

Expanding the space of phylogenetic trees to infer patterns predicted by shared processes of diversification

Evolution 2021

Jamie R. Oaks

Auburn University

phyletica.org

@jamoaks

Perry L. Wood, Jr.

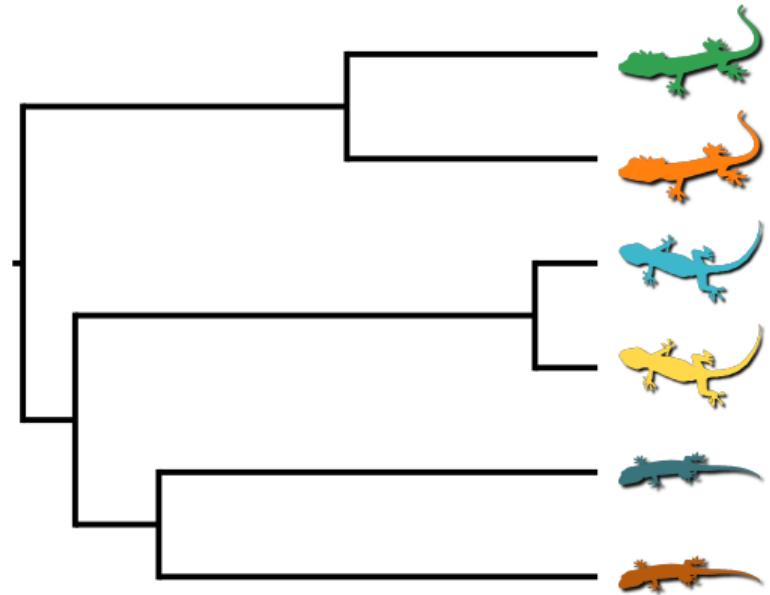
Auburn University

perryleewoodjr.com

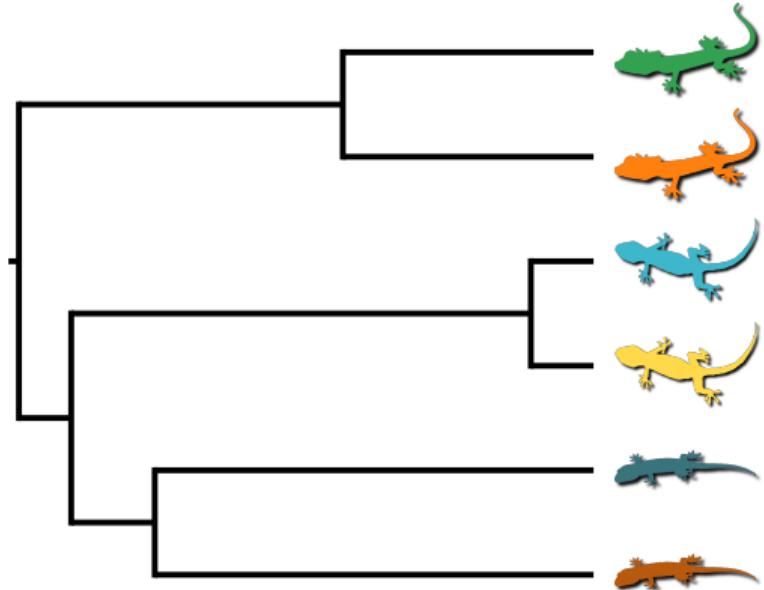
@perryleewoodjr

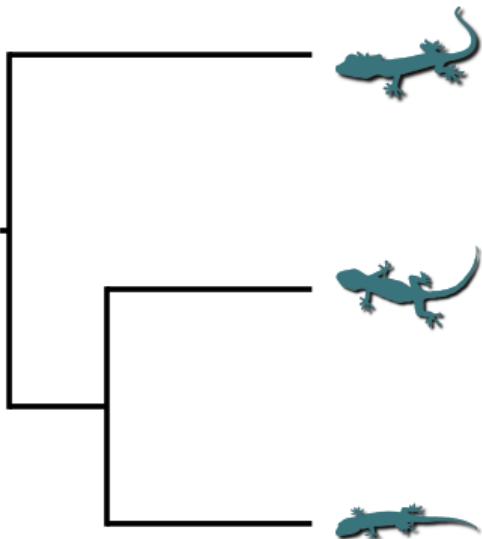


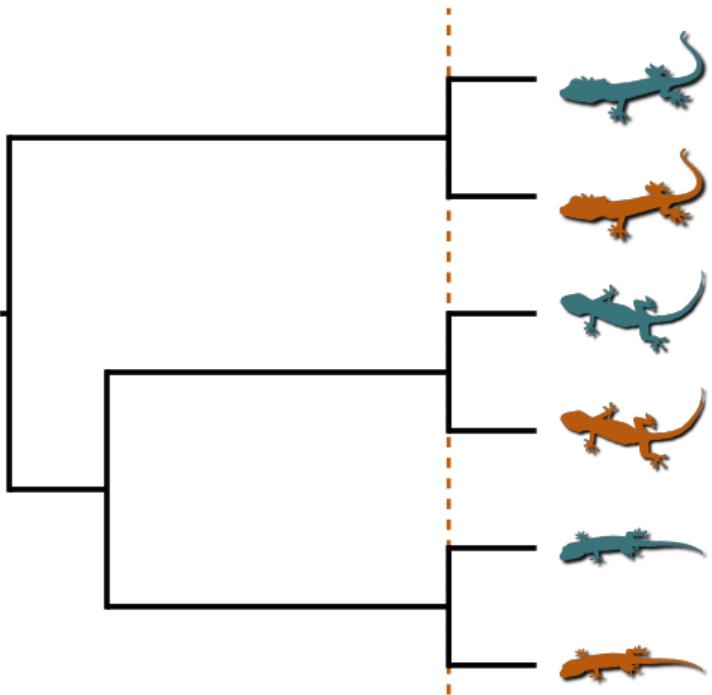
Thanks to Evolution organizers!

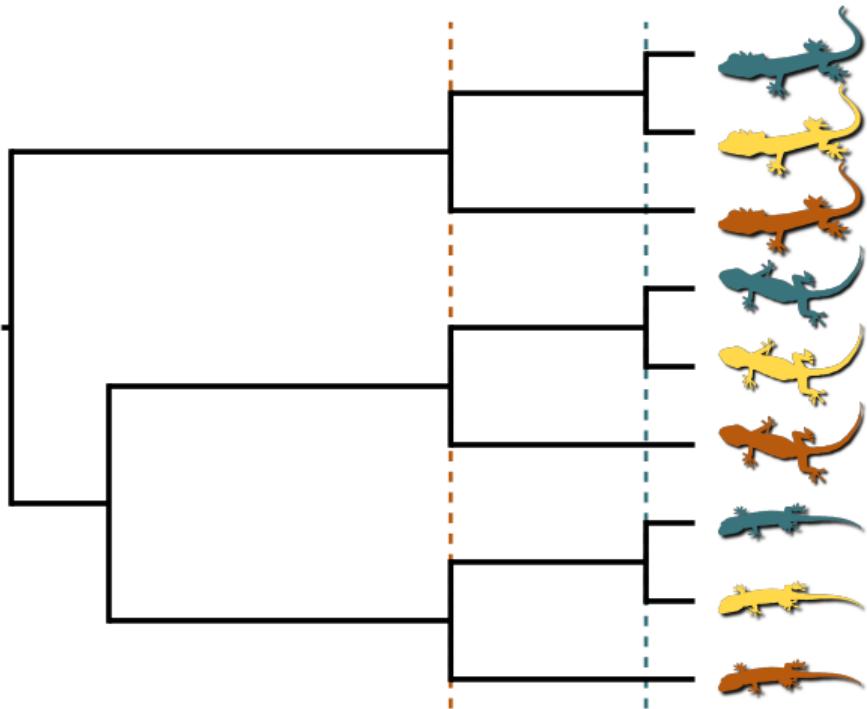


- ▶ **Assumption:** All processes of diversification affect each lineage independently and only cause bifurcating divergences.







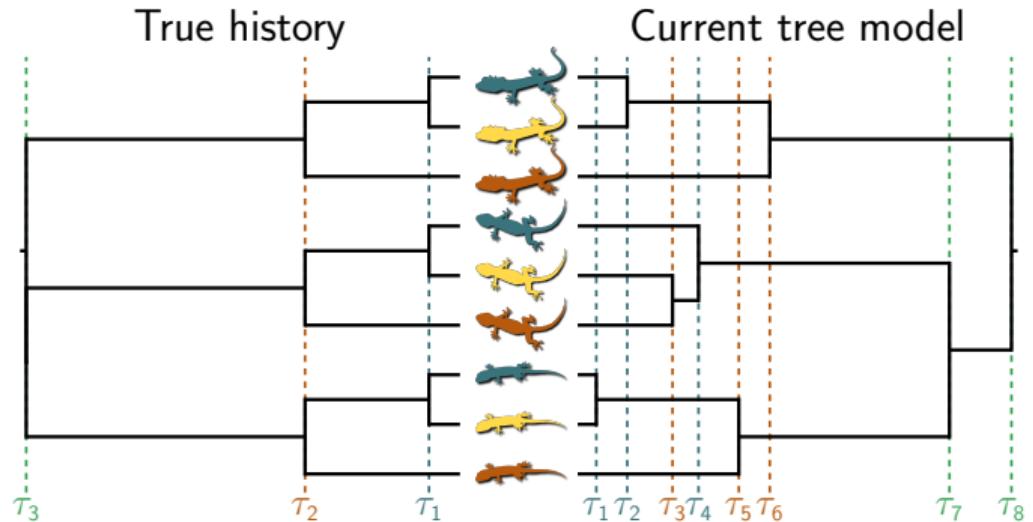


Why account for shared divergences?

