PBG 200A Notes

Sam Fleischer

November 4, 2016

1   Question 4 from Ch. 2 Notes

Assume that humans and chimps diverged around 5.5×10^6 years ago, have a generation time of ~ 20 years, that the speciation occured instantaneously in allopatry with no subsequent gene flow, and the ancestral effective population size of the human and chimp common ancestor population was $10,000$ individuals. Nachman and Crowell sequenced 12 pseudogeners in humans and chimps and found substitutions at 1.3% of sites.
***A)** What can you say about the mutation rate per site per generation at these genes, and how does it compare to other estimates of human mutation rate?*
***B)** All of the pseudogenes they sequenced are on the autosomes. What would your prediction be for pseudogenes on the X and Y chromosomes, given that there are fewer rounds of replication in the female germline than in the male germline.*

A) Set $0.013 = 2\mu(1 - C)T$. where μ is the mutation rate we are trying to find, C is the fraction of mutations which are neutral, and T is the number of generations since the common ancestor. We can calculate T by $5.5 \times 10^6/20 = 2.75 \times 10^5$. Then solving for $\mu(1 - C)$ gives us $\mu(1 - C) = \dfrac{0.013}{2 \times 2.75 \times 10^5} \approx 2.\overline{36} \times 10^{-8}$.

B)

2   From Last Time

- An ancestral population may split in to two separately drifting populations.

- Suppose both populations have current effective population size N_e and the ancestral effective population size is N_e. Note $F_{ST} = 1 - \dfrac{H_S}{H_T}$. The heterozygosity of the subpopulation,

$$H_S = 2N_e \times 2\mu. \tag{1}$$

  The heterozygosity of the total population,

$$H_T = \frac{1}{2}H_S + \frac{1}{2}N_B \tag{2}$$

  where H_B is the heterozygosity between subpopulations.

$$H_B = 2\mu(T + 2N_e) \tag{3}$$

  Thus

$$F_{ST} = \frac{\mu T}{\mu T + 4N_e\mu} = \frac{T}{T + 4N_e} \approx \frac{T}{4N_e} \qquad \text{given } \frac{T}{N_e} \ll 1 \tag{4}$$

- F_{ST} between Africans and Europeans is $F_{ST} \approx 0.095$. Assuming $N_e = 10,000$, we can find

$$T \approx 3,800 \text{ years} \approx 114,000 \text{ generations} \tag{5}$$

3   Island Mainland Model

- Mean heterozygosity: hae a set of 3 loci, biallelic. If frequency of one of them is 0.2, then heterozygosity is $2 \times 0.2 \times 0.8$. Do the same for 0.7, 0.8, take arithmetic mean, get $H_I \approx 0.35$. For the mainland, $H_M \approx 0.47$. We plug these in to eq. 72 in notes and get $m \approx 0.0039$. This shows it may only takes a small number of migrations to get very low levels of heterozygosity.

- The rate of migration affects F_{ST}, as is shown in the correlation between pairwise distance between populations and F_{ST}.

4 Population genetics of divergence between species

The rate of neutral mutations μ is equal to $(1 - C)\mu_T$. The number of neural mutations at a give time point is

$$2N\mu \tag{6}$$

And the probability that a particular mutation fixates is

$$2N\mu \times \frac{1}{2N} = \mu \tag{7}$$

So the substitution rate is equal to the neutral mutation rate (completely independent with population size). So the expected number of neutral substitutions in T time between two species with a common ancesotr is $2T\mu$.

We can use comparitive genomics to get an idea of which genes are functionally important.