Theor. Appl. Genet. 51, 133-137 (1977)

The Use Marker Alleles for the Introgression of Linked Quantitative Alleles

M. Soller and J. Plotkin-Hazan
Department of Genetics, The Hebrew University, Jerusalem (Israel)

Summary. It is shown that when an exotic strain and a commercial strain differ genetically at a quantitative locus and at an adjoining marker locus, repeated backcrosses to the commercial strain, retaining only backcross progeny carrying the exotic marker allele, will allow the effective introgression of the linked quantitative allele from the exotic to the commercial strain. The introgression procedure will be particularly effective when exotic and commercial strains differ at two nearby marker loci with the quantitative locus bracketed between them. The simultaneous introgression of a number of quantitative alleles from different exotic strains, and appropriate selection procedures in the intercross generations that follow are also considered.

Key words: Introgression - Quantitative-locus - Marker-locus - Linkage - Gene-pool

There is considerable interest in the preservation of exotic plant and animal gene pools as a store of potentially valuable genetic variation. Soller, Genizi and Brody (1976) have shown that in crosses between inbred lines differing at one or more marker loci, relatively modest experiments can enable the quantitative value of chromosomal segments adjoining the differentiating markers to be determined. Similar procedures can readily be adapted to characterize the quantitative value of marked chromosomal segments of exotic strains, and in this way identify chromosomal segments of potential usefulness for the improvement of quantitative traits in commercial strains. In this paper we show that when exotic and commercial strains differ in their alleles at a quantitative locus and at an adjoining marker locus, repeated backcrosses to the commercial strain, choosing all the while only those backcross progeny carrying the exotic marker allele, will allow the effective introgression of the linked quantitative allele from the exotic into the commercial strain.

Theory

We initially assume exotic and commercial strains, differentiated by alleles at a quantitative locus and at a nearby marker locus. Following Jayakar (1970) we will let A_1 and A_2 denote the alleles at the locus affecting the quantitative trait. The frequency of these alleles in backcross and intercross generations will be denoted a_1 and a_2, respectively. Within each of the three genotype classes A_1A_1, A_1A_2 and A_2A_2 the quantitative character, Y, is assumed to be normally distributed with variance σ^2 and means d, h and -d, respectively. M_1 and M_2 with frequencies m_1 and m_2, respectively, will denote the alleles at the marker locus. It is assumed that the three marker genotypes have distinguishable phenotypes. The probability of recombination between marker locus and quantitative locus will be denoted r_{MA}. Exotic and commercial strains will be arbitrarily assigned the genotypes M_1A_1/M_1A_1 and M_2A_2/M_2A_2, respectively. In some cases the exotic and commercial strains may also be differentiated at a second marker locus, S, having a probability of recombination r_{MS} with the first marker, and bracketing the quantitative locus A, between them. In subsequent sections r_{MA} and r_{MS} may also be treated as if they represent map distances, in Morgans, between the loci. The alleles at S will be denoted S_1 and S_2, with frequencies s_1 and s_2, respectively, and exotic and commercial strains will be assigned the genotypes $M_1A_1S_1/M_1A_1S_1$ and $M_2A_2S_2/M_2A_2S_2$, respectively.

In the absence of a screen for the favorable quantitative allele, A_1, the frequency of the allele in the

Table 1. The frequency of a favorable allele after k backcross generations in the absence of selection for a linked marker, and after k backcross and one intercross generations with selection for a linked marker or marker bracket

k	No marker	Single marker[a]			Marker bracket[b]	
		$r_{MA}=0.1$	$r_{MA}=0.2$	$r_{MA}=0.3$	$r_{MS}=0.2$	$r_{MS}=0.4$
1	0.25	0.81	0.64	0.49	0.98	0.92
2	0.12	0.73	0.51	0.34	0.97	0.88
3	0.06	0.66	0.41	0.24	0.96	0.85
4	0.03	0.59	0.33	0.17	0.95	0.82
5	0.02	0.53	0.26	0.12	0.94	0.78
6	0.01	0.48	0.21	0.08	0.93	0.75