1. Improve inference

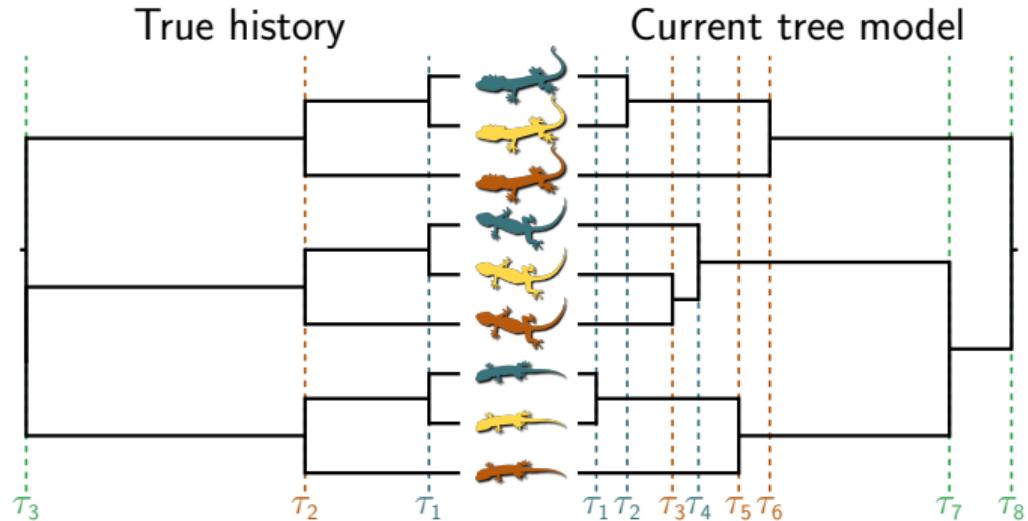
Why account for shared divergences?

1. Improve inference



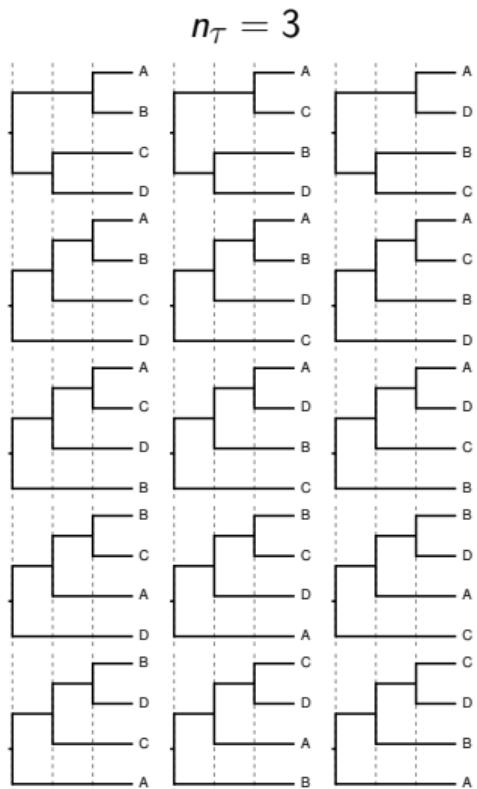
Why account for shared divergences?

