EFFECT OF LINKAGE ON THE HOMOZYGOSITY OF SELFED POPULATION

By M. Rajagopalan

Indian Council of Agricultural Research, New Delhi

When more than one gene pair is involved and if they are linked, the increase in the homozygosity from generation to generation will depend on the number of factors, the recombination fraction and the mating system used. The object of this note is to obtain the expression for the coefficient of inbreeding of an individual obtained by n-generations of selfing when two gene pairs are involved.

2. Let A — a and B — b be the two gene pairs and p the recombination fraction. Then there will be ten genotypes, namely, AABB, AABb, AAbb, AaBB, AB/ab, Ab/aB, Aabb, aaBB, aaBb and $aabb$. Let their frequencies in the n-th generation obtained by continued selfing be $f_{22}^{(n)}$, $f_{21}^{(n)}$, $f_{20}^{(n)}$, $f_{12}^{(n)}$, $f_{11c}^{(n)}$, $f_{11r}^{(n)}$, $f_{10}^{(n)}$, $f_{02}^{(n)}$, $f_{01}^{(n)}$ and $f_{00}^{(n)}$ respectively.

Then the following relations hold between frequencies in the n-th and $(n + 1)$-th generations:

$$\left.\begin{aligned}
f_{22}^{(n+1)} &= f_{22}^{(n)} + \tfrac{1}{4} f_{21}^{(n)} + \tfrac{1}{4} f_{12}^{(n)} + \tfrac{1}{4}(1-p)^2 f_{11c}^{(n)} \\
&\quad + \tfrac{1}{4} p^2 f_{11r}^{(n)} \\
f_{21}^{(n+1)} &= \tfrac{1}{2} f_{21}^{(n)} + \tfrac{1}{2} p(1-p) f_{11c}^{(n)} + \tfrac{1}{2} p(1-p) f_{11r}^{(n)} \\
f_{20}^{(n+1)} &= \tfrac{1}{4} f_{21}^{(n)} + f_{20}^{(n)} + \tfrac{1}{4} p^2 f_{11c}^{(n)} + \tfrac{1}{4}(1-p)^2 f_{11r}^{(n)} \\
&\quad + \tfrac{1}{4} f_{10}^{(n)} \\
f_{12}^{(n+1)} &= \tfrac{1}{2} f_{12}^{(n)} + \tfrac{1}{2} p(1-p) f_{11c}^{(n)} + \tfrac{1}{2} p(1-p) f_{11r}^{(n)} \\
f_{11c}^{(n+1)} &= \tfrac{1}{2}(1-p)^2 f_{11c}^{(n)} + \tfrac{1}{2} p^2 f_{11r}^{(n)} \\
f_{11r}^{(n+1)} &= \tfrac{1}{2} p^2 f_{11c}^{(n)} + \tfrac{1}{2}(1-p)^2 f_{11r}^{(n)} \\
f_{10}^{(n+1)} &= \tfrac{1}{2} p(1-p) f_{11c}^{(n)} + \tfrac{1}{2} p(1-p) f_{11r}^{(n)} + \tfrac{1}{2} f_{10}^{(n)} \\
f_{02}^{(n+1)} &= \tfrac{1}{4} f_{12}^{(n)} + \tfrac{1}{4} p^2 f_{11c}^{(n)} + \tfrac{1}{4}(1-p)^2 f_{11r}^{(n)} + f_{02}^{(n)} \\
&\quad + \tfrac{1}{4} f_{01}^{(n)} \\
f_{01}^{(n+1)} &= \tfrac{1}{2} p(1-p) f_{11c}^{(p)} + \tfrac{1}{2} p(1-p) f_{11r}^{(n)} + \tfrac{1}{2} f_{01}^{(n)} \\
f_{00}^{(n+1)} &= \tfrac{1}{4}(1-p)^2 f_{11c}^{(n)} + \tfrac{1}{4} p^2 f_{11r}^{(n)} + \tfrac{1}{4} f_{10}^{(n)} + \tfrac{1}{4} f_{01}^{(n)} \\
&\quad + f_{00}^{(n)}.
\end{aligned}\right\} \quad (1)$$

