Modeling Stratigraphic Ranges on Phylogenetic Trees

Tracy Heath

Ecology, Evolution, & Organismal Biology
Iowa State University

@trayc7
phyloworks.org

GSA 2019
Phoenix, AZ USA
September 22, 2019

Land Acknowledgement

We come together on *the unceded land of the O'odham, Yavapai, Akimel O'odham, and Hohokam Peoples*. I recognize that we are able to gather here because of a painful history of genocide and forced removal from this territory. This acknowledgement is to show respect for the many Indigenous Peoples still living on this land and to honor their ancestors.

Collaborators and Coauthors

Rachel Warnock

Walker Pett

Tanja Stadler (ETH Zürich)

Sasha Gavryushkina

Dan Ksepka

Paleontology & Neontology

"Except during the interlude of the [Modern] Synthesis, there has been limited communication historically among the disciplines of evolutionary biology, particularly between students of evolutionary history (paleontologists and systematists) and those of molecular, population, and organismal biology. There has been increasing realization that barriers between these subfields must be overcome if a complete theory of evolution and systematics is to be forged."

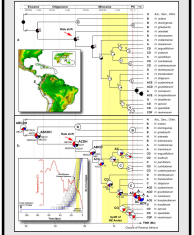
Reaka-Kudla, M.L. & Colwell, R.: in E.C. Dudley (ed.), *The Unity of Evolutionary Biology: Proceedings of the Fourth International Congress of Systematic & Evolutionary Biology*, Discorides Press, Portland, OR, p. 16.

A Time-Scale for Macroevolution

Phylogenetic divergence-time estimation

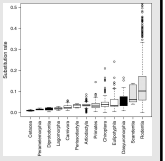
- What was the spacial and climatic environment of ancient angiosperms?
- How has mammalian body-size changed over time?
- Is diversification in Caribbean anoles correlated with ecological opportunity?
- How has the rate of molecular evolution changed across the Tree of Life?

Historical biogeography



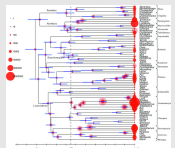
(Antonelli & Sanmartin. Syst. Biol. 2011)

Molecular evolution



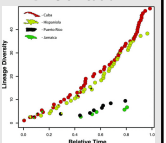
(Nabholz, Glemin, Galtier. MBE 2008)

Trait evolution



(Lartillot & Delsuc. Evolution 2012)

Diversification

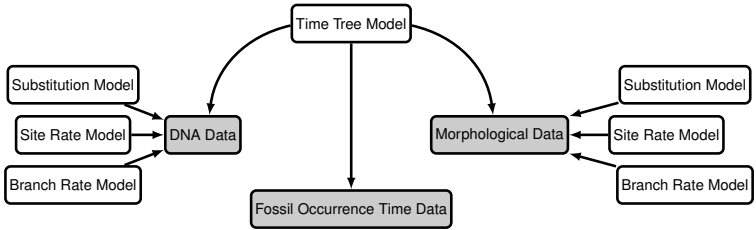


Anolis cristatus (image by L. Mahler)

(Mahler, Revell, Glor, & Losos. Evolution 2010)

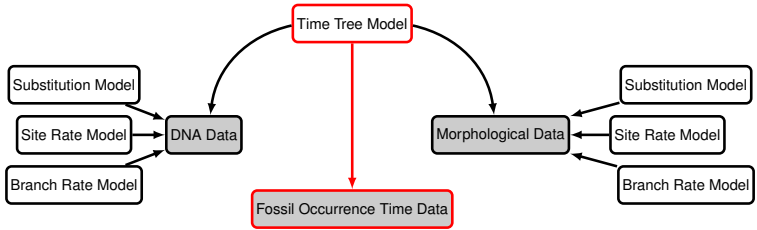
Integrative Analysis of Fossil & Extant Data

Combine models for sequence evolution, morphological change, & fossil recovery to jointly estimate the tree topology, divergence times, & lineage diversification rates



