Theoretical and Applied Genetics 41, 306—311 (1971)

A Multiple Regression Analysis of Hybrid Vigour in Single Crosses of *Dactylis glomerata L.*

R. Knight

Waite Agricultural Research Institute, The University of Adelaide, (Australia)

Summary. Thirty-six crosses of *Dactylis glomerata* together with their parents were assessed for seasonal yield at twelve successive harvests during a 16 month period. The yields varied with the environment provided by the climatic seasons of the year. For some crosses, the hybrid approximated to its mid-parental value at one environment but exceeded both parents in another, showing hybrid vigour.

The twelve yield values for a hybrid and for each of its parents were analysed as a multiple regression in the form —

$$y_{ij} = a + b_i P_i + b_j P_j$$

where y_{ij}, P_i and P_j are the yields of the hybrid and parents at any one harvest, b_i and b_j are regression coefficients pertaining to the respective parents and a is a constant. It was found that such a regression accounted for a great part of the seasonal yield variation of each cross, including the situation of a hybrid approximating to the mid-parental value or exceeding both parents.

The b value pertaining to an individual parent varied according to the parent with which it was crossed. In the discussion the values of b and a were considered in relation to the incidence of hybrid vigour, and the use of a multiple regression is compared with other analyses of quantitative inheritance.

1. Introduction

In the majority of analyses undertaken in quantitative inheritance the estimated variances, covariances and effects are obtained as the means of several hybrid combinations. This is true of parent-offspring regressions, heritability estimates and diallel tests. It is rarely that a single cross, on its own, provides sufficient data for estimation of a quantitative parameter and there is therefore little information on such parameters. An estimate becomes feasible when two parents and their hybrid are grown in several different environments and their yields in these environments are obtained. This approach has been followed in recent studies of genotype-environment interactions (Bucio Alanis and Hill, 1966; Perkins and Jinks, 1968; Breese, 1969). It is the purpose of this paper to present results obtained on single crosses and analyses of each of them in the form of multiple regressions.

The experimental material was the pasture grass *Dactylis glomerata L.* Yield was the herbage harvested in the different seasons of the year and therefore in different environments. Yield was determined by cutting to near ground level and weighing the herbage that had grown since the previous cut. By repetition of this procedure many assessments of two parents and their hybrid were made in the environments experienced during a year.

Within *Dactylis glomerata* two major groups can be distinguished on a basis of differences in seasonal growth rhythm. One group, composed of populations from northern Europe, is, in its natural habitat, summer growing and winter dormant, whereas the other group, of Mediterranean origin is indigenously summer dormant and winter growing. In southern Australia, with its Mediterranean-type climate, both groups will grow with more or less vigour at all seasons of the year provided the summer drought is alleviated by irrigation (Knight, 1966) thus peımitting many assessments of yield.

The results to be presented are for parents from both groups, and hybrids within and between the groups when grown under irrigation.

2. Materials and Methods

Nine parents were crossed in all combinations to give, after bulking of reciprocals, 36 single cross F_1 families. Details of the nine parents are given by Knight (1966) but the only feature relevant to the present study is that seven (designated 1 to 7) of the nine parents were of Mediterranean origin and are predominantly winter growing, while two (designated 8 and 9) were from northern Europe and are predominantly summer growing. The parents were not inbreds; they were single individuals chosen at random from seven wild ecotypes of the Mediterranean group, or from two bred strains of the northern European group. Parents and hybrids were planted in a sward trial with a randomised block layout having four replicates. *Dactylis glomerata* is cross-pollinated so that to obtain many individuals of the parental genotypes it was necessary to vegetatively propagate the parents. In this way 256 offsets were obtained for each parent, providing 64 offsets for each of the four plots. Each plot was a metre square, the individuals offsets being at a 12.5 × 12.5 cm spacing. To be comparable, the F_1 hybrids were also vegetatively propagated into the trial, each family being represented by approximately 80 genotypes. The 80 genotypes were propagated to produce the required 256 offsets. The trial was planted between June 26 and 28, 1961. At each harvest, made with a mower cutting to 3 cm (1¼ in) above ground level, the herbage of all offsets in a plot was harvested together and a yield of dry matter obtained. The yields from 12 harvests are to be considered.

NPK fertiliser was applied liberally before planting and thereafter at 3 monthly intervals. During the summer months (November-March inclusive) the trial was watered with a precision irrigator.

