Theor Appl Genet (1988) 76: 33–38

The use of cross prediction methods in a practical potato breeding programme

J. Brown*, P. D. S. Caligari, M. F. B. Dale, G. E. L. Swan and G. R. Mackay**
Scottish Crop Research Institute, Pentlandfield, Roslin, Midlothian EH25 9RF, Scotland

Received October 10, 1987; Accepted February 7, 1988
Communicated by A. R. Hallauer

Summary. Most previous studies on cross prediction methods have examined relatively few crosses, particularly in relation to the numbers involved in most breeding programmes. In this paper the feasibility of using cross prediction methods was examined in a practical potato (*Solanum tuberosum*) breeding scheme by the analyses of progeny from 52 crosses. The variate considered was breeder's preference, a visual assessment made of the harvested tubers to estimate their commercial potential. The results showed that it was possible to identify the superior crosses. Cross prediction based simply on the mean preference scores, averaged over scorers and clones within progenies, estimated on seedlings or first clonal year plants, provided the best estimate of a progeny's performance in the third clonal generation. Predictions based on the expected proportion of clones that would transgress a given target value also provided a good indication of a progeny's potential. The poorest prediction was obtained by using the observed frequency of desirable clones in a progeny sample. The implications for potato breeding are discussed.

Key words: Potatoes – *Solanum tuberosum* – Cross prediction – Breeding.

Introduction

The selection of genotypes in the early generations of a potato (*Solanum tuberosum*) breeding scheme is usually by breeders' preference, a visual assessment of tubers which takes account of a number of traits together determining commercial worth. It has been shown (Pfeffer 1963; Anderson and Howard 1981; Tai and Young 1984; Brown et al. 1984) that this type of selection in the seedling generation is inefficient when practised on an individual clone basis. Most current potato breeding schemes do not discard genotypes at the seedling stage although many select only a small percentage of clones in the first clonal year (Tai and Young 1984; Fischen 1984; Scholz 1986). Selection of individual genotypes in the first clonal generation has also been shown, however, to be no more effective than selection of seedlings (Maris 1966; Tai 1975; Brown et al. 1987b).

It has been suggested that negative selection, where only very inferior genotypes are discarded, should be used in the first two generations of a potato breeding programme (Maris 1966; Tai 1975). It has, however, also been recommended that no selection be carried out at these stages because at best, selection resulted in a random reduction in the number of clones and there was even a suggestion that selection had a negative effect (Brown et al. 1987b). If very few, or no genotypes are to be discarded in the early generations then breeders must begin their programmes with a greatly reduced number of clones or adopt a different breeding strategy.

Brown et al. (1987a) found that the mean performance of a cross was relatively repeatable over environments and generations especially in comparison with assessments of individual clones. This feature was used experimentally to examine the feasibility of introducing cross prediction methods into a potato breeding scheme. Caligari and Brown (1986) examined univariate cross prediction methods based on those

---

* Present address: ESCA genetics Corp., 830 Bransten Rd., San Carlos, CA 94070, USA

** Present aadress: Department of Agricultural Botany, Plant Sciences Laboratories, The University of Reading, Whiteknights, P.O. Box 221, Reading RG6 2AS, UK proposed by Jinks and Pooni (1976). Later, following the method given by Pooni and Jinks (1978) they extended this approach to take into account more than a single character (Brown and Caligari 1988). In both cases it was shown to be possible, under experimental conditions, to predict the frequency of clones from each cross that would transgress predefined values of one or more variate. Based on the probabilities obtained, it was possible to rank the crosses and thus determine which ones had the highest chance of producing a commercially desirable cultivar.

In common with virtually all other studies of cross prediction methods (Jinks and Pooni 1986; Snape 1982; Caligari et al. 1985; Powell et al. 1985) the work of Caligari and Brown (1986) and Brown and Caligari (1988) examined few crosses relative to the number handled in a full scale breeding programme. If, however, such cross prediction methods are actually to be useful to plant breeders they must be shown to be effective in handling large numbers of crosses. In the present study the feasibility of cross prediction methods is examined in an actual potato breeding programme and the effectiveness of such a breeding strategy in comparison to more traditional breeding methods is discussed.

