The use of doubled haploids in barley breeding
2. An assessment of univariate cross prediction methods

P. D. S. Caligari,*
W. Powell* and
J. L. Jinks†

* Scottish Crop Research Institute, Pentlandfield, Roslin, Midlothian, EH25 9RF, Scotland.
† Department of Genetics, University of Birmingham, P.O. Box 363, Birmingham, B15 2TT, U.K.

Methods of cross prediction have been developed by workers in the Genetics Department of the University of Birmingham. The frequency of families whose mean deviate from m by any specified amount can be estimated from m and $\sqrt{D}$ using the normal probability integral (Jinks and Pooni, 1976).

In the present paper it is shown that satisfactory estimates of the genetical parameters for cross prediction in spring barley can be obtained from either F3 or doubled haploid families. Furthermore, both systems provide a reliable way of ranking crosses but on average the doubled haploid system gave closer agreement between observed and expected. It is shown that estimates of both m and $\sqrt{D}$ are needed for accurate predictions. However, when data from an earlier TTC experiment were examined it was clear that genotype × environment interactions need to be taken into account if such predictions are to be of value in plant breeding.

INTRODUCTION

Jinks, Pooni and co-workers (1976, 1978) have shown that estimates of genetical parameters, from mating schemes involving early generations of crosses between inbred lines, can be used to predict the distributions of the recombinant inbred lines that can be obtained. The F2 Triple Test Cross (Kearsey and Jinks, 1968; Jinks and Perkins, 1970) provides satisfactory estimates of the additive genetical variance necessary for cross prediction and has been used both in wheat (Snape, 1982) and barley (Thomas and Tapsell, 1983; Tapsell and Thomas, 1983). However, this mating design is labour intensive and therefore its practical use in plant breeding is somewhat limited. An alternative approach to an extensive crossing programme has been suggested by Jinks and Pooni (1980). They advocate growing samples of random F3 lines and using twice the component of variance of F3 family means as a measure of the additive genetical variance. In barley this scheme would have the advantage that no cross pollinations would be required after the development of the F1 and plots rather than individual plants would form the basic unit of assessment. Other workers (Reinbergs, Park and Song, 1976; Simpson and Snape, 1979) have indicated that small numbers of doubled haploid lines produced from different crosses may be used for cross prediction. Doubled haploid generations provide a way of estimating the population mean and additive genetical variance and hence, the distribution of inbred lines that would be expected from a cross. In this paper the usefulness of the F3 and doubled haploid systems for predicting the frequency of transgressive segregants is compared.

MATERIALS AND METHODS

Twenty doubled haploids (DH) were produced from F1 hybrids of the following five spring barley crosses:
Golden Promise × Mazurka (TT1)
Universe × Mazurka (TT2)
Golden Promise × Ark Royal (TT3)
BH4/143/2 × Ark Royal (TT4)
Clipper × Ymer (TT5)

In addition 20 F3 families were produced together with a sample of 40 single seed descent lines (SSD) from each cross. The five pair-crosses examined were those originally included in the Triple Test cross experiment described by Tapsell (1984).

The seven parents, F3 families, DH lines and SSD lines were grown in a replicated experiment in 1983 at the Murrays Farm, East Lothian. The experimental design used was a randomised complete block with two replicates. Within a block each family was represented by a row of up to 10 seeds, sown at 5 cm spacings, with a wheat guard at each end of the row. Rows were spaced 22·5 cm apart and the whole experiment was netted to prevent bird damage.

Nine characters of agronomic importance were scored on the material. Details of the scoring procedures can be found in Powell, Hayter, Wood, Dunwell and Huang (1984). The nine characters scored were:

1) Final plant height in cm (Ht).
2) Yield of grain on the main stem in $g \times 10^{-2}$ (MSW).
3) Number of grains on the main stem (GN).
4) Thousand grain weight (TGW) obtained from 2 and 3.
5) Number of fertile tillers (TN).
6) Grain yield of the whole plant, single plant yield (SPY).
7) Awn emergence, measured as days from the 1st June until the awns emerged from the flag leaf sheath of the main stem (AE).
8) Maturity scored on a 1 to 9 scale (1 = early, 9 = late maturity) (Mat).
9) The length of the ear in cm (EL).

