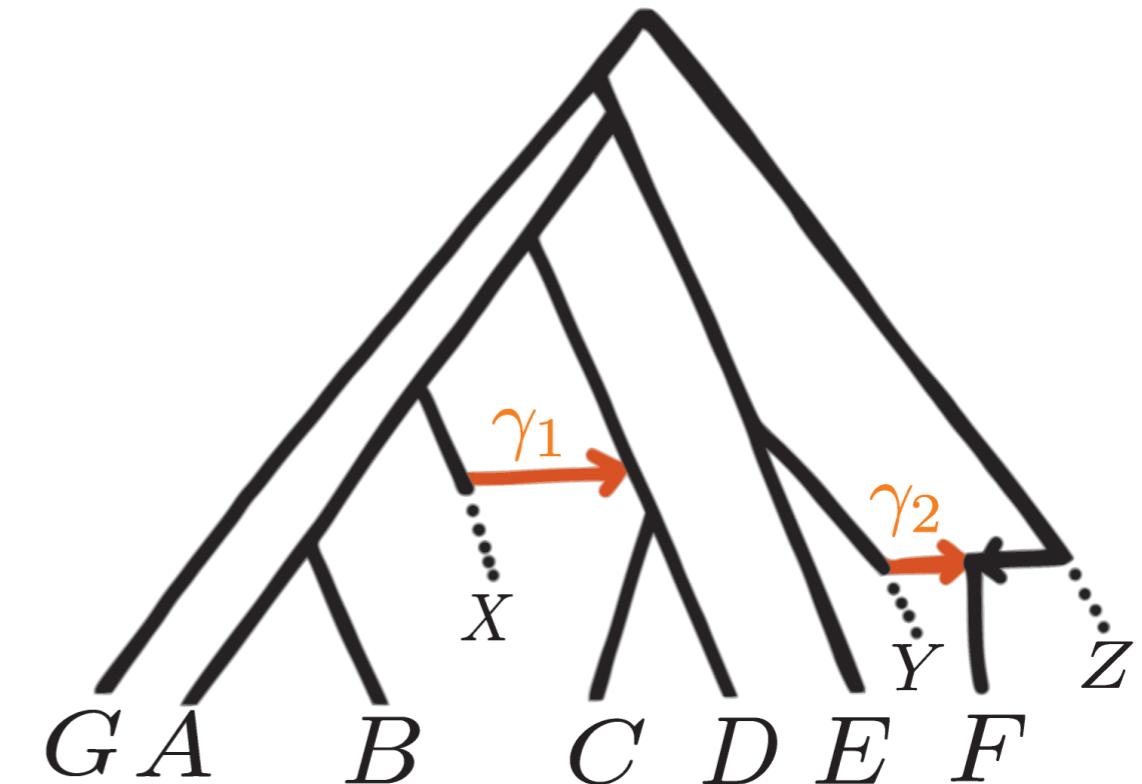
What if my
outgroup is F ?

Interpretation

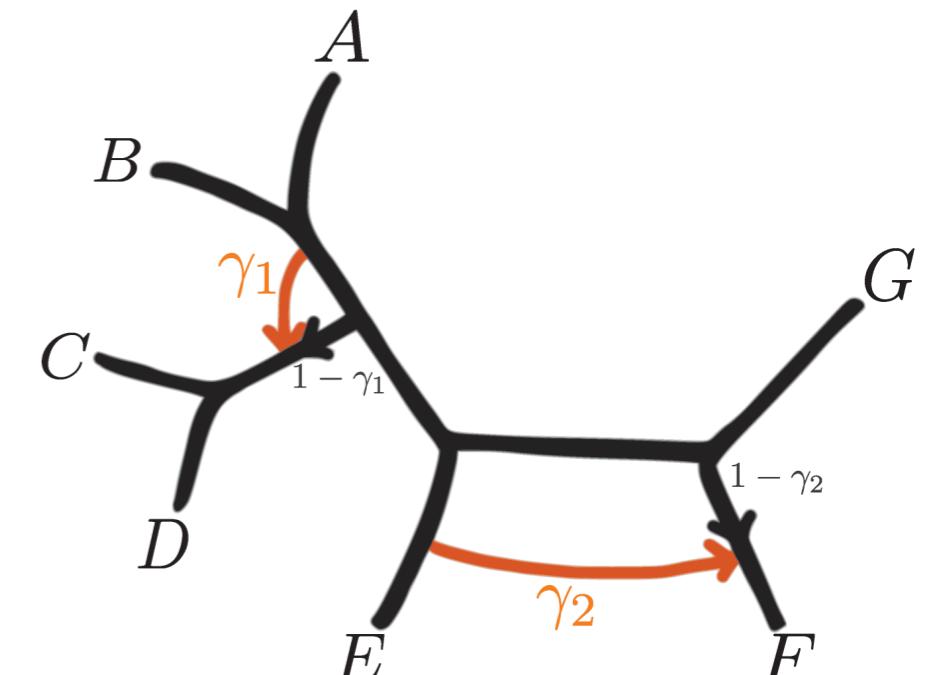
Phylogenetic networks



rooted network
(outgroup G)



rooted network
(outgroup F)



semi-directed network

Network thinking

Phylogenetic network

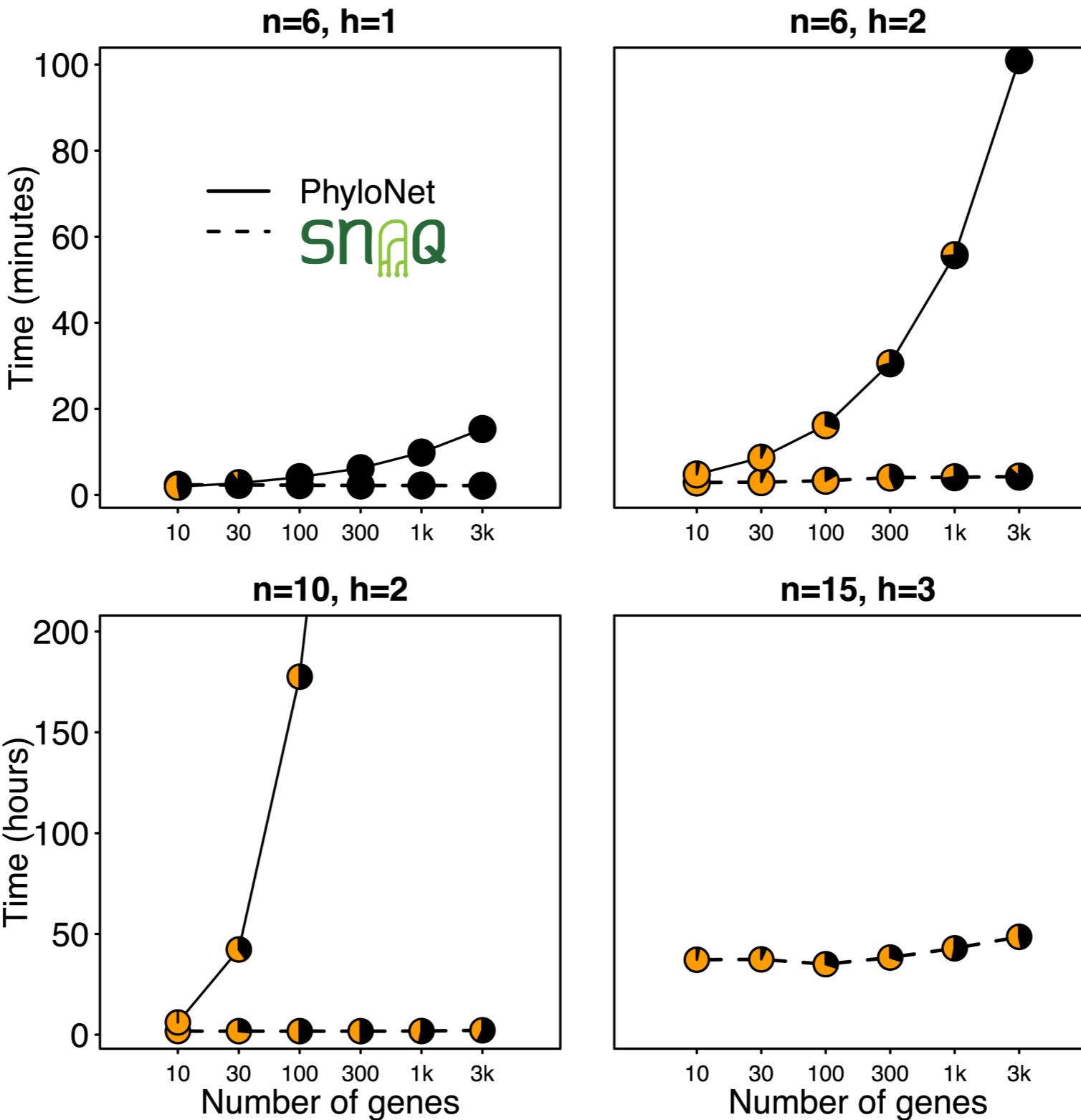
Interpretation challenges

Take home message:

- Networks are useful to represent non-tree-like evolution
- It is not as straight-forward to interpret networks compared to trees (history of speciation events)
- Methods to infer networks do not know biology
- We need other sources of information for accurate interpretation

Scalability

Phylogenetic networks

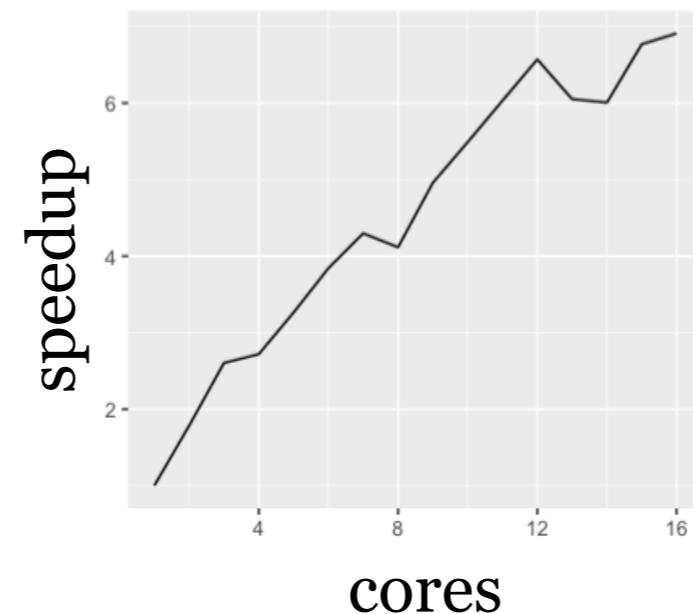
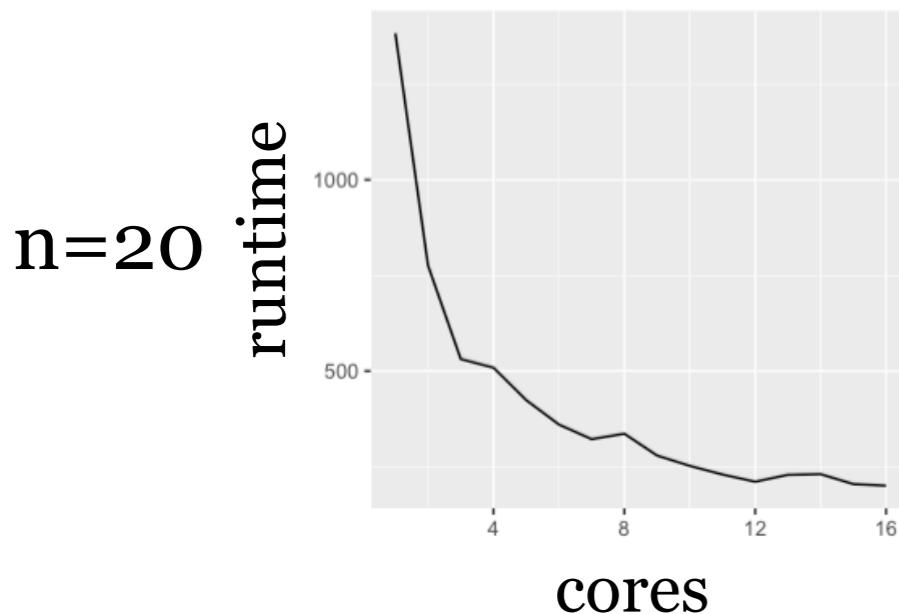
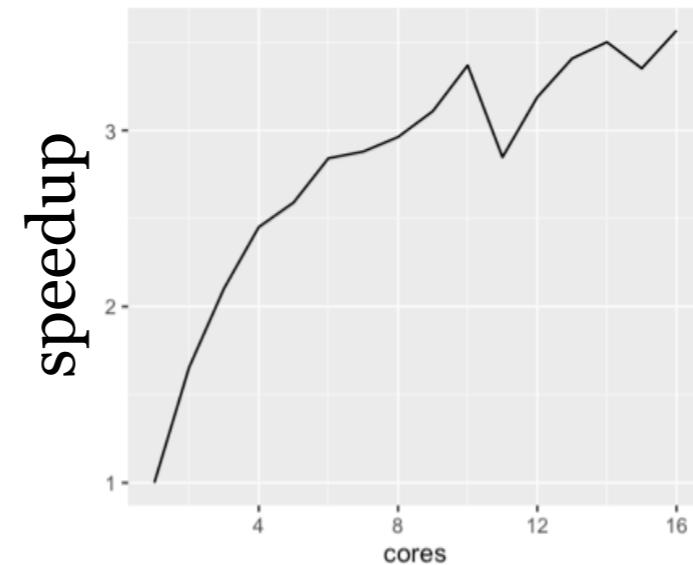
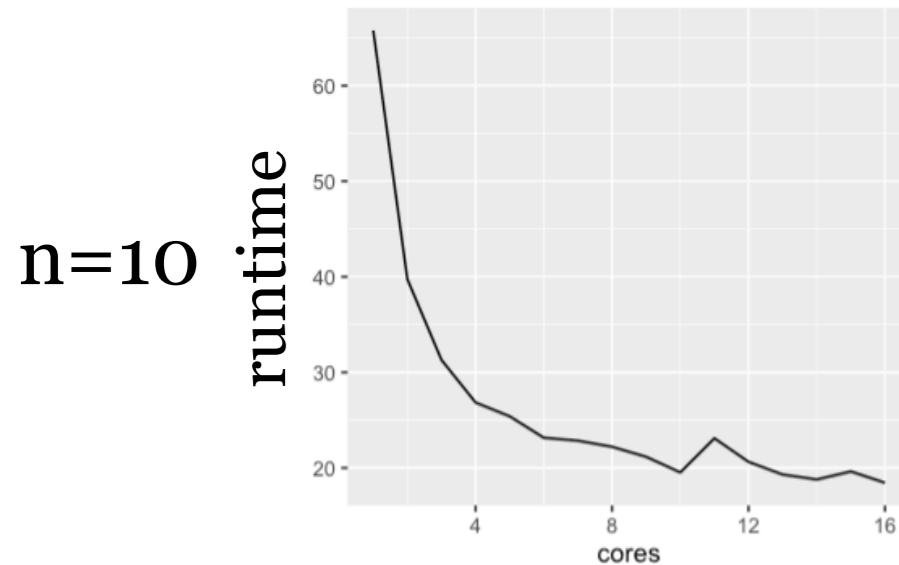


(Solís-Lemus, Ané, 2016, PLoS Genetics)

SNaQ is faster, but is it fast?



Tyler Chafin



Sungsik (Kevin)
Kong



Nathan
Kolbow

- Parallel quartet likelihood computation
- Weights to quartets: deviations from expected CFs
- Sampling of quartets without biasing inference

SNaQ 2.0
coming up soon!

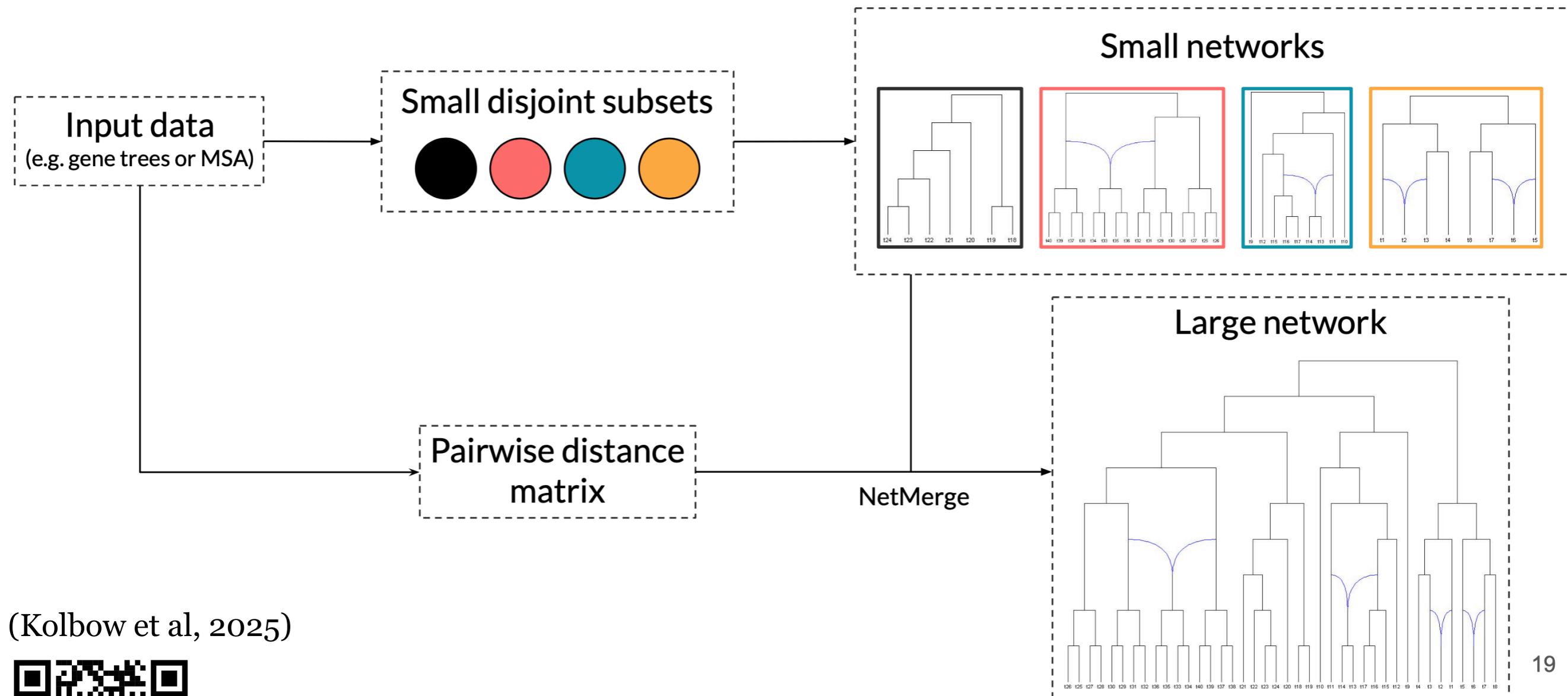
InPhyNet: Larger and larger networks



Nathan
Kolbow



Sungsik (Kevin) Kong



(Kolbow et al, 2025)



up to 1000 taxa!