1. Improve inference
2. **Provide a framework for studying processes of co-diversification**



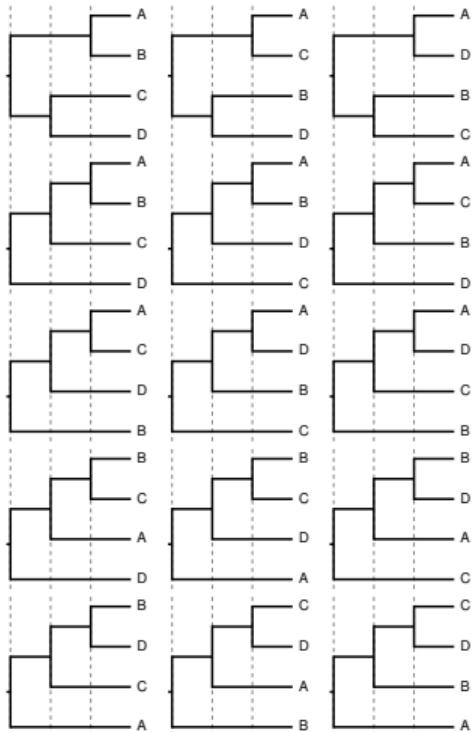
Generalizing tree space

Generalizing tree space

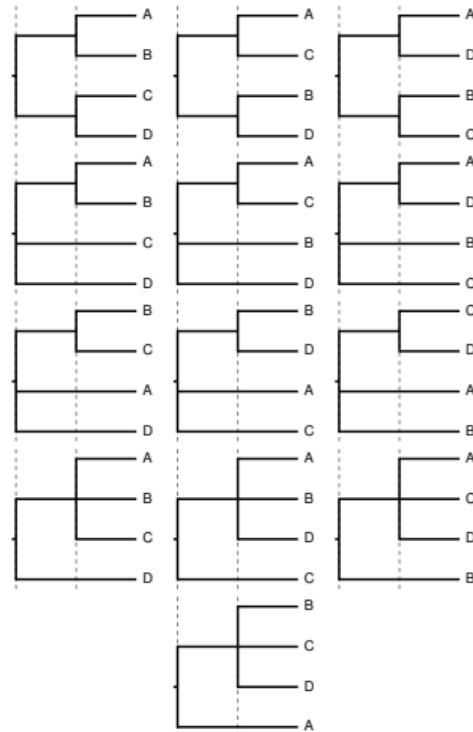


Generalizing tree space

$$n_T = 3$$

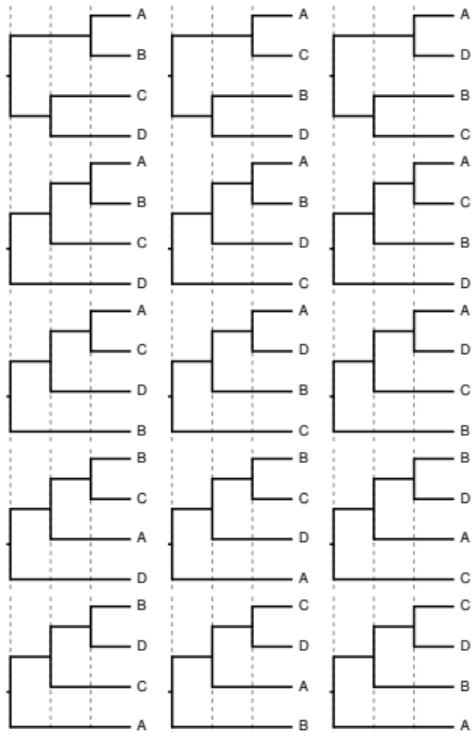


$$n_T = 2$$

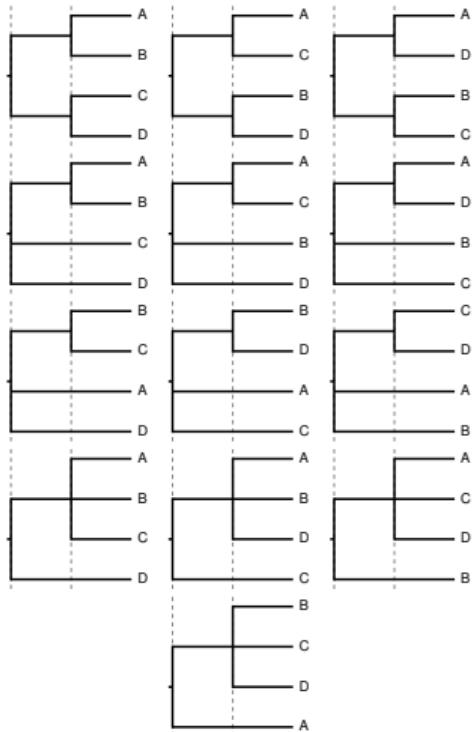


Generalizing tree space

$$n_T = 3$$



$$n_T = 2$$

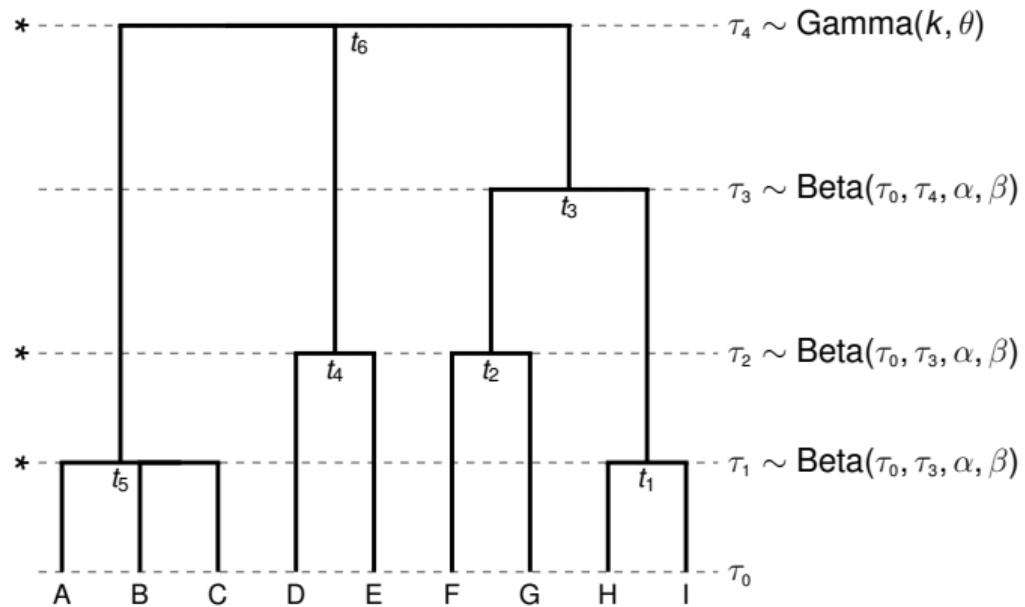


$$n_T = 1$$

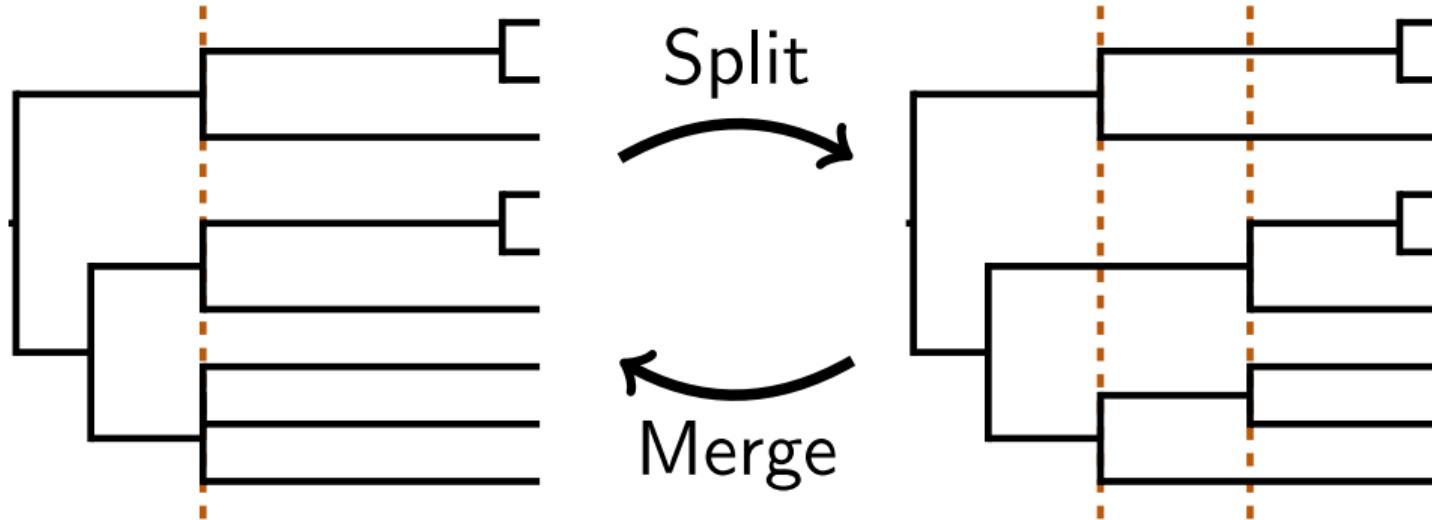


Generalized tree distribution

- ▶ All topologies equally probable
- ▶ Parametric distribution on age of root
- ▶ Beta distributions on other div times

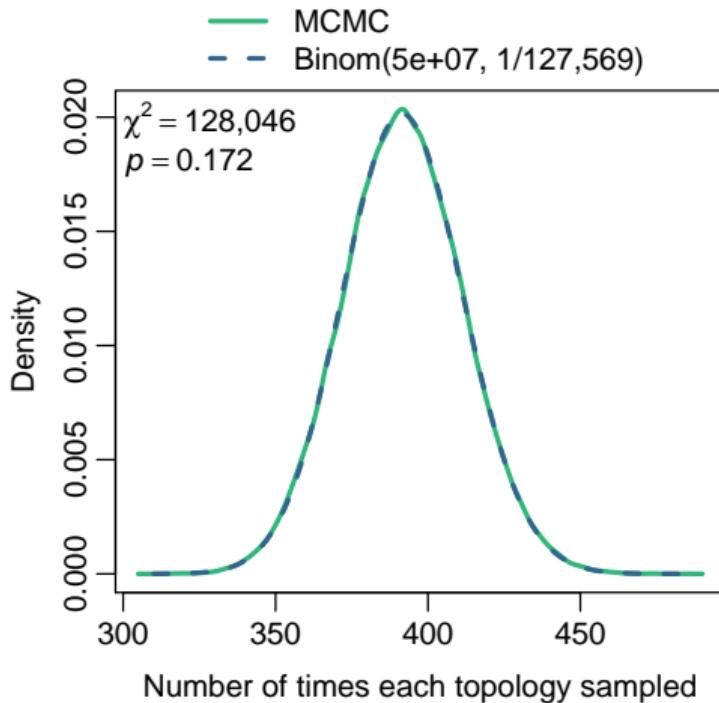


Inferring trees with shared divergences



Reversible-jump MCMC

Validating rjMCMC with 7-leaf tree



The rjMCMC algorithms sample the expected generalized tree distribution

Phycoeval (part of **Ecoevolity**²)

¹ D. Bryant et al. (2012). *Molecular Biology and Evolution* 29: 1917–1932

² J. R. Oaks (2019). *Systematic Biology* 68: 371–395

Phycoeval (part of **Ecoevolity**²)

- ▶ Tree model
 - ▶ rjMCMC sampling of generalized tree distribution

¹ D. Bryant et al. (2012). *Molecular Biology and Evolution* 29: 1917–1932

² J. R. Oaks (2019). *Systematic Biology* 68: 371–395

Phycoeval (part of **Ecoevolvity**²)

- ▶ Tree model
 - ▶ rjMCMC sampling of generalized tree distribution
- ▶ Likelihood model
 - ▶ CTMC model of characters evolving along genealogies
 - ▶ Infer species trees by analytically integrate over genealogies¹

¹ D. Bryant et al. (2012). *Molecular Biology and Evolution* 29: 1917–1932

² J. R. Oaks (2019). *Systematic Biology* 68: 371–395

Phycoeval (part of **Ecoevolvity**²)

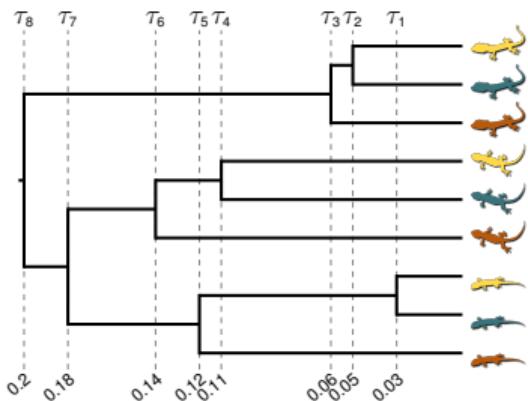
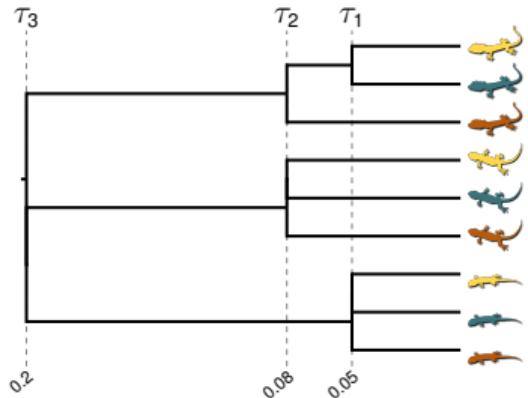
- ▶ Tree model
 - ▶ rjMCMC sampling of generalized tree distribution
- ▶ Likelihood model
 - ▶ CTMC model of characters evolving along genealogies
 - ▶ Infer species trees by analytically integrate over genealogies¹
- ▶ *Goal: Co-estimation of phylogeny and shared divergences from genomic data*

