Simulation of the Backcross Breeding Method. I. Effects of Heritability and Gene Number on Fixation of Desired Alleles[1]

Belum V. S. Reddy and Ralph E. Comstock[2]

ABSTRACT

The potential of backcross breeding for improvement of a pure line by addition of favorable alleles from a donor line was investigated by computer simulation coupled with quantitative genetics theory. Attention was focused on the effects of heritability (among individual F_2 plants) and of the number of genes of which the allele present in the donor line is more favorable than the one present in the recipient. The specific programs investigated were ones involving a large amount of effort, a minimum of 1,000 pollinations per backcross generation and selection among families in three or more generations. The criterion employed for effectiveness was the probability of fixation in the product of the program of favorable alleles derived from the donor line. Effectiveness was found to be greater when heritability is greater but the effect of increased heritability was not as great as might have been expected. Number of favorable alleles available from the donor line had greater impact, particularly when success was defined in terms of probability that all the available favorable alleles would be transferred from donor to recipient. Success, so defined, was limited to one allele in the case of the least laborious of the three programs studied and no more than five in the case of the most costly. On the other hand, when success was measured in terms of percent improvement in the selected trait, it appeared substantial change is possible with heritability as low as 15 percent and favorable alleles available in the range from 1 to 16. Additive effects and independent assortment were assumed for genes simulated.

Additional index words: **Selection, Genetic improvement.**

THE germplasm of a species exists in separate populations, varieties or genotypes. The prime problem of the breeder is to put together in one reproducible genotype the favorable alleles that can be drawn from different sources. It is obvious that this is a difficult task and evident from reports by Bliss and Gates (1968), Bailey (1972) and Bailey and Comstock (1976) that the difficulty is extreme in self-fertilized organisms.

However, the backcross method enables further improvement of an already superior genotype by transfer of one or more useful alleles from another source without otherwise significant change in the recipient genotypes unless unfortunate gene linkage is encountered. For examples of the effective use of this procedure see Allard (1960), Briggs and Knowles (1967), Briggle (1969) and Nagai et al. (1973). Rinke (1960) argued that sequential applications of the backcross method is a preferred procedure for aggregating the useful alleles of a species into a single genotype.

It is well known that this method works best and requires the least effort when the alleles to be transferred have their effects on traits that are highly heritable and that the transfer of one allele is much easier than the simultaneous transfer of a larger number of alleles. However, successful applications for the improvement of quantitative traits having imperfect heritability have been reported by Knott and Talukdar (1971) and Duvick (1974).

This paper will report results from a computer simulation exploration of the effects of heritability (H) and gene number (n) on the efficiency of the backcross method.

DETAILS OF THE INVESTIGATION

Backcross systems. In the case of all recurrent parent alleles that have no effect on the selection criterion and are not linked with segregating genes that do affect the selection criterion, expected frequency after t backcrosses is $[1 - (1/2)^{t+1}]$. When $t = 5$ this is 0.984 which indicates that almost all of the desired portion of the recurrent parent genotype would be automatically retrieved during five backcrosses. With this in mind our study was focused on programs involving a total of five backcrosses.

The several backcrosses of a program can be distributed in various ways relative to the total of selection practiced. Let B symbolize a series of backcrosses with no intervening selection and S symbolize a selection scheme that may involve both selection among individual plants and selection among families produced by self-fertilization. We use subscripts to indicate position in a sequence and superscripts to indicate number of backcrosses in a series. Thus $B^2_1S_1B^3_2S_2$ specifies a program involving two successive backcrosses without intervening selection followed by selection according to scheme 1 and then three further backcrosses followed by selection according to scheme 2 (which may or may not be the same as scheme 1).

This study considered three types of program: B^5_1S, $B^2_1S_1B^3_2S_2$ and $B^1_1S_1B^2_2S_2B^2_3S_3$, with $S_1 = S_2$ in the second type and $S_1 = S_2 = S_3$ in the third. Our actual simulations were restricted to the $B^5_1S_1$ program and the $B^2_1S_1$ and the $B^1_1S_1$ segments of the other two programs. These will be denoted as the B^5, B^2 and B^1 *systems,* respectively.

The computer program simulated production of each backcross individual by *single seed descent.* For example, in the B^2 system the program simulated production of each second backcross plant from a different first backcross parent.

