

## mcbette: model comparison using babette

Richèl J.C. Bilderbeek<sup>1</sup>

<sup>1</sup> Theoretical & Evolutionary Community Ecology, TRES, GELIFES, University of Groningen

DOI: [10.21105/joss.02762](https://doi.org/10.21105/joss.02762)

### Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Editor: [Kristen Thyng](#) ↗

### Reviewers:

- [@kthyng](#)

Submitted: 09 October 2020

Published: 26 October 2020

### License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).

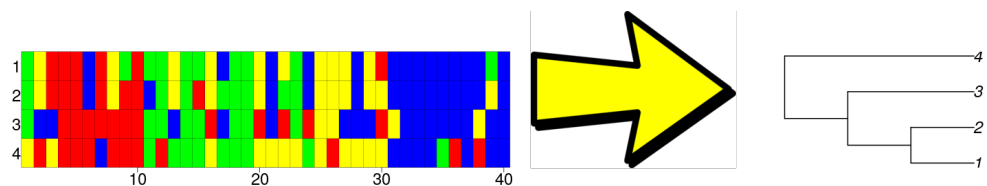


The mcbette logo.

## Abstract

One can generate a phylogeny from a DNA alignment and a model of evolution. Selecting an evolutionary model is non-trivial, as there are many. mcbette is an R package that determines the model that has most evidence for having generated the alignment, from a set of models. In this way, the model that is ‘simple enough, but not simpler’ can be used to generate a phylogeny.

## Statement of need



Constructing a species phylogeny (at the right) from a DNA alignment (at the left) using an evolutionary model (the arrow). mcbette allows for selecting an evolutionary model from a set of models.

mcbette is an R package to do model comparison between a set of evolutionary models on a DNA alignment, which allows to select that model that is closest to the process consistent with the DNA alignment and species tree.

Unlike other methods, mcbette can both be installed and run from an R script, allowing one to run many analyses using different models, examine the results directly from R and integrate mcbette into an existing R pipeline.

## Getting started

`mcbette` is aimed at being used by anyone interested in phylogenetics and assumes some basic knowledge about the field. The BEAST book (Drummond & Bouckaert, 2015) serves as an excellent starting point about the field of phylogenetics, where the `mcbette` README and vignette show a simplified worked-out example. The evolutionary models are those allowed by the `babette` R package (Bilderbeek & Etienne, 2018), which consist of (among others) a site model, clock model and tree model (see ‘Supported models’ below for an overview). `babette` is an R package to work with the phylogenetic tool BEAST2 (Bouckaert et al., 2019). Additionally, `mcbette` uses the novel ‘NS’ ‘BEAST2’ package (Russel, Brewer, Klaere, & Bouckaert, 2019) to do the actual model comparison.

To see a demo of `mcbette`, see the vignette:

```
vignette(topic = "demo", package = "mcbette")
```

## Quirks

`mcbette` has two quirks. First, `mcbette` only works under Linux and Mac, because BEAST2 packages only work under Linux and Mac (that is, without using a GUI). Second, `mcbette` uses the `rJava` package, because BEAST2 is written in Java. Getting `rJava` properly installed is the hardest part to get `mcbette` working.

## Supported models

At the time of writing, these are the BEAST2 models that `babette` supports:

- 1 site model: gamma site model
- 4 nucleotide substitution models: JC (after Jukes and Cantor), HKY (after Hasegawa, Kishino and Yano), TN (after Tamura and Nei), generalized time-reversible model
- 2 clock models: strict, relaxed log-normal
- 5 tree models: birth-death, coalescent Bayesian skyline, coalescent constant-population, coalescent exponential-population, Yule

To see these:

```
vignette(topic = "inference_models", package = "beautier")
```

## References

- Bilderbeek, R. J., & Etienne, R. S. (2018). `babette`: BEAUti 2, BEAST2 and Tracer for R. *Methods in Ecology and Evolution*, 9(9), 2034–2040. doi:[10.1111/2041-210X.13032](https://doi.org/10.1111/2041-210X.13032)
- Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., et al. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS computational biology*, 15(4), e1006650. doi:[10.1371/journal.pcbi.1006650](https://doi.org/10.1371/journal.pcbi.1006650)
- Drummond, A. J., & Bouckaert, R. R. (2015). *Bayesian evolutionary analysis with BEAST*. Cambridge University Press.

Russel, P. M., Brewer, B. J., Klaere, S., & Bouckaert, R. R. (2019). Model selection and parameter inference in phylogenetics using nested sampling. *Systematic biology*, 68(2), 219–233. doi:[10.1093/sysbio/syy050](https://doi.org/10.1093/sysbio/syy050)