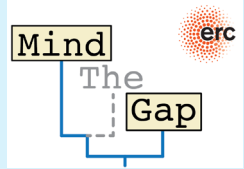


# Phylogenetic Inference from an Incomplete Fossil Record

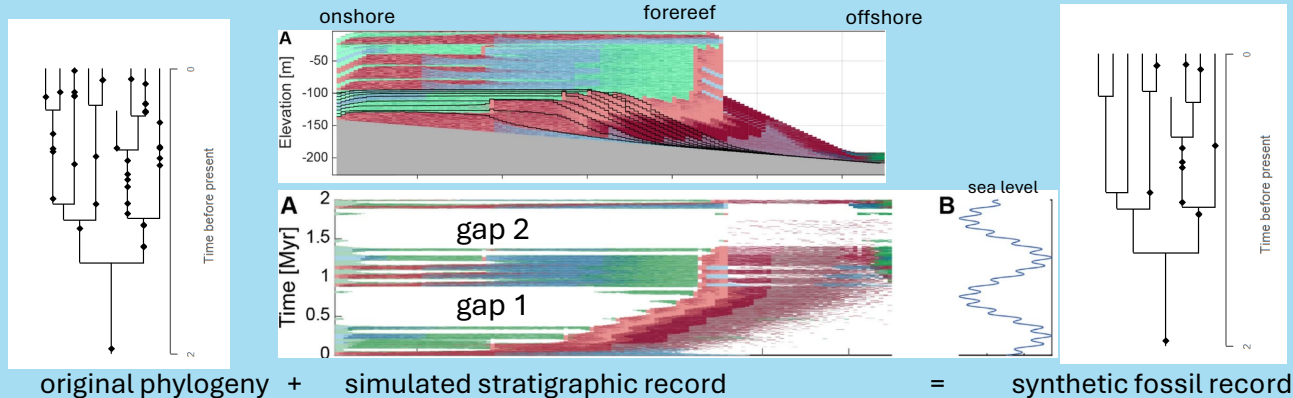


Niklas Hohmann<sup>1\*</sup>, Rachel Warnock<sup>2</sup>, Emilia Jarochowska<sup>1</sup>  
<sup>1</sup>Utrecht University, The Netherlands <sup>2</sup>FAU Erlangen-Nürnberg, Germany \*email: N.H.Hohmann@uu.nl



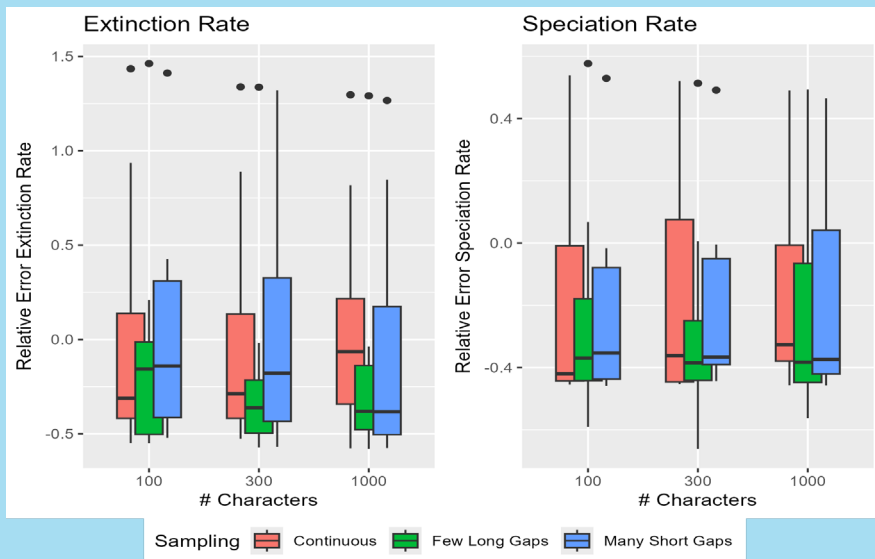
## Motivation and Workflow

Fossil data is crucial for the inference of time trees, however, the fossil record is incomplete. Here, we explore the impact of stratigraphy on phylogenetic inference using novel simulations



We generate a synthetic fossil record using the R packages *admttools* [1] and *StratPal* [2], combining (1) phylogenetic trees, molecular data & fossils from *FossilSim* [3] and *TreeSim* [4] with (2) stratigraphic architectures from *CarboCAT* [5,6]. We use *RevBayes* to estimate trees in a total evidence inference under the Fossilized Birth-Death (FBD) model.