¹ D. Bryant et al. (2012). *Molecular Biology and Evolution* 29: 1917–1932

² J. R. Oaks (2019). *Systematic Biology* 68: 371–395

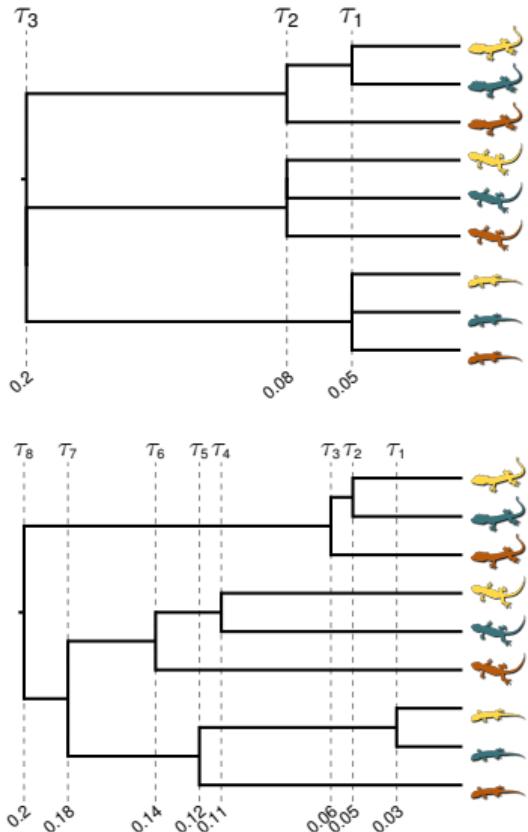
Methods: Simulations

- ▶ Simulated 100 data sets with 50,000 characters
- ▶ Strict clock



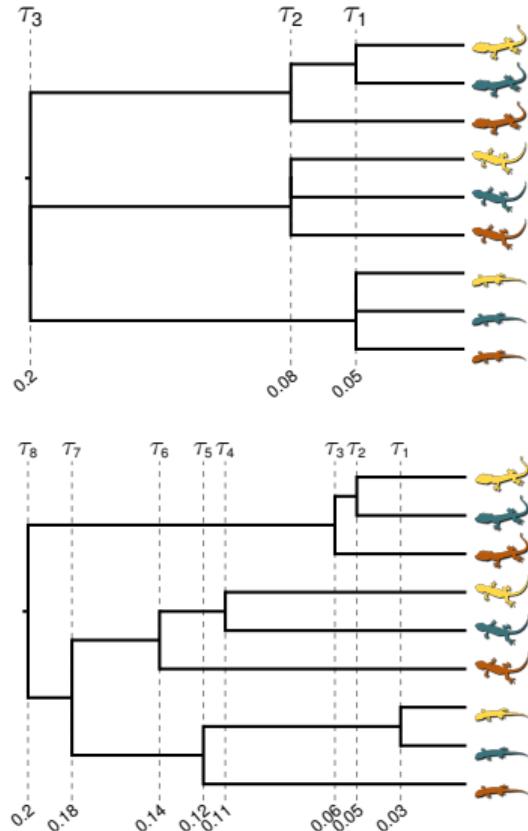
Methods: Simulations

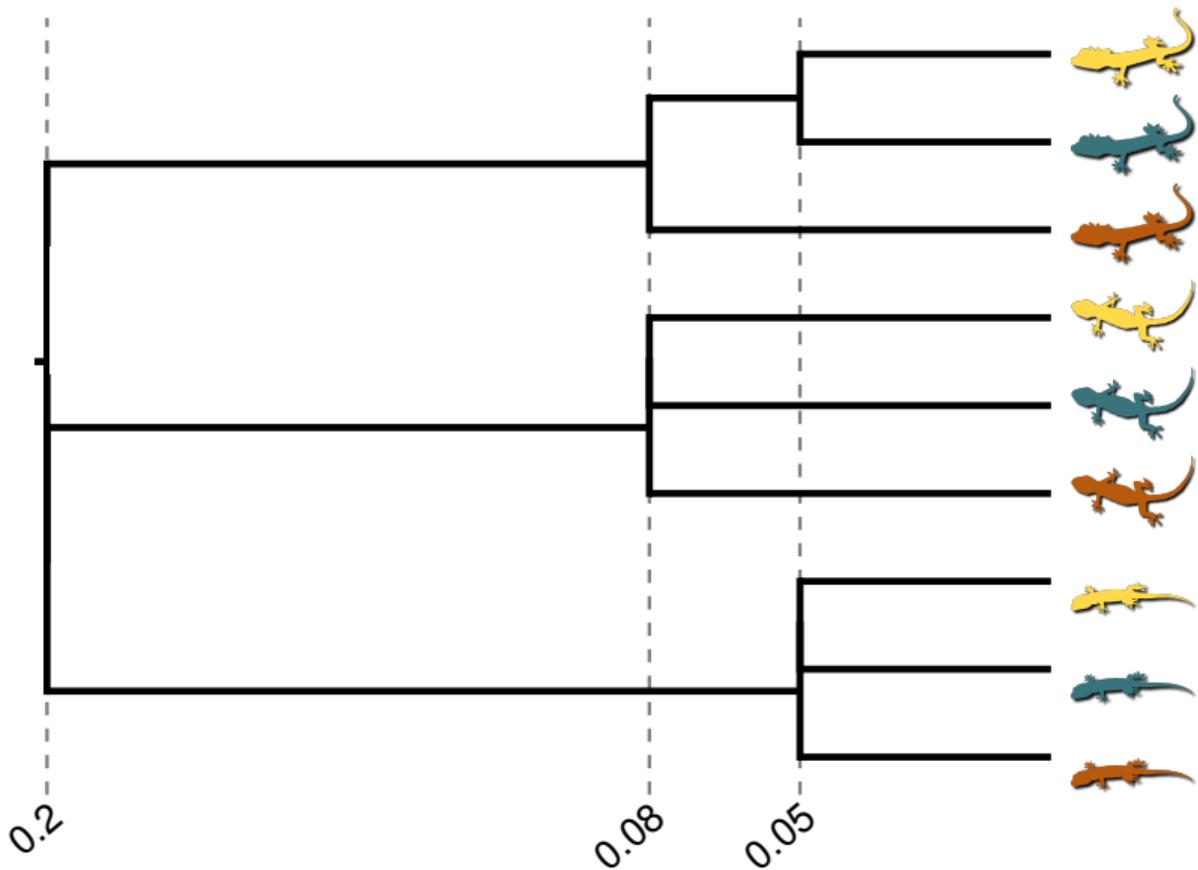
- ▶ Simulated 100 data sets with 50,000 characters
- ▶ Strict clock
- ▶ Analyze each data set with:
 - ▶ M_G = Generalized tree model
 - ▶ M_{IB} = Independent-bifurcating tree model

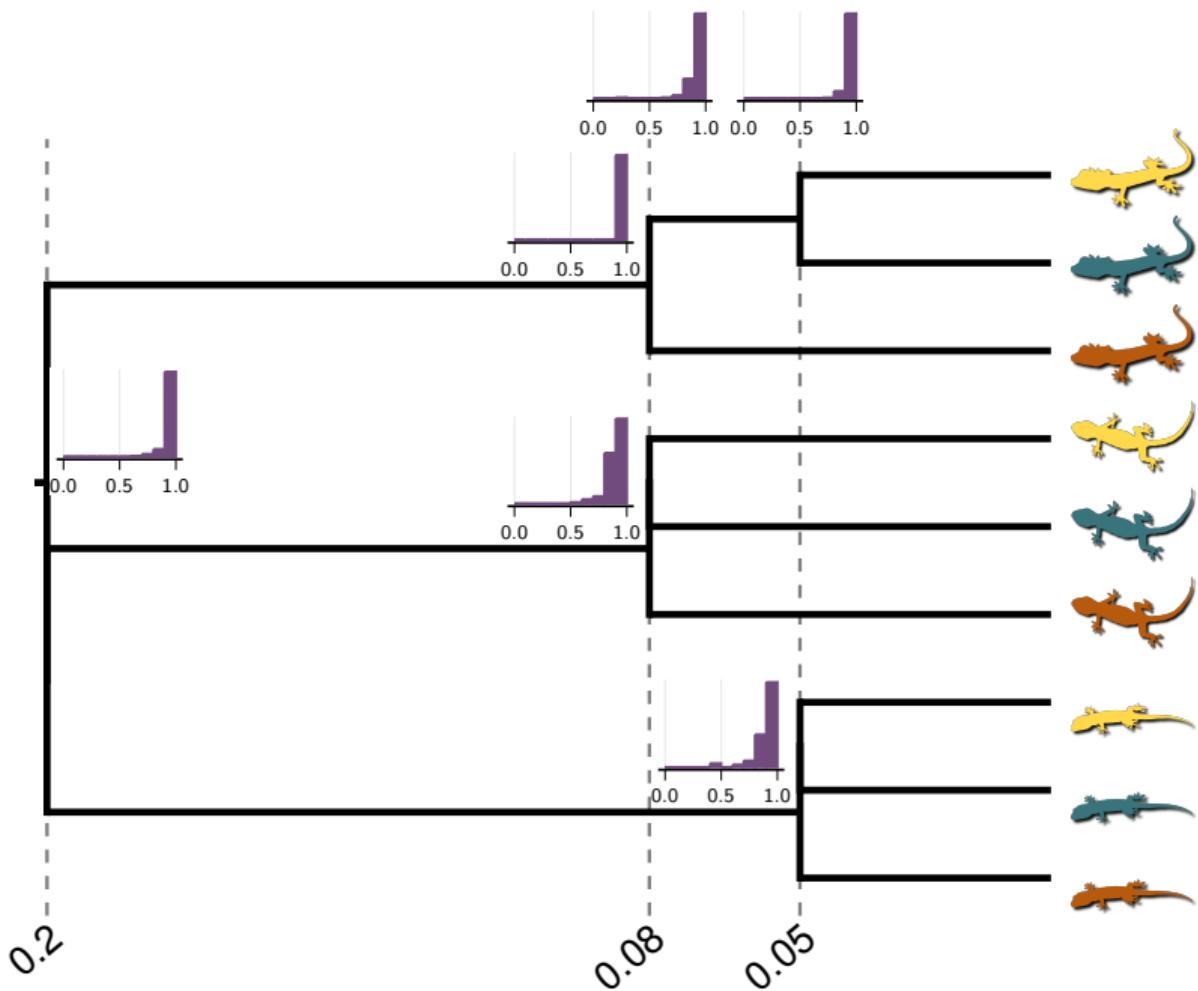


Methods: Simulations

- ▶ Simulated 100 data sets with 50,000 characters
- ▶ Strict clock
- ▶ Analyze each data set with:
 - ▶ M_G = Generalized tree model
 - ▶ M_{IB} = Independent-bifurcating tree model
- ▶ We also did simulations where topology and div times drawn from M_G and M_{IB} distributions