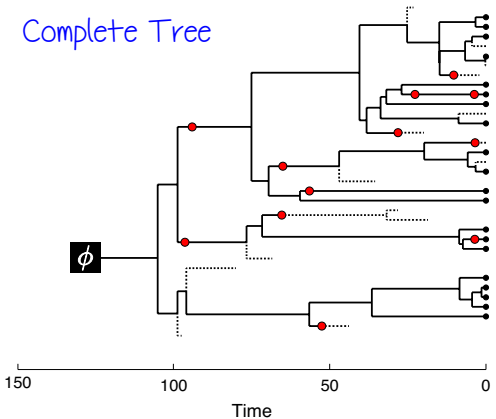
Integrative Analysis of Fossil & Extant Data

Today I will focus on work we have done to develop and extend diversification models that describe the generation of phylogenetic trees and fossil occurrent times



Modeling the Tree & Fossil Occurrence Times

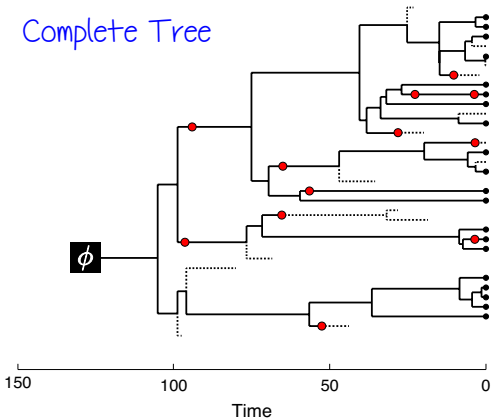
Stadler (2010) introduced a generating model for a serially sampled time tree — this is the *fossilized birth-death process*.

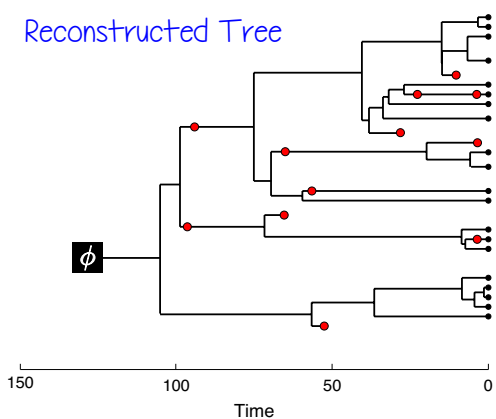


(Stadler. *Journal of Theoretical Biology* 2010)

Modeling the Tree & Fossil Occurrence Times

Stadler (2010) introduced a generating model for a serially sampled time tree — this is the *fossilized birth-death process*.





(Stadler. *Journal of Theoretical Biology* 2010)

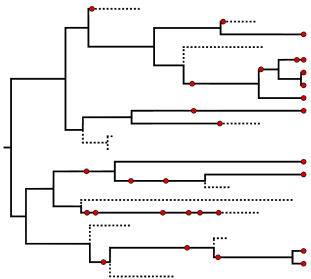
The Fossilized Birth-Death Process (FBD)

Improving statistical inference of absolute node ages

Eliminates the need to specify arbitrary calibration densities

Better capture our statistical uncertainty in species divergence dates

All reliable fossils associated with a clade are used



The fossilized birth–death process for coherent calibration of divergence-time estimates

Tracy A. Heath[a,b], John P. Huelsenbeck[a,c], and Tanja Stadler[d,e,1]

(Heath, Huelsenbeck, Stadler. *PNAS* 2014)

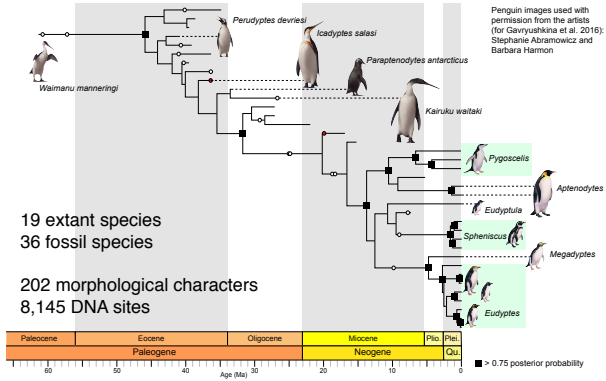
The Fossilized Birth-Death Process (FBD)

