Expectations of means and genetic variances in backcross populations*

T. S. Cox
Research Geneticist, USDA-ARS, Room 354, Throckmorton Hall, Kansas State University, Manhattan, KS 66506, USA

Received September 12, 1983
Communicated by A. L. Kahler

Summary. The genetic variance among F2-derived lines of backcrosses (BCgF2-derived lines) depends on the backcross generation (g), the number of F1 plants crossed and selfed in generations 1 through g, and the number of BCgF2-derived lines evaluated. Additive genetic variance decreases linearly with backcrossing when one BCF1 plant per generation is crossed and selfed. The relationship is curvilinear if more than one BCF1 plant is used; as the number of BCF1 plants increases, additive genetic variance among BC1F2-derived lines approaches that among BC0F2-derived lines. The effect of population size on genetic variance is due both to fixation of alleles in previous generations and to sampling of genotypes in the population being evaluated. Dominance and repulsion linkage can cause small increases in genetic variance from BC0 to BC1.

Key words: Introgression - Population size - Sampling - Backcrossing - Linkage

Introduction

Utilization of exotic germplasm often involves backcrossing to an adapted parent before starting selection in order to restore agronomic suitability (MacKey 1963). Several recent theoretical (Bailey 1977; Baker 1976), computer-simulation (Dudley 1982; Reddy and Comstock 1976), and empirical (Lawrence and Frey 1975; Rodgers 1982) studies have investigated the distribution of progeny in populations derived from backcrosses and the implications for plant breeding for quantitative traits.

To improve a quantitative trait when one parent is inferior to the other, a plant breeder must make enough backcrosses to increase the mean sufficiently but not so many that genetic variance is reduced below the level required for successful selection. In this paper, the theoretical means and genetic variances of populations of random F2-derived lines from successive backcrosses are expressed in terms of the number of backcrosses, the number of F1 plants crossed and selfed in each backcross generation, and the number of F2-derived lines evaluated per F1 family.

Population structure

The types of populations considered in this paper have been utilized experimentally by Lawrence and Frey (1975), Rodgers (1982), and Cox and Frey (198).

Populations are developed as follows. A cross is made between recurrent and donor homozygous parents, the F1 is backcrossed to produce BC1F1 and also selfed to produce the F2 (herein called BC0F2) generation (Fig. 1). Individual BC0F2 plants are self-pollinated to produce BC0F2-derived lines. Pollen from a number (n1) or BC1F1 plants is bulked and used to make the second backcross, producing the BC2F1; the same n1 BC1F1 plants are selfed to produce the BC1F2. BC1F2 seed is bulked and planted, and BC1F2-derived lines produced as in the BC0. The process is repeated through any number of backcrosses (Fig. 1).

BCgF2-derived lines may be propagated by self-pollination through any number of generations with bulkling of seed of selfed plants within lines. Thus, for example, we refer to BC2F2-derived lines in the F3, BC3F2-derived lines in the F4, etc.
Means of backcross populations

The mean of the BC2F2 generation in the absence of epistasis is

$$\mu_g = \{1 - (\frac{1}{2})^g\} R + (\frac{1}{2})^g (\frac{1}{2})^{g-1} H$$

(1)

where R is the mean of the recurrent parent and H the mean of the F1. Depending on which parent has the higher value, the progeny mean will increase or decrease linearly with the percentage of cultivated germplasm in the population (Fig. 2).

Genetic variances of backcross populations

The genetic variance among BC2F2-derived lines was obtained for two cases: (1) an additive model with dominance and (2) a two-locus model with dominance, repulsion linkage, and no epistasis.

Additive model with dominance

The genetic variance among BC2F2-derived lines may be divided into two components: (1) variance among BC2F1 families and (2) mean variance among BC2F2-derived lines within BC2F1 families. Symbolically,

$$V_g = V_g(F_1) + V_g(F_2/F_1).$$

(2)

The genetic variance among F1 families is zero (i.e., $V_0(F_1) = 0$). For $g > 0$, $V_g(F_1)$ depends on the number of F1 plants selfed to produce generation g and on the number of F1 plants backcrossed in generations 1 through $(g-1)$ but is independent of the number of BC2F2-derived lines evaluated. $V_g(F_1)$ may be derived as follows.

Let $n_g$ be the number of BC2F1 plants selfed to produce BC2F2-derived lines, and let k be the number of loci segregating.

Consider locus i at which $x_g$ out of $n_g$ BC2F1 plants are heterozygous. Let families derived from BC2F1 plants of the following genotypes have the indicated means:

<table>
<thead>
<tr>
<th>BC2F1 genotype</th>
<th>Frequency</th>
<th>Family mean (BC2F2-derived lines in F3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_iB_i$</td>
<td>$\frac{x_g}{n_g}$</td>
<td>$a_i$</td>
</tr>
<tr>
<td>$B_ib_i$</td>
<td>$\frac{x_g}{n_g}$</td>
<td>$d_i (\frac{1}{2})^{i-1}$</td>
</tr>
</tbody>
</table>

The variance due to locus i is

$$\frac{n_g!}{x_g!(n_g-x_g)!} \left( \frac{x_g}{2n_g-1} \right)^{x_g} \left( 1 - \frac{x_g-1}{2n_g-1} \right)^{n_g-x_g}$$

(4)

The variance among families, in the case of no dominance ($d_i = 0$), is

$$V_g(F_1) = n_g \left( 1 - \frac{x_g}{n_g} \right) a_i^2.$$  

(3)

The probability that $x_g$ BC2F1 plants are heterozygous at locus i is given by the binomial formula:

$$P(x_g) = \frac{n_g!}{x_g!(n_g-x_g)!} \left( \frac{x_g-1}{2n_g-1} \right)^{x_g} \left( 1 - \frac{x_g-1}{2n_g-1} \right)^{n_g-x_g}$$

where $x_g-1$ plants were heterozygous in the BC1F1. The variance due to locus i is

$$V_g(F_1) = \sum_{x_g=1}^{n_g} (P(x_g))(V_g(F_1)|x_g)$$

(4)

$$= \sum_{x_g=1}^{n_g} \frac{n_g!}{x_g!(n_g-x_g)!} \left( \frac{x_g-1}{2n_g-1} \right)^{x_g} \left( 1 - \frac{x_g-1}{2n_g-1} \right)^{n_g-x_g} \left( \frac{x_g}{n_g} \right)^{2} \left( \frac{x_g}{n_g} \right)^{2} a_i^2.$$