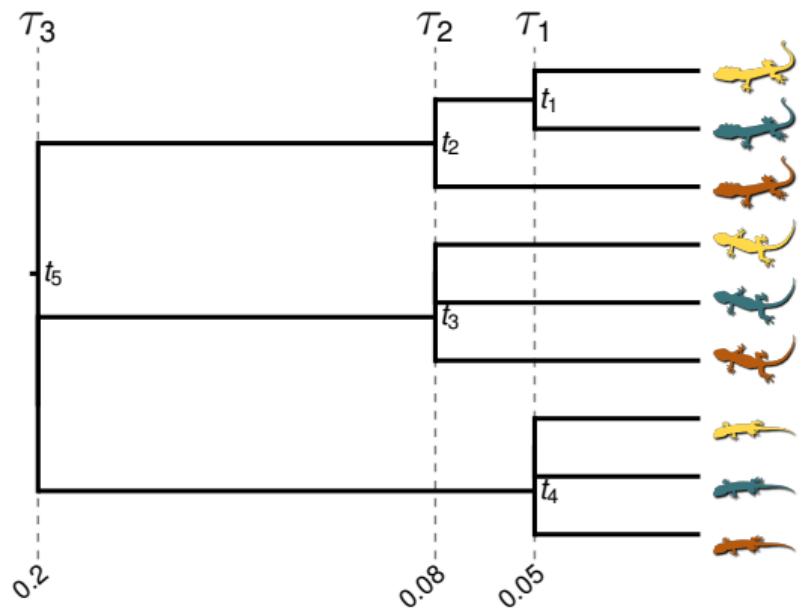




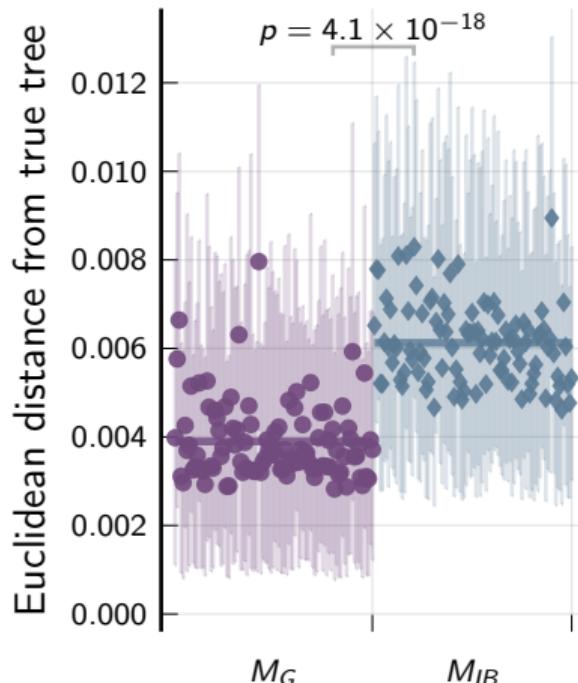


● M_G = Generalized model

◆ M_{IB} = Independent-bifurcating model

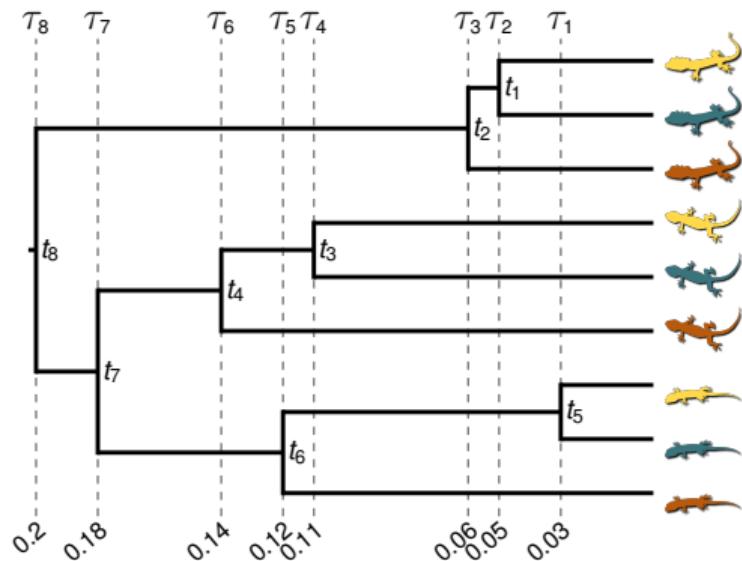


Wilcoxon signed-rank test P-value = $4.08e^{-18}$



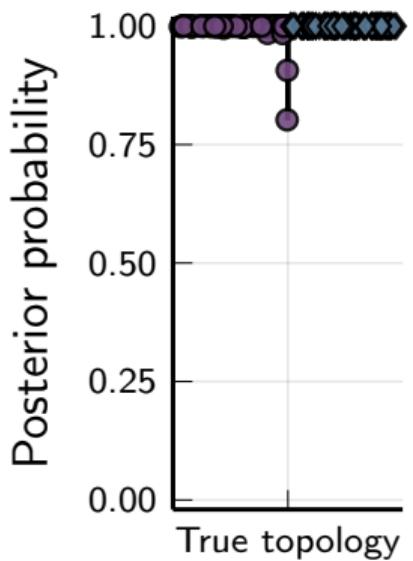
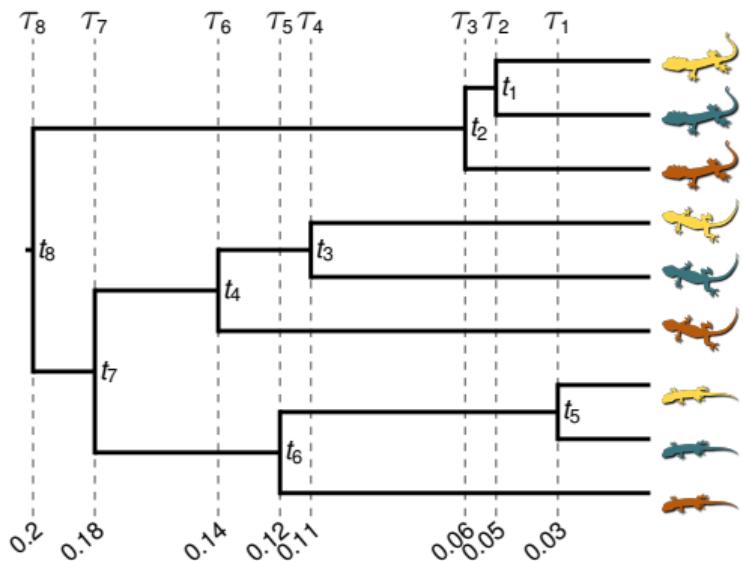
● M_G = Generalized model

◆ M_{IB} = Independent-bifurcating model



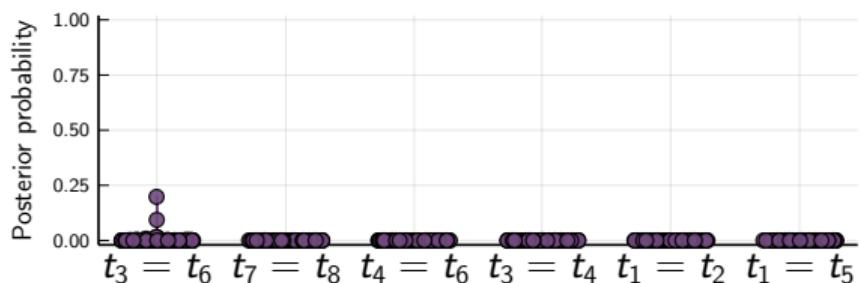
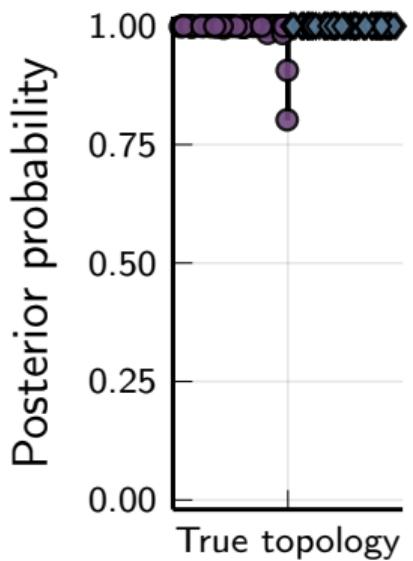
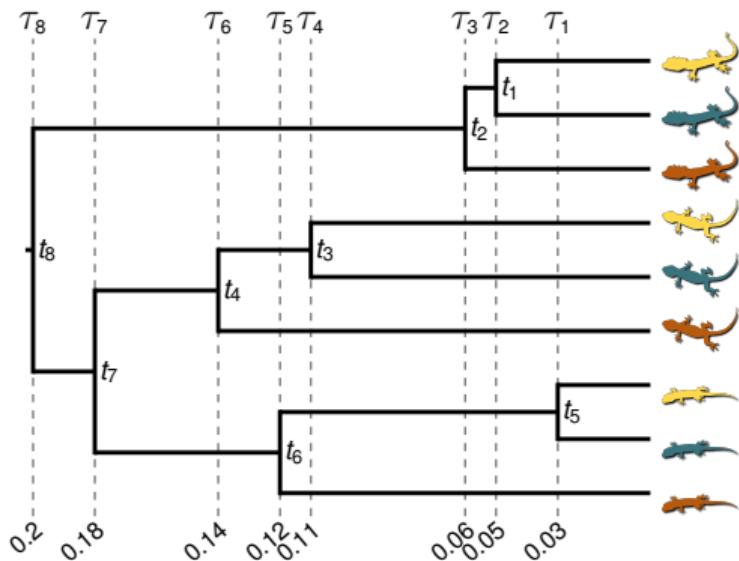
● M_G = Generalized model