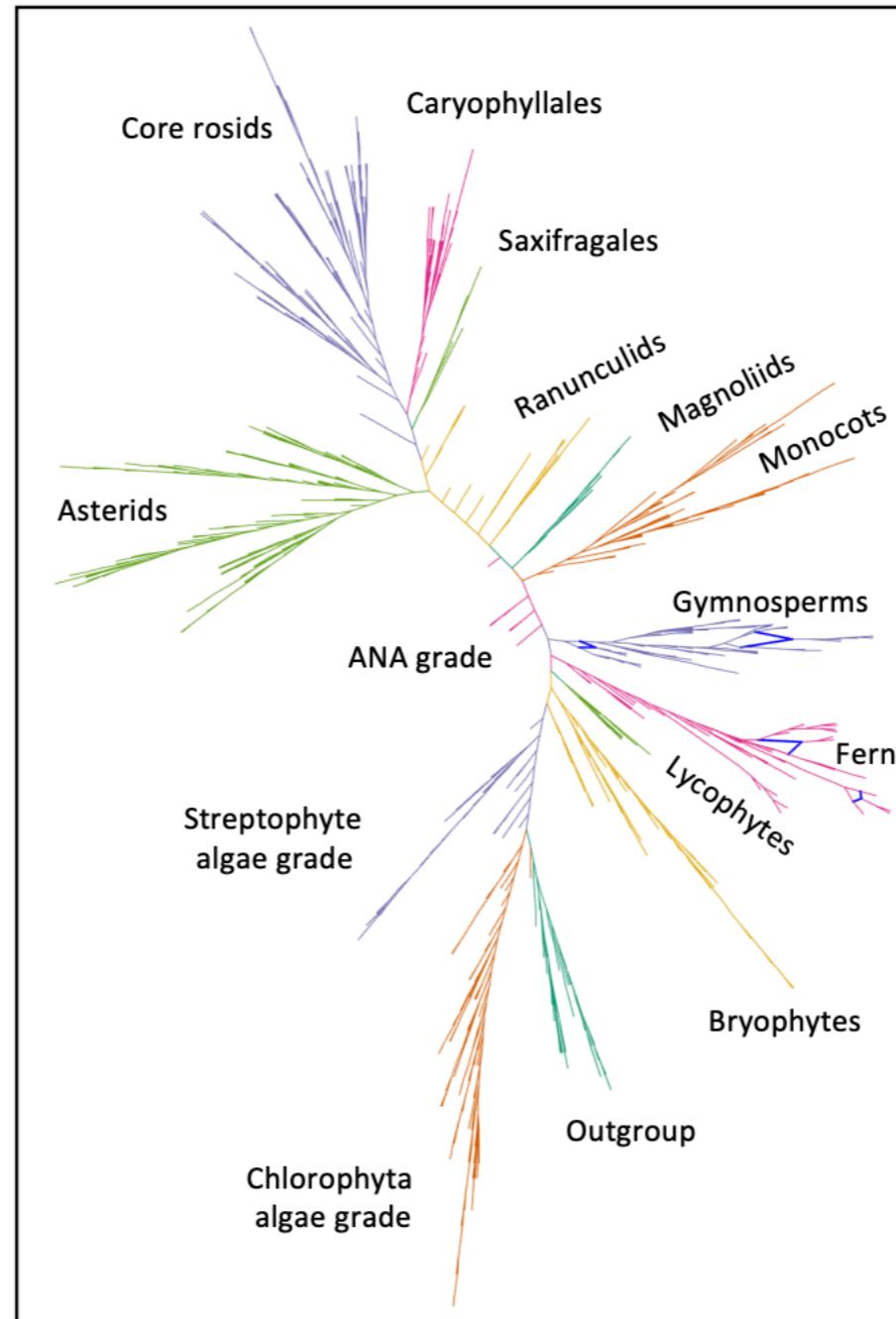
InPhyNet: Larger and larger networks



Nathan
Kolbow

Sungsik
(Kevin) Kong

(A) Species network



Article | [Open access](#) | Published: 23 October 2019

One thousand plant transcriptomes and the phylogenomics of green plants

[One Thousand Plant Transcriptomes Initiative](#)

[Nature](#) 574, 679–685 (2019) | [Cite this article](#)

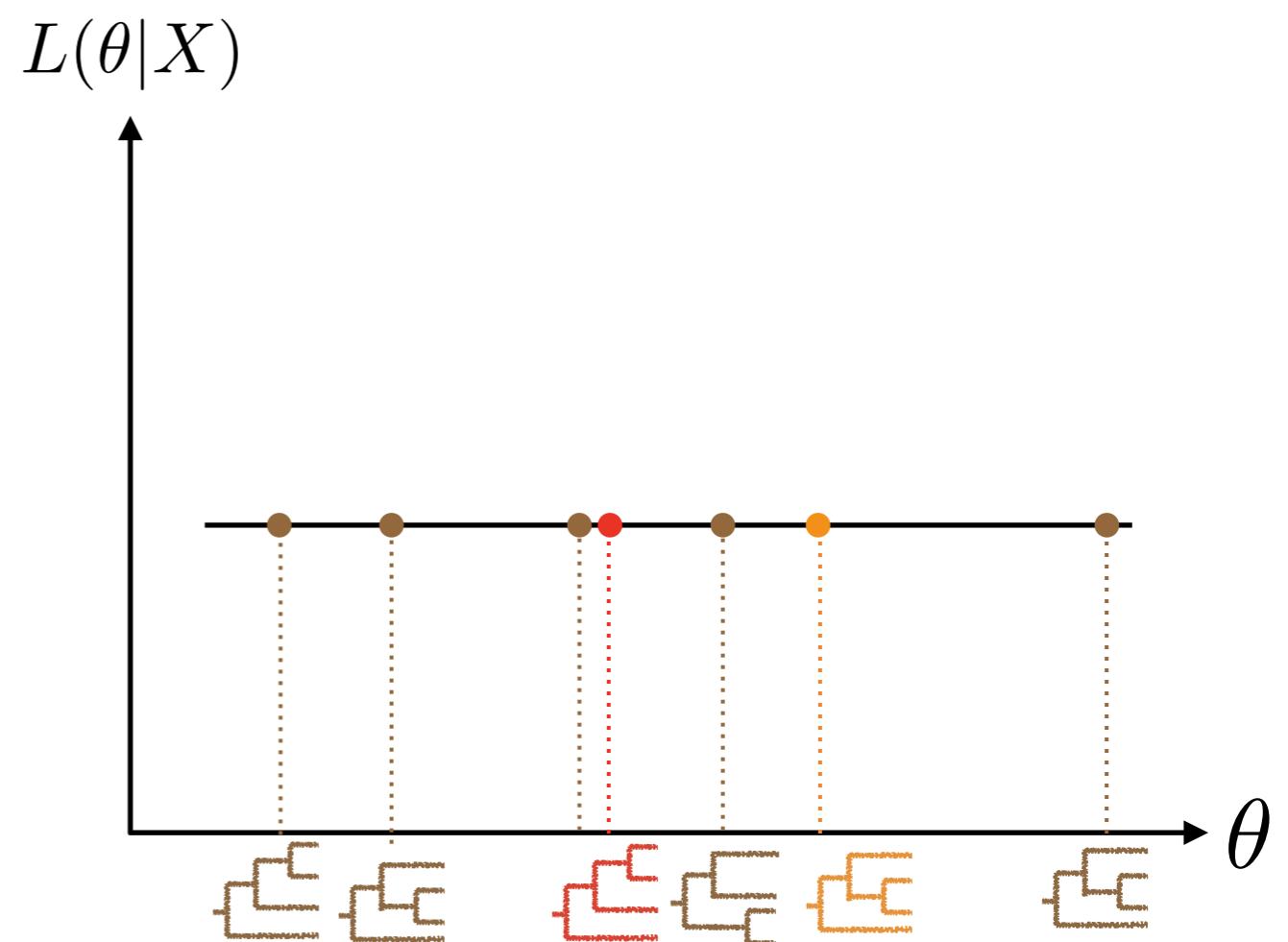
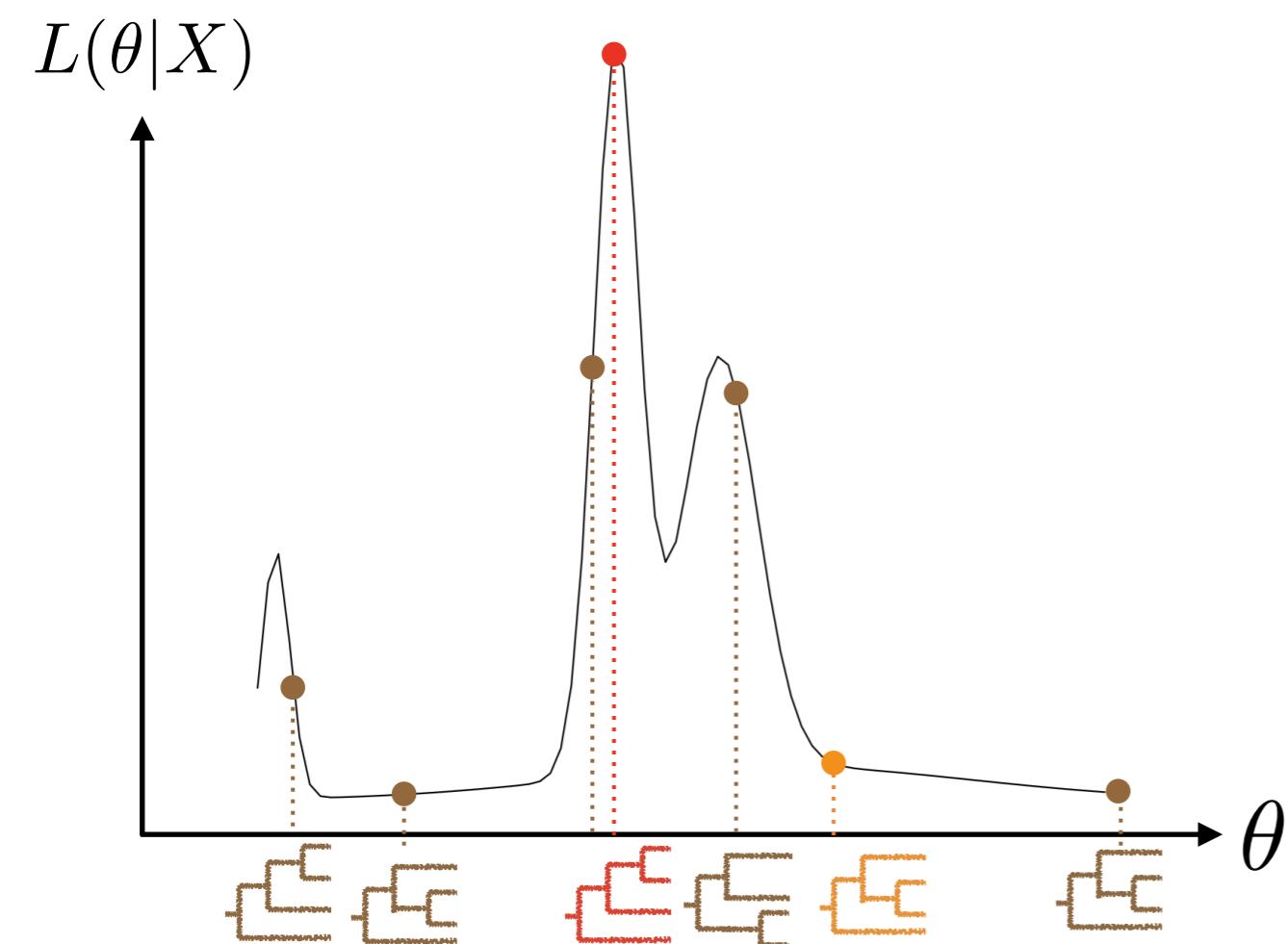
142k Accesses | 1066 Citations | 613 Altmetric | [Metrics](#)

(Kolbow et al, 2025)



Identifiability

Phylogenetic networks



Identifiability

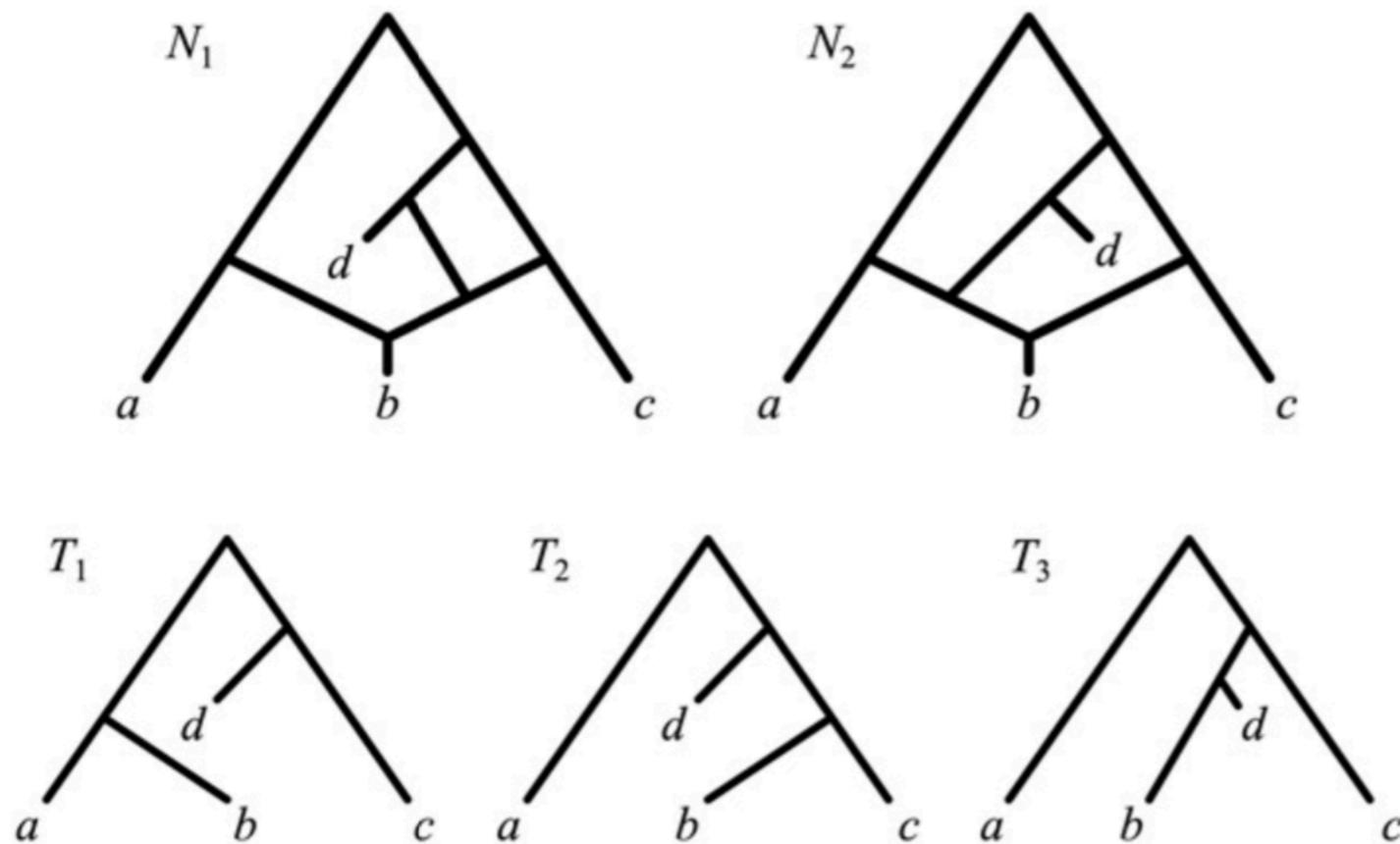
Phylogenetic network

RESEARCH ARTICLE

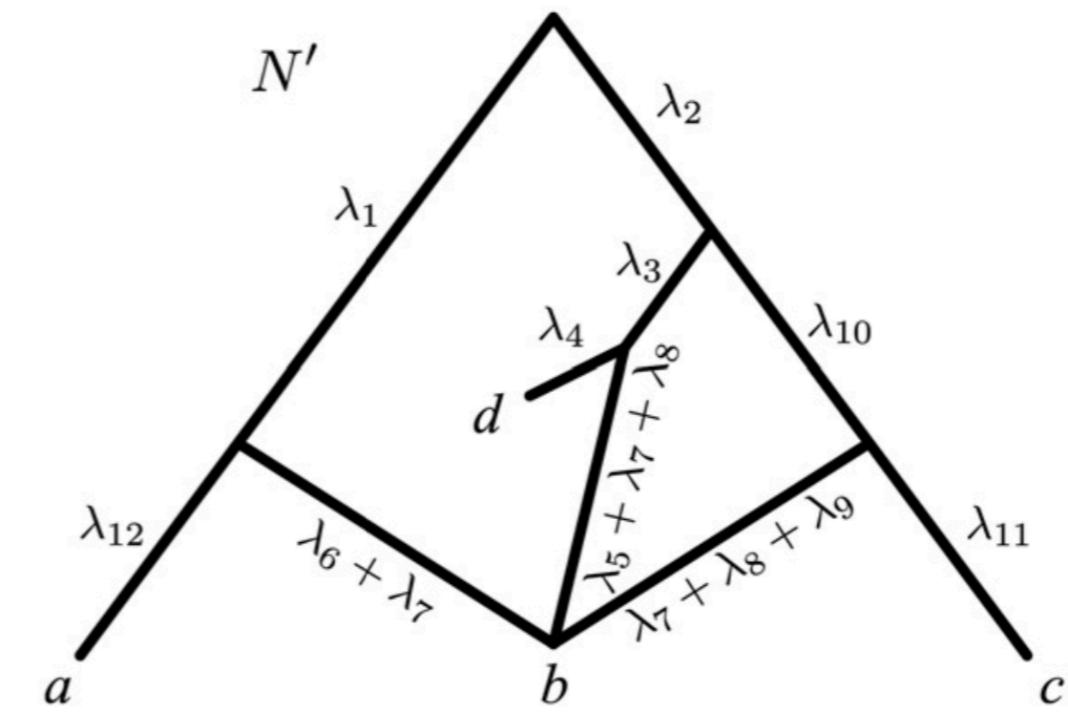
Reconstructible Phylogenetic Networks: Do Not Distinguish the Indistinguishable