Joint estimation of the tree topology, divergence times, and fossil species ages using discrete morphological data and fossil age ranges

Bayesian Total-Evidence Dating Reveals the Recent Crown Radiation of Penguins

Alexandra Gavryushkina ✉, Tracy A. Heath, Daniel T. Ksepka, Tanja Stadler, David Welch, Alexei J. Drummond Author Notes

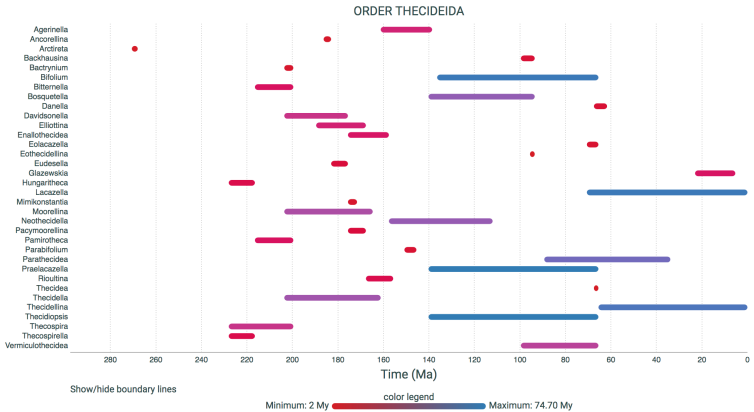
Systematic Biology, Volume 66, Issue 1, January 2017, Pages 57–73,
https://doi.org/10.1093/sysbio/syw060



(Gavryushkina et al. *Syst. Biol.* 2017)

Stratigraphic Range Data

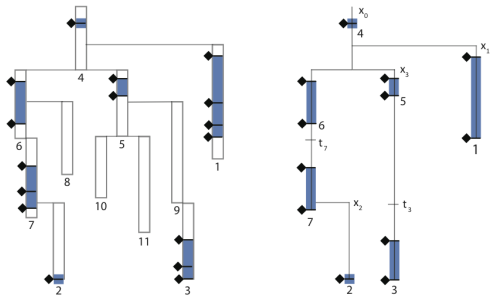
Practical approaches to analyzing the fossil record need to accommodate fossil information represented as lineage first and last appearances over time.



(figure from `https://gerardofurtado.com/sr/sr.html`, using data from Sepkoski 2002)

The Fossilized Birth-Death Range Process

Models to account for the structure of the fossil record and the nature of paleontological data



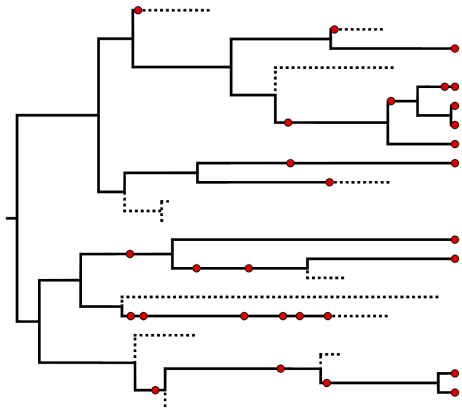
The fossilized birth-death model for the analysis of stratigraphic range data under different speciation modes

Tanja Stadler[a,b,*], Alexandra Gavryushkina[a,b], Rachel C.M. Warnock[a,b], Alexei J. Drummond[c], Tracy A. Heath[d]

(Stadler et al., *J. Ther. Biol.* 2018)

Lineage Diversification & Fossil Preservation

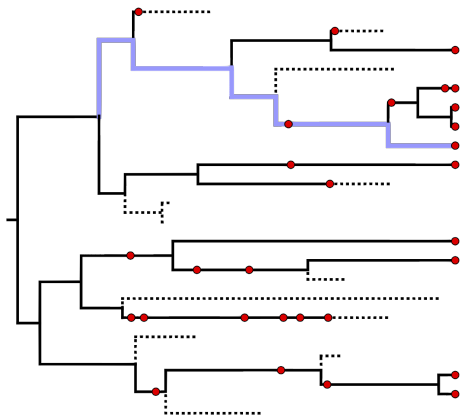
Under the FBD model, lineages speciate, go extinct, and are preserved, but much of the process may unobservable.