[a] r_{MA} is the probability of recombination between the favorable allele and the marker.
[b] r_{MS} is the probability of recombination between the two markers of the bracket

k'th backcross generation will be $(1/2)^{k+1}$. If only backcross or intercross progeny carrying the linked marker allele M_1 are chosen to continue the backcrossing procedure or to start the synthetic population, it can readily be shown that the proportion of M_1 chromosomes at the k'th backcross generation that still carry the favorable allele, A_1, will be $(1 - r_{MA})^{k+1}$. If the favorable allele is bracketed between two marker loci, M_1 and S_1, then even in the worst case (quantitative allele equidistant from the marker loci) the proportion of M_1S_1 chromosomes at the k'th generation that still carry the favorable A_1 allele will be $[1 - (1/4)r_{MS}^2]^k$, and the frequency of A_1 among M_1S_1/M_1S_1 progeny of the first intercross will be $[1 - (1/4)\, r_{MS}^2]^{k+1}$.

If the M_1M_2 backcross progeny are subjected to selection for the quantitative trait affected by A_1 during the backcrossing procedure, the frequency of A_1 among M_1 marker chromosomes will be increased somewhat. Following the method of Falconer (1960, p. 205) it is easy to show that the regression of a_1 on phenotype (b_{a_1P}) among the M_1M_2 backcross progeny will be

$$b_{a_1P} = \frac{a_1a_2(d + h)}{\sigma^2}$$

and the increase in gene frequency among the selected backcross progeny, Δa_1, will be

$$\Delta a_1 = \frac{a_1a_2(d + h)}{\sigma^2}\, i\sigma ,$$

where i is the standardized selection differential. More accurate discrimination of M_1A_1 genotypes can be had by progeny testing the M_1M_2 backcross progeny, and selecting those having the highest quantitative value. Progeny testing will probably be most economically carried out by crossing the backcross progeny to the commercial strain. In this way the progeny test offspring also form the next backcross generation. In this case the regression of a_1 on the mean of n offspring, $b_{a_1\overline{O}_n}$, again using Falconer's approximation, will be

$$b_{a_1\overline{O}_n} = \frac{\frac{1}{2}a_1a_2(d + h)}{\sigma^2/n} ,$$

and the increase in gene frequency in this case among the selected backcross progeny, $\Delta' a_1$, will be

$$\Delta' a_1 = \frac{\frac{n}{2}a_1a_2(d + h)}{\sigma^2}\, i\sigma/\sqrt{n} = \Delta a_1\left(\frac{\sqrt{n}}{2}\right) .$$

The effects of selection on gene frequencies in the synthetic population derived from the M_1M_1 intercross progeny are the same as above, with $\alpha = d + (a_2 - a_1)h$ in place of $(d + h)$. It should be noted that the effects of selection in any one of the backcross generations will be dissipated to a greater or lesser extent in subsequent generations, depending on r and on the number of generations remaining until the founding of the initial synthetic population.

Table 2. The effect of mass selection during the backcross or intercross generations on the frequency of a favorable allele as a function of allele frequency (a_1) and the proportional effect of the locus (d/σ) assuming no dominance

Initial allele frequency (a_1)	Proportional effect (d/σ)		
	0.1	0.2	0.3
0.5	0.025	0.050	0.075
0.8	0.016	0.032	0.048

Results

Table 1 shows the frequency of a favorable allele A_1 after k generations of backcrossing. It can be anticipated that five or six generations of backcrossing will probably be necessary in order to bring the synthetic population to the level of the original commercial strain. In this case the frequency of the favorable allele in the synthetic population will be very low (~0.01) At this frequency selection is relatively ineffective in increasing gene frequencies, so that it will require many generations of selection to bring the frequency of A_1 to the point where it will make a significant contribution to the mean of the synthetic population.