3. Multiple regression analysis

For each hybrid and its two parents a multiple regression was fitted to 12 sets of values corresponding to 12 harvests. Each set consisted of a hybrid yield, y, and the yield of its parents i and j, namely x_i and x_j. The regression was

$$y_{ij} = a + b_i x_i + b_j x_j \quad (1)$$

or using deviations

$$y_{ij} - \bar{y} = b_i (x_i - \bar{x}_i) + b_j (x_j - \bar{x}_j) \quad (2)$$

where $\bar{y}$, $\bar{x}_i$ and $\bar{x}_j$ are the mean yields of the hybrid and parents over the whole period of the experiment.

Fitting of a multiple regression to the data of a hybrid and its two parents was found invariably to account for a great deal of the variation, but it was evident that this was sometimes due to the correlated behaviour of the two independent variables x_i and x_j, for although the parents were not related in the genetic sense some of them had similar growth rhythms. The Mediterranean parents for example all had low yields in summer and high yields in winter. This problem was met (G. N. Wilkinson, private comm.) by including in the regression a sum of squares derived from variation in the mid-parent value. Midparental values were taken as a measure of the temporal effect, reflecting seasonal changes; the multiple regression was then evaluated for significance after this effect was removed.

The normal regression equation $y_{ij} = a + b_i x_i + b_j x_j$ may be expanded to

$$y_{ij} = a + (b_i + b_j) \frac{(x_i + x_j)}{2} + (b_i - b_j) \frac{(x_i - x_j)}{2} .$$

The second term on the right hand side of the equation is then the regression on the mid-parent.

The analysis of regression then takes the form given in Table 1, where C_{11}, C_{22} and C_{12} are the elements in the matrix of multipliers.

Table 1. *Regression analysis to establish a significant difference between the two b values*

	D.F.	S.S.
Regression on mid-parent	1	by subtraction
Difference between b_i and b_j	1	$\frac{(b_i - b_j)^2}{(C_{11} + C_{22} - 2\,C_{12})}$
Multiple regression	2	$b_i\, \Sigma x_i\, y + b_j\, \Sigma x_j\, y$

4. Results

It is intended to consider in detail representative analyses and then the statistics for all 36 F_1 hybrids. The results for parent 5, parent 8, and their F_1 hybrid are given in Figure 1. Successive yields have been joined to show the contrast in growth rhythms of parents and their hybrid. Parent 5 is of Mediterranean origin and had low summer yields (January, February) and high winter yields (July), whereas parent 8, of North European origin, had a complementary growth rhythm. The hybrid between them showed the growth rhythm of parent 5 but at a higher level of yield; as a consequence it was intermediate between its parents during the summer but exceeded its parents in yield during the winter. When a multiple regression is fitted to these data, the b values are 0.96 for parent 5 and 0.30 for parent 8. Both these coefficients are significant when the other is accounted for, and further, the coefficients are significantly different from one another. The constant a, calculated from $a = \bar{y} - b_i \bar{x}_i - b_j \bar{x}_j$, was only 30 kg/ha, a negligible quantity in the context of the present yields, so that the estimation equation is reduced sensibly to $y = b_i x_i + b_j x_j$ or specifically $y = 0.96\, x_i + 0.30\, x_j$. The calculated yield of the hybrid in any one environment (i.e. at any one harvest) was then a weighted derivative of its two parental yields. This calculated yield was a good estimate both when the observed hybrid yield was at the mid-parental value (e.g. March 29) and when hybrid vigour was manifest (July 17). The standard error of the parental and F_1 values were calculated on a logarithmic scale. At a harvest, two yields were significantly different at the 5% level if one of the yields was more than 1.3 times larger than the other.

A second cross with a similar result is presented in Figure 2 where the b values were 2.36 for parent 2 and 0.41 for parent 8, and the constant a was −80 kg/ha. These b values were significantly different from each other at the 0.1% level, while the constant a was again not significantly different from zero.

The two crosses considered so far have been the combinations 5 × 8 and 2 × 8. The third possible combination between these parents, namely 2 × 5, is given in Figure 3. Both parents are of Mediterranean origin and had low yields during the summer (January, February). In this cross the estimates of b_i and b_j are 0.52 and 0.64. They are not significantly different from each other, and their sum is not significantly greater than unity. There was no suggestion of hybrid vigour, and it is with values like these of b and a that the calculated hybrid yields approximate to the mid-parental yields. A still closer approach to the mid-parental values (b values each of 0.5 and a constant a value of zero) was obtained for the cross 3 × 4 (Figure 4), which had b values of 0.50 and 0.57 and an a value of 10 kg/ha.

