

# INPHEST Development Journal

December 9, 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  $(\tau, \eta)$ , where  $\tau$  is the simulation time of an event and  $\eta$  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,  $\psi$ , conditioned on  $\mathcal{H}$  and other parameters.
    - (c) Calculate summary statistics on the symbiont tree,  $\psi$ .

## 2 Event Dynamics

### Host lineage speciation

- Daughter host lineages inherit parent symbiont load:
  - Both host daughters inherit full symbiont load of parent?
  - Some other scheme?

### Host lineage extinction

- Symbiont lineages loses host. If this results in any symbiont lineage host set being reduced to the null set, the symbiont lineage goes extinct.

### Host lineage area gain

- No action (passive area gain by proxy by symbiont lineages associated with this host).

### Host lineage area loss

- If host lineage does not go extinct (host lineage range not reduced to the null set): no action (passive area loss by proxy by symbiont lineages associated with this host).
- If host lineage goes extinct (host lineage range reduced to the null set): as with host lineage extinction (above).

### Symbiont lineage speciation

- Daughter symbiont lineages inherit parent host set pseudo-allopatrically: one daughter lineage inherits one host, other inherit remaining.
- Daughter symbiont lineages inherit parent host set pseudo-subset-sympatrically: one daughter lineage inherits one host, other inherits entire host set.
- Daughter symbiont lineages inherit parent host set by random sampling with replacement.
- Daughter symbiont lineages partition parent host set (randomly).

#### **Symbiont lineage extinction**

- Symbiont lineage removed from tree.

#### **Symbiont lineage host gain**

- Symbiont lineage gains host.
- Note: potential (new) hosts restricted to hosts that share one or more biogeographical areas with a current host.

#### **Symbiont lineage host loss**

- Symbiont lineage loses host, going extinct if host set reduced to the null set.

#### **Symbiont lineage host jumping**

- Symbiont lineage gains host.
- Note: potential (new) hosts restricted to hosts that share one or more biogeographical areas with a current host.

### **3 Questions**

- Should symbionts/parasites track geography independently of hosts?
  - If so, then symbiont range would be restricted to a subset of the host range. Each host lineage would probably need to be modeled independently in each area wrt to parasite load.