RESULTS AND DISCUSSION

The means of the F3 and DH generations were used to estimate m, the mean of all possible homozygotes. In the case of the F3 system it was necessary to assume that $\frac{1}{4}[h]$ and $\frac{1}{16}[l]$ were negligible. The additive genetical variance D or $D + I$ for the DH population was calculated from the between lines mean square using the appropriate expected mean squares for the experiment. Estimates of $\sqrt{D}$ were also obtained for the F3's but in this case it must be assumed that $\frac{1}{8}H$ is negligible. Estimates of m and $\sqrt{D}$ for the F3 and DH generations are given in table 1 together with the parental means for the nine variates scored. From these estimates it was possible, using the method of Jinks and Pooni (1976), to calculate the proportion of

recombinant inbred lines expected to transgress the parental range. The probability of such transgressive lines was estimated as the sum of the normal probability integrals:

$$\int_{-\infty}^{\bar{P}_2} : \int_{\bar{P}_1}^{\infty}$$

corresponding with the value of

$$\frac{\overline{P1} - m}{\sqrt{D}} \quad \text{or} \quad \frac{m - \overline{P2}}{\sqrt{D}}.$$

The SSD lines used in this experiment were at the F7 generation and since the material was bulked from a single F4 plant the expected variance of recombinant inbred lines in $\frac{7}{8}D$ or $\frac{7}{8}(D + I)$. Therefore, when calculating the expected frequency of transgressive segregants $\frac{7}{8}D$ was substituted in the above formulae. The calculated frequencies for each character in the five crosses are given in table 2. It should be noted that where no significant D was detected the expected frequencies were obviously not calculated.

Using the expected frequencies shown in table 2, the expected numbers were calculated by multiplying the frequency by the number of lines grown i.e. 40 and rounding off to a whole number. These are presented in table 3 along with the observed number of transgressive lines observed among the 40 SSD lines. In general the observed numbers of lines agree reasonably well with those predicted by the F3 and DH systems. However, there are some cases where no significant additive genetical variation was detected in the DH while there was in the F3 and, of course, vice versa. In such cases no expected numbers could be calculated for the one system, for example, single plant yield in TT2, ear length in TT3 and thousand grain weight in TT4. The results would appear to confirm that both the F3 and DH systems can be used for cross prediction. Simple inspection of table 3, however, does not give a clear impression of one system being superior to the other. Correlating the observed and expected numbers of lines transgressing the parental range gave coefficients for the F3 and DH systems of 0·5184 and 0·6091 respectively.

While a plant breeder is interested in being able to predict the proportion or number of inbreds which will be superior to the initial pair of parents from which they were derived, a more important comparison is usually between the possible inbred lines and a standard variety which it is wished to better. Thus, it would be desirable to rank the crosses on this criterion i.e., discriminate between

Table 1 Parental means, F3 and DH means for the nine variates scored together with estimates of $\sqrt{D}$ from DH and F3 families