◆ M_{IB} = Independent-bifurcating model



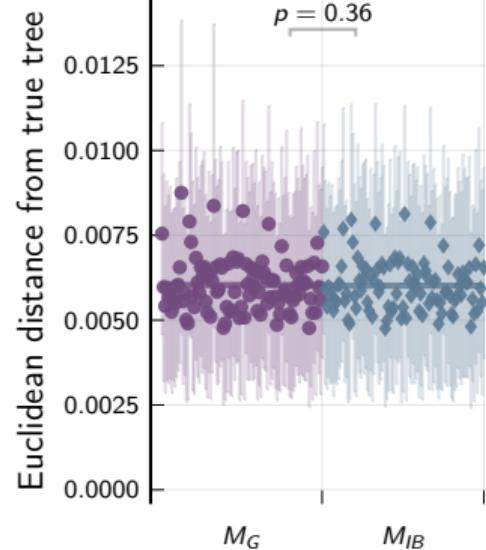
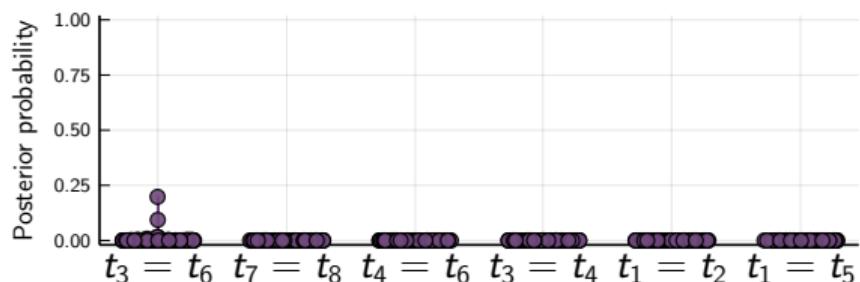
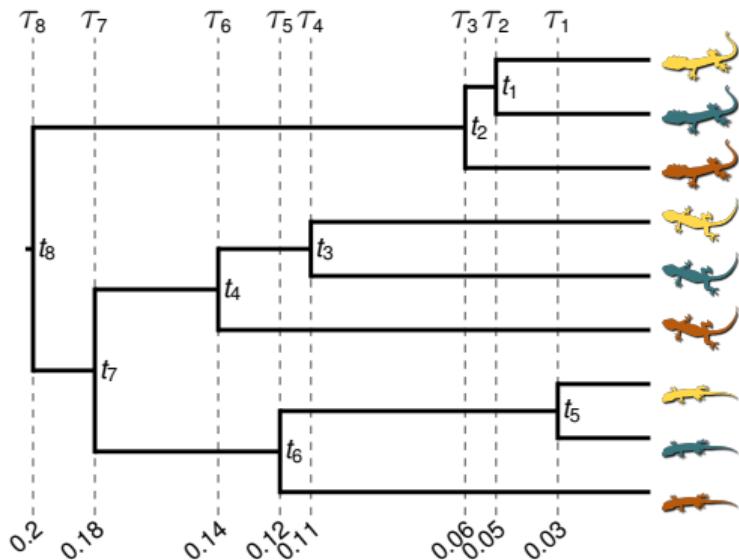
● M_G = Generalized model

◆ M_{IB} = Independent-bifurcating model



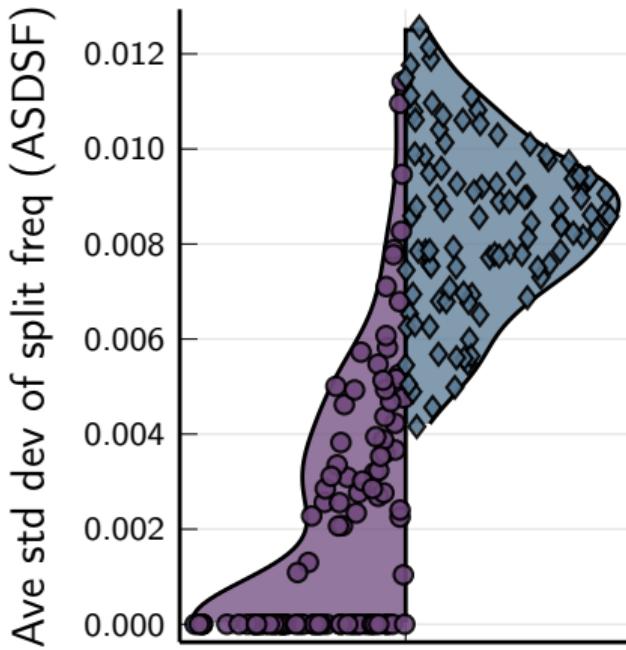
● M_G = Generalized model

◆ M_{IB} = Independent-bifurcating model



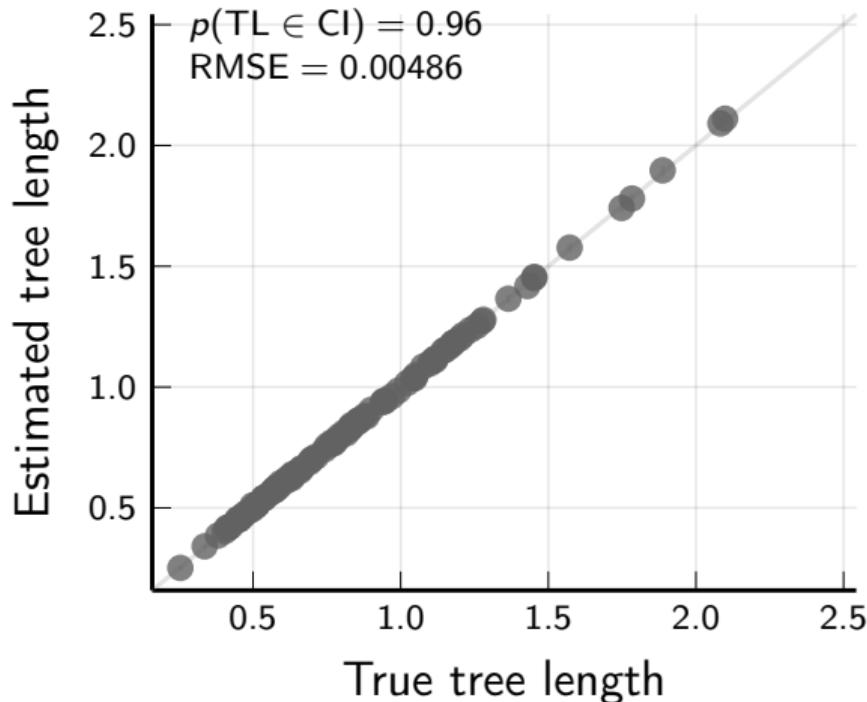
● M_G = Generalized model

◆ M_{IB} = Independent-bifurcating model

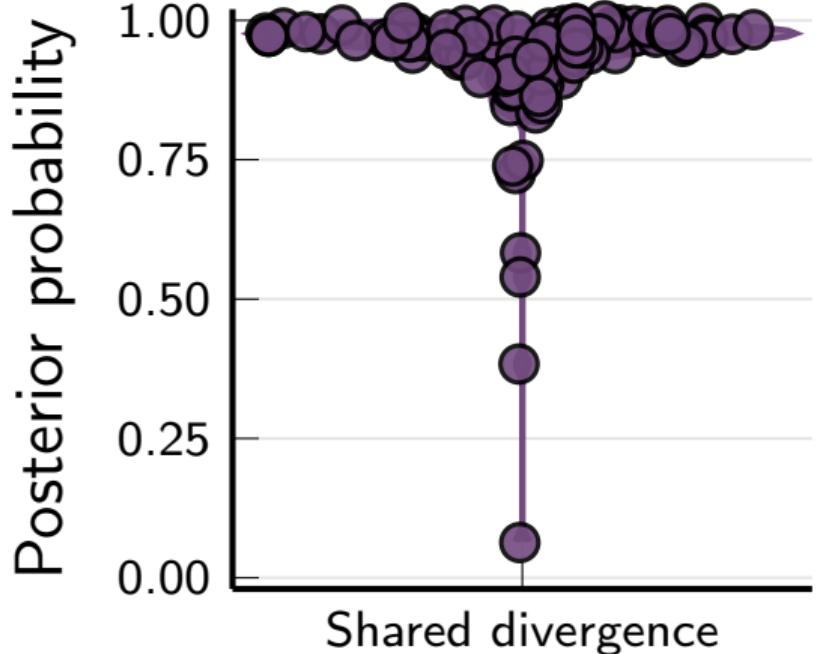
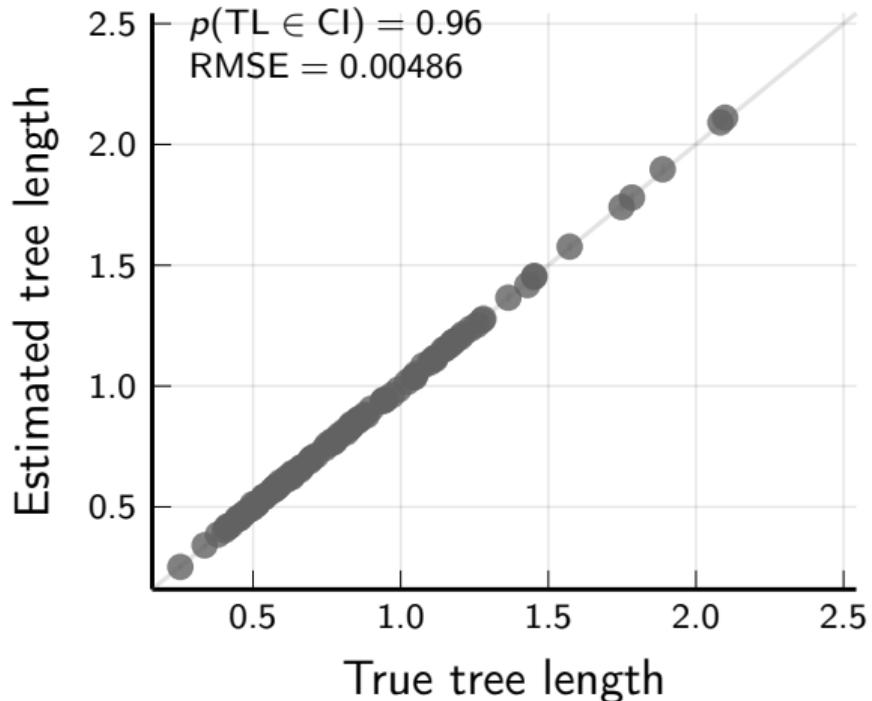