(Stadler et al., *J. Ther. Biol.* 2018)

Lineage Diversification & Fossil Preservation

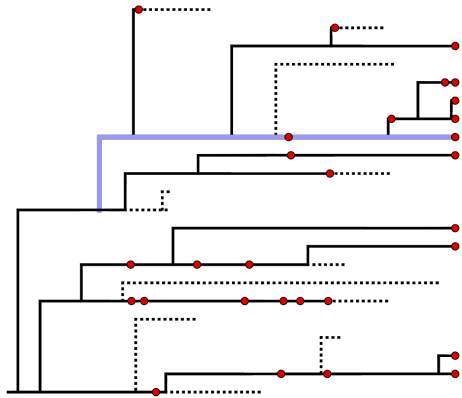
Often fossils are represented as lineages with multiple observations over time.



(Stadler et al., *J. Ther. Biol.* 2018)

Lineage Diversification & Fossil Preservation

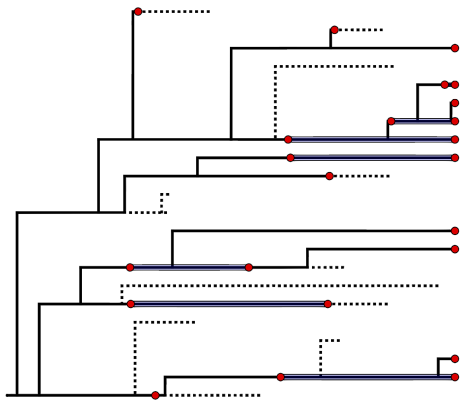
We can consider a process of budding speciation and fossils lying along a path are considered a lineage if they share morphological characters.



(Stadler et al., *J. Ther. Biol.* 2018)

Lineage Diversification & Fossil Preservation

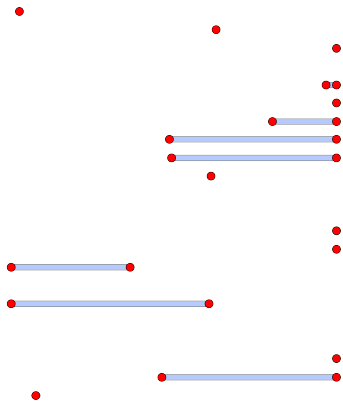
This gives us the sampled tree that unites these lineages as stratigraphic ranges with singletons and extant taxa.



(Stadler et al., *J. Ther. Biol.* 2018)

Lineage Diversification & Fossil Preservation

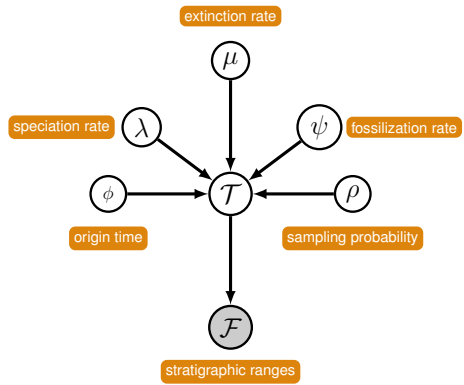
Before inferring the underlying phylogeny, our data represent stratigraphic ranges.



(Stadler et al., *J. Ther. Biol.* 2018)

Lineage Diversification & Fossil Preservation

For the FBD family of models the free parameters are the speciation rate, extinction rate, fossil recovery rate, origin time of the process, and probability of sampling a taxon in the present.