		Ht	MSW	EL	GN	SPY	TGW	AE	Mat	TN
TT1	Golden promise	45·70	144·39	9·99	29·99	8·61	47·80	10·90	4·40	8·87
	Mazurka	59·91	172·89	11·66	32·13	9·38	53·34	12·44	5·25	8·03
	F3 *m*	51·12	161·12	10·48	30·75	8·79	51·80	11·54	4·56	7·62
	DH *m*	55·29	169·87	10·67	31·53	10·81	53·51	10·81	4·20	8·48
	F3 $\sqrt{D}$	6·67	8·07	1·49	1·63	—	3·29	1·14	0·31	—
	DH $\sqrt{D}$	4·57	25·23	1·47	3·11	—	3·57	1·51	1·34	—
TT2	Universe	51·19	153·60	10·54	29·58	10·82	51·96	13·70	4·56	8·62
	Mazurka	59·91	172·89	11·66	32·13	9·38	53·44	12·44	5·25	8·03
	F3 *m*	55·97	164·91	11·30	31·68	9·52	51·30	11·36	4·08	7·78
	DH *m*	51·91	145·60	10·74	28·29	9·77	49·15	12·66	4·42	9·41
	F3 $\sqrt{D}$	4·86	—	—	—	—	—	2·34	1·69	—
	DH $\sqrt{D}$	5·52	—	—	—	—	—	1·31	0·35	—
TT3	Golden promise	45·70	144·39	9·99	29·99	8·61	47·80	10·90	4·40	8·87
	Ark Royal	66·58	161·79	10·90	30·06	13·06	53·88	15·90	7·30	10·54
	F3 *m*	54·63	161·98	10·34	20·76	10·63	52·74	15·03	5·87	8·94
	DH *m*	57·03	160·65	11·01	30·90	11·09	51·64	15·22	5·18	9·61
	F3 $\sqrt{D}$	6·84	10·36	—	—	2·78	3·08	2·13	0·96	2·71
	DH $\sqrt{D}$	7·81	17·16	1·09	—	2·44	17·16	3·10	1·80	2·69
TT4	BH4/143/2	43·99	142·99	10·29	29·17	8·37	46·90	11·90	4·50	8·17
	Ark Royal	66·58	161·79	10·90	30·06	13·06	53·88	15·90	7·30	10·54
	F3 *m*	55·31	152·34	10·48	30·23	9·84	50·21	14·51	5·31	10·75
	DH *m*	61·34	163·93	10·91	31·10	12·95	55·63	14·11	5·58	10·02
	F3 $\sqrt{D}$	8·53	27·55	1·55	3·94	4·61	4·39	0·95	1·02	3·02
	DH $\sqrt{D}$	7·57	17·16	1·06	0·84	—	—	2·38	0·91	—
TT5	Clipper	48·36	86·50	6·38	16·15	6·98	56·71	1·00	2·88	9·85
	Ymer	59·19	168·28	12·12	31·22	11·24	53·76	12·70	4·30	9·31
	F3 *m*	52·57	110·51	10·07	19·31	6·92	56·47	4·05	2·15	6·76
	DH *m*	56·59	131·02	9·42	24·53	9·91	53·19	8·69	3·61	9·81
	F3 $\sqrt{D}$	11·89	38·70	2·19	5·62	0·86	2·29	1·59	1·30	—
	DH $\sqrt{D}$	7·36	24·40	1·81	4·37	2·32	1·73	5·48	1·07	—

Table 2 Expected frequency of transgressive segregants

		Ht		MSW		EL		GN		SPY		TGW		AE		Mat		TN	
		$>P_1$	$<P_2$	$>P_1$	$<P_2$	$>P_1$	$<P_2$	$>P_1$	$<P_2$	$>P_1$	$<P_2$	$>P_1$	$<P_2$	$>P_1$	$<P_2$	$>P_1$	$<P_2$	$>P_1$	$<P_2$
TT1	F3	0·08	0·19	0·06	0·01	0·26	0·40	0·18	0·31	—	—	0·31	0·10	0·06	0·27	0·09	0·29	—	—
	DH	0·14	0·01	0·45	0·14	0·24	0·31	0·42	0·30	—	—	0·50	0·04	0·19	0·31	0·20	0·50	—	—
TT2	F3	0·01	0·15	—	—	—	—	—	—	—	—	—	—	0·14	0·50	0·23	0·50	—	—
	DH	0·06	0·44	—	—	—	—	—	—	—	—	—	—	0·19	0·43	0·01	0·50	—	—
TT3	F3	0·03	0·10	0·50	0·04	—	—	—	—	0·19	0·22	0·34	0·04	0·26	0·02	0·06	0·05	0·26	0·49
	DH	0·14	0·07	0·47	0·16	0·50	0·16	—	—	0·21	0·14	0·44	0·40	0·40	0·07	0·10	0·32	0·36	0·39
TT4	F3	0·08	0·08	0·37	0·37	0·39	0·45	0·50	0·39	0·24	0·37	0·20	0·23	0·06	0·00	0·01	0·20	—	—
	DH	0·23	0·01	0·50	0·22	0·50	0·26	0·50	0·01	—	—	—	—	0·21	0·16	0·02	0·10	—	—
TT5	F3	0·27	0·35	0·05	0·27	0·16	0·04	0·01	0·27	0·00	0·50	0·46	0·10	0·00	0·02	0·04	0·05	—	—
	DH	0·35	0·12	0·05	0·03	0·05	0·04	0·05	0·02	0·27	0·09	0·01	0·50	0·22	0·07	0·24	0·23	—	—