New method improves MCMC convergence and mixing

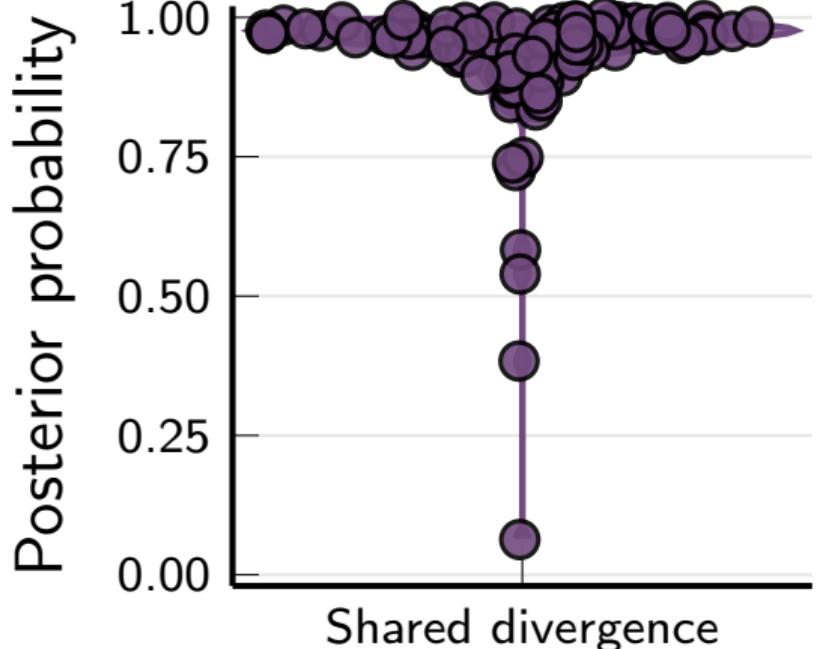
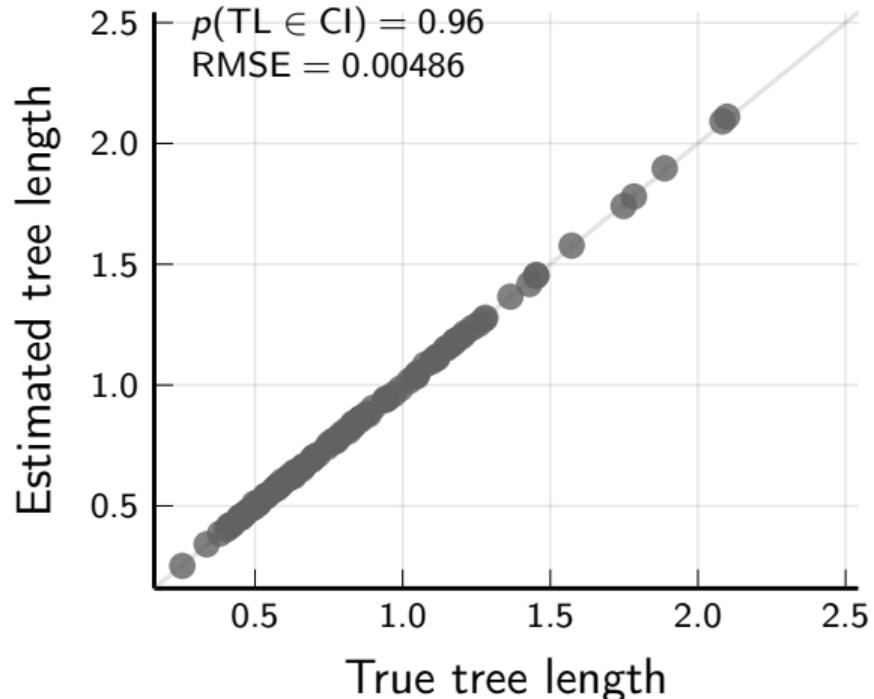
Tree-varying simulation results



Tree-varying simulation results



Tree-varying simulation results



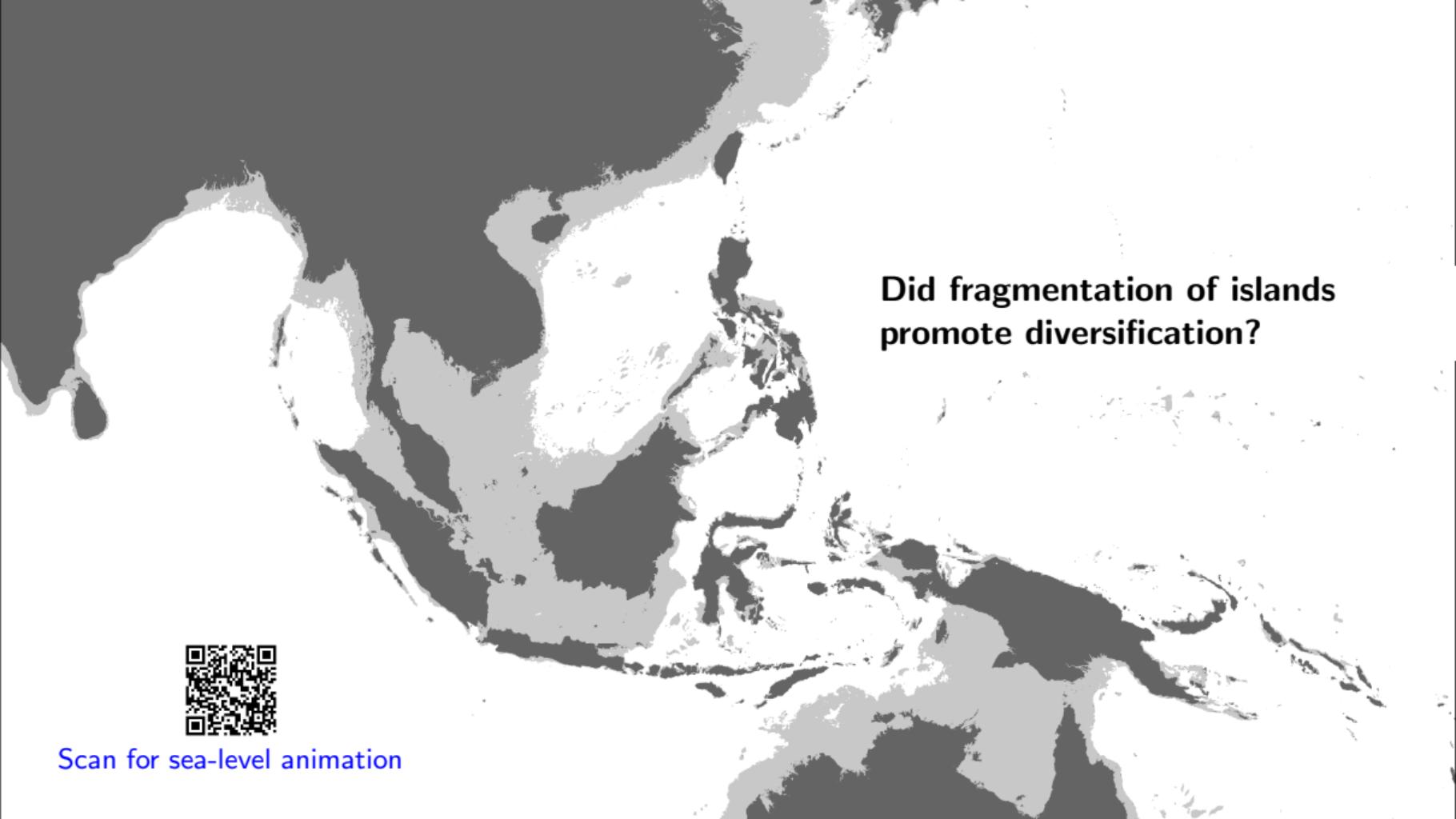
New method performs well with data simulated on random trees