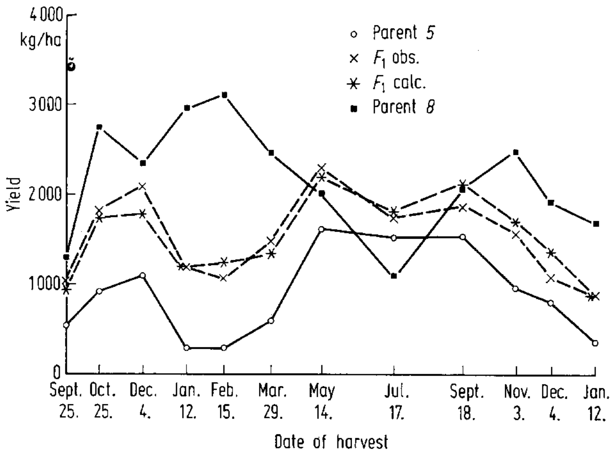

Fig. 1. The yields at each harvest for parents 5 and 8, their F_1 hybrid and the calculated value for the F_1. The yields have been joined by lines to emphasize the contrasts in growth rhythm

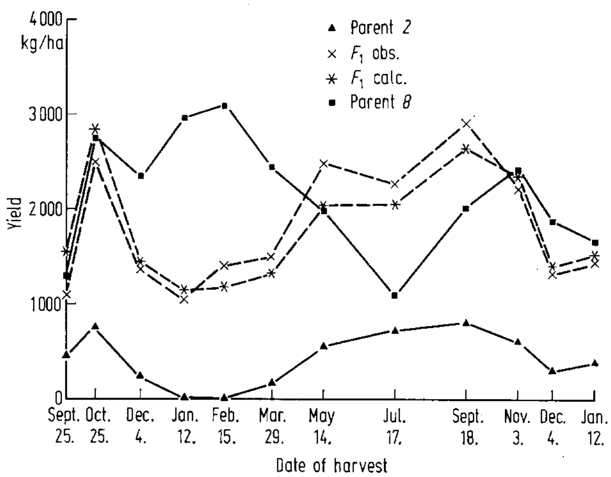

Fig. 2. The yield at each harvest for parents 2 and 8, their F_1 hybrid and the calculated value for the F_1

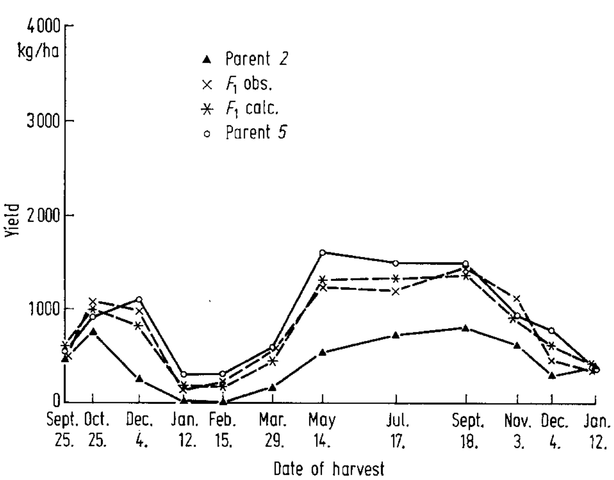

Fig. 3. The yields at each harvest for parents 2 and 5, their F_1 hybrid and the calculated value for the F_1

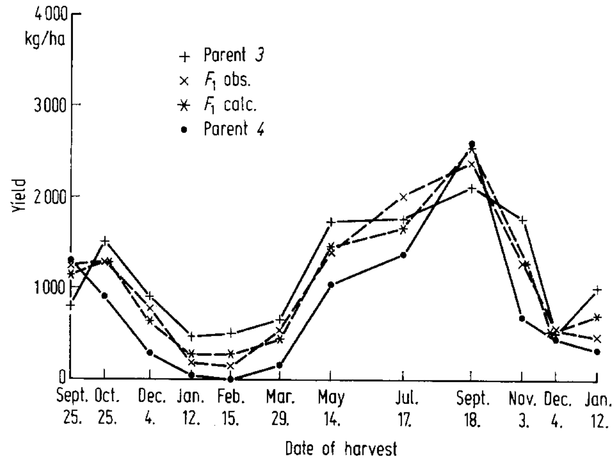

Fig. 4. The yields at each harvest for parents 3 and 4 and their F_1 hybrid and the calculated value for the F_1

