

# Priors & Troubleshooting

Veronika Bošková

& Joëlle Barido-Sottani & Denise Kühnert & Jūlija Pečerska  
& David Rasmussen & Chi Zhang

February 07, 2017



Priors and starting values

Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

Word of caution

References

- ▶ Distribution of a parameter before the data is collected and analysed
- ▶ as opposed to POSTERIOR distribution which combines the information from the prior and the data

Priors and starting values

Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

Word of caution

References

- ▶ Should not be and is not universal for all the analyses you will ever do in your research
- ▶ Should incorporate all the knowledge about the parameter/underlying process we have before looking at the data
  - ▶ use results of previous independent experiments
  - ▶ use other independent evidence
  - ▶ use model comparison to choose the model best fitting the data
- ▶ Should not be too restrictive if prior knowledge/assumptions are weak
  - ▶ One can use diffuse priors
- ▶ Prior distribution does not have to, and is not expected to, be exactly the same as the posterior
- ▶ May not be adjusted after the run, to give higher and higher posterior support

## Priors and starting values

### Priors

#### Prior distribution

#### Tree prior

#### Substitution model prior

#### Clock prior

#### Parameter prior

#### Think twice

#### Starting values

## Troubleshooting

### Possible issues

#### Initialization failed

#### Nothing mixed

#### One parameter did not mix

## Word of caution

## References

# What is a prior in Bayesian phylogenetics?

- Using Bayes theorem, we can decompose the posterior:

$$P(\text{genetic sequences} | \text{genealogy}, \text{demographic model}, \text{substitution model}, \text{molecular clock model}) = P(\text{genetic sequences} | \text{genealogy}, \text{demographic model}, \text{substitution model}, \text{molecular clock model}) \underbrace{P(\text{genealogy} | \text{demographic model})}_{\text{Tree prior}} \underbrace{P(\text{demographic model})}_{\text{Prior information}} P(\text{substitution model}) P(\text{molecular clock model})$$

genetic sequences      genealogy      demographic model      substitution model      molecular clock model

Figure adapted from [du Plessis and Stadler, 2015]

## Priors and starting values

### Priors

#### Prior distribution

#### Tree prior

#### Substitution model prior

#### Clock prior

#### Parameter prior

#### Think twice

### Starting values

## Troubleshooting

### Possible issues

#### Initialization failed

#### Nothing mixed

#### One parameter did not mix

## Word of caution

## References

- ▶ Is a choice of
  - ▶ model
    - ▶ tree-generating models, nucleotide/AA/codon substitution models, ...
- and of
  - ▶ distribution of plausible values for a parameter of interest
    - ▶ Uniform, Normal, Beta,...

## Priors and starting values

Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

## Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

## Word of caution

## References

- ▶ Have to pick one from Coalescent or Birth-death process framework
- ▶ Have to put priors on parameters of the chosen model
  - ▶ e.g. growth-rate of the population,  $R_0$ , extinction rate, ...

Priors and starting values

Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

Word of caution

References

- ▶ The selection is big: JC69, HKY85, ..., GTR
- ▶ Have to put priors on parameters of the chosen model
- ▶ Use model which has been previously identified to be best for your type of data
  - ▶ e.g. HKY85
    - ▶ Prior for transition/transversion rate ratio ( $\kappa$ )
    - ▶ Prior for base frequencies
- ▶ To choose the best model
  - ▶ Use model comparison
  - ▶ Use rjMCMC directly in BEAST2 to sample from the posterior distribution including different substitution models. The model where rjMCMC spends the most time (samples the most from), is the best fitting model (e.g. bModelTest [Bouckaert, 2016])

Priors and starting values

Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

Word of caution

References

- ▶ Strict clock: all branches have the same clock rate
- ▶ Relaxed clock
  - ▶ Uncorrelated: branches have independent clock rate distributions
  - ▶ Correlated: child branch has clock rate distribution correlated to distribution of the parent branch

## Priors and starting values

### Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

## Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

## Word of caution

## References



- ▶ Can be fixed to a given value  
(though this is generally not recommended)
- ▶ Can have upper and lower limits
  - ▶ If we know that any infected individual recovers after 5-10 days, we can set the distribution of infectious period to be e.g. min 4 days and max 11 days
- ▶ If specified by a parametric distribution, the parameters of this distribution can also be assigned a prior (hyperprior)
- ▶ You can visualise the distribution in BEAUti

## Priors and starting values

### Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

## Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

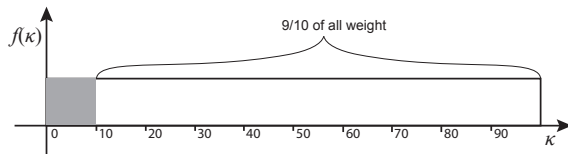
## Word of caution

## References

# Is uniform distribution a non-informative prior?

- Not really

- Imagine setting a  $\text{Uniform}(0, 100)$  prior for the transition/transversion rate ratio ( $\kappa$ ). You also know that the most likely values for  $\kappa$  are between 0 and 10. But you now put 9/10 of the weight to values  $> 10$ .