Fabio Pardi^{1,3*}, Celine Scornavacca^{2,3}

1 Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier (LIRMM, UMR 5506) CNRS, Université de Montpellier, France, **2** Institut des Sciences de l'Evolution de Montpellier (ISE-M, UMR 5554) CNRS, IRD, Université de Montpellier, France, **3** Institut de Biologie Computationnelle, Montpellier, France



Undistinguishable with the
“displayed trees” criterion



Solution: Canonical
network (“unzipped”)



<https://solislemuslab.github.io/>



mstdn.social/@solislemuslab



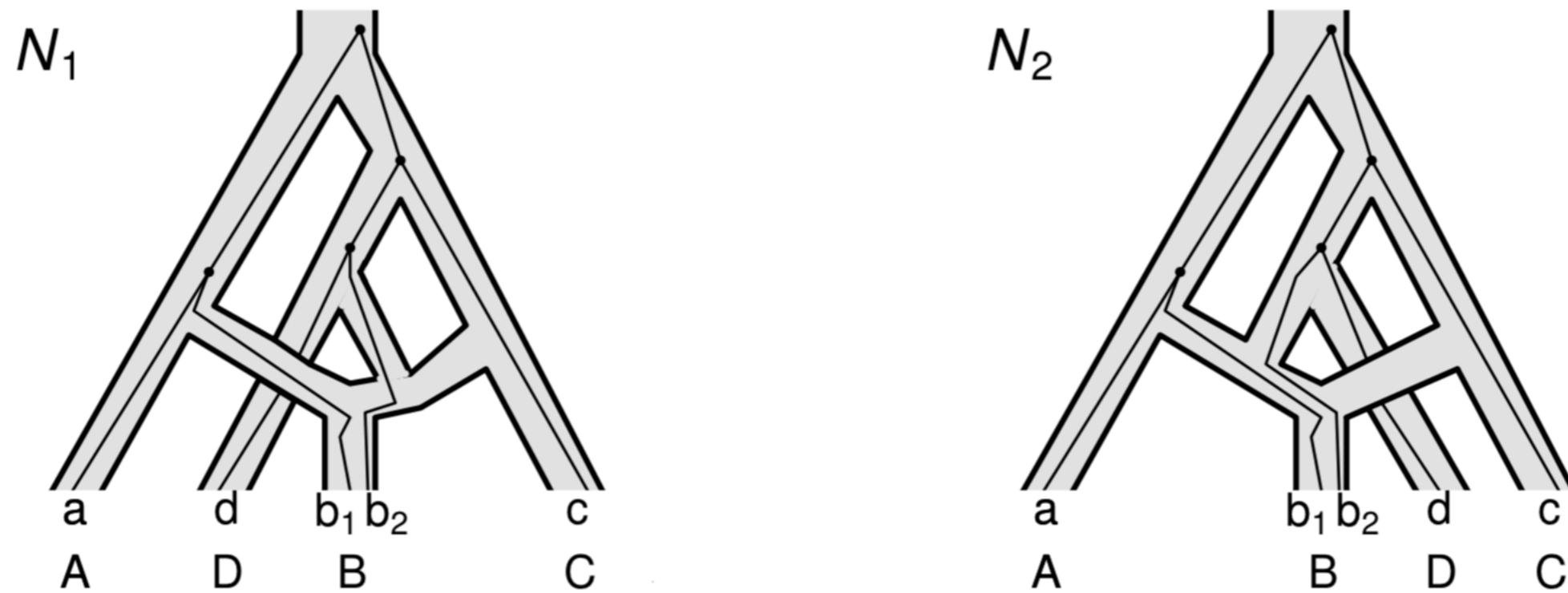
crsl4



@thestatistician

Displayed Trees Do Not Determine Distinguishability Under the Network Multispecies Coalescent

Sha Zhu¹, James H. Degnan²



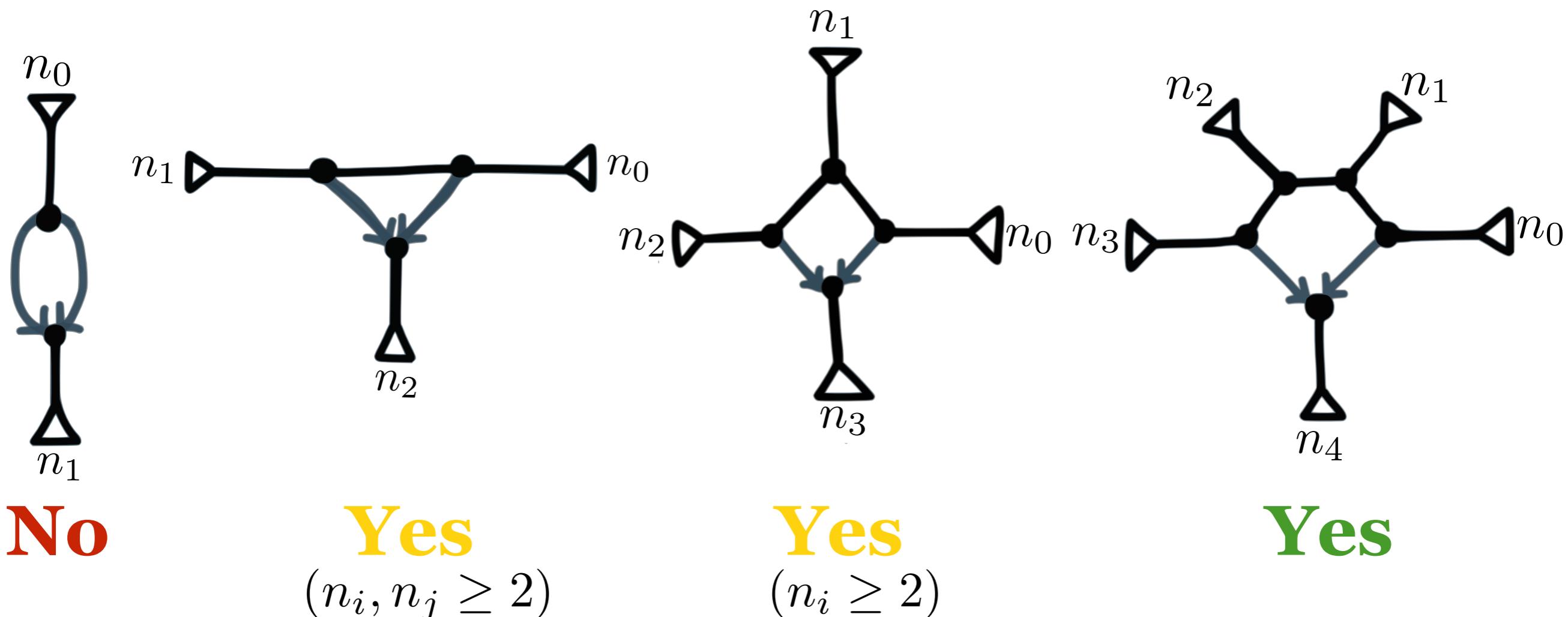
Distinguishable under the MSC

Identifiability

Level-1 Phylogenetic network

RESEARCH ARTICLE

Infering Phylogenetic Networks with
Maximum Pseudolikelihood under
Incomplete Lineage Sorting

Claudia Solís-Lemus^{1*}, Cécile Ané^{1,2}

Generic Identifiability $t_i \in (0, \infty), \gamma \in (0, 1)$

Identifiability

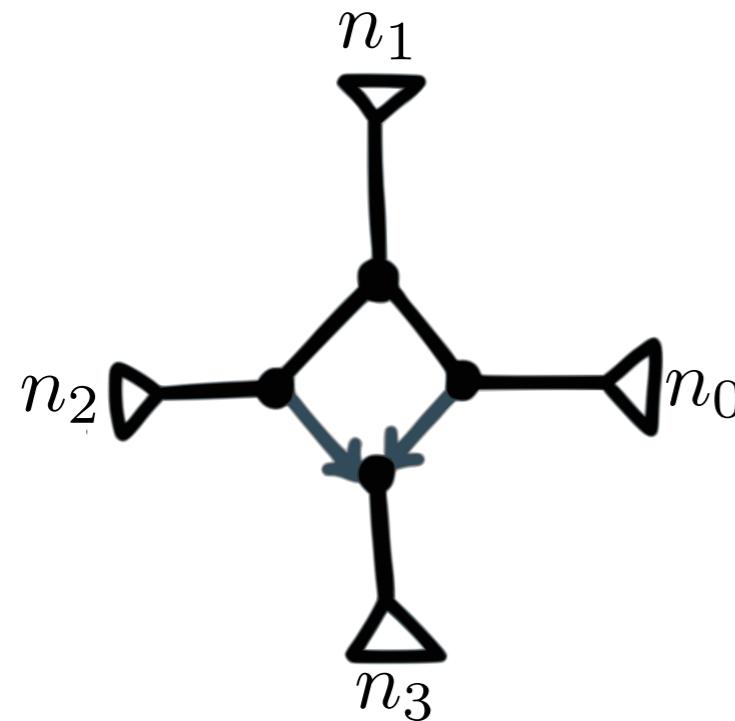
Level-1 Phylogenetic network

RESEARCH ARTICLE

Infering Phylogenetic Networks with
Maximum Pseudolikelihood under
Incomplete Lineage Sorting

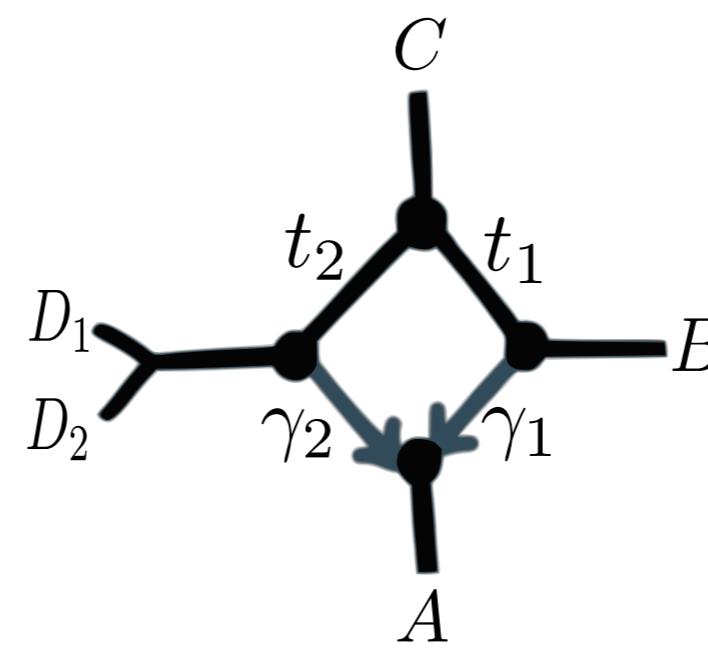
Claudia Solís-Lemus^{1*}, Cécile Ané^{1,2}

In theory

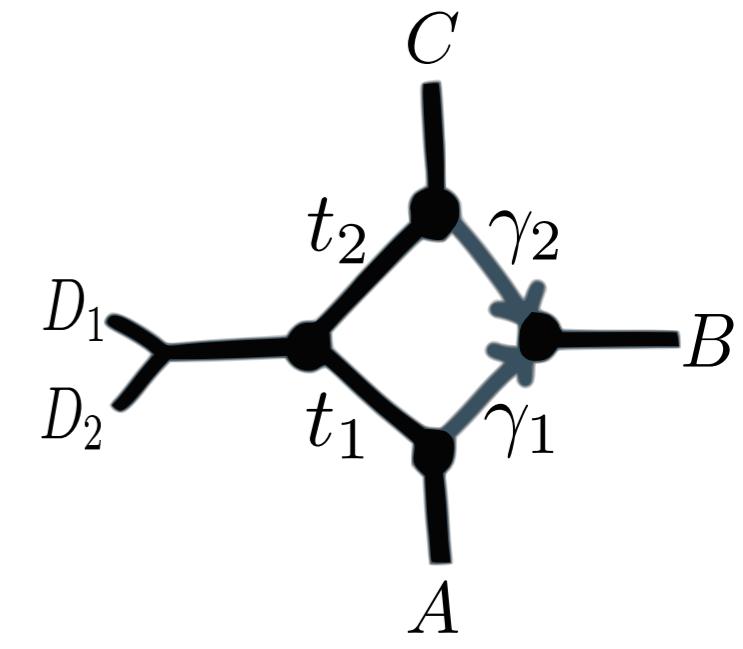


Yes
 $(n_i \geq 2)$

In practice



Sometimes



<https://solislemuslab.github.io/>



<mstdn.social/@solislemuslab>



crsl4



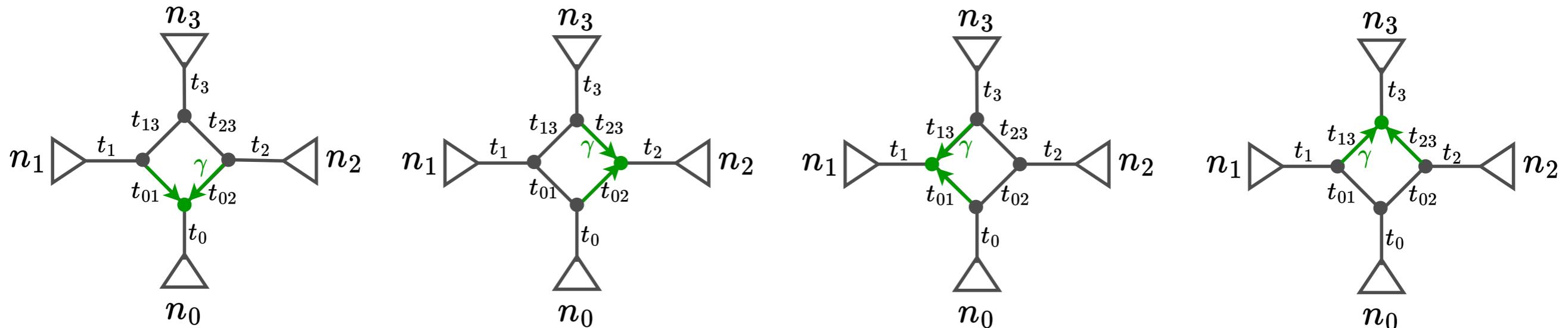
@thestatistician

Identifiability

4-node hybridization cycle (diamond)



George Tiley



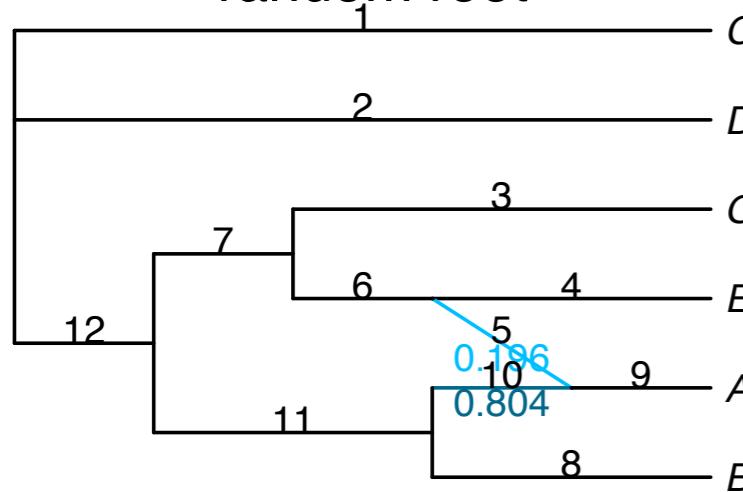
Generic Identifiability $t_i \in (0, \infty), \gamma \in (0, 1)$

(Tiley et al, 2025)

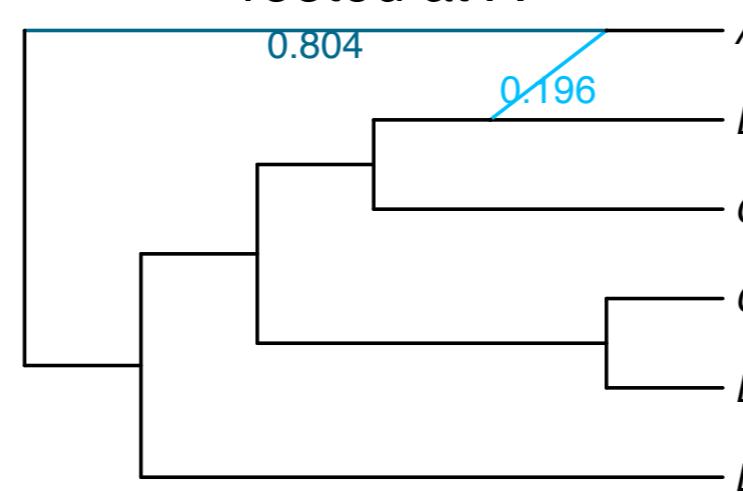
Practical identifiability

Conflicts with the outgroup!