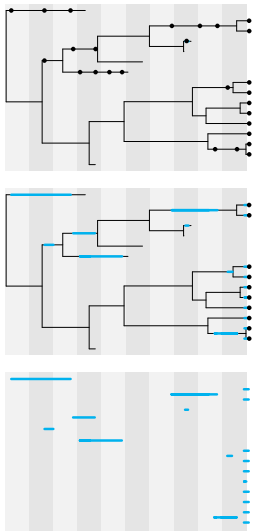


(Stadler et al., *J. Ther. Biol.* 2018)

The FBD Family of Models

FBD models can accommodate different kinds of paleontological data

- specimen-level sampling
- when the fossil age data are coded for first and last occurrences (stratigraphic ranges)
- when only stratigraphic range data are available



(Stadler *J. Ther. Biol.* 2010; Heath et al. *PNAS* 2014; Stadler et al., *J. Ther. Biol.* 2018)

Inference under FBD Models

The various FBD models have been implemented in Bayesian inference software for phylogenetics

RevBayes

https://revbayes.com

BEAST 2

http://www.beast2.org

Allows us to average over uncertainty in parameters using Markov chain Monte Carlo (MCMC)

(RevBayes: Höhna et al. *Syst Biol* 2016)

(BEAST 2.5: Bouckaert et al. *PLoS Comp Bio* 2019)

Simulation under FBD Models

FossilSim

https://github.com/fossilsim/fossilsim

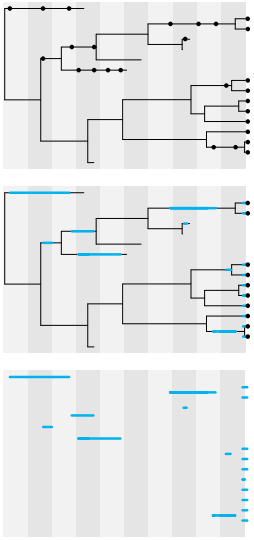
Methods in Ecology and Evolution | British Ecological Society

APPLICATION | Open Access

FossilSim: An R package for simulating fossil occurrence data under mechanistic models of preservation and recovery

Joëlle Barido-Sottani, Walker Pett, Joseph E. O'Reilly, Rachel C. M. Warnock

First published: 04 March 2019 | https://doi.org/10.1111/2041-210X.13170



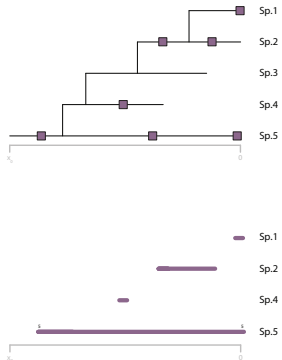
FBD for Stratigraphic Range Data

Testing the accuracy of parameter estimation under the FBD range process

We simulated stratigraphic range data using FossilSim and allowed for incomplete sampling.

Simulated data were analyzed under the FBD model to examine the accuracy and precision of parameter estimates: speciation (λ), extinction (μ), and the fossil sampling proportion.

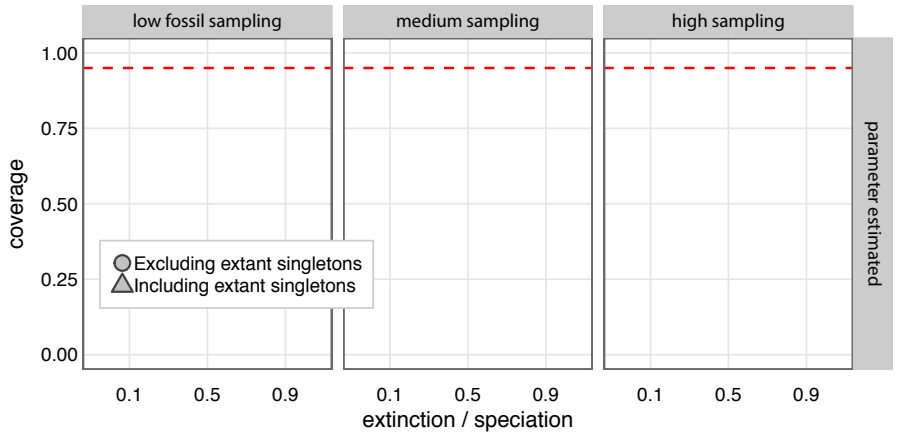
The FBD model allows us to use singleton taxa, including extant singletons



(Warnock, Heath, Stadler, *in review*)