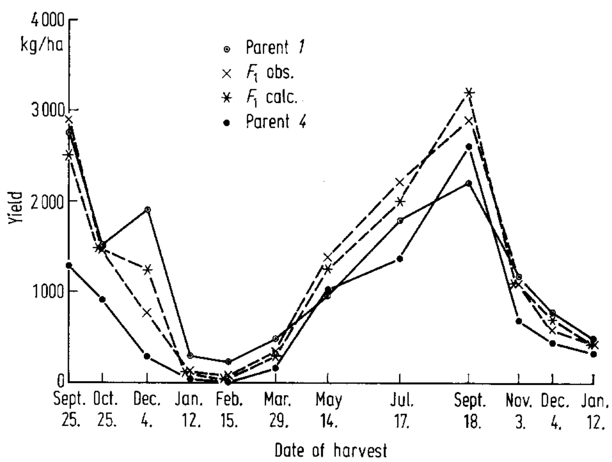

Fig. 5. The yields at each harvest for parents 1 and 4, their F_1 hybrid and the calculated value for the F_1

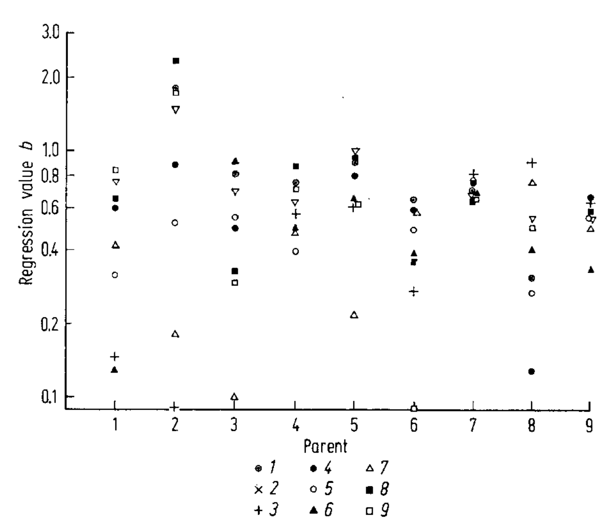

Fig. 6. Variation in multiple regression values (log scale) for each parent when crossed to the other 8 parents

The final example is a cross in which the b values are not significantly different from each other but their sum is significantly greater than unity. The cross between parents 1 and 4 exhibited hybrid vigour during the winter (Figure 5). The b values for this cross were 0.59 and 0.77 for parents 1 and 4 respectively and the constant a was -10 kg/ha. These b values were not significantly different from each other and the a value was non-significant; however, the sum of the b values is significantly greater than 1.00, a situation which leads to estimates of the hybrid value above either parent.

Table 2 contains the results for all 36 crosses. In the upper right half of this table, each cell shows at the top the b value pertaining to the parent in the row; on the second line the b value pertaining to the parent in the column; on the third line the value of the constant a of the regression is shown on the left, the multiple correlation coefficient on the right. The statistical significance level is given for each b value, when the other b value in the multiple regression has been accounted for. This is the usual significance for a multiple regression but a more rigorous analysis of significance was made when the correlated temporal effect was removed by taking out a sum of squares derived from variation in the mid-parental value (section 3). Those crosses in which significance occurs after the mid-parental effect has been removed are given in the lower left half of Table 2, either when b_1 is significantly different from b_2, or when the sum of b_1 and b_2 is significantly greater than unity. The relation of these two criteria to the occurrence of hybrid vigour is considered in the discussion.

The parents differed in the variation in their b values as shown in Figure 6, where the 8 values obtained for a parent over an array of crosses is plotted as an ordinate on a logarithmic scale. Extreme examples are parents 2 and 7; parent 2 gave very different values according to the parent with which it was crossed whereas parent 7 showed negligible variation. It is also evident from this graph that in general the ranking of parents was not similar, the only slight similarity being between parents 1, 2 and 4.

Discussion

A feature of the results was that, within the environmental fluctuations provided by the climatic seasons at a single site, the yield of a hybrid could vary from a value close to the mid-parent to a heterotic value exceeding either parent. This relative performance of a hybrid and its parents in different environments may be a common phenomenon but is not revealed by most experiments as neither parents nor environments are as diverse as those considered here. Instances in which similar results were obtained have been documented (Clausen *et al.*, 1955; Hiesey, 1956).