The term, *backcross population,* will always refer to plants produced by the final backcross of a series. Thus the backcross population of the B^5 system consists of the plants produced by the fifth backcross.

Selection scheme. A single scheme, involving four cycles of selection and four plant generations, was employed. It began with selection, based on phenotype of the individual, of 500 from 1,000 plants of the backcross population (first backcross plants in the B^1 system, second backcross plants in the B^2 system or fifth backcross plants in the B^5 system). This was followed by three cycles of selection between and within families. Each cycle involved production by self-fertilization of a family from each selection of the preceding cycle, selection among families and then selection of one plant from each selected family. Approximately 12.6% of the families were selected in each of the

[1] Contribution from the Minnesota Agric. Exp. Stn., St. Paul, MN 55108. Journal paper No. 9390. Partof a thesis submitted by the senior author to the Univ. of Minnesota Graduate School in partial fulfillment of the Ph.D. degree requirements. Received 15 Jan. 1976.

[2] Research associate, International Crops Research Institute for the Semi-Arid Tropics, Hyderabad-500016, A.P., India, and professor of genetics, Univ. of Minnesota.

Table 1. Fractional increase in the genotypic mean resulting from transfer of all *n* favorable alleles from a donor line (assuming n/m = 1.0).†

n	H = 0.05	H = 0.15	H = 0.25
1	0.63 C‡	1.10 C	1.41 C
4	1.26 C	2.19 C	2.83 C
8	1.79 C	3.10 C	4.00 C
16	2.53 C	4.38 C	5.66 C

† Multiply listed value by $\sqrt{n/m}$ when $n/m < 1.0$.
‡ C = coefficient of variation among F_2 individuals divided by 100.

three cycles so that there was one selected family and from that one selected plant at the end of the scheme. Our simulation assumed that 100 individuals per family contributed to the family phenotypic mean and that individual plant data were obtained on 25 plants per family. Because of the selection among individual plants, this scheme would not be applicable to all traits of all economic species.

Genetic model. Genes of which the favorable allele is homozygous in the donor line and the less favorable allele is homozygous in the recipient line will be identified as *class I genes.* Those for which the reverse is true will be referred to as *class II genes.* With respect to class I genes we assumed equal and additive effects and independent assortment. Assumptions that applied to all genes were no mutation, no meiotic drive and homozygosity in both the donor and the recipient line.

Heritability, gene number and gene effects. Let $2u$ be the difference in genetic effect between the two homozygous genotypes in the case of class I genes. Then, given the additivity assumption of our model, the genetic variance contributed in F_2 by each of these genes is $u^2/_2$ and the total contributed by n of them is $nu^2/_2$. If this is taken to be n/m times the total additive genetic variance in F_2, the latter can be written

$$\sigma^2_g = mu^2/_2 \qquad [1]$$

Then, if the F_2 phenotypic variance (σ^2_p) is visualized in terms of the population mean and the coefficient of variation among individuals, we have

$$\sigma^2_p = C^2\overline{P}^2 \approx C^2\overline{Y}^2 \qquad [2]$$

where $\overline{P}$ = the phenotypic mean,
$\overline{Y}$ = the genotypic mean, and
C = the coefficient of variation divided by 100.

It follows that the heritability (H) of individual phenotype in the F_2 is

$$H = mu^2/2C^2\overline{Y}^2 \qquad [3]$$

$$\text{and } u = C\overline{Y}\sqrt{2H}/\sqrt{m} \qquad [4]$$

Finally, if we define p as the fraction by which the recipient line would be improved by incorporation of the better allele of n genes from the donor line, it is apparent that

$$p = 2un/\overline{Y} = nC\sqrt{8H/m} = C\sqrt{n/m}\ \sqrt{8nH} \qquad [5]$$

This expression enables us to identify combinations of H, n and n/m that are worthy of attention. Some values of p are shown in Table 1. They indicate that, in terms of p (potential for improvement), even very low values of H deserve attention. For example, given n = 4, n/m = 1/4 and H = 0.05 we find that p = 0.63C which is 0.10 if C = 0.16. We concluded that heritabilities as low as 0.05 deserved attention and actually investigated the range from 0.05 to 0.45 and values of n from 1 to 16.

Details of the simulation program. Digital computer programming for simulation of genetic events has been discussed by Fraser and Burnell (1970) and various specifics of the computer program employed in the work being reported were outlined by Reddy (1974). We will confine ourselves here to stating that we tested our program in several ways and are confident that it was doing what we intended and what herein it is described as doing. There are, however, some aspects of what our program was written to do that should be clarified.