Table 1 also shows the frequency of a favorable allele, A_1, linked to a marker allele, M_1, or brackted between two marker alleles, $M_1 - S_1$, after k generations of backcrossing in which only backcross progeny carrying the linked marker or the marker bracket are chosen to continue the backcrossing program. For single marker selection, the frequency of the favorable allele in the synthetic population, even after five or six backcrosses, will generally be greater than 0.2. At this frequency selection can be quite effective in increasing gene frequencies, so that subsequent selection should quite rapidly enable the new favorable allele to have a significant effect on the population mean. The efficiency of a marker bracket, in maintaining high frequencies of the bracketed favorable allele over a series of backcrosses, is astonishing. For example, a 40 cM bracket will bring an enclosed favorable allele across five generations of backcrossing and one intercross to a frequency of 0.78 in the M_1S_1/M_1S_1 progeny of the first intercross. At this frequency the introgressed favorable allele will already have a significant impact on the mean of the synthetic population.

Table 2 shows values of Δa_1 for mass selection, on the assumption h = 0, as a function of a_1 and d/σ. Progeny testing with n = 20 would be 2.2 times as effective as mass selection in maintaining the frequency of the favorable allele during the backcrossing program and increasing it thereafter. Progeny testing with n = 50 would be 3.5 times as effective as mass selection. For example, with $a_1 = 0.5$, and $d/\sigma = 0.1$, a_1 for mass selection will be 0.02, while with progeny testing at n = 50 it will be about 0.07. Assuming five backcross steps and one intercross, the final gene frequency with selection for the case $r_{MA} = 0.2$ will be about 0.30 for mass selection and 0.43 for progeny testing. The latter figure is a respectable improvement over the value of 0.26 expected without selection. In general the values of Table 2 indicate that for the values of d/σ likely to be obtained in practice, progeny testing is a necessity if a rapid increase in the frequency of the introgressed allele in the synthetic population is to be attained. It should be pointed out that the economic value of a quantitative allele having effects of the magnitude postulated ($d/\sigma = 0.1$) are by no means negligible. Considering egg production, for example, such an allele might increase production by 5 eggs/bird/year. This would enable the same total egg production to be obtained with 2 % fewer birds. Saving in feed would be commensurate.

Discussion

The results of this analysis suggest that a complete introgression procedure, utilizing a linked marker, but not a bracket, would take about 10 to 12 generations, namely: 6 to 7 generations for the F - 1, backcrosses and intercross, and 4 to 5 generations of subsequent selection to increase the frequency of the introgressed allele. A bracket would shorten the process by 3 to 4 generations. Thus introgression procecedures of the short described would be feasible for most plants and some animals.

Multiple introgression of a number of quantitative alleles simultaneously into the same commercial strain could be carried out in two ways: (1) independent introgression of each favorable allele into a different subline of the commercial strain, followed by intercrossing of the various sublines to produce a synthetic population containing all of the favorable alleles, or (2) intercrossing the exotic strains to produce a syn-

thetic strain containing all of the favored alleles, followed by introgression of the alleles en bloc into the commercial strain, each favorable allele monitored by its own marker. In the latter procedure only $(1/2)^t$ of the progeny in each backcross generation will have all favorable markers, where t is the number of marked segments being introgressed, and in the case of introgressed brackets only $[1/2\cdot(1-r_{MS})]^t$ will have all favorable brackets intact. Thus this procedure will be suitable only for small t, or plants and animals with high fecundity. In both cases the additional intercrossings required in order to produce the synthetic strains will attenuate the frequency of favorable alleles among the marked chromosomes to the same degree as additional backcrosses. For example, for $r_{MA}=0.2$, simultaneous introgression of four quantitative alleles (requiring two additional intercrosses) would reduce the frequency of favorable alleles among marked chromosomes from 0.26 to 0.17, and in the case of brackets with $r_{MS} = 0.4$, the frequency of favorable alleles among bracketed segments would be reduced from 0.78 to 0.72. In neither instance would the reduction be critical to the success of the introgression program, and for genes at these frequencies the difference could be made up by an additional generation or two of selection. In contrast, the effectiveness of selection in increasing further the frequency of alleles that are already at a high frequency (i.e. the alleles of the commercial strain at other quantitative loci) is poor compared to the effectiveness of backcrossing. Thus, it would not appear to be useful to reduce the number of backcross generations in order to increase the frequency of introgressed alleles in the initial synthetic population.