The multiple regression analysis was attempted when, as in Figure 1, the hybrid performance appeared as a weighted vector of its two parents, both influencing the growth rhythm of the hybrid but not to the same degree. The depiction of a hybrid as a weighted mean of its two parents has encompassed all the hybrid situations ranging from an approximation to the mid-parent value to hybrid vigour exceeding both parents.

An analysis in the form of a multiple regression implies a model that is different to that recently developed in quantitative inheritance by Bucio Alanis and Hill (1966) and Perkins and Jinks (1968) for the performance of hybrids in several environments. They suggest that the yield of an F_1 in a particular environment may be considered as

$$F_1 = \mu + (h) + (1 + \beta)\,\varepsilon + \delta + e$$

where (h) is the average deviation of the F_1 from the overall mean μ and is constant over the several environments; ε is an effect of the particular environment and is calculated as the deviation of the mid-parental value from μ; $(1 + \beta)$ is a regression coefficient; δ is a deviation from regression and e is an error effect. They further suggest that the term $(1 + \beta)\,\varepsilon$ is a depiction of the genotype-environment interaction as a linear function of the environment. It is to be noted that the effect ε is not a physical measure of the environment but is calculated from the mean value of the two parents in that particular environment, thus the term $(1 + \beta)\,\varepsilon$ is derived from

$$\beta\left[\frac{P_i + P_j}{2} - \mu\right].$$

This method of calculating the environmental effect differs from one they have used in other studies (Perkins and Jinks, 1968) where again a physical measure was not used but where the mean value of many genotypes was taken as a measure of the environment (Knight, 1970).

The multiple regression analysis of the Dactylis data differs in that specific regression values were calculated for each parent. The superficial similarity to the above model is evident if the multiple regression equation presented earlier

$$y_{ij} = a + (b_i + b_j)\,\frac{(x_i + x_j)}{2} + (b_i - b_j)\,\frac{(x_i - x_j)}{2}$$

is rewritten as

$$F_1 = a + (\beta_i + \beta_j)\,\frac{(P_i + P_j)}{2} + (\beta_i - \beta_j)\,\frac{(P_i - P_j)}{2}\,.$$

The present approach differs in that it considers the direct relation of the hybrid to each of its two parents and separate b values are calculated. In addition there is a conceptual difference in that the parental yields are not considered as providing a measure of the environment.

From Table 2 it is evident that the contribution of a parent to the performance of its hybrid, as measured by the b values may vary with the other parent with which it is combined. It is possible to consider this

Table 2. *Multiple regression coefficients (b values) and their significance, the multiple regression constant (a) and the multiple correlation value of each cross*

Parent	Parent 1	2	3	4	5	6	7	8	9
1		0.13 1.81*** −40 97	0.15 0.82*** 60 98	0.59** 0.77*** −100 98	0.32* 0.92*** 50 94	0.75*** 0.65 210 87	0.42*** 0.71*** 160 97	0.65** 0.32 660 82	0.84*** 1.21** −1010 89
2	$b_1 - b_2$*** $b_1 + b_2 > 1$***		− 0.23 0.92** −200 93	0.89* 0.49** − 50 97	0.52* 0.64*** −10 96	1.50*** 0.40* 10 93	0.18 0.70** 40 98	2.36*** 0.41* − 80 91	1.76*** 0.34 80 88
3	$b_1 - b_2$**			0.50** 0.57*** 10 97	0.54** 0.62** −80 98	0.70*** 0.28 80 94	0.10 0.84* 210 96	0.33 0.94** −600 81	0.30 0.66* 160 74
4	$b_1 + b_2 > 1$**				0.40** 0.83** −90 96	0.63*** 0.59*** 130 98	0.48* 0.79** − 20 98	0.89*** 0.13 980 91	0.72*** 0.68* 70 89
5	$b_1 - b_2$**					1.01*** 0.50** 0 96	0.22 0.81** 150 94	0.96*** 0.30* 20 93	0.63** 0.56* 140 88
6		$b_1 - b_2$* $b_1 + b_2 > 1$**			$b_1 + b_2 > 1$**		0.58* 0.69*** 360 92	0.37 0.55* 660 69	0.02 0.56 950 50
7				$b_1 + b_2 > 1$*				0.65* 0.79** −200 81	0.66*** 0.51* 500 88
8		$b_1 - b_2$*** $b_1 + b_2 > 1$**		$b_1 - b_2$**	$b_1 - b_2$***				0.51** 0.59* −110 92
9	$b_1 + b_2 > 1$*	$b_1 - b_2$** $b_1 + b_2 > 1$*							