A phenotypic value, or mean value in the case of families, provided the basis for selection. Each of these was the sum of effects of genes simulated and a random value employed to reflect the sum of all other effects on phenotype. Symbolically,

Table 2. Effects of gene number (n) and heritability on probability of fixation.

Backcross system	n	Heritability 0.05	0.15	0.25	0.35	0.45
B^5	1	0.59 (0.12)†	1.00 (0.00)	1.00 (0.00)		
	2	0.25 (0.06)	0.58 (0.04)	0.62 (0.06)	0.75 (0.06)	0.66 (0.06)
B^2	2	0.92 (0.04)	0.97 (0.03)	1.00 (0.00)		
	3	0.61 (0.07)	0.89 (0.04)	0.95 (0.03)	0.95 (0.03)	1.00 (0.00)
	4	0.65 (0.05)	0.77 (0.03)	0.86 (0.04)	0.87 (0.03)	0.92 (0.03)
	6	0.40 (0.04)	0.67 (0.03)	0.68 (0.03)	0.73 (0.02)	0.79 (0.03)
B^1	4	0.67 (0.07)	0.92 (0.03)	0.94 (0.03)	0.99 (0.01)	1.00 (0.00)
	6	0.60 (0.05)	0.84 (0.04)	0.84 (0.02)	0.87 (0.03)	0.96 (0.02)
	8	0.54 (0.03)	0.71 (0.02)	0.78 (0.03)	0.82 (0.02)	0.84 (0.02)
	12	0.51 (0.04)	0.64 (0.02)	0.68 (0.02)	0.71 (0.02)	0.79 (0.02)
	16	0.49 (0.03)	0.59 (0.03)	0.62 (0.02)	0.68 (0.02)	0.69 (0.02)

† Figures in parentheses are standard errors of the estimates. These were computed from variation among frequencies observed in different replications.

$$P = y + r \qquad [6]$$

where P = phenotypic value (or mean phenotypic value),
y = total genetic effect of the simulated genes, and
r = a random quantity.

The variance of r is clearly critical. It should properly reflect heritability, family size, variance of plot effects, variance of genotype-environment interaction effects, etc.

Of the genes affecting the trait under selection, only those of class I were simulated. Each favorable allele of each of these genes was assigned unit value and the value, y, for each individual was determined by counting the number of favorable alleles. This means that we set $u = 1.0$. As an example, if we use *1* and *0* to symbolize more favorable and less favorable alleles, respectively, and if the number (n) of class I genes is 3.0, the y-value for the 101/000 genotype is 2.0.

Consider first the variance of r in the case of single plant phenotypic values when n/m = 1.0; i.e., when all segregating genes that affect the selected trait belong to class I. Then r is completely nongenetic and its variance in F_2 must be

$$\sigma^2_e = (1 - H)\sigma^2_p = \qquad [7]$$
$$(1 - H)\sigma^2_p\sigma^2_g/\sigma^2_g = (1 - H)nu^2/2H$$

This variance was assumed the same in all backcross generations as in the F_2. Each value of r was obtained as the product of σ_e and a random value (R) drawn from a normal distribution with mean equal to zero and variance equal to 1.0 so that the phenotypic value for the i-th individual was

$$P_i = y_i + R_i\sqrt{(1 - H)nu^2/2H} \qquad [8]$$

In the case of family means when n/m = 1, the variance around the genotypic value of the parent was assumed to be

$$\sigma^2_r = E + (\sigma^2_e + \sigma^2_{gw})/a \qquad [9]$$

where E = the sum of family × macroenvironment interaction variance and variance due to plot effects,
a = number of individuals per family, and
$\sigma^2_{gw} = u^2n'/2$ = genetic variance among individuals within families where $n' \leq n$ is number of class I genes heterozygous in the parent.

The phenotypic means for families were obtained as follows:

$$\overline{P}_i = y_i + R_i\sigma_r \qquad [10]$$

where y_i = the value of y for the parent of the i-th family, and R is used as described for equation [8].

In all work being reported E was set equal to $0.2\sigma^2_e$. The rationale for this value and effects of varying it will be discussed later.