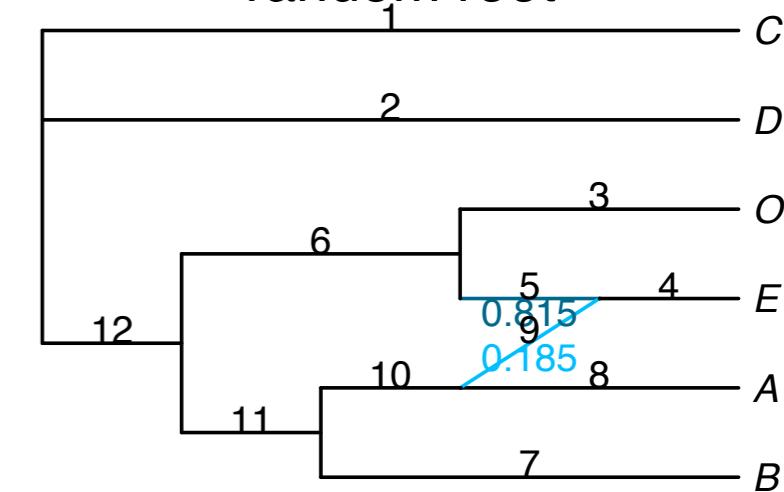
best net, score=28.3,
random root



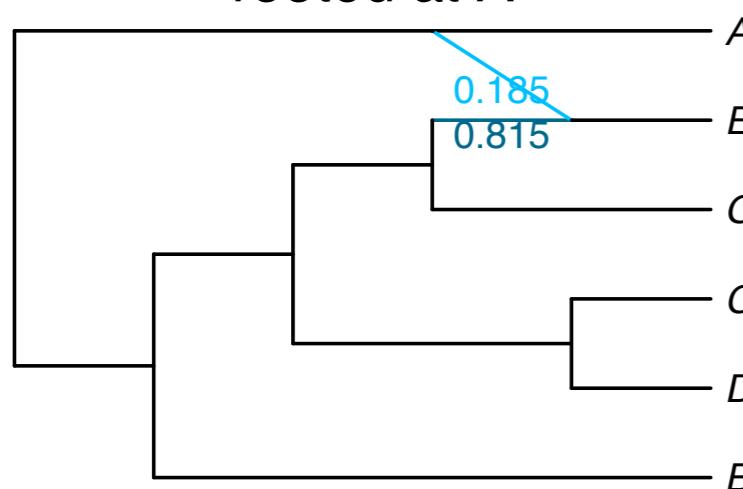
best net, score=28.3,
rooted at A



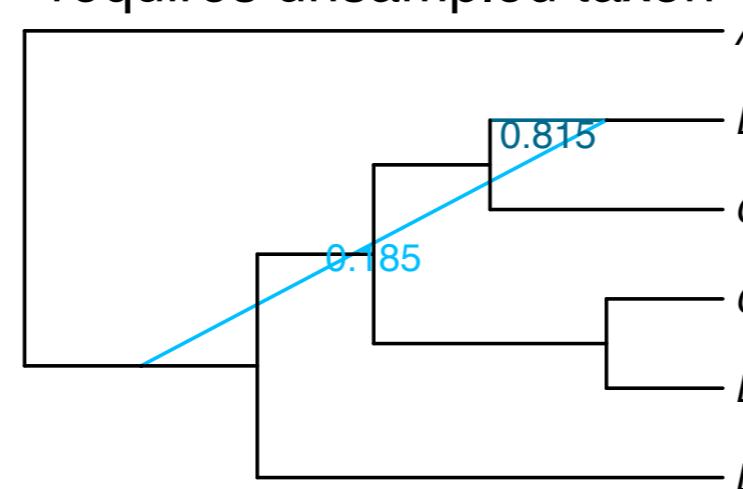
direction modified, score=31.5,
random root



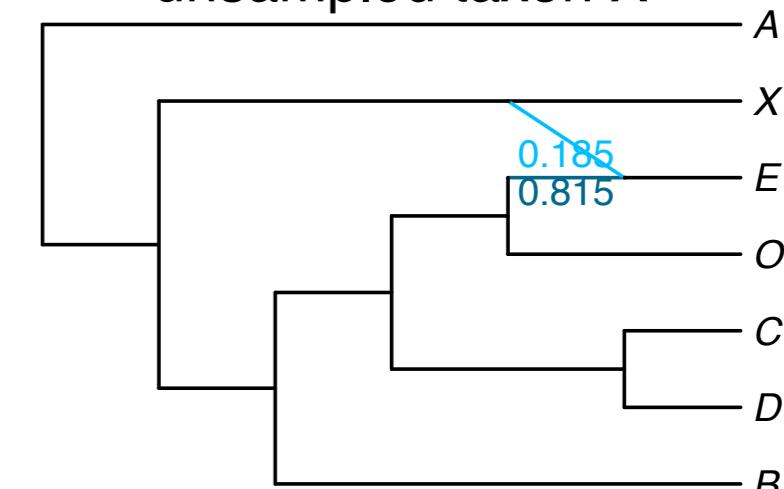
direction modified, score=31.5,
rooted at A



direction modified, score=31.5,
requires unsampled taxon

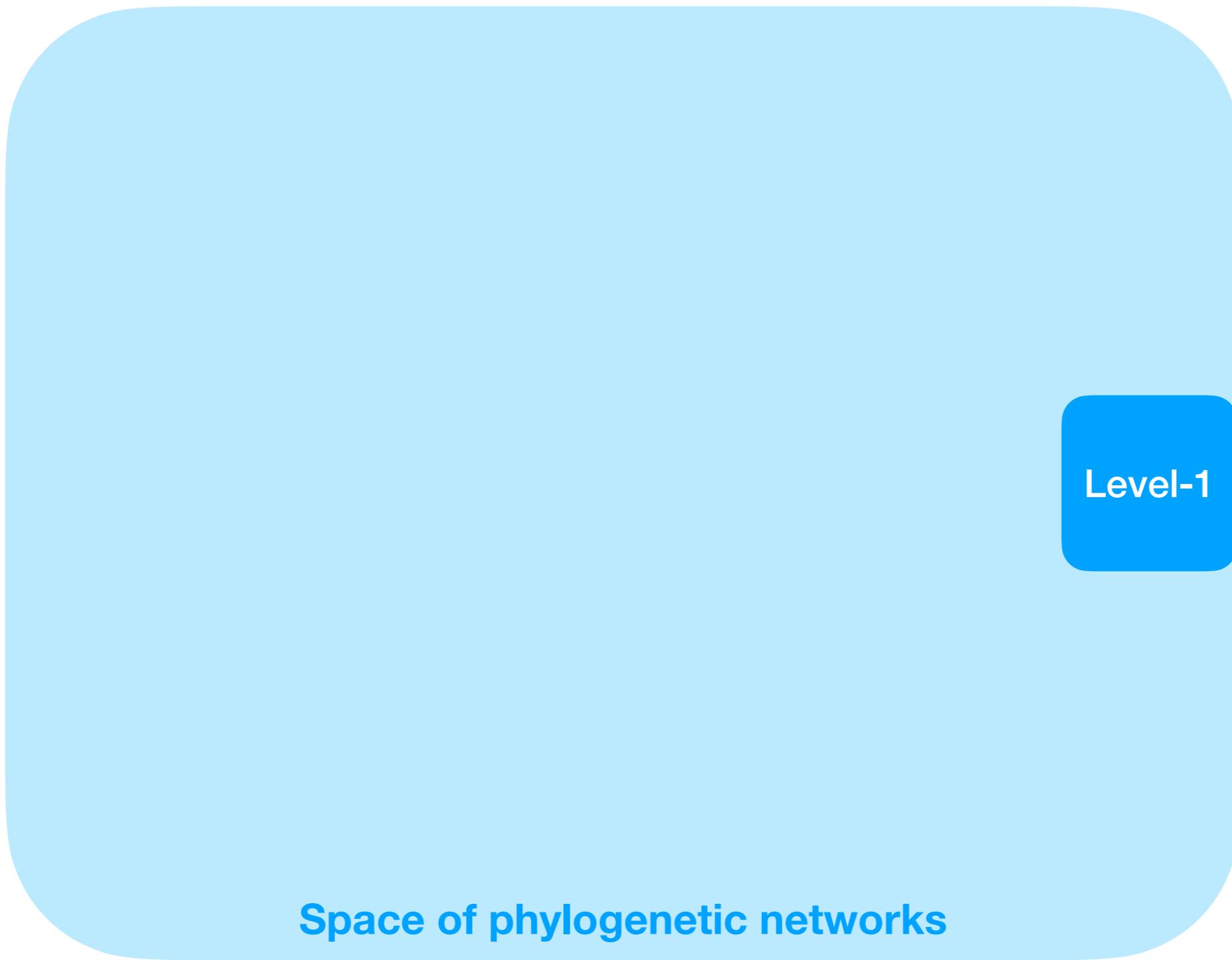


direction modified, score=31.5,
unsampled taxon X



Identifiability

Phylogenetic networks



- ✓ Pseudolikelihood
- ? Likelihood



<https://solislemuslab.github.io/>



mstdn.social/@solislemuslab



crsl4



@thestatistician

Identifiability

Phylogenetic networks

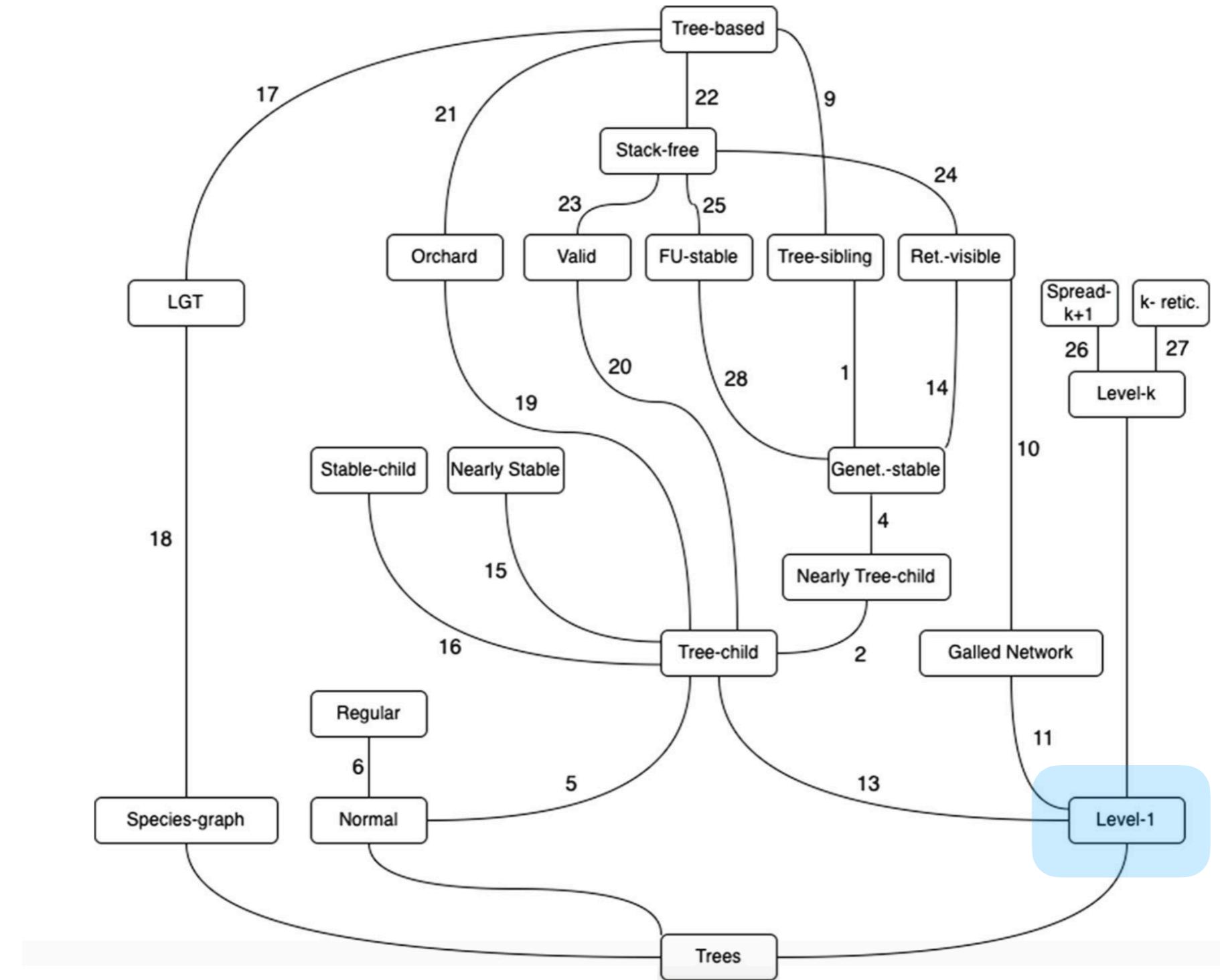


Classes of explicit phylogenetic networks and their biological and mathematical significance

Sungsik Kong¹ · Joan Carles Pons² · Laura Kubatko^{1,3} · Kristina Wicke⁴

Received: 21 September 2021 / Revised: 18 January 2022 / Accepted: 31 March 2022 /

Sungsik (Kevin)
Kong



Upcoming results on identifiability of **level-2**, **galled** and **tree-child** networks (Collaborative efforts among network community)



<https://solislemuslab.github.io/>



mstdn.social/@solislemuslab



[crsl4](https://crsl4.github.io)

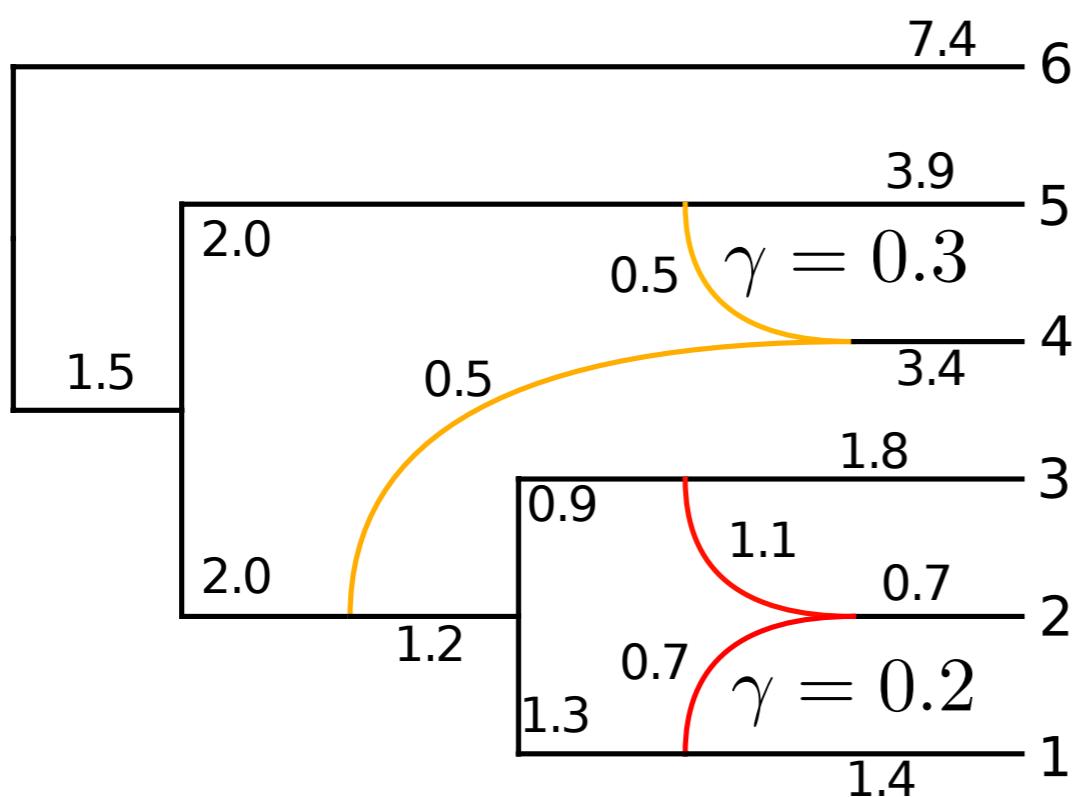


@thestatistician

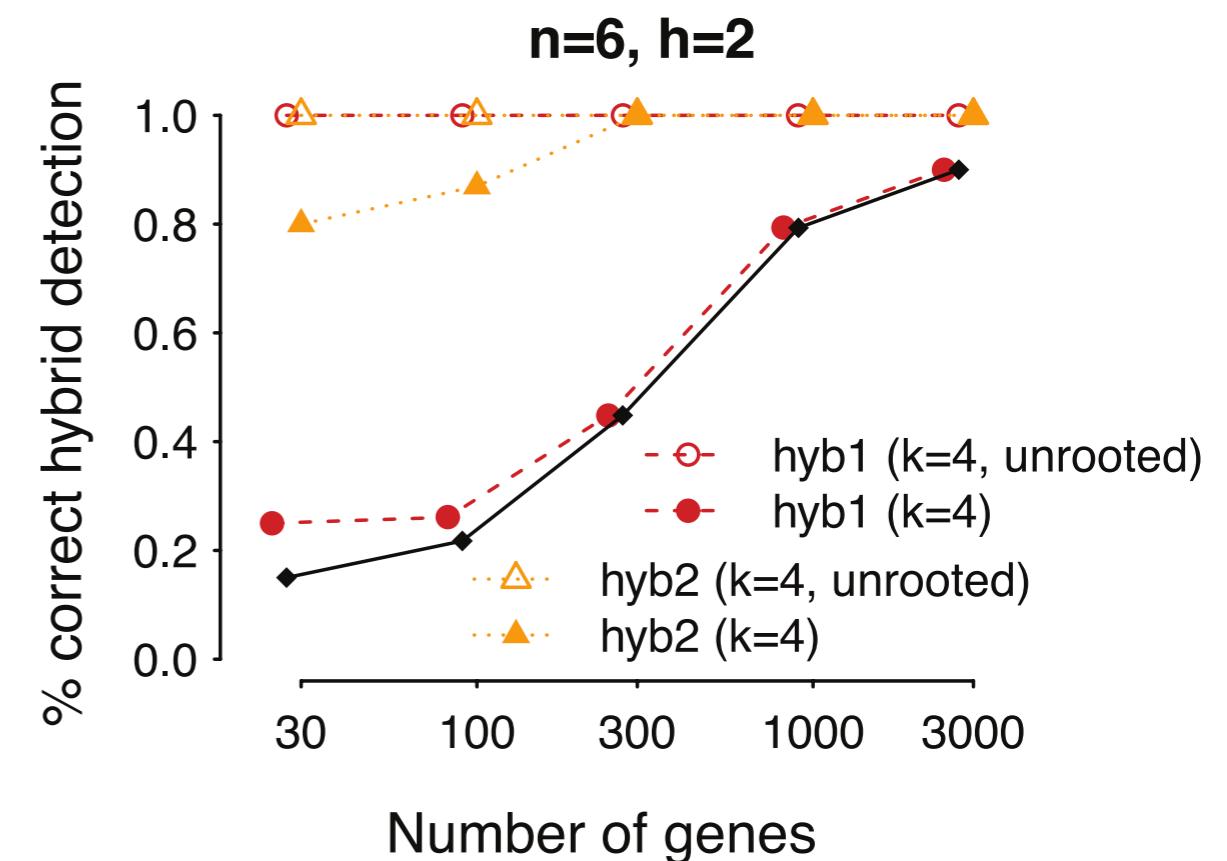
Identifiability

Phylogenetic networks

Good diamond



Bad diamond



Practical identifiability

What does this have to do with me?