Estimation of Speciation, Extinction, & Sampling

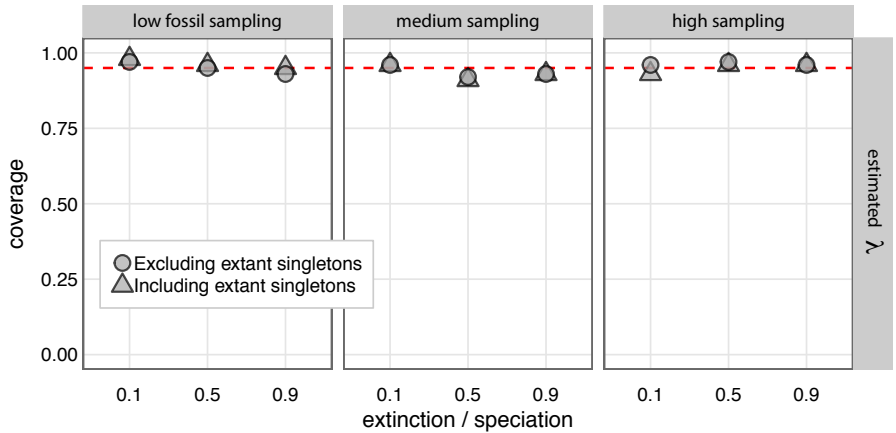
Simulations varied in the level of fossils sampled and the rate of turnover (extinction / speciation)



(Warnock, Heath, Stadler, *in review*)

Estimation of Speciation, Extinction, & Sampling

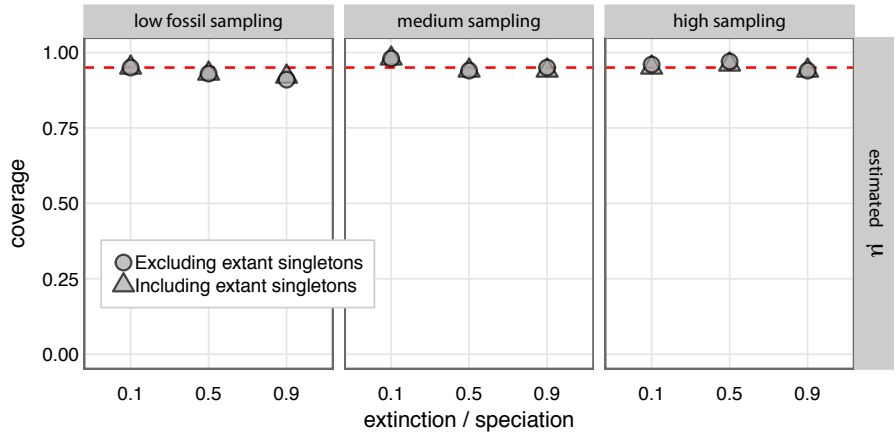
With only fossil occurrence times, the FBD model can accurately estimate the **speciation rate**



(Warnock, Heath, Stadler, *in review*)

Estimation of Speciation, Extinction, & Sampling

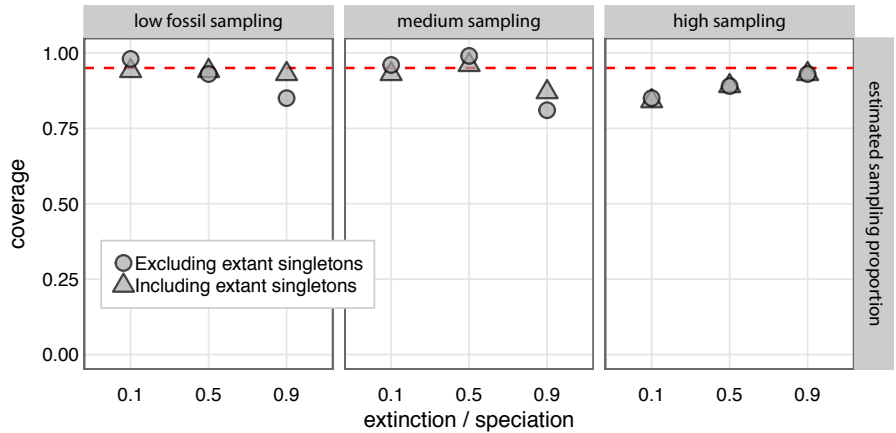
With only fossil occurrence times, the FBD model can accurately estimate the **extinction rate**