crosses having the highest and lowest potential to exceed a standard variety. With this in mind the five crosses used have been ranked in terms of their ability, as predicted by the F3 and DH systems, to generate inbreds exceeding the highest scoring parent amongst the seven used. Furthermore, using the data from the Triple Test Cross experiment which was carried out in 1978 with the same crosses, predictions were also made (Tapsell, 1984).

One way to compare the three systems is to correlate the predicted rankings from F3, DH and TTC with those observed in the SSD lines. For this purpose Spearman's rank correlation was used

Table 3 The observed and expected numbers of lines predicted by the F3 and DH methods

		Ht			MSW			EL			GN			SPY		
		OBS	ExF3	ExDH	OBS	ExF3	ExDH	OBS	ExF3	ExDH	OBS	ExF3	ExDH	OBS	ExF3	ExDH
TT1	$> P_1$	5	3	6	10	2	18	8	10	9	13	7	17	—	—	—
	$< P_2$	8	8	1	13	1	6	16	16	12	14	12	12	—	—	—
	$< \bar{P}_1 > \bar{P}_2$	27	29	34	17	37	17	16	14	18	13	20	11	—	—	—
TT2	$> P_1$	11	8	2	—	—	—	—	—	—	—	—	—	—	—	—
	$< P_2$	8	6	18	—	—	—	—	—	—	—	—	—	—	—	—
	$< \bar{P}_1 > \bar{P}_2$	21	26	20	—	—	—	—	—	—	—	—	—	—	—	—
TT3	$> P_1$	3	1	5	15	20	19	13	—	20	—	—	—	2	8	8
	$< P_2$	4	4	3	15	1	6	16	—	6	—	—	—	12	9	6
	$< \bar{P}_1 > \bar{P}_2$	33	30	17	10	19	15	11	40	14	—	—	—	26	23	26
TT4	$> P_1$	3	3	9	14	15	20	19	15	20	17	20	20	8	10	—
	$< P_2$	3	3	0	11	15	9	12	18	11	16	15	0	14	15	—
	$< \bar{P}_1 > \bar{P}_2$	34	34	31	15	10	11	9	7	9	7	5	20	22	22	—
TT5	$> P_1$	12	11	14	3	2	2	1	6	2	1	1	2	6	0	11
	$< P_2$	15	14	5	5	11	1	3	1	1	2	11	1	10	20	4
	$< \bar{P}_1 > \bar{P}_2$	13	15	21	32	27	37	36	33	37	37	28	37	24	20	25

		TGW			AE			Mat			TN		
		OBS	ExF3	ExF3	OBS	ExF3	ExDH	OBS	ExF3	ExDH	OBS	ExF3	ExDH
TT1	$> P_1$	18	12	20	11	3	8	7	4	8	—	—	—
	$< P_2$	14	4	2	10	11	12	27	12	20	—	—	—
	$< \bar{P}_1 > \bar{P}_2$	8	24	18	19	27	20	6	24	12	—	—	—
TT2	$> P_1$	—	—	—	15	6	8	18	9	0	—	—	—
	$< P_2$	—	—	—	15	20	17	18	20	20	—	—	—
	$< \bar{P}_1 > \bar{P}_2$	—	—	—	10	4	15	4	11	20	—	—	—
TT3	$> P_1$	12	14	18	6	11	16	7	2	4	9	11	13
	$< P_2$	12	2	16	6	1	3	8	2	12	23	20	15
	$< \bar{P}_1 > \bar{P}_2$	16	24	8	28	28	21	25	36	24	8	9	11
TT4	$> P_1$	16	8	—	10	2	9	4	1	1	—	—	—
	$< P_2$	5	9	—	9	0	6	9	8	4	—	—	—
	$< \bar{P}_1 > \bar{P}_2$	19	23	—	21	38	25	27	31	35	—	—	—
TT5	$> P_1$	8	18	1	9	0	9	12	2	10	—	—	—
	$< P_2$	18	4	20	10	1	3	13	20	9	—	—	—
	$< \bar{P}_1 > \bar{P}_2$	14	18	19	21	39	28	15	18	21	—	—	—