- In fact there is nothing such as an non-informative prior
- If little or no information on the parameter is available, use diffuse priors
- Try to avoid  $\text{Uniform}(-\infty, \infty)$  or  $\text{Uniform}(0, \infty)$

## Priors and starting values

### Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

## Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

## Word of caution

## References

- ▶ Sometimes the prior distribution is such that the sum or the integral of the prior values does not converge, this is called an IMPROPER prior
- ▶ Examples
  - ▶  $1/x$
  - ▶  $\text{Uniform}(-\infty, \infty)$

## Priors and starting values

### Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

## Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

## Word of caution

## References

# Are my priors what I set them to be?

- ▶ Not always
  - ▶ Induced priors may change the picture, i.e. if the parameters interact, the marginal prior distribution for each individual parameter may be different from the originally specified prior
- ▶ Use sampling from the prior, to see what your 'real' prior is

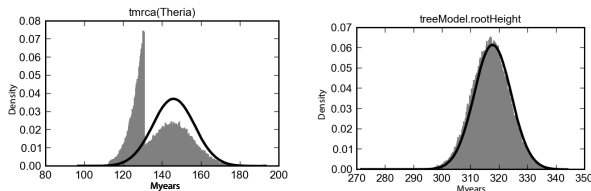


Figure adapted from [Heled and Drummond, 2012]

The marginal prior distributions that result from the multiplicative construction (gray) versus calibration densities (black line) specified for the calibrated nodes.

## Priors and starting values

### Priors

- Prior distribution
- Tree prior
- Substitution model prior
- Clock prior
- Parameter prior
- Think twice
- Starting values

## Troubleshooting

- Possible issues
- Initialization failed
- Nothing mixed
- One parameter did not mix

## Word of caution

## References

- ▶ Use all the prior knowledge you have to choose models and set appropriate parameter priors
- ▶ Sample from the prior distribution before using your data to check you really have the priors you want

Priors and starting values

Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

Word of caution

References

- ▶ Are just starting values
- ▶ Have to be within the prior distribution, and its upper and lower limits, you chose for the parameter
- ▶ Use your best guess
  - ▶ BEAST2 attempts 10 times at most (can be changed) to initialize the run, but if the starting values are unreasonable, the runs may keep failing
- ▶ Start from different starting values and make sure the chains converge to the same distribution

## Priors and starting values

### Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

### Starting values

## Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

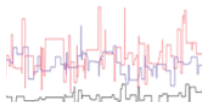
## Word of caution

## References

## ► Initialization failed

Start likelihood: -Infinity after 1000  
Fatal exception: Could not find a pr

## ► Nothing mixed



## ► One parameter did not mix

kappa.noncoding	13.143	191
kappa.1stpos	6.28	376
kappa.2ndpos	8.643	372
kappa.3rdpos	27.988	93
mutationRate.noncodi...	0.347	273
mutationRate.1stpos	0.459	238
mutationRate.2ndpos	0.185	237

### Priors and starting values

#### Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

### Troubleshooting

#### Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

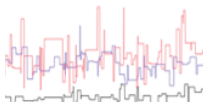
### Word of caution

### References

- Initialization failed

Start likelihood: -Infinity after 1000  
Fatal exception: Could not find a pr

- Nothing mixed



- One parameter did not mix

kappa.noncoding	13.143	191
kappa.1stpos	6.28	376
kappa.2ndpos	8.643	372
kappa.3rdpos	27.988	93
mutationRate.noncodi...	0.347	273
mutationRate.1stpos	0.459	238
mutationRate.2ndpos	0.185	237

## Priors and starting values

### Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

### Starting values

## Troubleshooting

### Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

## Word of caution

## References



Start likelihood: -Infinity after 1000 initialisation attempts

Fatal exception: Could not find a proper state to initialise. Perhaps try another seed.

P(posterior) = -Infinity (was -Infinity)

P(prior) = -Infinity (was -Infinity)

P(BDMM) = -Infinity (was -Infinity)

P(R0Prior) = -0.5586849541070393 (was -0.5586849541070393)

P(rPrior) = -11.46042136866474 (was -11.46042136866474)

P(rateMatrixPrior) = -0.14088025499381485 (was -0.14088025499381485)

P(samplingProportionPrior) = -10.049507225748343 (was -10.049507225748343)

P(becomeUninfectiousRatePrior) = -0.7811241751317991 (was -0.7811241751317991)

java.lang.RuntimeException: Could not find a proper state to initialise. Perhaps try another seed.

at beast.core.MCMC.run(Unknown Source)

at beast.app.BeastMCMC.run(Unknown Source)

at beast.app.beastapp.BeastMain.<init>(Unknown Source)

at beast.app.beastapp.BeastMain.main(Unknown Source)

at beast.app.beastapp.BeastLauncher.main(Unknown Source)

Fatal exception: Could not find a proper state to initialise. Perhaps try another seed.