Take home message:

- Different networks might have similar (pseudo)likelihood score (especially as complexity increases), so statistically indistinguishable!
- Do you have the necessary data (and signal) to infer reticulation events?
- Placement of hybridization cycle might be more accurate than direction of gene flow.

Practical identifiability

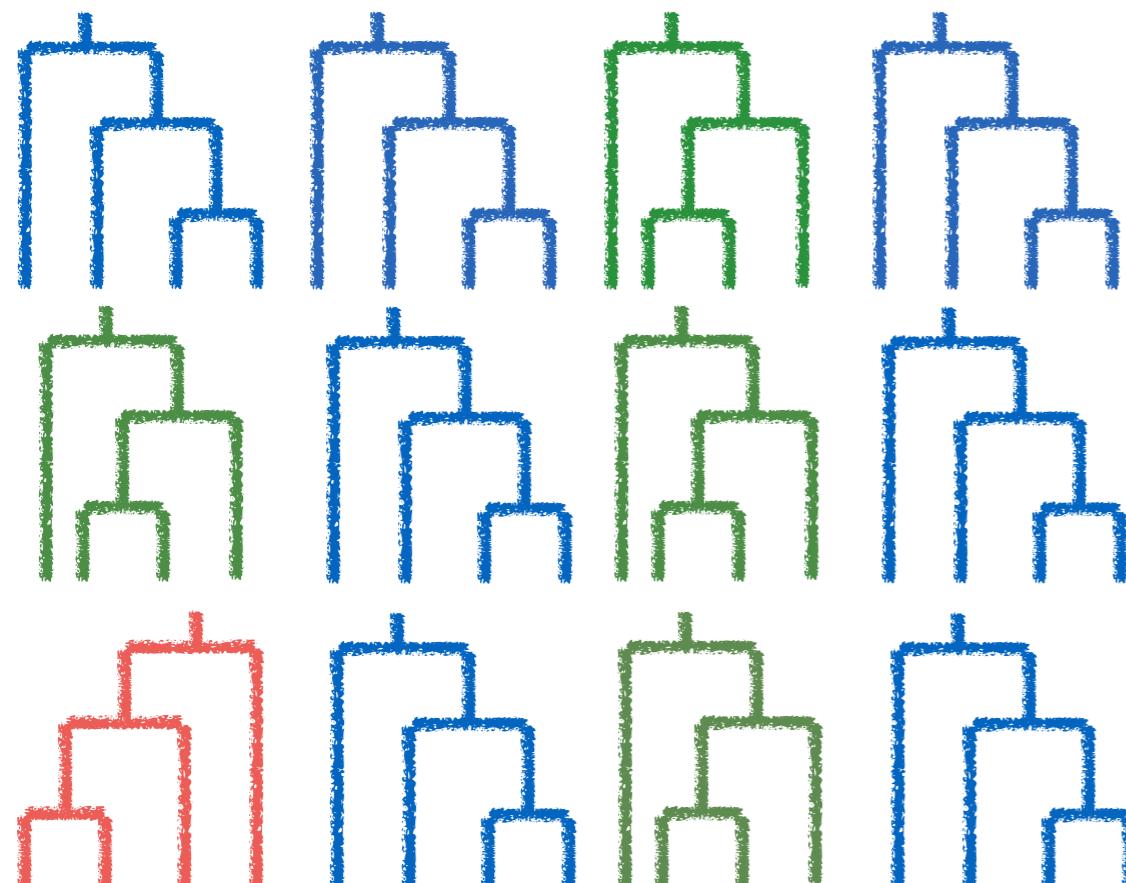
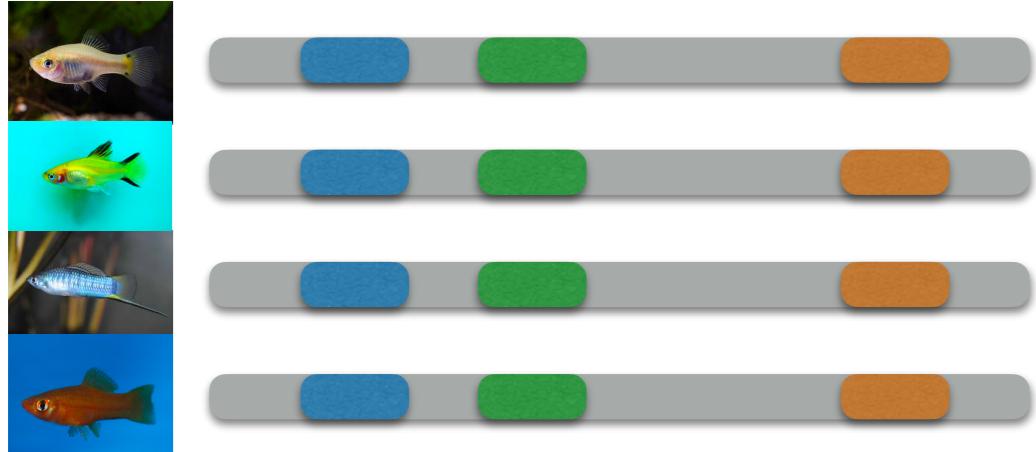
What does this have to do with me?

Take home message:

- Different networks might have similar (pseudo)likelihood score (especially as complexity increases), so statistically indistinguishable!
- **Do you have the necessary data (and signal) to infer reticulation events?**
- Placement of hybridization cycle might be more accurate than direction of gene flow.

When?

Phylogenetic network



Data

Goodness-of-fit test
Hypothesis test:
Is a tree a good fit?

TICR
→
GitHub



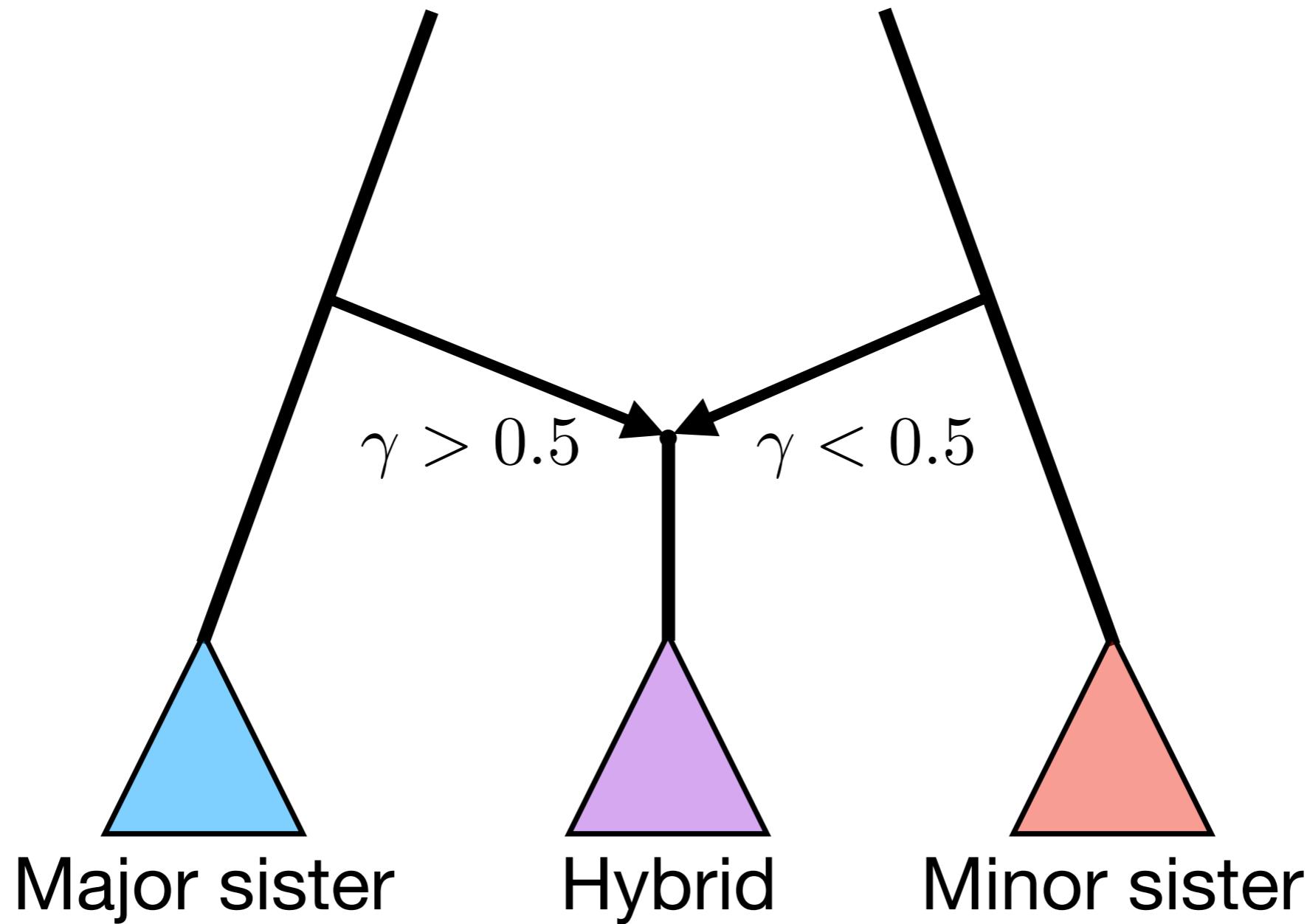
<https://github.com/nstenz/TICR>
(Stenz et al, 2015, Syst Bio)

Practical advice

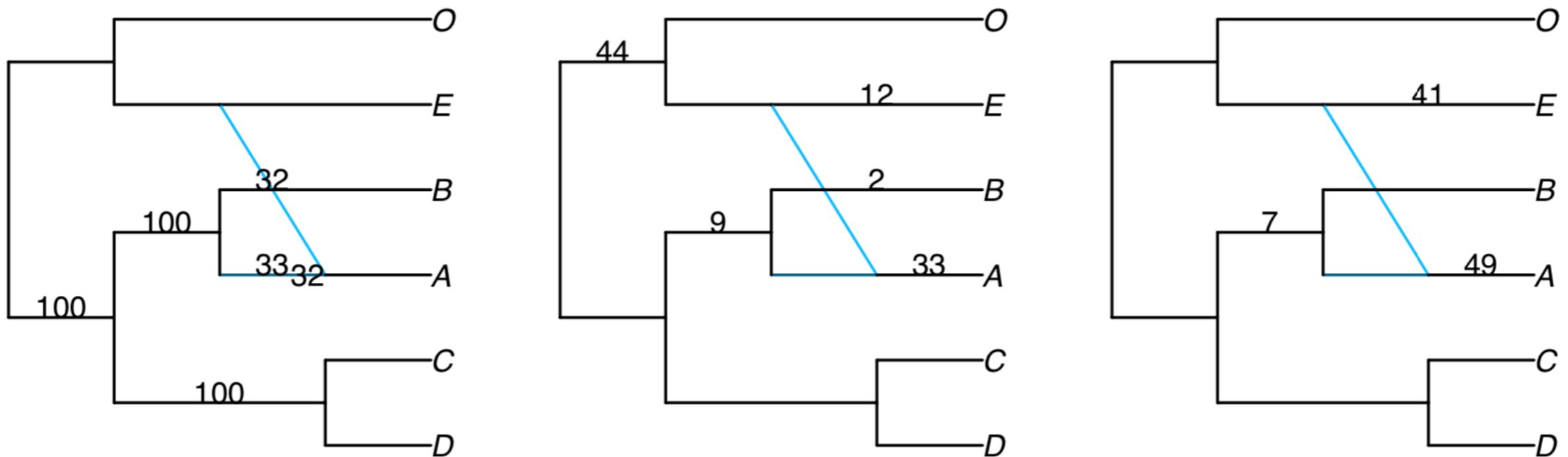
- Do multiple runs
- Do bootstrap
- Check the .networks output file (especially if hybridization conflicts with outgroup)
- Interpretation of ghost lineages
- What is the quality of my input data (gene trees/CFs)?
- Run SNaQ sequentially: $h=0, h=1, h=2, \dots$

When to stop?
(Cai and Ané, 2020)