Let W_n, X_n and Y_n denote the frequencies of the double homozygotes, single heterozygotes and double heterozygotes respectively. These can be seen to be:

$$\left.\begin{aligned} W_n &= f_{22}^{(n)} + f_{20}^{(n)} + f_{02}^{(n)} + f_{00}^{(n)} \\ X_n &= f_{21}^{(n)} + f_{12}^{(n)} + f_{10}^{(n)} + f_{01}^{(n)} \\ \text{and} \\ Y_n &= f_{11c}^{(n)} + f_{11r}^{(n)} \end{aligned}\right\} \quad (2)$$

Making use of (2), the equations in (1) can be expressed as:

$$\left.\begin{aligned} W_{n+1} &= W_n + \tfrac{1}{2} X_n + \tfrac{1}{2} kY_n \\ X_{n+1} &= \tfrac{1}{2} X_n + (1 - k)\, Y_n \\ Y_{n+1} &= \tfrac{1}{2} kY_n \end{aligned}\right\} \quad (3)$$

where

$$k = \{p^2 + (1 - p)^2\}.$$

3. Let P_n denote $(X_n + Y_n)$, the frequency of heterozygotes in the n-th generation. From (3) we obtain the following recurrence relation for P_n:

$$P_{n+2} - \tfrac{1}{2}(1 + k)\, P_{n+1} + \tfrac{1}{4} kP_n = 0 \quad (4)$$

The solution for P_n can be shown to be

$$P_n = (P_0 + Y_0)\, (\tfrac{1}{2})^n - Y_0\, (\tfrac{1}{2} k)^n \quad (5)$$

Defining the inbreeding coefficient F_n, in the n-th generation as the reduction in the heterozygosity relative to the initial population, we have

$$F_n = \frac{P_0 - P_n}{P_0} = \left(1 - \frac{1}{2^n}\right) - \frac{1}{2^n}\left(1 - k^n\right)\frac{Y_0}{P_0} \quad (6)$$

It may be noted that the first term, $(1-1/2^n)$ is the expression for the inbreeding coefficient when a single gene pair is involved. Further, when two gene pairs are involved F_n depends on the constitution of the initial population which is not the case for a single gene pair.

4. Let us take the special case when the initial population consists of one or both single and double heterozygotes or either of them only. Then the expression for the inbreeding coefficient will be

$$F_n = 1 - 2\,(\tfrac{1}{2})^n + (\tfrac{1}{2}k)^n \quad (7)$$

When there is no linkage $p = \frac{1}{2}$ therefore, $k = \frac{1}{2}$ and the value of the inbreeding coefficient will be

$$F_n\,(p = \tfrac{1}{2}) = 1 - 2\,(\tfrac{1}{2})^n + (\tfrac{1}{4})^n = \left(1 - \frac{1}{2^n}\right)^2 \tag{8}$$

This can be seen to be the square of the inbreeding coefficient when a single gene pair is involved.

When linkage is tight $p = 0$ therefore, $k = 1$ and the expression for the inbreeding coefficient will be

$$F_n\,(p = 0) = 1 - \frac{1}{2^n}\,. \tag{9}$$

This is same as for a single factor and is expected as both the gene pairs behave as one.

The following table gives the values of the inbreeding coefficient up to five generations for different recombination fractions for the population considered in this section. It is obvious that with closer linkage there is a greater increase in homozygosity.

p \ n	1	2	3	4	5
0 ..	0·5000	0·7500	0·8750	0·9375	0·9688
0·1 ..	0·4100	0·6681	0·8189	0·9033	0·9491
0·2 ..	0·3400	0·6156	0·7893	0·8884	0·9420
0·3 ..	0·2900	0·5841	0·7744	0·8821	0·9496
0·4 ..	0·2600	0·5676	0·7676	0·8796	0·9387
0·5 ..	0·2500	0·5625	0·7656	0·8789	0·9385

Summary

The expression for the coefficient of inbreeding of an individual obtained by n-generations of selfing when two gene pairs are involved is obtained.

The author is grateful to Dr. V. G. Panse and Shri V. N. Amble for the encouragement given by them.