variation as a measure of h in each hybrid, the dominance contribution of each parent to the particular cross. The value of h may be positive or negative and it would be calculated as

$$h = (b_i - {}^1/_2)\, P_i + (b_j - {}^1/_2)\, P_j \,.$$

The h in a cross is thus predictable from the parental yields. Calculation of h for each of the crosses in Table 2 would show that it differed widely from cross to cross within an array. Also implicit in this analysis is that the actual value of h within a cross will vary from one environment to another reflecting variation in the parental values. In some environments h will be positive in others negative.

In relation to the incidence of hybrid vigour, reference was made to the values of b and a in the regression equations. If hybrid vigour is defined as the capacity to exceed the yield of either parent then it will occur when the following conditions are met.

(i) The b values sum to a value greater than unity and the constant a is not negative.

(ii) The environment permits a relatively high yield by the parent with the higher b value.

Alternatively hybrid vigour may occur when the constant a is large and positive. This results in a uniform advantage of the hybrid over the mid-parental value, irrespective of the environment. In these circumstances the hybrid will outyield either parent when the parental yields are similar.

Although the parents and environments considered in this study were diverse there is no reason to believe the analysis is not generally applicable. Unfortunately there are still relatively few studies in which sufficient data on parents and hybrids have been obtained to permit any adequate test of this new approach.

Zusammenfassung

Bei *Dactylis glomerata* wurden 36 Kreuzungen zusammen mit den Elterlinien hinsichtlich des jahreszeitlichen Ertrages in 12 aufeinanderfolgenden Ernten innerhalb einer 16monatigen Periode beurteilt. Die Erträge variierten mit der Umwelt, die durch die jahreszeitlichen Klimaschwankungen bestimmt wurde. Bei einigen Kreuzungen ist der Ertrag der Bastarde dem Mittelelterwert in einer Umwelt angenähert, überschreitet jedoch in einer anderen den Meßwert beider Eltern und zeigt somit Heterosis.

Die 12 Ertragswerte eines Bastards und jeder seiner Eltern wurden mit Hilfe einer multiplen Regression der Form

$$y_{ij} = a + b_i\, P_i + b_j\, P_j$$

analysiert, wobei y_{ij}, P_i und P_j die Erträge des Bastards und der Eltern in jeder einzelnen Ernte und b_i und b_j die Regressionskoeffizienten der entsprechenden Eltern sind und a eine Konstante darstellt. Es wurde beobachtet, daß eine derartige Regression einen großen Teil der Variation der jahreszeitlichen Erträge jeder Kreuzung erklärte, einschließlich der Situation, daß sich ein Bastard dem Mittelelterwert nähert oder beide Eltern übertrifft.

Der b-Wert, der einem bestimmten Elter zugeordnet wird, variierte in Abhängigkeit vom anderen Elter, mit dem gekreuzt wurde. In der Diskussion werden die Werte von b und a in Beziehung zum Vorkommen der Heterosis betrachtet und es wird die Anwendung der multiplen Regression mit anderen Analysen der quantitativen Vererbung verglichen.

References

1. Breese, E. L.: The measurement and significance of genotype-environment interactions in grasses. Heredity **24**, 27–44 (1969). — 2. Bucio Alanis, L., Hill, J.: Environmental and genotype-environmental components of variability. Heredity **21**, 399–405 (1966). — 3. Clausen, J., Hiesey, W. M., Nobs, M. A.: Diploid, tetraploid and hexaploid hybrids of Achillea. Yearb. Carnegie Inst. **54**, 182–183 (1955). — 4. Hiesey, W. M.: Responses of hybrid bluegrass to contrasting environments. Proc. 7th Int. Grassld. Congr. 212–219 (1956). — 5. Knight, R.: The growth of cocksfoot in a winter rainfall environment. Proc. 10th Int. Grassld. Congr. 308–311 (1966). — 6. Knight, R.: The measurement and interpretation of genotype-environment interactions. Euphytica **19**, 225 to 235 (1970). — 7. Perkins, J. M., Jinks, J. L.: Environmental and genotype-environmental components of variability. III. Multiple lines and crosses. Heredity **23**, 339–356 (1968).

Received March 13, 1971
Communicated by W. Seyffert

R. Knight
Waite Agricultural Research Institute
The University of Adelaide
Adelaide (Australia)