BEAST has terminated with an error. Please select QUIT from the menu.

## Priors and starting values

### Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

## Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

## Word of caution

## References

- ▶ **Example:**  $P(\text{RateMatrixPrior}) = -\text{Infinity}$  (was  $-\text{Infinity}$ )
- ▶ **Possible solutions:**
  - ▶ Change seed
  - ▶ Adjust initial conditions
  - ▶ Use excludable/slicing priors
  - ▶ Check for silly/incompatible priors
  - ▶ Check for underflow (too low values)
  - ▶ Talk to the BEAST2 support group

## Priors and starting values

### Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

## Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

## Word of caution

## References

- ▶ **Example:**  $P(\text{BDMM}) = -\text{Infinity}$  (was  $-\text{Infinity}$ )
- ▶ **Possible solutions:**
  - ▶ Change seed
  - ▶ Talk to the BEAST2 support group

## Priors and starting values

### Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

## Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

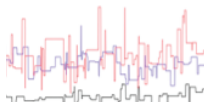
## Word of caution

## References

- Initialization failed

Start likelihood: -Infinity after 1000  
Fatal exception: Could not find a pr

- Nothing mixed



- One parameter did not mix

kappa.noncoding	13.143	191
kappa.1stpos	6.28	376
kappa.2ndpos	8.643	372
kappa.3rdpos	27.988	93
mutationRate.noncodi...	0.347	273
mutationRate.1stpos	0.459	238
mutationRate.2ndpos	0.185	237

## Priors and starting values

### Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

### Starting values

## Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

## Word of caution

## References

- ▶ Possible solutions:
  - ▶ Increase chain length
  - ▶ Run multiple independent chains
  - ▶ Increase sampling frequency (if ACT permits)
  - ▶ Check identifiability
  - ▶ Check if model is misspecified

## Priors and starting values

### Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

## Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

## Word of caution

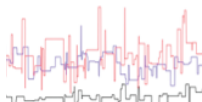
## References

# One parameter did not mix

## ► Initialization failed

Start likelihood: -Infinity after 1000  
Fatal exception: Could not find a pr

## ► Nothing mixed



## ► One parameter did not mix

kappa.noncoding	13.143	191
kappa.1stpos	6.28	376
kappa.2ndpos	8.643	372
kappa.3rdpos	27.988	93
mutationRate.noncodi...	0.347	273
mutationRate.1stpos	0.459	238
mutationRate.2ndpos	0.185	237

### Priors and starting values

#### Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

#### Starting values

### Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

### Word of caution

### References

- ▶ Possible solutions:
  - ▶ Run longer (or combine several independent chains)
  - ▶ Tweak the operators
    - ▶ Increase operator weight for low ESS parameters
    - ▶ Add upDown operator for correlated parameters

Priors and starting values

Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

Word of caution

References

## Priors and starting values

### Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

## Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

## Word of caution

## References

- Possible reasons:
  - A statistic is present in the model, for which the model cares too much, but the data says nothing about



- ▶ **Bad news:**
  - ▶ At the moment we cannot directly examine the ESS of the trees in BEAST2 (except for the tree height)
- ▶ **Good news:**
  - ▶ Mixing of the continuous parameters and likelihoods is indicative of mixing of trees
  - ▶ Can also use AWTY to see if the clade probabilities stabilised

## Priors and starting values

### Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

## Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

## Word of caution

## References

- ▶ Bayesian analysis always gives an answer!
- ▶ The answer is how uncertain we are
- ▶ In practice, it is important to evaluate the impact of the prior on the posterior in a Bayesian robustness analysis
  - ▶ Ideally, the posterior should be dominated by your data, such that the choice of the prior has little influence on the result
  - ▶ If this is not the case, the choice of prior is very important, and should be reported

## Priors and starting values

### Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

## Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

## Word of caution

## References

- du Plessis, L. and Stadler, T. (2015). Getting to the root of epidemic spread with phylodynamic analysis of genomic data. *Trends in microbiology*, 23(7):383–386.
- Heled, J. and Drummond, A. J. (2012). Calibrated tree priors for relaxed phylogenetics and divergence time estimation. *Systematic Biology*, 61(1):138–149.

## Priors and starting values

### Priors

Prior distribution

Tree prior

Substitution model prior

Clock prior

Parameter prior

Think twice

Starting values

## Troubleshooting

Possible issues

Initialization failed

Nothing mixed

One parameter did not mix

## Word of caution

## References