All parameter sets employed in computer simulation runs assumed n/m = 1.0 but it will be shown later that the data obtained can be interpreted with reference to other values of n/m.

RESULTS AND DISCUSSION

Data on the frequencies of favorable alleles are summarized in Table 2. Each value shown is based on 16 replicate simulation runs and is the average

($\overline{Q}$) over replications and loci (class I) of the frequencies of favorable alleles in final selections. Symbolically,

$$\overline{Q} = \sum_{i=1}^{s} x_i/2ns \qquad [11]$$

where x_i = number of favorable alleles in the genotype of the selection in the i-th replication, s = number of replications, and n = number o^c class I genes. $\overline{Q}$ consitiutes an estimate of the probability, P(A), of fixation of favorable alleles in genotypes obtained via the backcross system in question. The general pattern of the data shown in Table 2 is unambiguous, and its implications can be summarized as follows:

1. P(A) is greater when H is greater.

2. Within backcross systems and at all values of H, P(A) is smaller when n is greater.

3. When n and H are the same, P(A) is larger for B^2 than for B^5 and for B^1 than for B^2.

All of this was expected but quantitative relationships were not so readily apparent in advance.

Given perfect heritability (H = 1), the individual with the best genotype always has the best phenotype and can therefore be identified for selection. As the nongenetic standard deviation, σ_e, becomes large relative to the positive effect of a single favorable allele the correlation between phenotypic and genotypic values decreases, and hence, the chance that the individuals with the best genotypes will be those selected also decreases. It is apparent from equation [7] that in our simulation the ratio of the plus effect of a favorable allele to σ_e was $\sqrt{2H/n(1 - H)}$ which is 3.94 times as large when H = 0.45 as when H = 0.05.

The primary effect of n has to do with the probability of the required genotype in the backcross population. Successful transfer of n favorable alleles from the donor to the recipient line requires occurrence of the n-fold heterozygote in the backcross population from which selection is practiced. Let q_t be the probability of that genotype in a system involving t backcrosses before selection is practiced. Then $q_t = (1/2)^{tn}$ and the probability that there will be one or more n-fold heterozygotes among k individuals is $1 - (1 - q_t)^k$. Since k = 1,000 was used in our simulations, some values pertinent to the data of Table 2 are:

tn	5	8	10	12	16
$1 - (1 - q_t)^{1,000}$	1.0	0.98	0.62	0.22	0.015

These make it readily apparent that complete success is not to be expected, even with high heritability, with B^5 when n $\geqslant$ 2, with B^2 when n $\geqslant$ 5 or with B^1 when n $\geqslant$ 10.

Interpretation for cases where n/m < 1.0. As noted earlier, all of our data were obtained assuming n/m = 1.0; i.e., that all genes of which different alleles were present in the donor and recipient lines were class I genes. We will now show that information pertinent to other values of n/m are provided by the data shown in Table 2.

In the case of selection among backcross individuals we used

$$\sigma^2_r = (1 - H)nu^2/2H \qquad [7a]$$

which assumes that all genetic variance in the F_2 generation would have been contributed by class I genes. Given n/m < 1.0, the variance of r should have been partly due to class II genes and, assuming additive effects of those genes, should have been

$$\sigma^2_r = \frac{(1 - H)\, mu^2}{2 \quad H} + [(\frac{1}{2})^{t-1} - (\frac{1}{2})^{2t-1}]\,(m - n)\frac{u^2}{2} \qquad [12]$$

We may now ask what value of H used in (12) with a value of n/m < 1.0 would yield the same value of σ^2_r as any specified value of H used in [7a]. Distinguish the value of heritability in [12] by $\overline{H}$ instead of H. Then equating [7a] and [12] and solving for $\overline{H}$, we obtain

$$\overline{H}_t = H/[\frac{n}{m} + H(1 - \frac{n}{m})\{1 - (\frac{1}{2})^{t-1} + (\frac{1}{2})^{2t-1}\}] \qquad [13]$$

The extremes of this expression occur when t = 1 and when t is large. When t = 1,

$$\overline{H}_1 = H/[\frac{n}{m} + \frac{H}{2}(1 - \frac{n}{m})] \qquad [14]$$

and when t = 5,

$$\overline{H}_5 \approx H/[\frac{n}{m} + H(1 - \frac{n}{m})] \qquad [15]$$

It is not difficult to show that these results apply also in the case of within family selection among individuals.