Bootstrap on Networks



Bootstrap on Networks



Hybrid
clades

Minor
sister
clades

I want to find signal of hybridization in my data

Hybrid detection
methods

HyDe

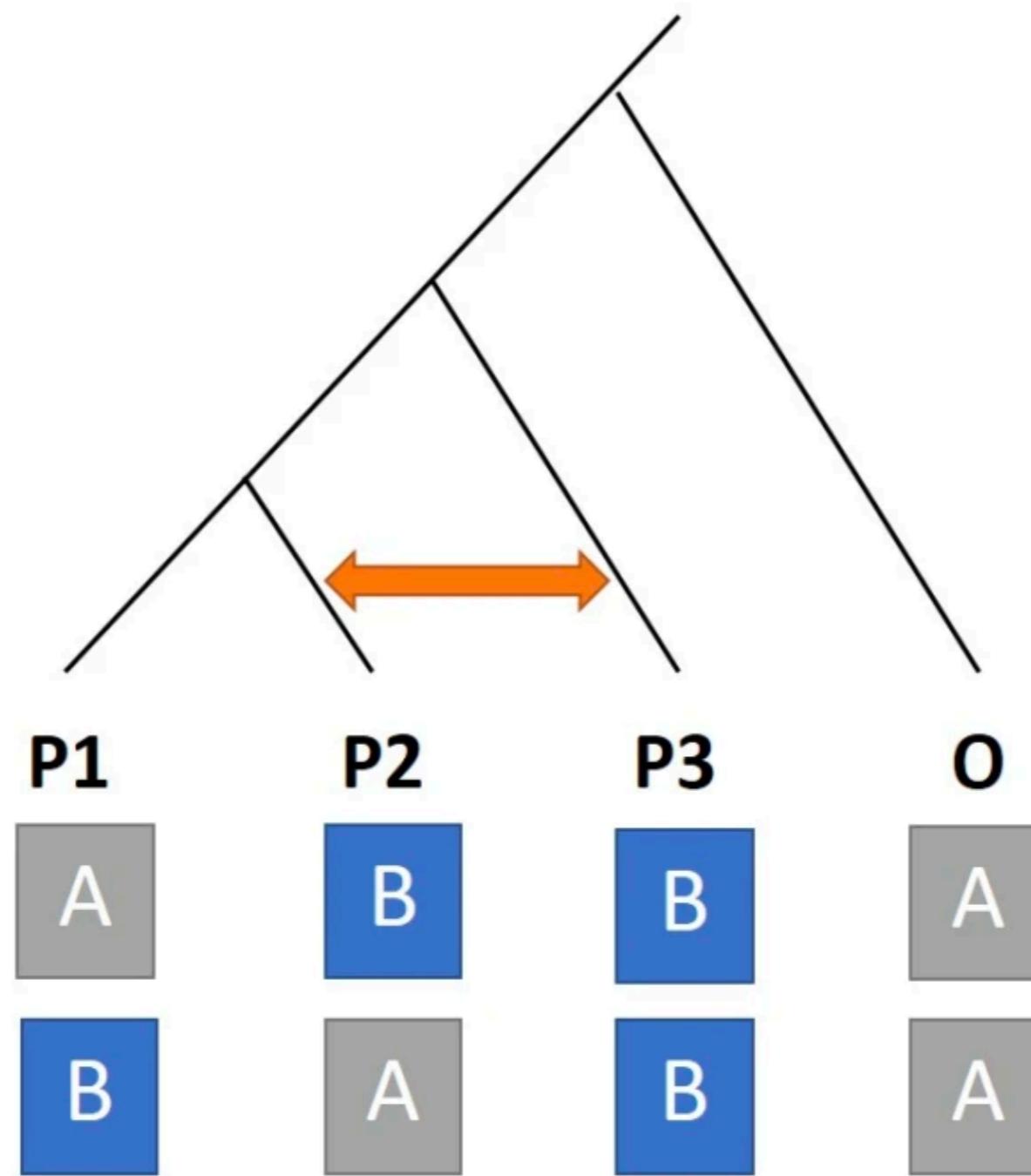
ABBA-BABA

MSCQuartets

Network inference
methods

PhyloNet
SNaQ
BEAST2
NANUQ
RF-Net
PhyNEST

D statistics





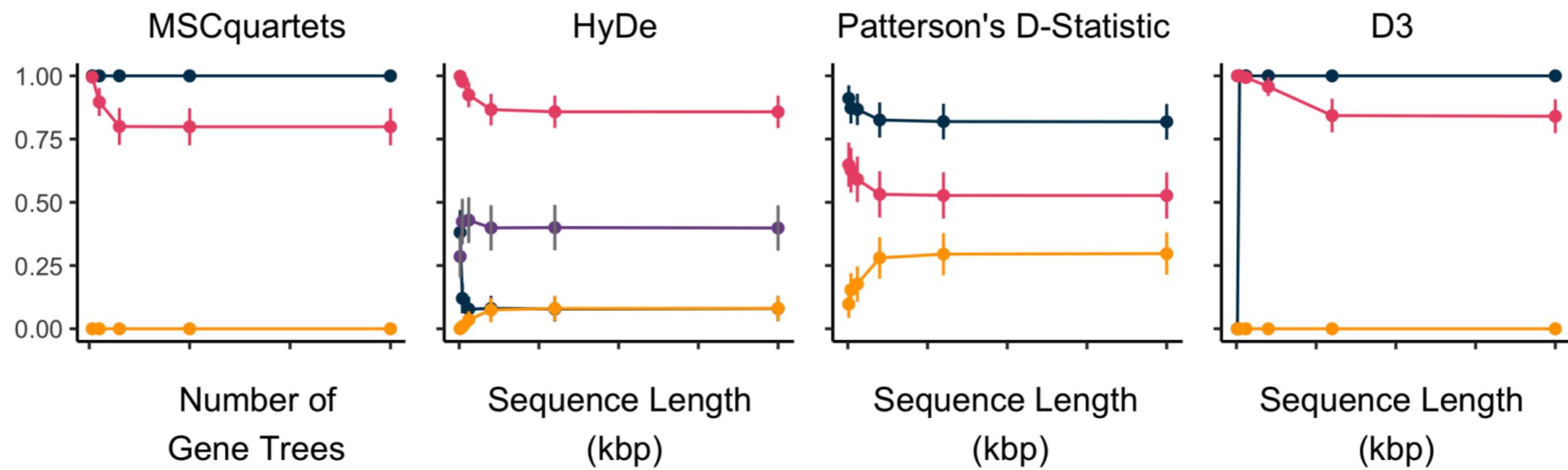
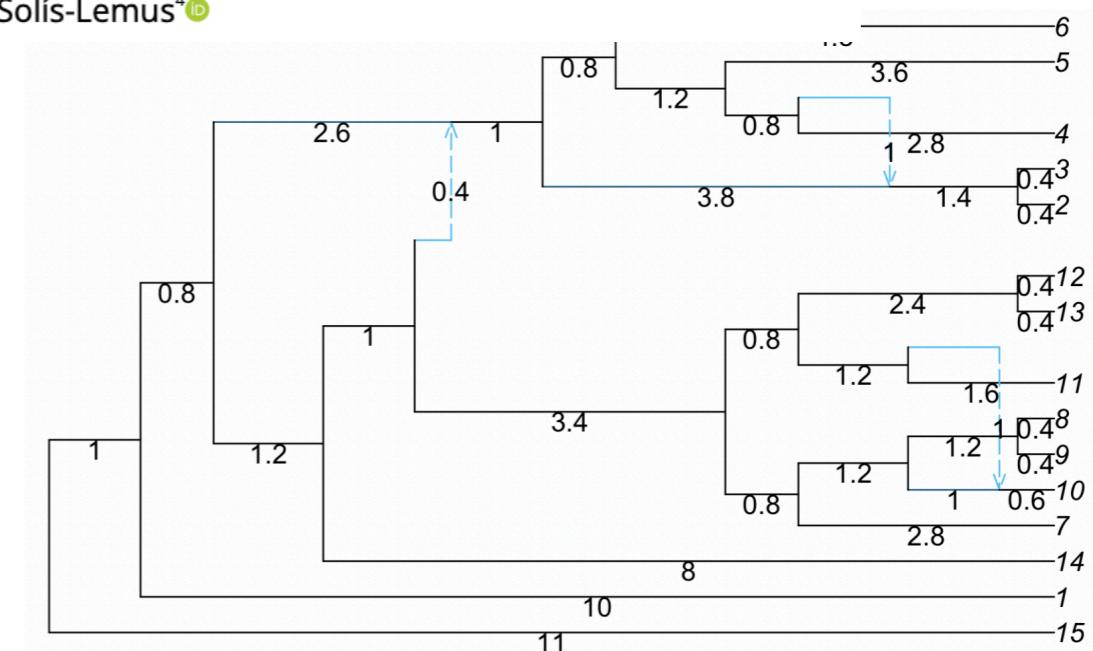
Marianne
Bjørner

Descriptions

Detectability of Varied Hybridization Scenarios Using Genome-Scale Hybrid Detection Methods

Marianne B Bjorner¹, Erin K Molloy², Colin N Dewey³, Claudia Solís-Lemus⁴

Ancient gene flow distorts discovery rates for hybrid detection methods



(See also Kong & Kubatko, 2021)



Laura 3:15 PM

An improvement to ABBA-BABA from Kevin and me (forthcoming in the Bulletin of the Society of Systematic Biologists):

<https://www.biorxiv.org/content/10.1101/2023.06.20.545699v1.full.pdf>



A Likelihood Ratio Test for Hybridization Under the Multispecies Coalescent

Jing Peng¹, Sungsik Kong^{2,3}, Laura Kubatko^{3,4*}



Sungsik (Kevin)
Kong

I want to find signal of hybridization in my data

Hybrid detection methods

HyDe

ABBA-BABA

MSCQuartets

LRT

Network inference methods

PhyloNet

SNaQ

BEAST2

NANUQ

RF-Net

PhyNEST

Finding hybridization requires integration of approaches

Hybrid detection + Network estimation

STEM-hy	gene trees rooted, BL	likelihood	hybridization b/w sister lineages
PhyloNet InferNetwork_ML	gene trees rooted	likelihood	
PhyloNet InferNetwork_MPL	gene trees rooted	triplet likelihood	
PhyloNetworks SNaQ	gene trees or quartet CFs	quartet likelihood	level-1 network
PhyloNet MCMC_GT	gene trees rooted	Bayesian	compound prior
PhyloNet MCMC_SEQ	alignments	Bayesian	compound prior no rate variation
BEAST2 SpeciesNetwork	alignments	Bayesian	birth-hyb prior
PhyloNet MLE_BiMarkers	biallelic sites	likelihood	compound prior
PhyloNet MCMC_BiMarkers	biallelic sites	Bayesian	compound prior
HyDe	sites	invariants	4 taxa, 1 hyb.
D-statistics	sites	freq. patterns	4 taxa, 1 hyb.
MSCQuartets	gene trees or quartet CF	quartet likelihood	4 taxa, 1 hyb.
NANUQ	gene trees or quartet CF	quartet likelihood	split (level-1) network
RF-Net	gene trees	min RF distance	No ILS
PhyNEST.jl	sites	quartet likelihood	level-1 network
PhyloDiamond.jl	gene trees or quartet CF	quartet likelihood	level-1 4-cycle network

Estimate explicit network under the coalescent

Hybrid detection methods

Estimate split network



JOURNAL ARTICLE

Ultrafast learning of four-node hybridization cycles in phylogenetic networks using algebraic invariants

✉

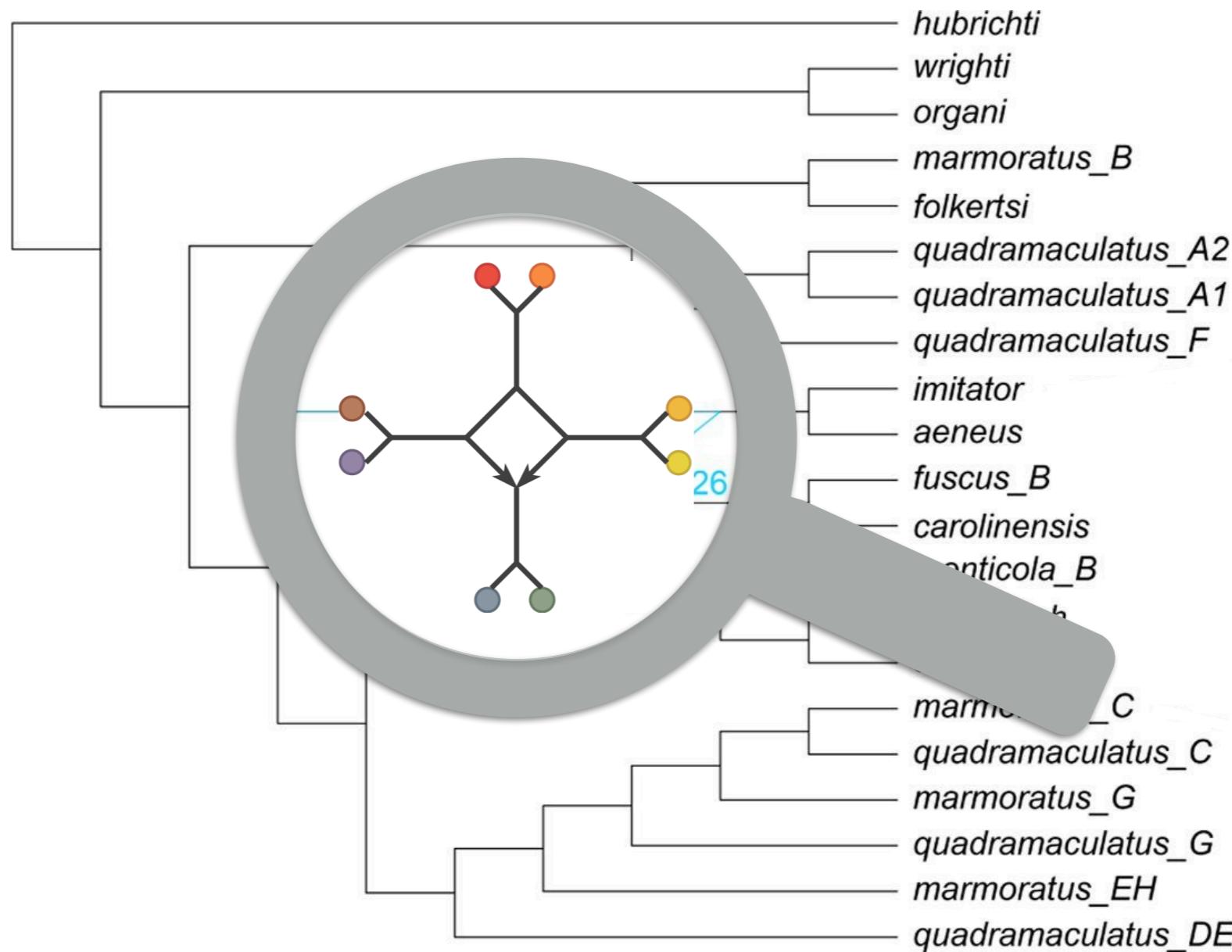
Zhaoxing Wu, Claudia Solís-Lemus ✉

Zhaoxing
(Bella) Wu

Bioinformatics Advances, Volume 4, Issue 1, 2024, vbae014,

Identifies hybridization cycles of 4 nodes

**Phylo
Diamond.jl**





JOURNAL ARTICLE

Ultrafast learning of four-node hybridization cycles in phylogenetic networks using algebraic invariants



Zhaoxing Wu, Claudia Solís-Lemus ✉

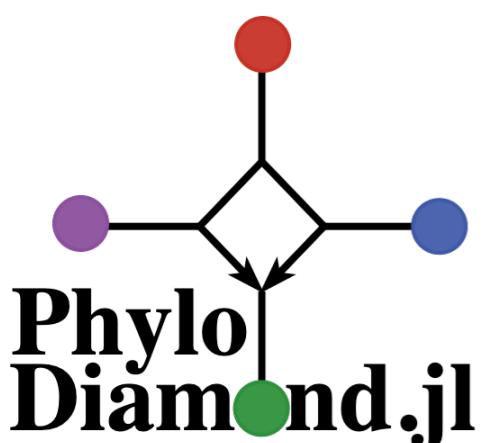
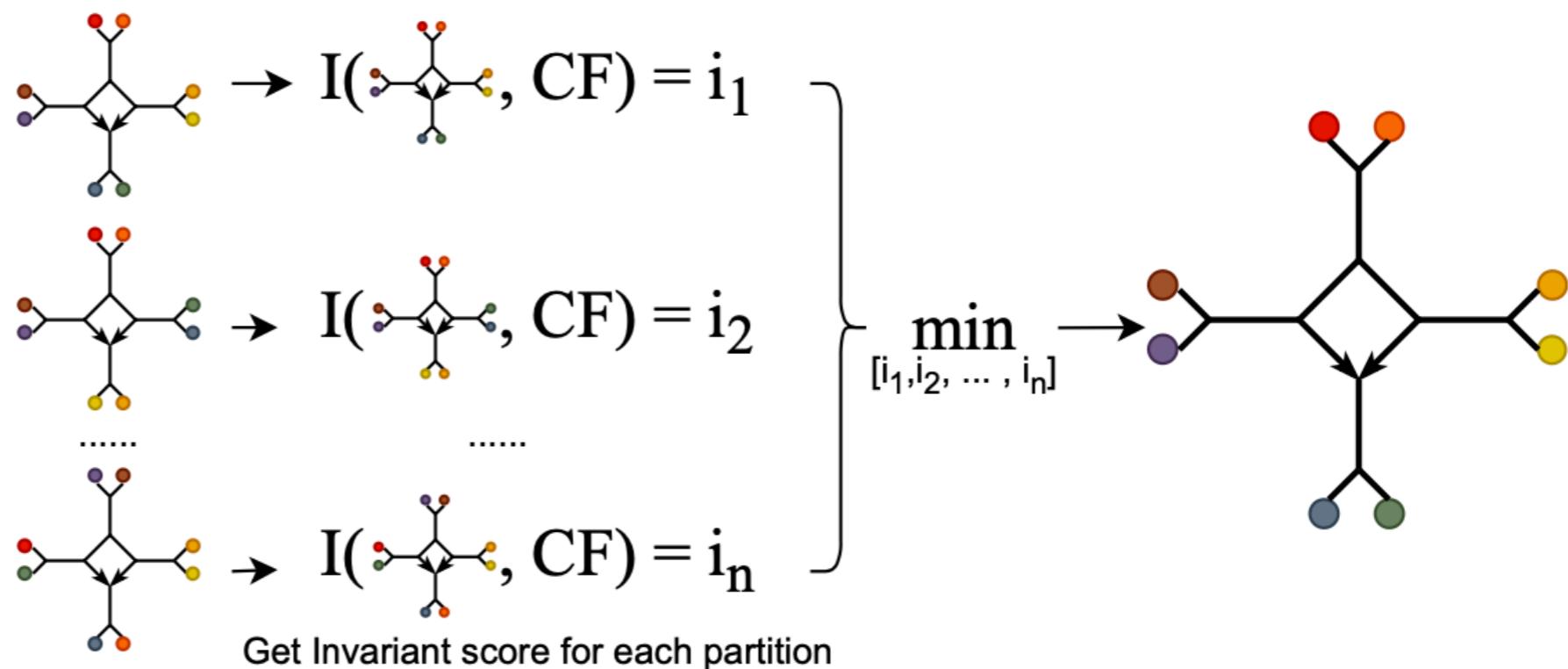
Bioinformatics Advances, Volume 4, Issue 1, 2024, vbae014,

Zhaoxing
(Bella) Wu

Iterate through partitions of taxa

Concordance Factor Table (CF)

	oo	o o	o o o	o o o o	o o o o o	o o o o o o
oo	red	orange	yellow	green	blue	brown
o o	orange	grey	black			
o o o		grey	black			
o o o o		grey	black	green		
o o o o o		grey	black	green	blue	
o o o o o o		grey	black	green	blue	brown





JOURNAL ARTICLE

Ultrafast learning of four-node hybridization cycles in phylogenetic networks using algebraic invariants