Scan for sea-level animation



Scan for sea-level animation



**Did fragmentation of islands
promote diversification?**

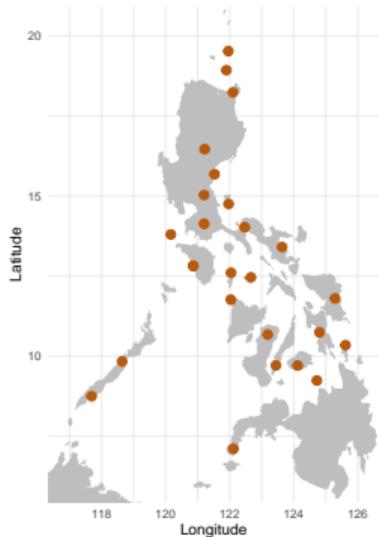


Scan for sea-level animation

Cyrtodactylus



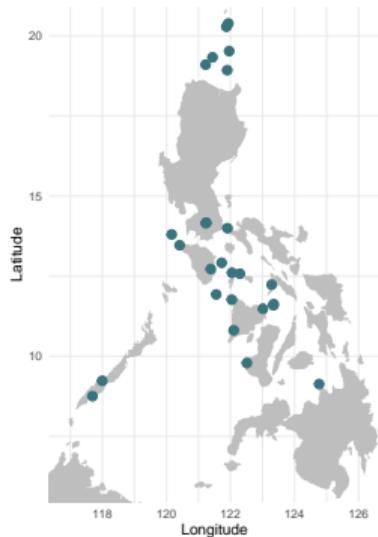
©Rafe M. Brown



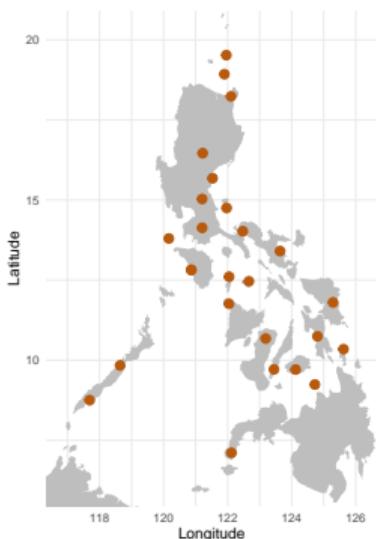
Gekko



©Rafe M. Brown



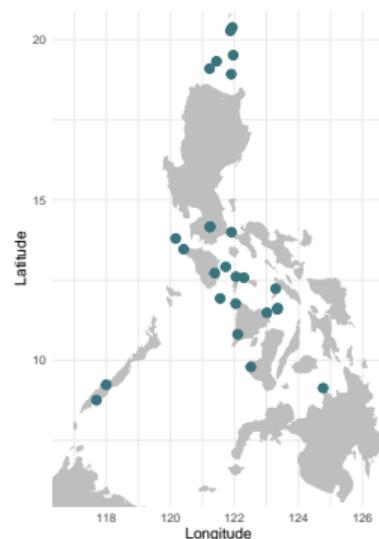
Cyrtodactylus



©Rafe M. Brown

1702 loci
155,887 sites

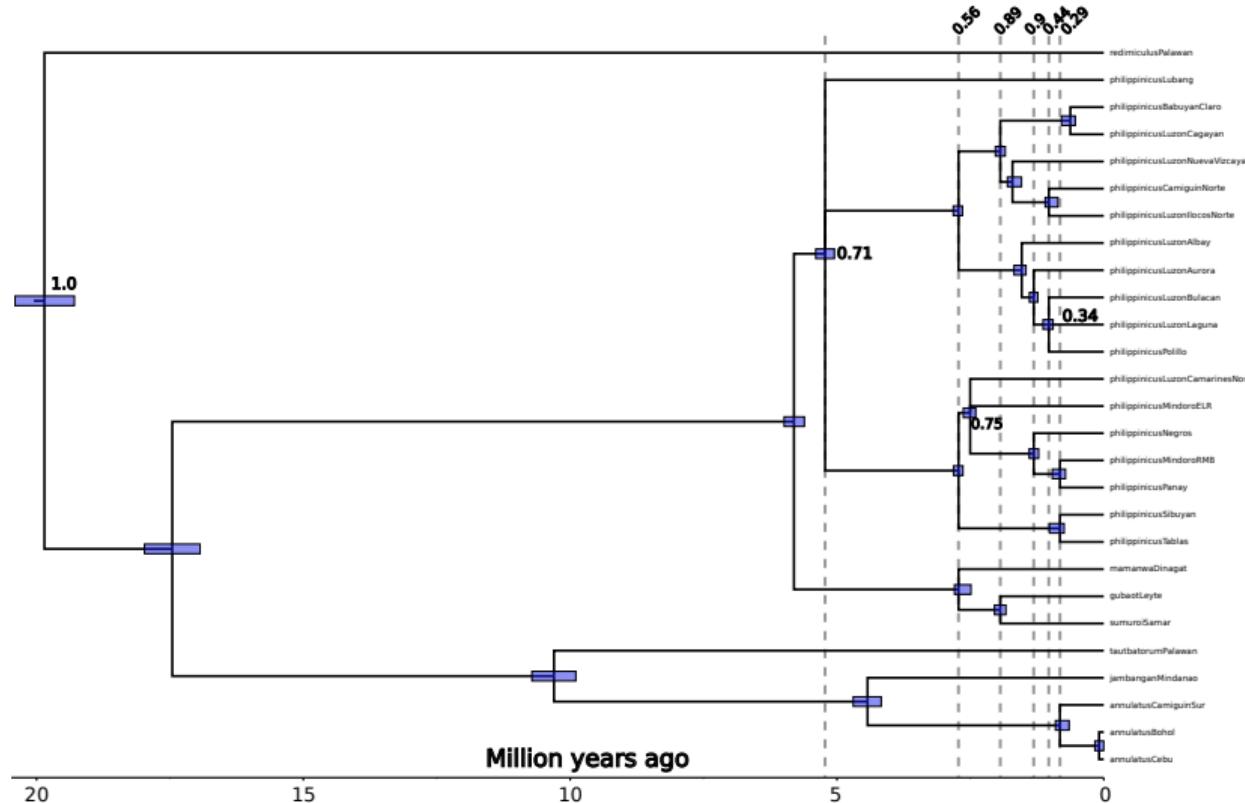
Gekko



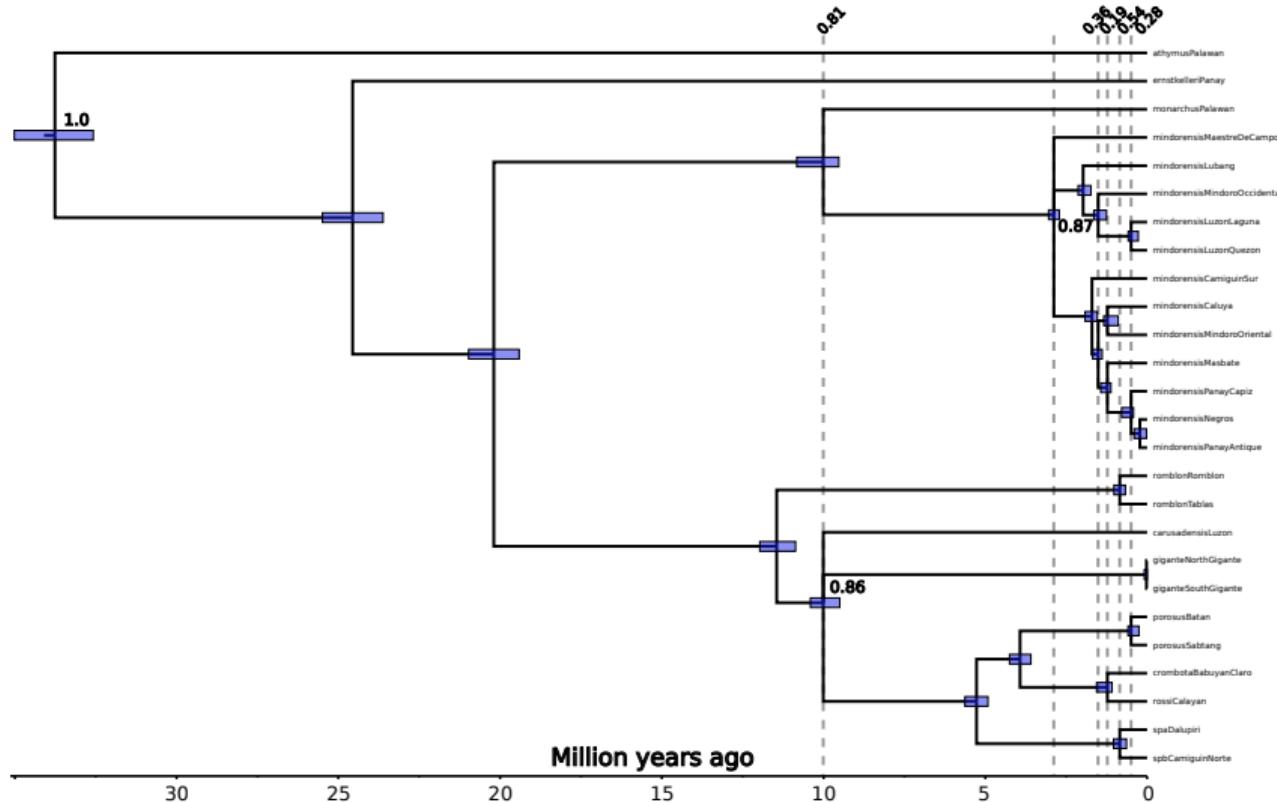
©Rafe M. Brown

1033 loci
94,813 sites

Cyrtodactylus



Gekko



Take-home points

- ▶ We can accurately infer phylogenies with shared divergences with moderately sized data sets

Take-home points

- ▶ We can accurately infer phylogenies with shared divergences with moderately sized data sets
- ▶ Among Philippine gekkonids, we found support for shared divergences predicted by sea-level changes

So much to do...

- ▶ Theory on generalized tree space
- ▶ Better algorithms to explore this new tree space
- ▶ Couple with other data models, relaxed-clock models, biogeographic models, etc.
- ▶ Develop process-based priors

Everything is available...

Software:

- ▶ Phycoeval: <https://github.com/phyletica/ecoevolity>
(release coming soon)

Open-Science Notebooks:

- ▶ Phycoeval simulations: <https://github.com/phyletica/phycoeval-experiments>
- ▶ Gecko RADseq: <https://github.com/phyletica/gekgo>

Preprint coming soon...

Acknowledgments

- ▶ Phyletica Lab (the Phyleticians)
- ▶ Mark Holder
- ▶ Rafe Brown
- ▶ Cam Siler

Computation:

- ▶ Alabama Supercomputer Authority
- ▶ Auburn University Hopper Cluster

Funding:



DEB 1656004

Photo credits:

- ▶ Rafe Brown
- ▶ [PhyloPic](#)

Thanks to the organizers of Evolution 2021!

Questions?

joaks@auburn.edu

phyletica.org



© 2007 Boris Kulikov boris-kulikov.blogspot.com