In the case of selection among families we used equation [9]

where $E = 0.2\sigma^2_e$ and $\sigma^2_e = (1 - H)nu^2/2H$. The value of σ^2_{gw} varies but could never exceed $nu^2/2$. Substituting this maximum and a = 100 (used in all simulation runs),

$$\sigma^2_r = 0.21(1 - H)nu^2/2H + 0.01(nu^2/2) \qquad [16]$$

The second term (σ^2_{gw}/a) is trival relative to the first and can be ignored. Assuming n/m < 1.0 and again that E equals 0.2 of the nongenetic variance, it would be approximate to employ

$$\sigma^2_r = 0.21(1 - H)mu^2/2H + \sigma^2_{gw}/a \qquad [17]$$

where the second term would be slightly larger than when n/m = 1.0 but still trivial relative to the first. Proceeding roughly as before, we can distinguish the heritability in [17] by $\overline{H}$ instead of H, equate the significant portions of [16] and [17] and solve for $\overline{H}$ to obtain the value of H which, used in [17] with a value of n/m < 1.0, would yield the same

value of σ^2_r as any specified value of H used in [16]. We obtain

$$\overline{H} = H/[\frac{n}{m} + H(1 - \frac{n}{m})], \qquad [18]$$

the value given by [15] for selection among individuals when t is large. Values of $\overline{H}_1$ and $\overline{H}_5$ obtained using equations [14] and [15] are listed below for n/m = 1/2 and 1/4.

n/m	H	$\overline{H}_1$	$\overline{H}_5$
1/2	0 05	0.098	0.095
	0.15	0.28	0.26
	0.25	0.44	0.40
	0.35	0.60	0.52
	0.45	0.75	0.62
1/4	0.05	0.19	0.17
	0.15	0.49	0.41
	0.25	0.73	0.57
	0.35	0.91	0.68
	0.45		0.77

The following kind of information is provided. The data (listed in Table 2) obtained for H = 0.05 and n/m = 1.0 are also applicable to the situation in which n/m = 1/2 and H $\approx$ 0.095 and to the situation where n/m = 1/4 and H $\approx$ 0.17. As another example, P(A) for n/m = 1/2 and H = 0.40 is approximately that for n/m $\approx$ 1 and H = 0.25 but P(A) for n/m = 1/4 and H = 0.40 is roughly that for n/m = 1 and H = 0.15.

The data in relation to complete programs. We must now reemphasize that the B^2 and B^1 systems cannot be viewed as complete backcross programs. Neither of them includes enough backcrosses to make the probability of fixation of favorable alleles at class II loci as large as required.

In contrast to the B^2 system, the 2 segment program $B^2{}_1S_1B^3{}_2S_2$, includes five backcrosses so that the probability of fixation of favorable alleles at class II loci is 0.984. We have an estimate of P (A) for B^2 but require an estimate $\hat{P}(A)_p$ of the corresponding probability of fixation for the $B^2{}_1S_1B^3{}_2S_2$ program. Consider a single class I gene and let

P (A2) = probability of fixation of the favorable allele at the end of the program given that it was homozygous in the end product of the first segment of the program, and

P (A1) = probability of fixation of the favorable allele at the end of the program given that it was heterozygous in the end product of the first segment of the program.

Then $P(A)_p = f_2P(A2) + f_1P(A1)$ [19]

where f_2 and f_1 = probabilities that the favorable allele will be homozygous and heterozygous, respectively, in the end product of the first segment of the program, and

$f_2 + f_1/2 = P(A)$.

Rearranging, $f_2 = P(A) - f_1/2$, and substituting in [19] we obtain

$$P(A)_p = P(A)P(A2) + f_1\,[P(A1) - \frac{1}{2}\,P(A2)] \qquad [20]$$

The absolute bounds of the final term of (20) are $\pm f^1/2\ P(A2)$. Hence the magnitude of $f_1/2\ P(A) = f_1/(2f_2 + f_1)$ deserves attention. Because the individuals in the final cycle of the selection scheme are produced by three successive self-fertilizations, this ratio would be 1/8 in the absence of any selection but the totality of selection practiced would make it much smaller. Clearly the final term of [20] is trivial relative to P (A)P (2A) so that for practical purposes