Table 4 The rank correlation coefficients for the three prediction systems

	Ht	MSW	GN	SPY	TGW	AE	Mat
F3	−0·325	0·100	0·575	0·125	−0·575	1·000	0·575
DH	0·675	0·500	0·675	−0·575	0·200	1·000	0·975
TTC	0·075	—	−0·100	0·575	0·225	1·000	—

(Snedecor and Cochran, 1980). The rank correlation coefficients for the three systems are given in table 4, but it should be noted that each coefficient is based on only four df. Comparison of the correlations for F3 and DH shows that for five of the seven characters the DH's have a higher coefficient than do the F3's, for one they are equal and smaller

only for single plant yield. It should be noted that for this character no significant D was detected in the DH populations for three of the crosses. Thus, on average, in this experiment the DH must be regarded as giving a better prediction. Turning to the correlations for the TTC it is clear they are generally very small and indicate a very poor agreement between observed and predicted rankings. However, the predictions from the TTC are based on data collected in 1978 while the observed rankings are from 1983. Thus, the poor agreement is almost certainly due to differences between the two environments which include seasons and husbandry, in other words to genotype × environment interactions.

Having established that both the F3 and DH systems may be used to successfully rank crosses within any one environment it is of both theoretical and practical significance to compare the relative importance of m and $\sqrt{D}$ in the prediction equation. For this purpose linear regression analyses were used. Estimates of m and $\sqrt{D}$ obtained from the F3 and DH generations were regressed against the predicted rankings of the crosses. In addition the multiple regression of m and $\sqrt{D}$ was also calculated. For ease of presentation the coefficients of determination (R^2) for eight of the characters scored are given in table 5. R^2 indicates the extent to which the rankings of the crosses can be predicted for any value of m or $\sqrt{D}$ and was calculated as the percentage of the regression item to the total sums of squares in the regression analysis of variance. It can be seen from table 5(a) that $\sqrt{D}$ is of far greater importance relative to m in terms of predicting the ranking of crosses. Furthermore, accurate estimates of both genetic parameters are needed from both prediction systems to provide the best method of ranking crosses. A number of barley breeding programmes now include the examination of F2 bulk populations. Such predictions rely simply on the mean as a predictor, which we have already seen for DH and F3's does not on average give a very powerful prediction. This can be examined here by estimating the coefficients of determination for the DH, F3 and F2 means and the observed rankings obtained from the SSD population. These coefficients are presented in table 5(b). As can be seen the DH and F3 are in general greater than the F2 and it is clear that an estimate of D is required for predictions to be useful in ranking the crosses used here.

CONCLUSIONS

The results indicated that both the F3 and DH systems gave reasonable agreement of predicted numbers of inbred lines transgressing the parental range with those actually observed among the SSD lines. When the crosses were ranked according to their ability to better a standard variety the DH system gave, on average, the closest correspondence between observed and predicted ranking. This was perhaps to be expected since, as already noted, the estimates from the DH's are based on D while the F3 estimates require the assumption that H is negligible. It has been established that accurate estimates of both genetical parameters (m and $\sqrt{D}$) are needed to provide a meaningful ranking of crosses and therefore caution should be exercised if only estimates of m are used in prediction. When results from a TTC experiment carried out in 1978 were used for predictions they correlated very poorly with the observed values in the 1983 experiment. Thus the results support the conclusion that both F3 and DH can be used in predictions within an environment. But the results from the earlier TTC show, very clearly, that genotype × environment interactions need to be taken into account if such predictions are going to be of value in plant breeding. In this context it is likely that DH's will have an advantage over the use of F3's. The DH's can readily be multiplied to enable trials to be carried out over different seasons and sites,