(Warnock, Heath, Stadler, *in review*)

Estimation of Speciation, Extinction, & Sampling

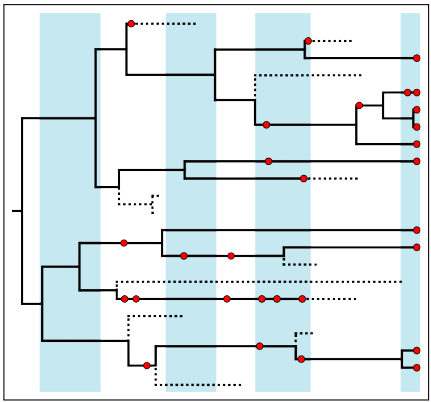
With only fossil occurrence times, the FBD model can accurately estimate the **fossil sampling proportion**

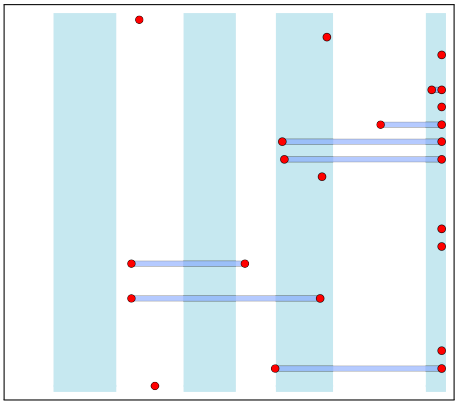


(Warnock, Heath, Stadler, *in review*)

Extending the FBD Process

The FBD skyline models allow for the FBD rates to change over time





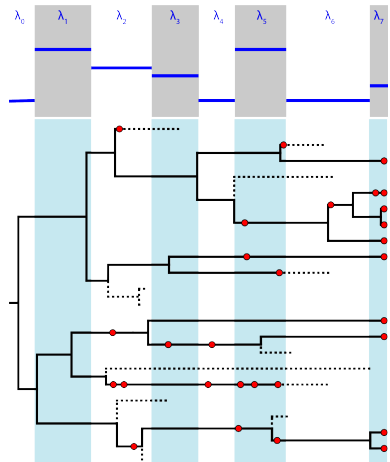
(Stadler et al. *PNAS* 2013; Pett, Warnock, et al. *in prep*; Warnock, Pett, et al. *in prep*;)

FBD Skyline Model

The rates of speciation, extinction, and/or fossil sampling can vary over the timeline of the FBD

The timeline can be divided into known intervals (e.g., geological epochs) or arbitrary intervals (i.e., 6 of equal length)

Additional processes can be used to model the change over time (e.g., geometric Brownian motion)

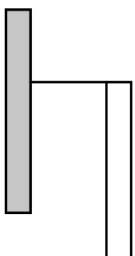


(Stadler et al. *PNAS* 2013; Pett, Warnock, et al. *in prep*; Warnock, Pett, et al. *in prep*;)

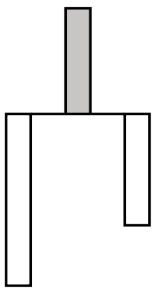
FBD with Morphospeciation

When coestimating the phylogeny using discrete morphological data and stratigraphic ranges, we have to account for different modes of speciation

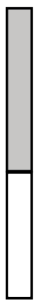
asymmetric speciation



symmetric speciation



anagenetic speciation



(Stadler et al. *PNAS* 2013; Pett, Warnock, et al. *in prep*; Warnock, Pett, et al. *in prep*;)

FBD with Morphospeciation

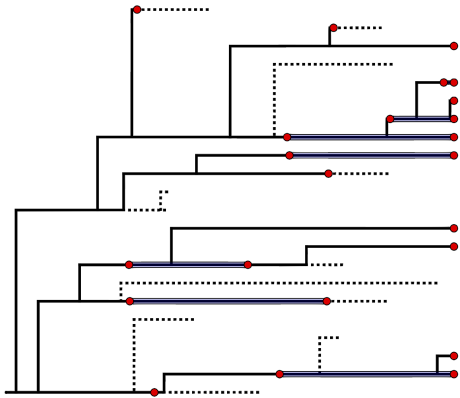
Under the FBD range model, we account for cladogenesis and anagenesis and compute the probability of a set of stratigraphic ranges united by a phylogeny

