

The probability of a gene on branch e going extinct is:

$$E_e = p_e^L + p_e^S E_f E_g + p_e^D E_e E_e + \left(\frac{1}{|S|} \sum_h p_h^T E_h \right) E_e, \quad (1)$$

where $|S|$ is the number of branches of S including the stem branch leading to the root, the sum goes over all of these, and f and g are descendants of e , if these do not exist the corresponding terms must be dropped. The event probabilities are derived from DTL "rates" according to:

$$\begin{aligned} p_e^S &= 1/(1 + \delta_e + \tau_e + \lambda_e), \\ p_e^D &= \delta_e/(1 + \delta_e + \tau_e + \lambda_e), \\ p_e^T &= \tau_e/(1 + \delta_e + \tau_e + \lambda_e), \\ p_e^L &= \lambda_e/(1 + \delta_e + \tau_e + \lambda_e). \end{aligned} \quad (2)$$

The same term $\bar{E}^T = 1/|S| \sum_h p_h^T E_h$ is present for all e , hence more compactly

$$E_e = p_e^L + p_e^S E_f E_g + p_e^D E_e^2 + \bar{E}^T E_e. \quad (3)$$

The equation has the problem of depending on itself (E_e) and the global average \bar{E}^T – we already know E_f and E_g by the time we get to branch e . What I do aside of the usual initial conditions is first set all E_e and \bar{E}^T to zero and run a initial round of calculation from the leaf toward the root during which I also calculate a first estimate of \bar{E}^T . Using these values I then repeat this to converge toward fix point values for the E_e .

The probability of gene tree branch u on branch e of S is:

$$\begin{aligned} P_{e,u} &= p_e^S (P_{g,v} P_{f,w} + P_{g,w} P_{f,v} + E_f P_{g,u} + P_{f,u} E_g) \\ &\quad + p_e^D (P_{e,v} P_{e,w} + 2P_{e,u} E_e) \\ &\quad + \left(\frac{1}{|S|} \sum_h p_h^T P_{h,w} \right) P_{e,v} + \left(\frac{1}{|S|} \sum_h p_h^T P_{h,v} \right) P_{e,w} \\ &\quad + \left(\frac{1}{|S|} \sum_h p_h^T E_h \right) P_{e,u} + \left(\frac{1}{|S|} \sum_h p_h^T P_{u,h} \right) E_e \end{aligned} \quad (4)$$

Here the recurring terms are $\bar{P}_u^T = 1/|S| \sum_h p_h^T P_{u,h}$, using which we have more compactly

$$\begin{aligned}
P_{e,u} = & \bar{P}_e^S (P_{g,v} P_{f,w} + P_{g,w} P_{f,v} + E_f P_{g,u} + P_{f,u} E_g) \\
& + \bar{P}_e^D (P_{e,v} P_{e,w} + 2P_{e,u} E_e) \\
& + \bar{P}_w^T P_{e,v} + \bar{P}_v^T P_{e,w} \\
& + \bar{E}^T P_{e,u} + \bar{P}_u^T E_e,
\end{aligned} \tag{5}$$

f and g are descendants of e , v and w are descendants of u , if any of these do not exist the corresponding terms must be dropped. The equation again has the problem of depending on itself ($P_{e,u}$) and the global average \bar{P}_u^T – we already know \bar{P}_v^T , \bar{P}_w^T , $P_{v,*}$ and $P_{w,*}$, as well as all the $P_{*,g}$ and $P_{*,f}$ that we need by the time we get to branch e of S and u of G . What I do aside of the usual initial conditions is first set all $P_{e,u}$ and the \bar{P}_u^T to zero and run a initial round of calculation from the leaf of both S toward the root and using double recursion along G , during which I also calculate a first estimate of \bar{P}_u^T . Using these values I then repeat this to converge toward fix point values for the $P_{e,u}$.

Incorporating amalgamation is straightforward. Reconciliations can be obtained by stochastic traceback along $P_{e,u}$ such that the \bar{P}_u^T terms are also expanded to decide where a transfer goes.