ParEvol

William R. Shoemaker and Jay T. Lennon

30 May, 2018

Figure out what Z represents, go back through paper, step-by-step

Probability of parallel evolution

$$\mathbf{E}(M_{ij}) = \mathbf{E}(Z_i Z_j) - \mathbf{E}_k(Z_i Z_k) - \mathbf{E}_k(Z_j Z_k) + \mathbf{E}_{kl}(Z_k Z_l)$$

where Z_i represents population i.

$\mathbf{E}(Z_i Z_j)$ is the expected value of the joint probability distributions for the number of mutations in each gene. Because the two populations are independant (i.e., no cross-contamination), this reduces to $\mathbf{E}(Z_i)\mathbf{E}(Z_j)$.

$\mathbf{E}_k(Z_i Z_k) = \frac{1}{n}\sum_{k=1}^{n} \mathbf{E}(Z_i Z_k) = \frac{1}{n}(\sum_{k=1}^{n} \mathbf{E}(Z_i) \sum_{k=1}^{n} \mathbf{E}(Z_k)) = \mathbf{E}(Z_i)(\frac{1}{n}\sum_{k=1}^{n} \mathbf{E}(Z_k))$ is the expected value for sample i after averaging over all populations. Again, assuming independance. The same principle applies for sample j.

$\mathbf{E}_{kl}(Z_k Z_l)$ is the expected value of the joint the probability distribution for two populations chosen at random with replacement. Again, assuming independance, we can re-write this as $\mathbf{E}_{kl}(Z_k Z_l) = \frac{1}{n^2}\sum_{k=1}^{n} \mathbf{E}(Z_k) \sum_{l=1}^{n} \mathbf{E}(Z_l)$

Can we say that Z_i is the mean of the poisson process describing the rate that mutations are acquired? Or is it just proportional and the PCA procedure rescales the data?

ignore this stuff. . . .

$$\frac{\partial \overline{M}_b(t)}{\partial t} = \int N U_b \rho_b(s,t) p_{fix}(s) ds$$

$$\frac{\partial \overline{M}(t)}{\partial t} = \frac{\partial \overline{M}_b(t)}{\partial t} + U_n$$

assume we're in the strong-selection, weak-mutation (SSWM) limit, where $p_{fix}(s) \approx 2s$ and that each mutation has the same fitness effect $\rho_b(s,t) = s$, giving us the simple toy model for the rate that a gene acquires beneficial mutations

$$\frac{\partial \overline{M}_b(t)}{\partial t} = \int 2 N U_b s^2 ds$$

after solving the integral, we get

$$\frac{\partial \overline{M}_b(t)}{\partial t} = \frac{2}{3} N U_b s^3 + C$$

assuming $\overline{M}_b(0) = 0$, we solve for the constant and get

Start by focusing on a scenario where the beneficial DFE evolves according to simple mean-field dynamics

$$L_b \frac{\partial \rho_b(s,t)}{\partial t} = -N U_b \rho_b(s,t) p_{fix}(s)$$

$$\rho_b(s,t) = p_0(s)e^{-2NU_b st/L_b}$$

and evaluate $\frac{\partial \overline{M}_b(t)}{\partial t}$ by assuming that $p_0(s)$ is an exponential distribution with a rate parameter of 1 (e^{-2}) and setting $2NU_b t/L_b$ equal to k, we evaluate the integral