

Derivation of Approximations for Low Infection Rates

Derivation of Equation (6)

As shown in the main text, the iteration equation for the frequency X_t of the resistance allele is

$$X_{t+1} = 2X_t \frac{\sum_{j=0}^n \binom{n}{j} j \left(\frac{1}{2}\right)^n (1 - \gamma_{j-1, n-j})}{n(1-m)}, \quad (1)$$

where n is the number of the children per family, m is the population disease mortality, and $\gamma_{j,k}$ is the disease mortality rate for an Aa child with j Aa sibs, one Aa parent, k aa sibs and one aa parent. This is given by

$$\gamma_{jk} = \left[\pi_1 - (1 - \pi_1) \left[1 - (1 - \pi_1 f_1)^{j+1} (1 - \pi_0 f_1)^{k+1} \right] \right] \mu. \quad (2)$$

If infection rates are low, then we make the approximation

$$(1 - \pi_1 f_1)^{j+1} = 1 - (j+1)\pi_1 f_1 + O(\pi_1^2 f_1^2), \quad (3)$$

$O(\pi^2 f^2)$ represents terms of order two and higher in πf . These terms are negligible

when π_0, π_1 , and f_1 are sufficiently small and n is not too large. Then (2) becomes

$$\gamma_{jk} = \left[\pi_1 - (1 - \pi_1) \left[(j+1)\pi_1 f_1 + (k+1)\pi_0 f_1 \right] \right] \mu + O(\pi_1^2 f_1^2) \quad (4)$$

Substituting this into (1), we get

$$X_{t+1} \approx \frac{2X_t}{n(1-m)} \left[\sum_{j=0}^n \binom{n}{j} j \left(\frac{1}{2}\right)^n - \sum_{j=0}^n \binom{n}{j} j \left(\frac{1}{2}\right)^n \left\{ \pi_1 - (1 - \pi_1) \left[(j)\pi_1 f_1 + (n-j+1)\pi_0 f_1 \right] \right\} \mu \right] \quad (5)$$

or

$$X_{t+1} \approx \frac{2X_t}{n(1-m)} \left[\frac{n}{2}(1 - \mu\pi_1) + (1 - \pi_1) f_1 \sum_{j=0}^n \binom{n}{j} j \left(\frac{1}{2}\right)^n \left\{ [(j)\pi_1 + (n-j+1)\pi_0] \right\} \mu \right] \quad (6)$$

This can be written as

$$X_{t+1} \approx \frac{2X_t}{n(1-m)} \left[\frac{n}{2} (1 - \mu\pi_1) + (1 - \pi_1) \mu f_1 \left(\pi_1 E(n_{Aa}^2) + \pi_0 E(n_{Aa} n_{aa}) \right) \right] \quad (7)$$

Where n_{Aa} and n_{aa} are the numbers of Aa and aa genotype individuals in the family, respectively (not including the Aa individual whose survival is being calculated). It can be easily shown that

$$E(n_{Aa}^2) = \sum_{j=0}^n \binom{n}{j} j^2 \left(\frac{1}{2} \right)^n = \frac{n(n+1)}{4} \quad (8)$$

and

$$E(n_{Aa} n_{aa}) = \sum_{j=0}^n \binom{n}{j} j(n+1-j) \left(\frac{1}{2} \right)^n = \frac{n(n+1)}{4} \quad (9)$$

Substituting these into (7), we get

$$X_{t+1} \approx \frac{2X_t}{n(1-m)} \left[\frac{n}{2} (1 - \mu\pi_1) + (1 - \pi_1) \mu f_1 \frac{n(n+1)}{4} (\pi_1 + \pi_0) \right] \quad (10)$$

or

$$X_{t+1} \approx \frac{X_t}{(1-m)} \left[(1 - \mu\pi_1) + (1 - \pi_1) \mu f_1 \frac{(n+1)}{2} (\pi_1 + \pi_0) \right] \quad (11)$$

Now, in the absence of family structure, (1) becomes

$$X_{t+1}^{(0)} \approx \frac{X_t (1 - \mu\pi_1 - (1 - \pi_1) \mu f_1 \pi_0 (n+1))}{(1-m)} \quad (12)$$

Rearranging (10), we get

$$X_{t+1} \approx \frac{X_t}{1-m} \left\{ (1 - \mu [\pi_1 - (1 - \pi_1) \pi_0 f_1 (n+1)]) + (1 - \pi_1) \mu f_1 \frac{(n+1)}{2} (\pi_1 - \pi_0) \right\} \quad (13)$$

or

$$X_{t+1} \approx X_{t+1}^{(0)} + \frac{(1-\pi_1)\mu f_1 \frac{(n+1)}{2}(\pi_1 - \pi_0)}{1-m} \quad (14)$$

or

$$X_{t+1} \approx X_{t+1}^{(0)} + \frac{\Delta M_F}{1-m} X_t \quad (15)$$

with $\Delta M_F = \frac{1}{2}(n+1) f_1 (\pi_0 - \pi_1) \mu$. This is equation (6) in the main text.

Derivation of equation (8)

The selection coefficient for a rare allele can be approximated by

$$s = \frac{X_{t+1} - X_t}{X_t} \quad (16)$$

The selection coefficient with family structure is then

$$s_1 = \frac{\frac{X_t (1 - \mu\pi_1 - (1 - \pi_1) \mu f_1 \pi_0 (n+1))}{(1-m)} + \frac{\Delta M_F}{(1-m)} X_t - X_t}{X_t} \quad (17)$$

After some rearranging, this becomes

$$s_1 = \frac{(1 - \mu\pi_1 - (1 - \pi_1) \mu f_1 \pi_0 (n+1))}{(1-m)} + \frac{\Delta M_F}{(1-m)} - \frac{1-m}{1-m} \quad (18)$$

or

$$s_1 = \frac{m - (\mu\pi_1 + (1 - \pi_1) \mu f_1 \pi_0 (n+1))}{(1-m)} + \frac{\Delta M_F}{(1-m)}$$

Taking $\Delta M_0 = m - \mu [\pi_1 + (1 - \pi_1) f_1 \pi_0 (n + 1)]$ ($\Delta M_0 = m - \mu [\pi_1 + (1 - \pi_1) f_1 \pi_0 (n + 1)]$)

(this is the approximate average mortality avoided by a heterozygote relative to a normal individual when there is no family structure), we get

$$s_1 = \frac{\Delta M_0}{(1 - m)} + \frac{\Delta M_F}{(1 - m)} \quad (19)$$

We can easily see that the selection coefficient s_0 in the absence of family structure is

$$s_1 = \frac{\Delta M_0}{(1 - m)} \quad (20)$$

From this, it follows directly that the ratio of selection coefficients is

$$\frac{s_1}{s_0} = 1 + \frac{\Delta M_F}{\Delta M_0} \quad (21)$$