

INPHEST Development Journal

October 29, 2015

1 Simulation of Summary Statistics

1. Generate source host tree \bar{T} .

Inputs

- Sequence data.
- Divergence time calibration priors.

Outputs

- \bar{T} : time-calibrated phylogeny.

Procedure

- Use BEAST to generate sample from the posterior distribution of trees given the data, \mathcal{T} .
- Summarize \mathcal{T} using MCCT or other criteria to produce \bar{T} .

2. Generate samples of biogeographical history augmented host trees, \mathbf{T}^* .

Inputs

- \bar{T} : time-calibrated phylogeny.
- Labeling of tips on \bar{T} by geography (distributional range, specified as presence/absence over set of areas, for each tip).

Outputs

- \mathbf{T}^* : samples of host tree augmented with biogeographical history.

Procedure

- Use BayArea or RevBayes to generate stochastic mapped samples.

3. Generate samples of schedule of host events, \mathbb{H} .

Inputs

- \mathbf{T}^* : samples of host tree augmented with biogeographical history.

Outputs

- \mathbb{H} : samples of a schedule of host events, where each distinct element $\mathcal{H}_i, \mathcal{H}_i \in \mathbb{H}$ bijectively corresponds to a distinct tree from \mathbf{T}^* , and consists of a vector of ordered pairs of (τ, η) , where τ is the simulation time of an event and η is the event (host speciation, host extinction, host dispersal, etc.) that occurs at that time on the particular tree to which it corresponds.
4. For *each* simulation replicate:
 - (a) Sample a schedule of host tree events, \mathcal{H}_i , from \mathbb{H} .
 - (b) Simulate symbiont tree, ψ , conditioned on \mathcal{H} and other parameters.
 - (c) Calculate summary statistics on the symbiont tree, ψ .