$$P(A)_p = P(A)P(A2). \qquad [21]$$

Note now that if the favorable allele is homozygous in the end product of the $B^2{}_1S_1$ segment of the program, the first backcross of the $B^3{}_2S_2$ segment will yield only heterozygous genotypes (as in the original F_1. Therefore the probability of heterozygous individuals will be 1/4 in the backcross population of the second segment as in the backcross population of the first segment. For that reason, and given the same nongenetic variance in all generations, it is clear that P (A) and P (A2) will be equal if genetic variances (from other loci) are equal and that P (A2) > P (A) if genetic variance is less in the second segment of the program. The latter will be so for genetic variance from class II genes because, in the second segment, there will always have been more backcrosses to the recurrent parent (those of the first segment in addition to those of the second). It will also be so for genetic variance from other class I genes if P (A) < 1.0 and $n > 1.0$. The importance of reduced genetic variance from other class I genes is indicated in Table 2 by the increases seen in $\overline{Q}$ as n decreases. In summary P (A2) = P (A) when n/m = 1.0 and P (A) = 1.0 or n = 1.0. Otherwise P (A2) > P (A). Thus

$$[P(A)]^2 \leqslant P(A)_p \leqslant P(A) \qquad [22]$$

$P(A)_p$ will of course be less than P (A) unless P (A2) = 1.0.

By arguments parallel to those detailed above it can be shown that for the $B^1{}_1S_1B^2{}_2S_2B^2{}_3S_2$ program

$$[P(A)]^3 \leqslant P(A)_p \leqslant P(A) \qquad [23]$$

Potential of backcross program. The information from this study can be interpreted in terms of (a) the genetic situations in which the program should be completely successful in the sense of transferring from donor to recipient all of the favorable alleles available in the donor but lacking in the recipient line or (b) the amount of genetic improvement in the recipient line whether all available favorable alleles are or are not transferred.

The probability, P (n), that each of n favorable alleles will be fixed in the product of a backcross program will be $[P(A)_p]^n$ if the n genes segregate independently. This, together with equations [22] and [23] and the fact that in the case of the B^5S program $P(A)_p = P(A)$, provides the basis for using the data

Table 3. Lowest heritability ($\hat{H}$) for which the estimate of the probability of fixation of the favorable alleles of all class I genes is equal to or greater than 0.50.†

Backcross program	n/m	n	H
B^5S	1	1	0.05
	1/2	1	0.10
$B^2_1S_1B^3_2S_2$	1	2	0.05
		3	0.15
		4	0.45
	1/2	2	0.10
		3	0.25
$B^1_1S_1B^2_2S_2B^2_3S_3$	1	4	0.25
		6	0.45
	1/2	4	0.40

† Results when n/m = 1/2 were obtained using information tabulated in the section dealing with n/m $<$ 1.0.

Table 4. Estimates of improvement in genotypic mean of the recipient line through use of the backcross programs when H = 0.25.

Program	n	n/m = 1	n/m = 1/2
B^5S	1	1.4 C†	1.0 C
	2	1.2 C	0.8 C
$B^2_1S_1B^3_2S_2$	2	2.0 C	1.3 C
	3	2.2 C	1.4 C
	4	2.1 C	1.2 C
	6	1.6 C	1.1 C
$B^1_1S_1B^2_2S_2B^2_3S_3$	4	2.4 C	1.6 C
	6	2.0 C	1.5 C
	8	1.9 C	1.0 C
	12	1.5 C	0.9 C
	16	1.3 C	0.8 C

† C = coefficient of variation among F_2 individuals ÷ 100.

in Table 2 to obtain estimates of P (n). Table 3 shows, for various combinations of n/m and *n*, the lowest heritability at which $\hat{P}$ (n) $\geq$ 0.50. If p, given by equation [5], is multiplied by $P(A)_p$, the expected improvement of the recipient line, as a fraction of the mean in the F_2 of the donor and recipient lines, is approximated. Values shown in Table 4 were obtained using the estimates of $P(A)_p$ provided by the lower limits of equations [22] and [23].

Table 3 indicates that the B^5S program is effective for transfer of one good allele but not for simultaneous transfer of two or more. Up to three favorable alleles available in a donor line can be transferred simultaneously in the $B^2_1S_1B^3_2S_2$ program and up to four (or five) using the $B^1_1S_1B^2_2S_2B^2_3S^3$ program. Table 4 indicates the potential for considerable improvement in a quantitative trait even when a large proportion of favorable alleles available cannot all be transferred together. Both tables indicate that more can be accomplished with the three-segment program than with the two-segment program and more with the latter than with the B^5S program. Of course the three-segment program requires the most effort and the B^5S program the least. Not to be lost sight of is the obvious fact that if all favorable alleles available from a donor line are not transferred, a closer approach to complete success can be achieved in a following program using the same donor but the product of the first program as the recipient.

