# Stochastic population genetics: homework 3
## To be returned on May 17

May 11, 2017

## 1 Coalescence with migration

In this problem we are interested in two populations of $N >> 1$ diploids living on two separate "islands". At each generation, for any given island a parent allele is picked at random (as in a W-F model) from the same island with probability $1 - M/4N$ and from the other island with probability $M/4N$ ($M = O(1)$), i.e. the probability that the two alleles come from the same island of the individual which is being considered is $(1 - M/4N)^2$. Mutations occur independently of parents origin at the rate $\theta/4N$. All mutations are assumed to be neutral and we take the limit of infinite size genome, i.e. mutations never occur twice at the same site and always form a unique new genotype).

As in the W-F model, we consider each allele of each individual independently effectively giving two subpopulations of $2N$ individuals. We denote by $P_c(\theta)$ the probability that two alleles are the same if they belong to the same island and by $P_d(\theta)$ the probability if they belong to two different islands.

a. Why can we write that at first order the probability for two alleles from two different islands to have the same parent is 0? Show that

$$P_c(\theta) = \frac{1}{1 + \theta + M} + \frac{M}{1 + \theta + M} P_d(\theta). \tag{1}$$

b. Show that

$$P_d(\theta) = \frac{M}{\theta + M} P_c(\theta). \tag{2}$$

Compute $P_c(\theta)$ and $P_d(\theta)$ as functions of the parameters. Check the limit $\theta = 0$.

c. How would the result of the previous question change if we had $q$ islands and that the destination of each migration was chosen at random with equal probability within the islands?

d. Let $t_c$ be the time to the first common ancestor between two individuals from the same island and $t_d$ that same time for two individuals from different

islands. Show that

$$P_c(\theta) = \int_0^\infty du\, e^{-u\frac{\theta}{2N}} P(t_c = u), \tag{3}$$

and equivalently

$$P_d(\theta) = \int_0^\infty du\, e^{-u\frac{\theta}{2N}} P(t_d = u). \tag{4}$$

e. From Eq. 3 and Eq. 4 show that

$$-2N P_d'(0) = \langle t_d \rangle \text{ and that } -2N P_c'(0) = \langle t_c \rangle. \tag{5}$$

Compute these two expectation values for $q$ islands. What is the expected number of differences between two alleles? Give an intuitive explanation for the fact that the average time to the most recent common ancestor between two alleles of the same subpopulation is independent of the migration rate.