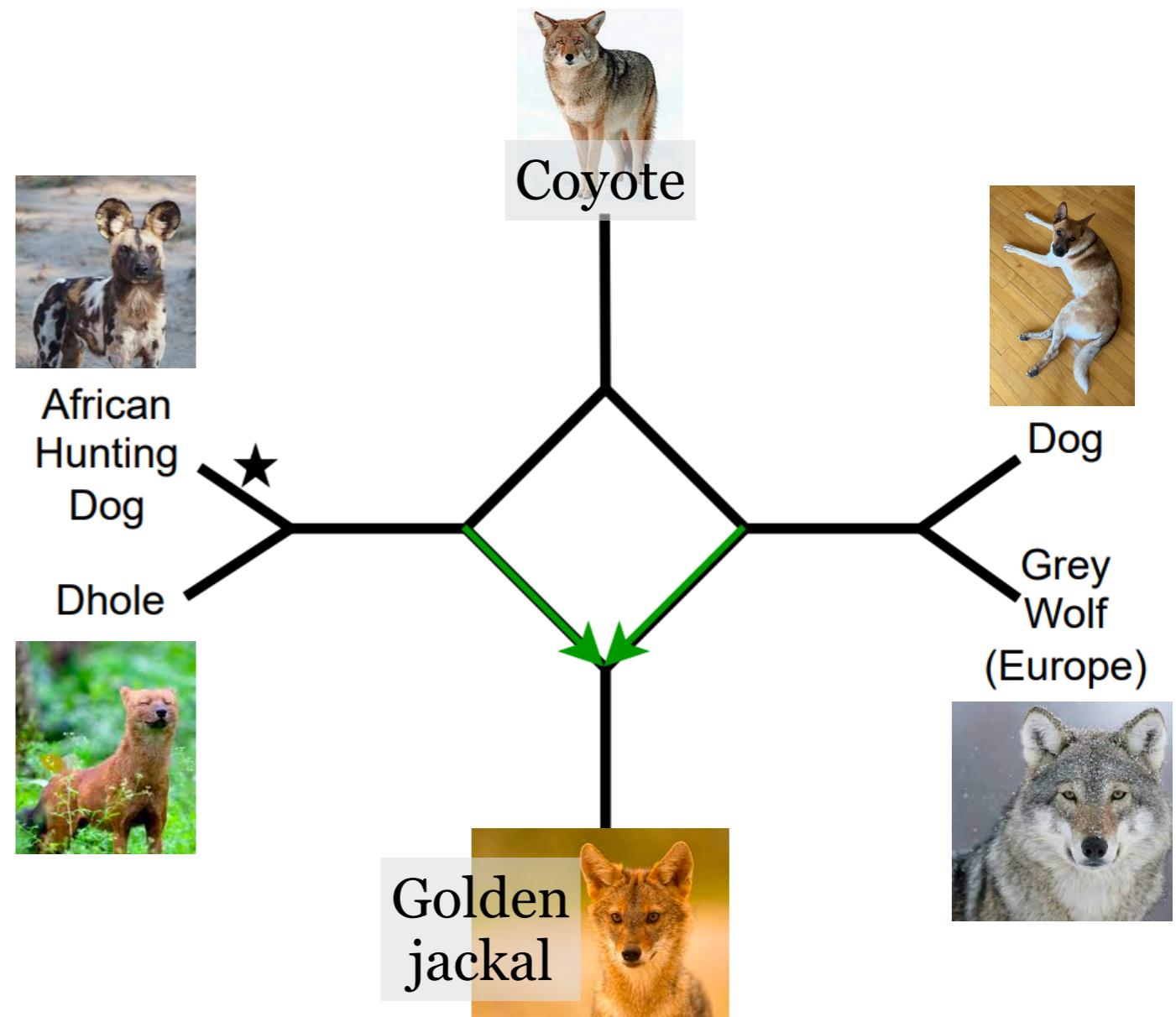
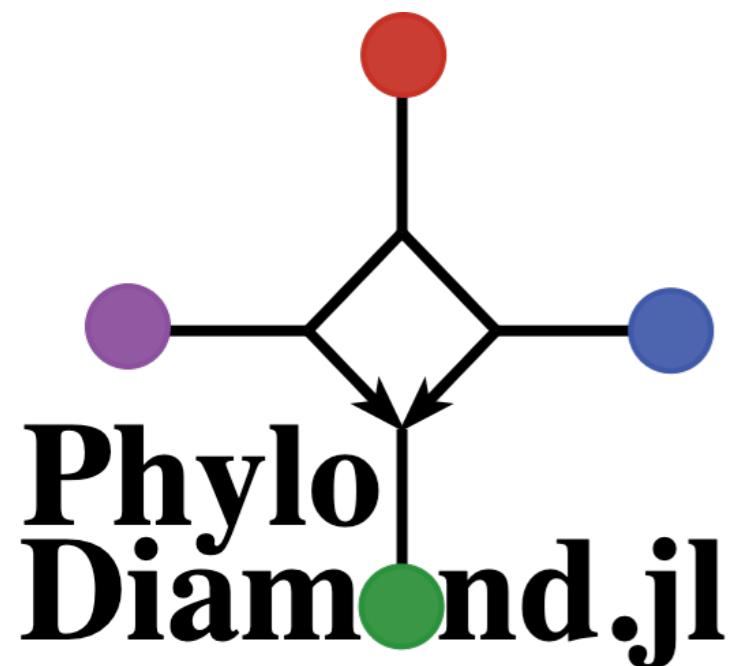
✉

Zhaoxing Wu, Claudia Solís-Lemus ✉

Zhaoxing
(Bella) Wu

Bioinformatics Advances, Volume 4, Issue 1, 2024, vbae014,

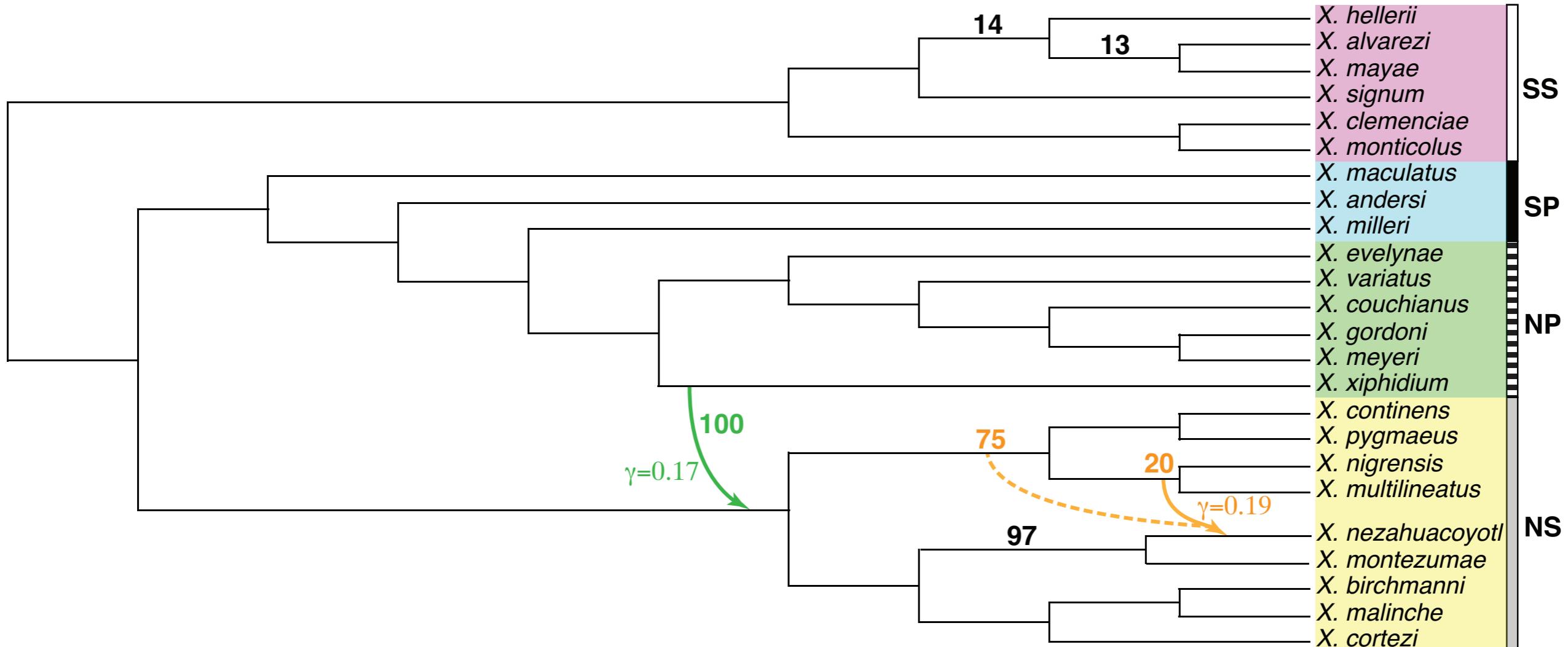
Method	Time (seconds)
Phylogenetic invariants (our method)	6.78
SNaQ	140.58
PhyloNet ML	2723.99
PhyloNet MPL	281.25



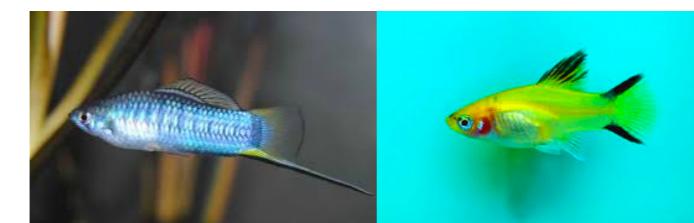
(Data from Gopalakrishnan et al, 2018)

Part II

I have the network, now what?



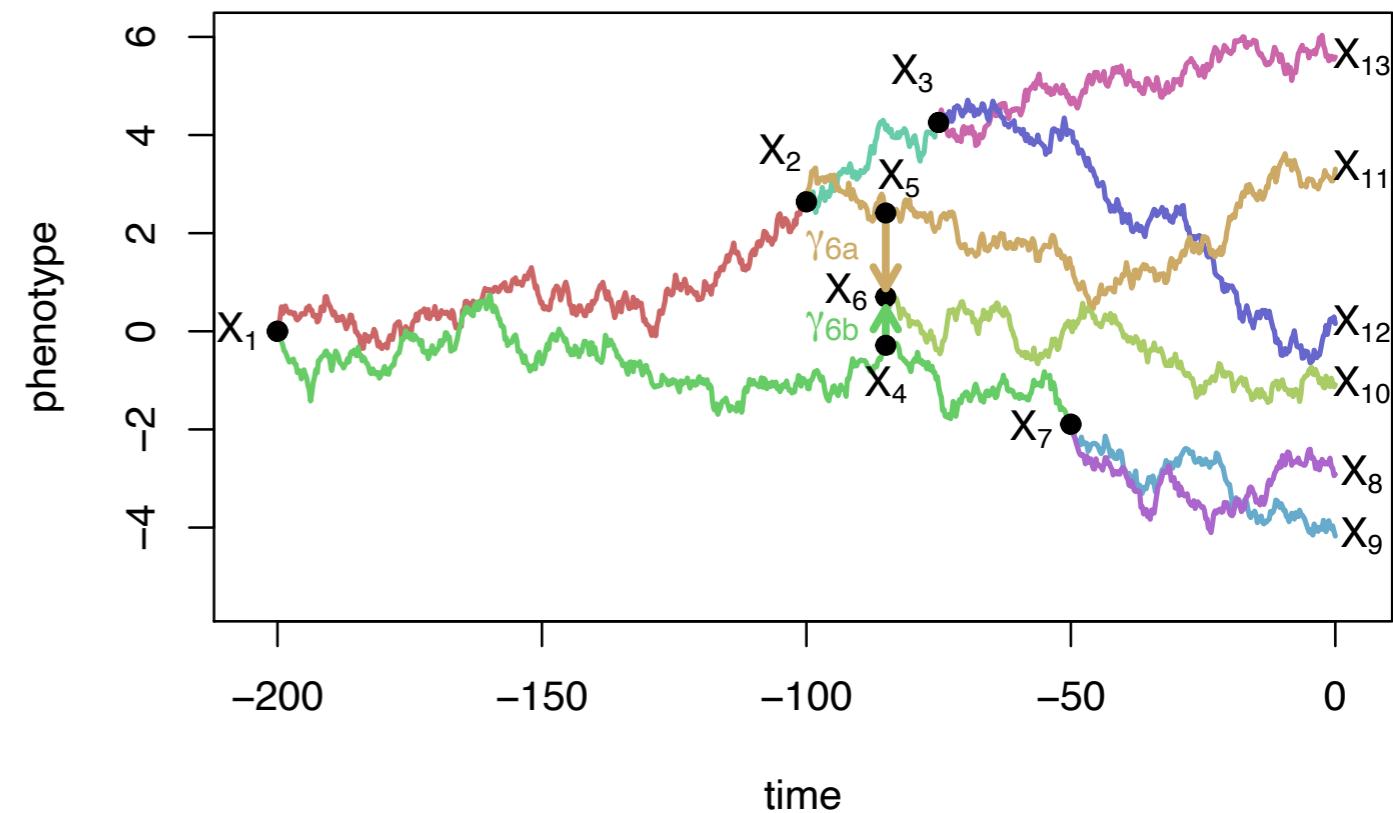
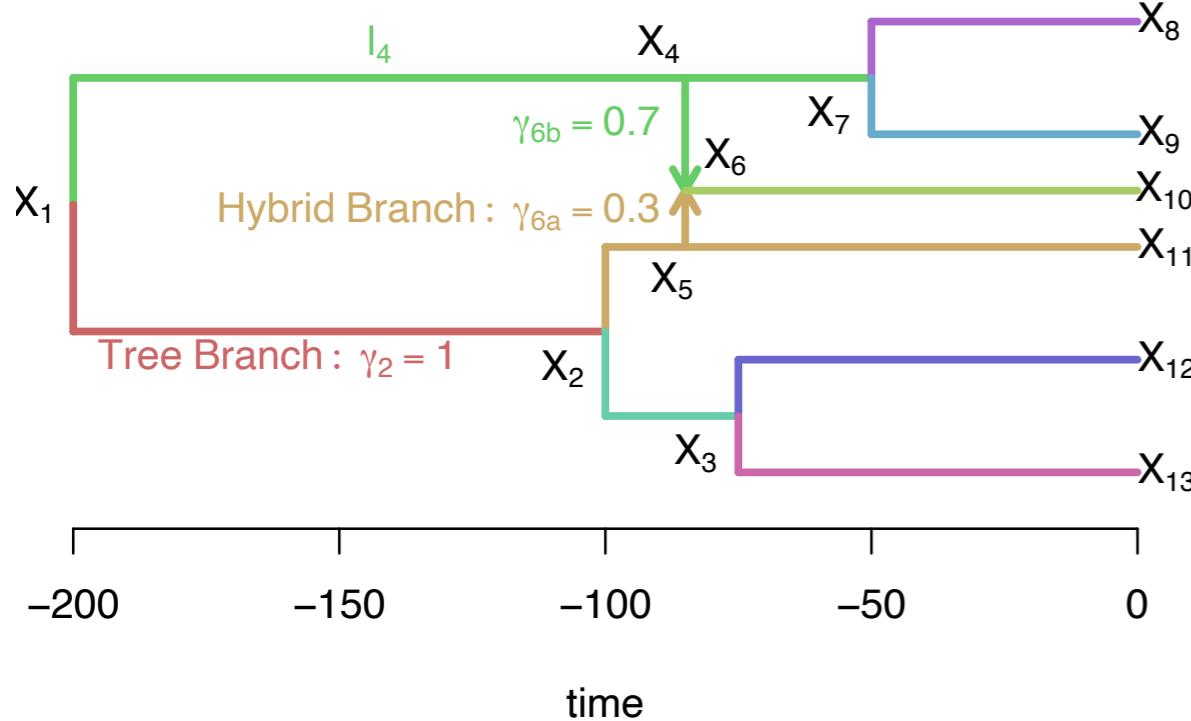
- Sword index
- Female preference



(Cui et al., 2013)

(Solís-Lemus, Ané, 2016, PLoS Genetics)

Trait models of evolution in networks



Brownian Motion
+ weighted
average in hybrid

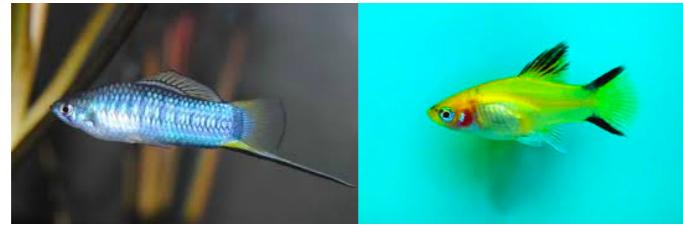
$$X_h = \gamma_1 X_{p_1} + \gamma_2 X_{p_2}$$

(Bastide et al, 2018, Syst Bio)

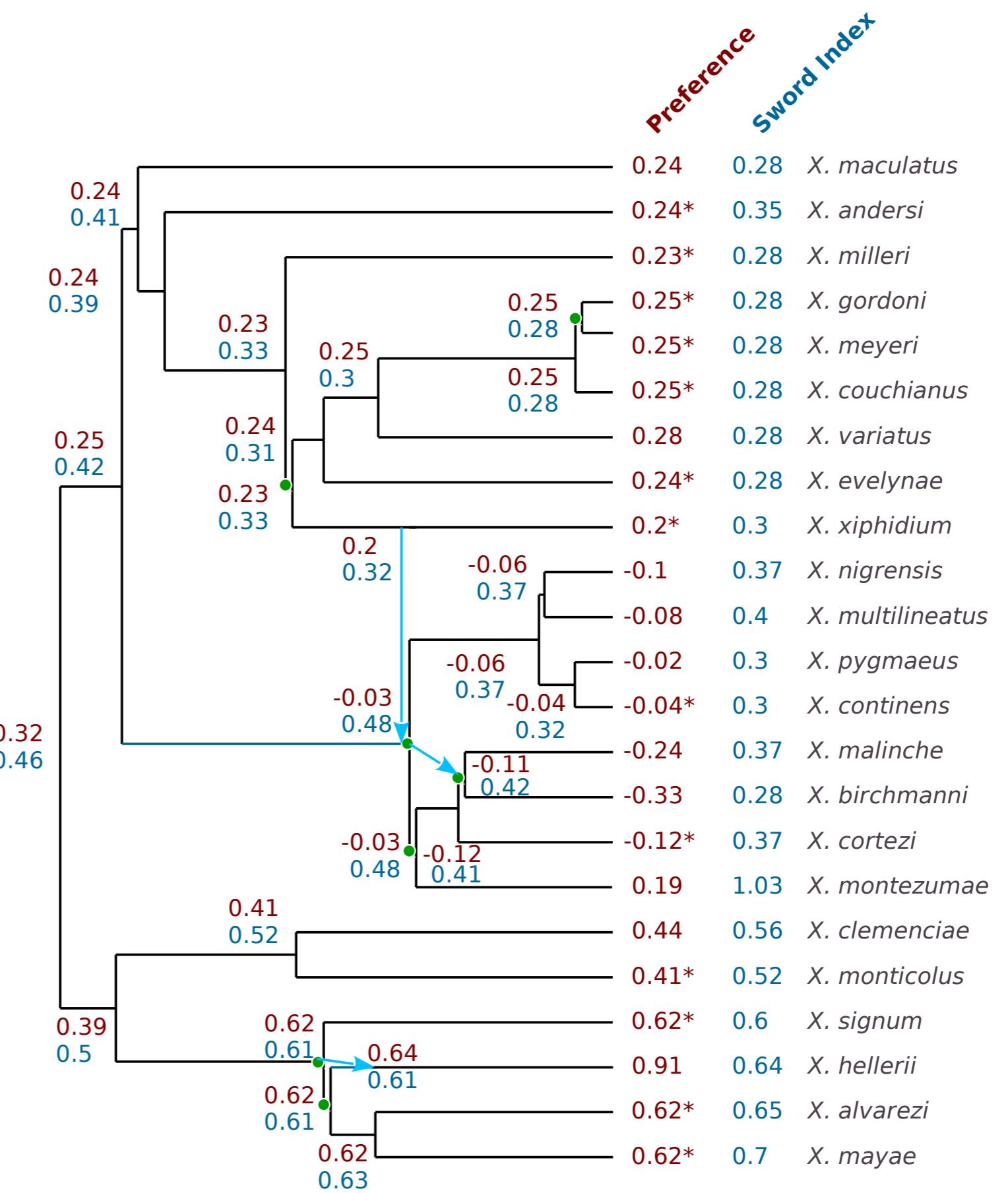
$$\mathbf{X} \sim N(X_{root}, \sigma^2 \mathbf{V})$$