In all cases effectiveness is less when n/m is less. This is primarily due to lower values of the ratio (u/σ_e), given the same heritability and number of class I genes, and only secondarily to presence of genetic variance arising from class II genes.

It is well known that genetic linkage can reduce the realized improvement in backcross programs from the potential, given independent assortment of class I genes from others affecting total value of genotype, indicated by our data. For discussion of this matter see, for example, Allard (1960). As noted by Allard, linkage of favorable alleles in the donor with unfavorable alleles of genes that affect total value but not the trait that is the basis for selection in the backcross program can prevent complete reconstitution of the valuable portion of the recipient parent. In such cases the actual selection criterion employed becomes critical. The value of evidence concerning genetic covariances in the F_2 as a guide in this connection could be investigated by the simulation procedure.

Magnitude of E. The choice of $0.2\sigma^2_e$ as the value substituted for E in the phenotypic variance of family means can only be defended in a general way. E was defined to include variance from plot effects and from family × macroenvironment interaction effects and will clearly be variable in relative magnitude depending on the trait or traits for which selection is practiced and the design of the field comparisons on which selection among families is based. Our choice reflected recognition that E will always be substantially lower than σ^2_e but that family × macroenvironment interaction variance is known to be substantial in the case of some of the quantitative traits important to the breeder.

In retrospect we think that the interaction portion of E should have been geared to the magnitude of genetic variance rather than nongenetic variance; there can be no interaction when there is no genetic variance, and the greatest amount of interaction variance is logically to be expected when genetic variance is greatest.

Making E larger or smaller has an effect on outcome; P (A) is inversely related to E but the impact is minor except when H is very small. This issue will be treated in greater detail in a later manuscript.

Variations in the selection scheme. It is readily apparent that the selection exerted in a backcross program can be varied in many ways. Some deviations from the scheme employed in work here reported are mandated in work with some traits in some species. For example, in the case of crops (e.g., small grains) where minimal spacing between plants is normal in commercial practice, selection among individual plants will be counter-productive if the trait is one for which genotype × plant density interaction is important.

Data concerning selection scheme variations will be reported subsequently.

REFERENCES

Allard, R. W. 1960. Principles of Plant Breeding. John Wiley and Sons, Inc., New York.

Bailey, T. B. 1972. Utilization of source populations in producing superior varieties in self-fertilizing species. Ph.D. thesis, Univ. of Minnesota.

———, and R. E. Comstock. 1976. Linkage and the synthesis of better genotypes in self-fertilizing species. Crop Sci. 16:363-370.

Bliss, F. A., and C. E. Gates. 1968. Directional selection in simulated populations of self-pollinated plants. Aust. J. Biol. Sci. 21:705-719.

Briggle, L. W. 1969. Near-isogenic lines of wheat with genes for resistance to *Erysiphe graminis* f. sp. *tritici*. Crop Sci. 9:70-72.

Briggs, F. N., and P. F. Knowles. 1967. Introduction to plant breeding. Reinhold Publishing Corporation, New York.

Duvick, D. N. 1974. Continuous backcrossing to transfer prolificacy to a single-eared inbred line of maize. Crop Sci. 14:69-71.

Fraser, A., and D. Burnell. 1970. Computer models in genetics. McGraw-Hill Book Co., New York.

Knott, D. R., and B. Talukdar. 1971. Increasing seed weight in wheat and its effects on yield, yield components and quality. Crop Sci. 11:280-283.

Nagai, K., H. Fujimaki, and M. Yokoo. 1973. Breeding and some genetical findings for multiracial resistance to blast by using *Japannica* × *Indica* hybrids of rice. Japan Agric. Res. Q. 7: 63-70.

Reddy, Belum V. S. 1974. Evaluation of the backcross breeding method. Ph.D. thesis, Univ. of Minnesota.

Rinke, E. H. 1960. Reaching for yield. Proc. 15th Annual Hybrid Corn Industry-Research Conference, p. 61-67.