In addition to the favorable quantitative alleles being introgressed, the chromosomal segment carrying a marker or included within a bracket may contain one or more unfavorable alleles affecting the trait in question, or others. The frequency of the unfavorable allele in the marked or bracketed chromosomal segment will also decrease in the course of the backcrossing and intercrossing procedure. Once the synthetic population is formed and selection begins, the fate of the two alleles, one favorable the other unfavorable will depend on the tightness of the linkage between them and on the degree of linkage disequilibrium in the initial synthetic population. The degree of linkage disequlibrium in turn, will depend on the relative locations of favorable allele, unfavorable allele and marker. If, relative to the marker, the unfavorable allele (say, b) is distal to the favorable allele, then the proportion of those chromosomes in the first intercross generation carrying the favorable allele that also carry the unfavorable allele will be $(1-r_{AB})^{k+1}$, where r_{AB} is the probability of recombination between favorable and unfavorable alleles, e.g. for $r_{AB}=0.1$, and $k = 5$, this proportion equals 0.53. If the unfavorable allele is found between favorable allele and marker, the corresponding proportion will be $(1-r_{AB}r_{MB})^{k+1}$, where r_{MB} is the probability of recombination between the unfavorable allele and the marker, e.g. if the unfavorable allele is midway between marker and favorable allele this proportion will equal 0.92 for the case $r_{MA} = 0.2$, $k = 5$. Clearly, if the two alleles enter the synthetic population in a state of linkage disequlibrium, their subsequent fate during the first few generations of selection will depend to some degree of mutual hitchhiker effects. A detailed analysis of this situation is beyond the scope of this paper, but will be published elsewhere (Soller, Marcus and Brody, 1977). In general, it will not place severe limitations on the usefulness of the introgression program.

A major limitation of the procedures described in this paper, however, is the requirement that exotic and commercial strains be differentiated with respect to the marker alleles adjoining the favorable quantitative locus. If either of the strains is segregating at the marker locus, then the introgression procedure could be carried out on suitable selected sublines. If both strains are homozygous for the same marker a convenient procedure is not readily apparent, particularly if linkage or marker and quantitative loci is tight. The ultimate solution in such cases will come as geneticists provide means for detecting increasing numbers of biochemical and other markers in economically useful plants and animals. Along the same lines, the striking effectiveness of marker brackets for the introgression of the enclosed chromosome segments provides an additional incentive to expand the number of polymorphic systems that can be recognized in commercially important species.

Acknowledgement

The authors thank T. Brody and Y. Weller for helpful discussions.

Literature

Falconer, D.L.: Introduction to Quantitative Genetics. Edinburgh: Oliver and Boyd 1960

Jayakar, S.D.: On the detection and estimation of linkage between a locus influencing a quantitative character and a marker locus. Biometrics 26, 451-464 (1970)

Soller, M.; Genizi, A.; Brody, T.: On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. Theor. Appl. Genet. 47, 35-39 (1976)

Soller, M.; Marcus, M.; Brody, T.: Mutual hitchhiker effects between favorable and unfavorable alleles. (To be published)

Received August 2, 1977
Communicated by R.W. Allard

M. Soller
J. Plotkin-Hazan
Dept. of Genetics
The Hebrew University
Jerusalem (Israel)