- Phylogenetic signal
- Ancestral reconstruction
- Phylogenetic regression
- Phylogenetic ANOVA

- Sword index
- Female preference

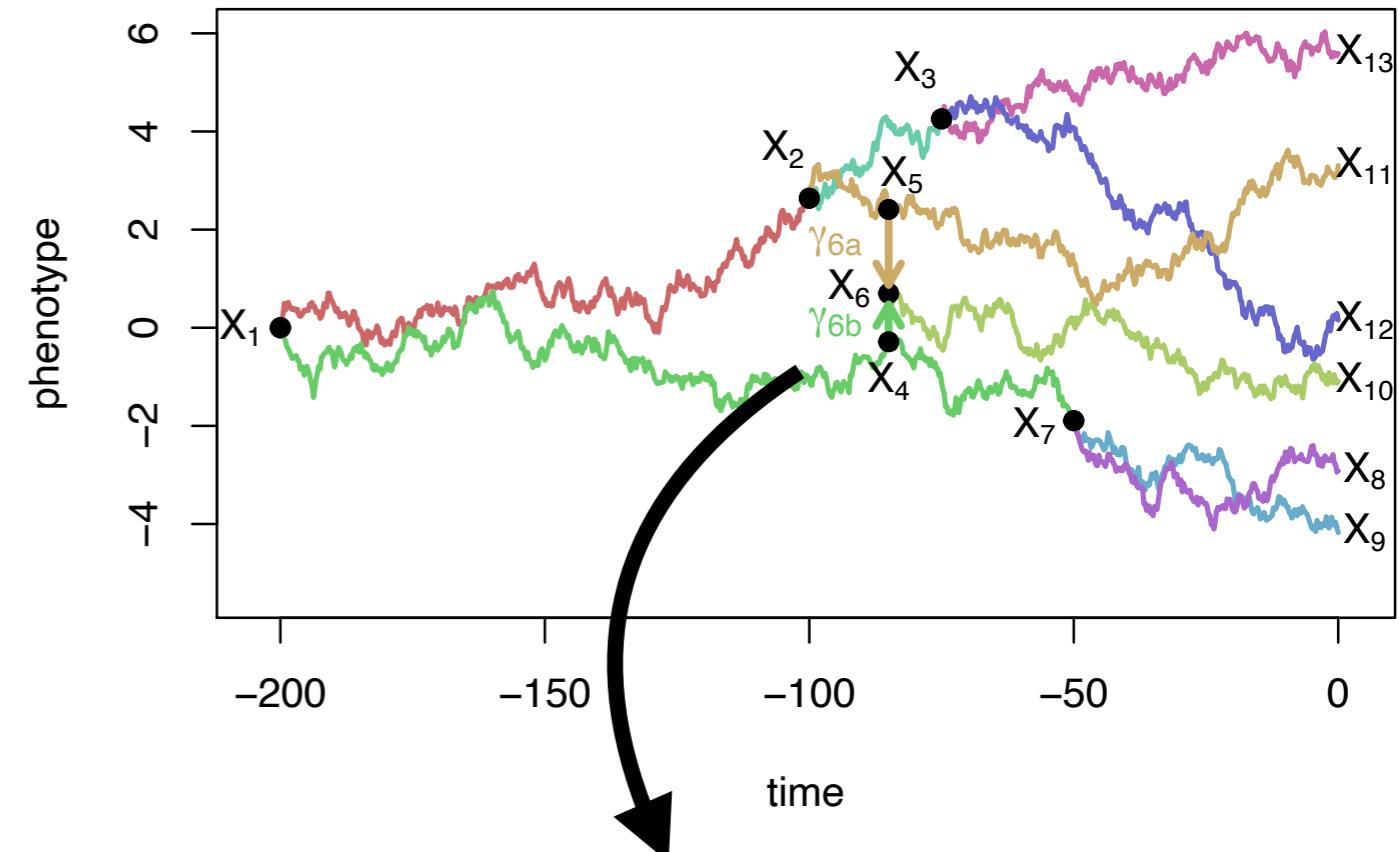
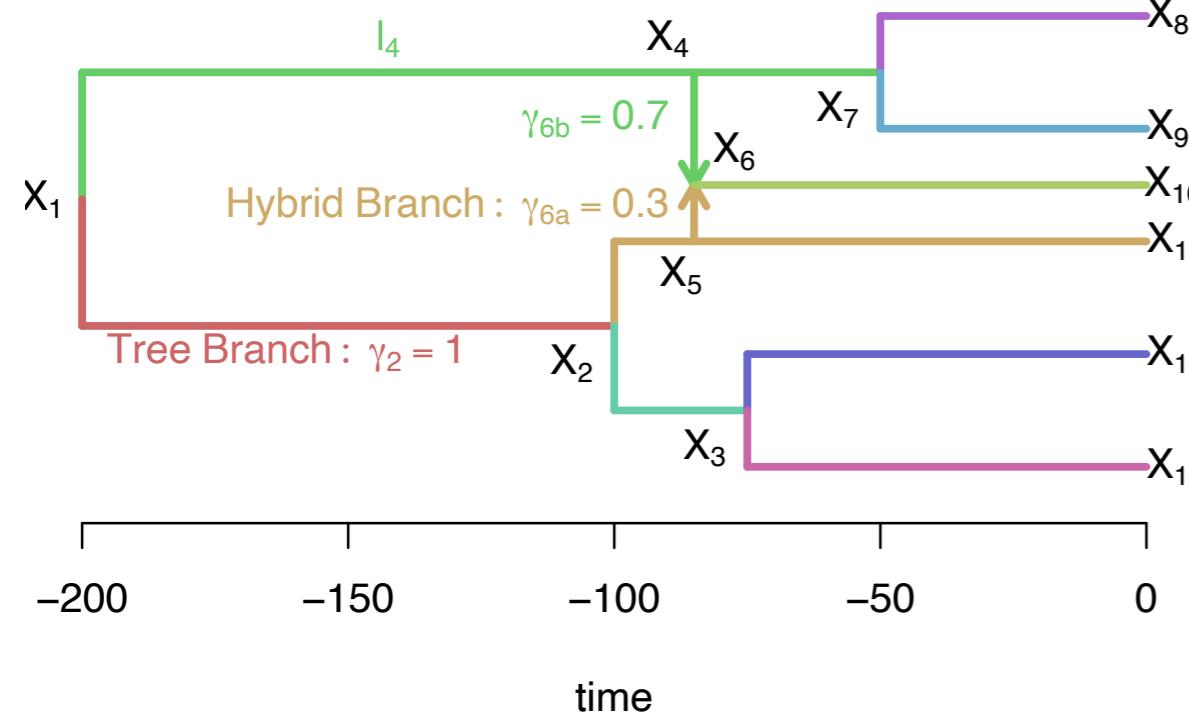


- **Ancestral reconstruction:** common ancestor likely had sword
- **Phylogenetic regression:** positive association between sword index and female preference but not significant ($p = 0.106$)



Test for transgressive evolution

$$X_h = \gamma_1 X_{p_1} + \gamma_2 X_{p_2} + \Delta_h$$



$\Delta_h = 0$ No transgressive evolution

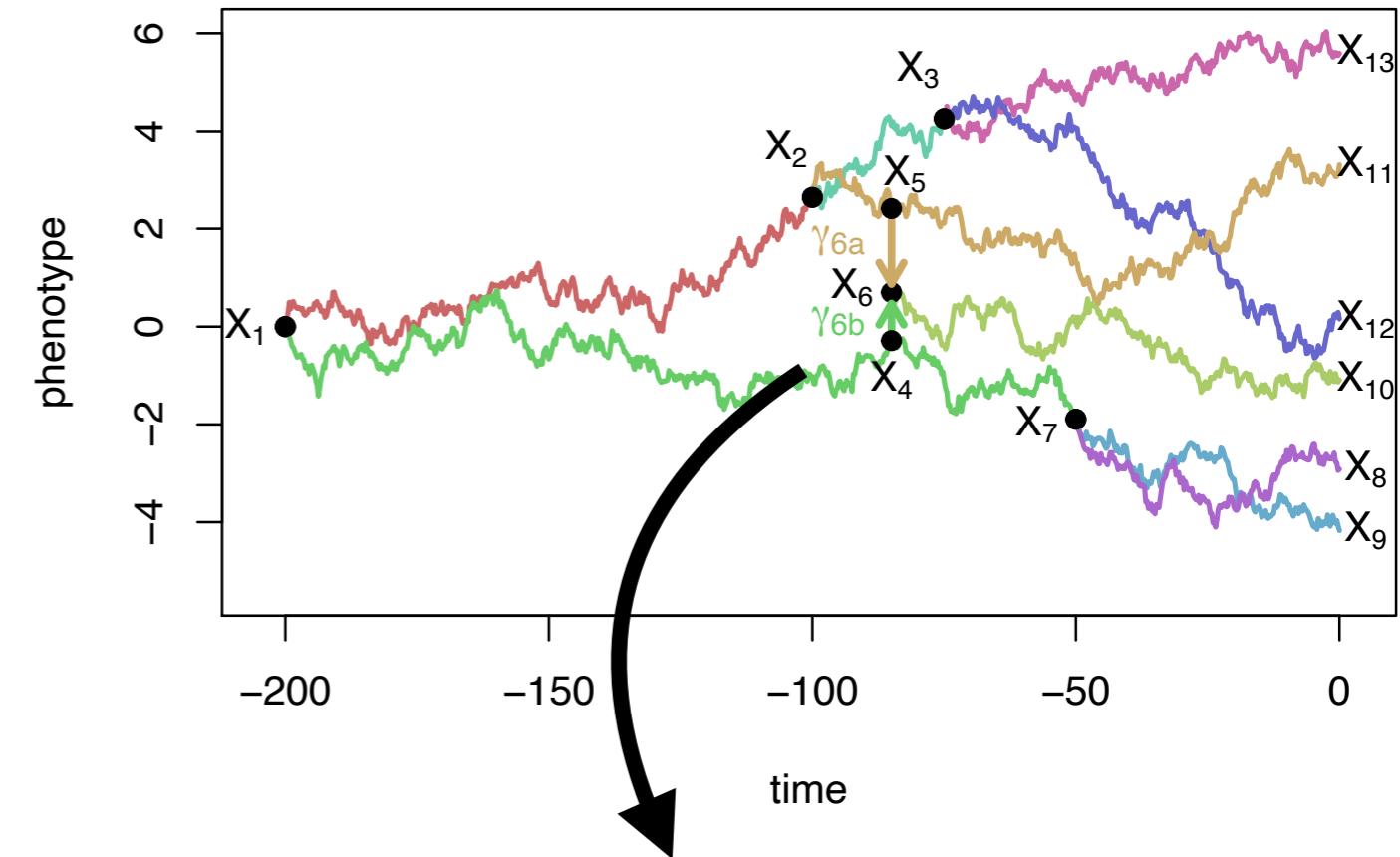
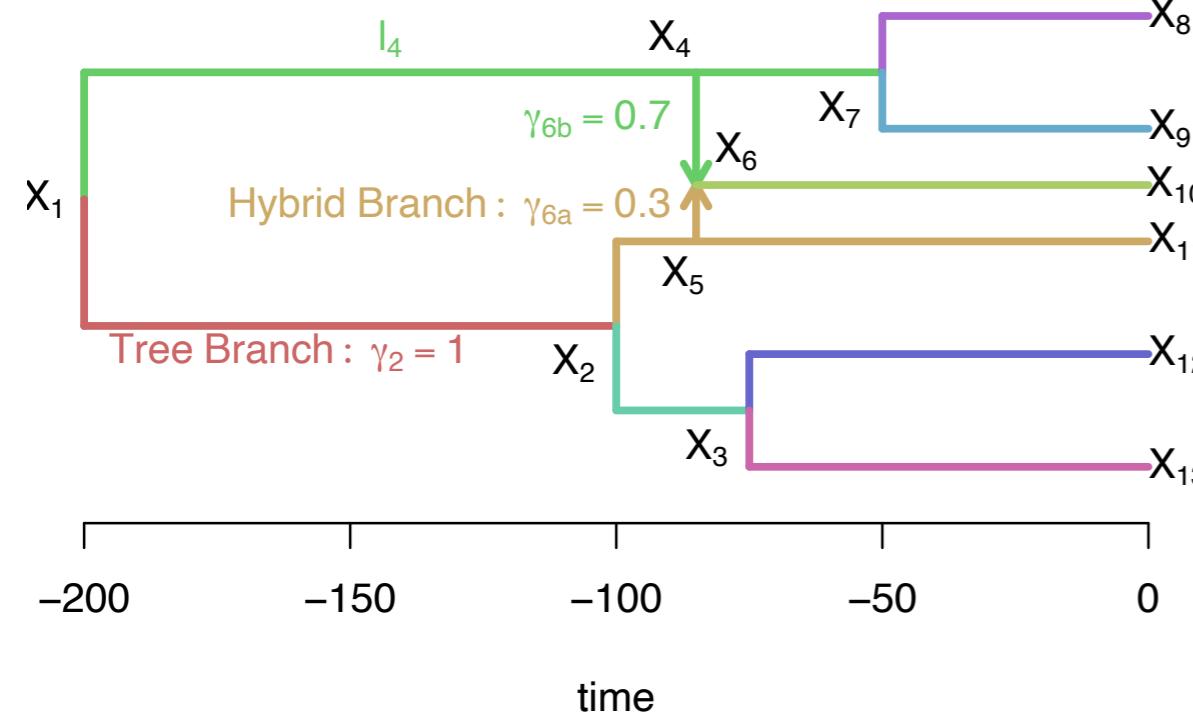
$\Delta_h = \Delta$ Single-effect transgressive evolution

Δ_h Multi-effect transgressive evolution

F tests

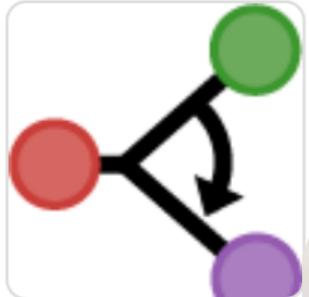
Hybrid value:
shift from
parents range

Test for transgressive evolution



- Sword index: $p=0.55$
- Female preference: $p=0.0064$

Hybrid value:
shift from
parents range



JuliaPhylo



Cécile Ané



Sungsik (Kevin)
Kong



Nathan Kolbow



Josh Justison



Ben Teo



Paul Bastide

<https://juliaphylo.github.io/JuliaPhyloWebsite/>



<https://solislemuslab.github.io/>



<mstdn.social/@solislemuslab>



crsl4



@thestatistician

julia

Machine-learning
black box



Ventilators for
COVID patients



Learning the Tree
of Life



Proteins in plant
viruses



Fungi and
fungus-farming
insects



Simulations of
HIV epidemics



Statistical
ecology of sharks



EL ZOOMINARIO

VISIBILITY OF LATINX IN STEM

1st Friday every month 3pm CT

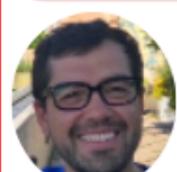


Plant membrane
lipids



@latinxinstem

Sustainable dairy
industry



A quest into the
origins of life



Cardiac
arrhythmias and
computers



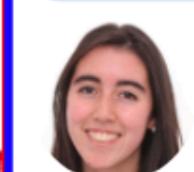
Plants and
herbivorous
insects



Does stress helps
us survive?



What makes
potatoes sick?



<https://solislemuslab.github.io/>



<mstdn.social/@solislemuslab>

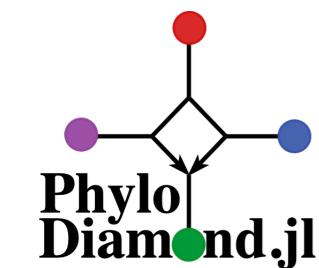


crsl4



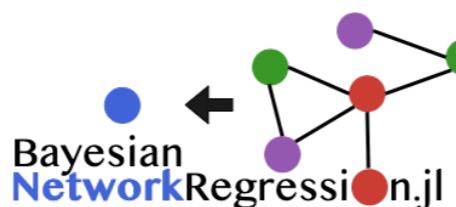
@thestatistician

Solis-Lemus lab



MiNAA

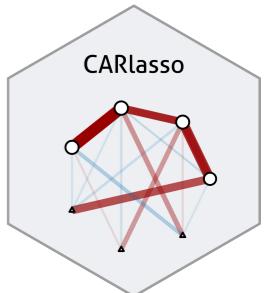
BioKlustering!



snaQ

WISCONSIN
FAST STATS

phylo
networks



WISCONSIN
UNIVERSITY OF WISCONSIN-MADISON



USDA

WID



<https://solislemuslab.github.io/>



<mstdn.social/@solislemuslab>



[solislemuslab](#)



@thestatistician

Questions?