λ = cladogenetic speciation rate

λ_a = anagenetic speciation rate

μ = extinction rate

ψ = fossil recovery rate

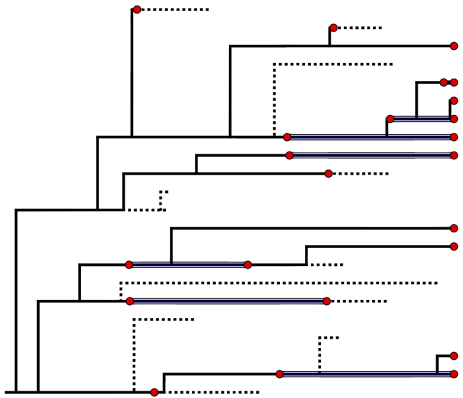


(Stadler et al. *JTB* 2018; Pett, Holder, Stadler, Heath *in prep*;)

FBD with Morphospeciation

If combining the morphospeciation with discrete character data, the commonly used model—Mk—is not appropriate.

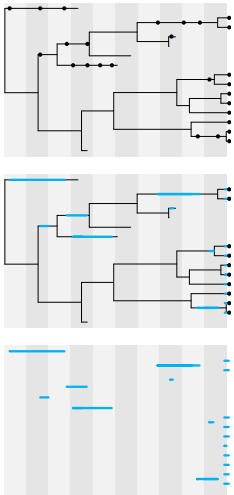
We are currently testing the properties of a new model that couples discrete morphological change with the FBD range model.



(Stadler et al. *JTB* 2018; Pett, Holder, Stadler, Heath *in prep*;)

The Synthesis of Paleontology & Neontology

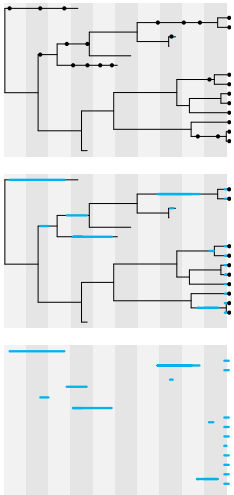
Our fields are becoming more integrated because of new data, larger datasets, and especially new statistical methods.



The Synthesis of Paleontology & Neontology

Our fields are becoming more integrated because of new data, larger datasets, and especially new statistical methods.

But these new methods are still very far from capturing the structure of the fossil record and dealing with the challenges of integrating paleontological data…

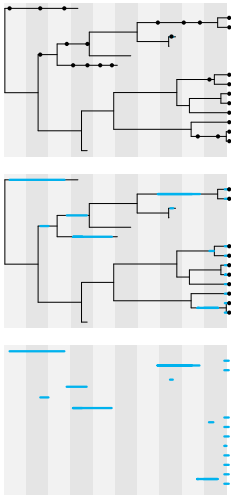


The Synthesis of Paleontology & Neontology

Our fields are becoming more integrated because of new data, larger datasets, and especially new statistical methods.

But these new methods are still very far from capturing the structure of the fossil record and dealing with the challenges of integrating paleontological data…

…so there's a lot more work ahead.



Acknowledgements

Thanks to the symposium organizers & the Paleontological Society!

Co-authors on FBD work: Walker Pett, Rachel Warnock, Alexandra Gavryushkina, Daniel Ksepka, Joshua Justison, David Welch, John Huelsenbeck, Alexei Drummond, Mark Holder, & Tanja Stadler

Thanks to other collaborators:

- Daniel Thomas
- Rob Meredith
- Joëlle Barido-Sottani
- Janet Buckner
- Dagmara Żyła
- April Wright
- Michael Landis
- Sebastian Höhna
- Kevin Quinteros
- Wade Dismukes
- Jordan Satler
- Lorenzo Chavarria

Paleontological Society

NSF

Funding from NSF Grants: DEB-1556853, DEB-1556701, DBI-1759909