Table 5(a)   Coefficients of determination for estimates of m and $\sqrt{D}$ obtained from the DH and F3 systems onto the predicted cross rankings

		Ht	MSW	EL	GN	SPY	TGW	AE	Mat
DH	m	64·13	66·80	12·60	8·80	3·99	0·00	59·86	91·19
	$\sqrt{D}$	83·02	45·43	26·57	33·61	95·20	72·21	81·95	0·00
	$m+\sqrt{D}$	90·29	96·57	87·06	83·11	95·79	73·91	97·78	93·89
F3	m	0·00	3·37	25·46	0·59	54·48	55·35	85·46	76·40
	$\sqrt{D}$	50·93	7·20	89·20	38·58	67·23	17·64	0·13	24·57
	$m+\sqrt{D}$	59·66	88·60	89·91	90·43	82·50	85·62	85·68	80·05

Table 5(b)   Coefficients of determination for F2, DH and F3 means onto the observed cross rankings

	Ht	MSW	EL	GN	SPY	TGW	AE	Mat
$\bar{F2}$	22·75	17·89	24·27	33·24	23·55	4·13	80·46	23·56
$\overline{DH}$	2·81	73·11	4·25	81·81	1·23	15·18	59·86	75·93
$\bar{F3}$	65·74	28·16	59·50	42·63	69·31	13·11	85·46	33·31

as is necessary, in order to take environmental sensitivity into account in such prediction work.

REFERENCES

JINKS, J. L. AND PERKINS, J. M. 1970. A general method for the detection of additive, dominance and epistatic components of variation. III. F2 and backcross populations. *Heredity*, *25*, 419–429.

JINKS, J. L. AND POONI. H. S. 1976. Predicting the properties of recombinant lines derived by single seed descent. *Heredity*, *36*, 253–266.

JINKS, J. L. AND POONI, H. S. 1980. Comparing predictions of mean performance and environmental sensitivity of recombinant inbred lines based upon F3 and triple test cross families. *Heredity*, *45*, 305–312.

KEARSEY, M. J. AND JINKS, J. L. 1968. A general method of detecting additive, dominance and epistatic variation for metrical traits. I. Theory. *Heredity*, *23*, 403–409.

POONI, H. S. AND JINKS, J. L. 1978. Predicting the properties of recombinant inbred lines derived by single seed descent for two or more characters simultaneously. *Heredity*, *40*, 349–361.

POWELL, W., HAYTER, A. M., WOOD, W., DUNWELL, J. M. AND HUANG, B. 1984. Variation in the agronomic characters of microspore derived plants of *Hordeum vulgare* cv. Sabarlis. *Heredity*, *52*, 19–23.

REINBERGS, E., PARK, S. J. AND SONG, L. S. P. 1976. Early identification of superior barley crosses by the doubled haploid technique. *Z. Planzenzucht.*, *76*, 215–224.

SIMPSON, E. AND SNAPE, J. W. 1979. Cross prediction for yield using doubled haploid lines. *Barley Genet. News.*, *9*, 95–97.

SNAPE, J. W. 1982. Predicting the frequencies of transgressive segregants for yield and yield components in wheat. *Theor. Appl. Genet.*, *62*, 127–134.

SNEDECOR, G. W. AND COCHRAN, W. G. 1980. *Statistical methods*, 7th edition, Iowa State Univ. Press.

TAPSELL, C. R. 1984. *Cross Prediction Studies on Spring Barley.* Ph.D. Thesis. Univ. of Birmingham.

TAPSELL, C. R. AND THOMAS, W. T. B. 1983. Cross prediction studies on spring barley 2. Estimation of genetical and environmental control of yield and its component characters. *Theor. Appl. Genet.*, *64*, 353–358.

THOMAS, W. T. B. AND TAPSELL, C. R. 1983. Cross prediction studies on spring barley. 1. Estimation of genetical and environmental control of morphological and maturity characters. *Theor. Appl. Genet.*, *64*, 345–352.