## Q1: Does the structure of the fossil record bias phylogenetic inference?



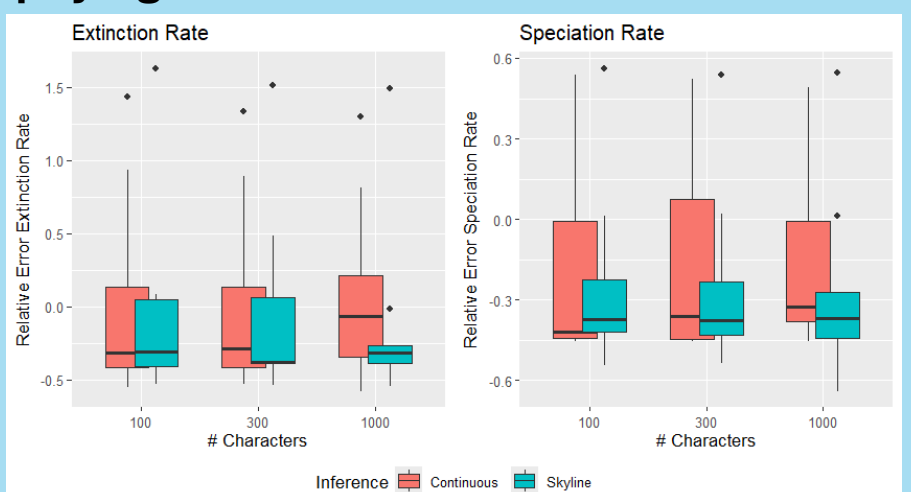
**No**, after accounting for number of fossil samples we find no systematic difference across different sampling scenarios.

Sampling scenarios are from simulated carbonate platforms from [6] scenarios A & B, with gap sized being determined by the frequency of sea level fluctuations.

## Open questions & further work

- Why are all estimates biased?
- Run more replicates & determine coverage
- Change stratigraphic model (e.g. to *CarboKitten.jl* [7])

## Q2: Does geological information improve phylogenetic inference?



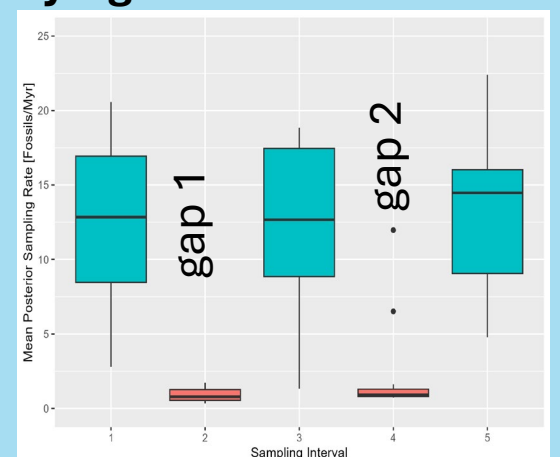
**Yes**, relative error is reduced when providing information on gaps to the FBD inference.

In the skyline inference, strong priors on low sampling rate during gaps were used. Continuous inference assumes no gaps and uses uninformative priors on sampling.

## Q3: Can geological information be recovered from phylogenetic inference?

**Yes**, we reliably recover gaps in the fossil record from the FBD skyline model

Sampling intervals 2 & 4 are gaps in the fossil record due to drops in sea level in scenario A from [6].



**References:** [1] Hohmann (2025): R package “admttools” [2] Hohmann & Jarochowska (2025): “StratPal: An R package for creating stratigraphic paleobiology modelling pipelines”; *Methods in Ecology and Evolution*. [3] Barido-Sottani et al. (2019): “FossilSim: An R package for simulating fossil occurrence data under mechanistic models of preservation and recovery”; *Methods in Ecology and Evolution* [4] Stadler (2019): R package “TreeSim”. [5] Burgess (2013): “CarboCAT: A cellular automata model of heterogeneous carbonate strata”; *Computers & Geosciences*. [6] Hohmann et al. (2024): “Identification of the mode of evolution in incomplete carbonate successions”; *BMC Ecology and Evolution*. [7] Hidding et al. (2025): “CarboKitten.jl”.