Materials and methods

Four samples of 25 seeds from each of 191 *S. tuberosum* crosses were sown into seed pans in the spring of 1984. These 191 crosses were representative of all the commercial breeding material to be grown in that year. The seed pans were randomly arranged into sample blocks. After three weeks, 20 randomly chosen seedlings from each pan (i.e., 80 seedlings per cross) were transplanted into 10 cm square pots. These were arranged in glasshouses on four beds, one sample from each cross on each bed. Within the beds, the sample groups were randomised and the seedlings grown to maturity. At harvest, the soil from each pot was removed and the tubers that each seedling produced were placed back into the appropriate empty pot. After all the seedlings were harvested (15, 280 genotypes) the tubers were independently assessed by four potato breeders for breeder's preference, using a 1–9 scale of increasing preference. Such an assessment automatically combines the visible tuber characters which go to determine the breeder's overall estimate of a clone's agronomic potential i.e., its commercial worth. Two tubers were then taken from each seedling and kept at 4 °C until the following spring.

In the first clonal year (1985) the 191 crosses were divided into two groups. The first group consisted of 52 crosses, chosen to represent the range of crosses as assessed by preference scores. These crosses were grown at two locations, Blythbank (BB), Peeblesshire and Murrays (MURR), East Lothian. The remaining crosses (139) were grown only at BB. The BB site is used by the Potato Breeding Department of The Scottish Crop Research Institute (SCRI) for the multiplication of disease-free clonal stocks. In comparison the MURR site is used for yield and other ware assessment trials. It should also be noted that BB is the only location that clones in the first and second clonal generations of the SCRI breeding scheme are grown. The MURR site is always planted earlier and harvested later than BB and hence has a longer growing season, making it more comparable with that of a commercial ware crop.

In the first clonal year trials each cross (191 at BB and 52 at MURR) was represented by four samples, each sample containing 15 plants. The samples were arranged in four completely randomised blocks with each cross represented in each block by a single-drill sample of 15 genetically unique clones. At harvest each plant was hand dug and the tubers laid on the soil surface. The produce from every plant was assessed, independently, by three potato breeders (three of those who had carried out the seedling assessments) for breeders' preference. After assessment the produce from each of the BB plants (the seed site) was taken into storage for use as seed in the following year's trial.

In 1986, only the 52 crosses that had previously been grown at both BB and MURR were grown. In this, the second clonal year, each of the crosses was represented by 40 genetically unique clones at BB and by 20 at MURR, these latter 20 being a sub-sample of the 40 at BB. At each site, clones were grown in single-drill, four plant plots, the plots being arranged in a completely randomised block at each site. The plots were mechanically harvested and the produce of each plot hand picked into boxes. While still in the field, the produce from each plot was again independently assessed for breeders' preference by three potato breeders. The three breeders who carried out this assessment had all scored the seedling stage but one, had not scored the first clonal year. Overall this meant that only two of the four breeders had assessed all three stages of the experiment.

The basic item of datum examined in this study was the mean preference, in other words the preference score of each plot, averaged over four breeders at the seedling stage and over three breeders in the first and second clonal years. It should be noted that before assessing each stage in the experiment the breeders had agreed that a preference score of 5, or more, would indicate that the particular plot, under normal selection criteria, would have been selected; a corollary of which is, of course, that a score of less than 5 would indicate that the plot would have been discarded.

Results

The clones that were grown in the first year at BB and also the second clonal year at BB and MURR were uniquely identified. In each of the trials, every clone was assessed for overall preference. The correlation between scores at BB in the first clonal year and BB in the second clonal year gave a significant ($P < 0.001$) coefficient of 0.55 based on 2079 degrees of freedom. A similar correlation, based on 1040 observations, between BB in the first clonal year and MURR in the second clonal year was also highly significant ($P < 0.001$) with a coefficient of 0.41. Despite the formal significance of these correlations, regression of first clonal year scores onto second clonal year scores only accounted for 30% and 17% of the total variation at BB and MURR, respectively. These correlations are similar, though slightly higher, than those previously reported by Maris (1966), Tai (1975) and Brown et al. (1987 b).

The effect of selecting individual clones at the first clonal year stage can be simulated since, as already noted, a score of 5, or more, indicated that a particular clone, under normal selection, would have been retained for re-trial. The normal practise at SCRI has been to grow the first and second clonal year trials only at BB so this comparison will be considered first (Table 1 a). From the 2080 clones that were grown at BB in both years, 222 were selected in the first clonal year and 181 were selected in the second. However, only 69 clones were selected in both years. Therefore, 62% of the second clonal year selections would have been discarded in the first year under normal conditions. Similarly when the 1040 clones that were grown at BB and MURR in the first and second clonal years respectively were considered (Table 1 b), 94 clones were selected at MURR in the second clonal year, 25 clones were selected at both. Thus, 73% of the selections made at MURR in the second clonal year would have been discarded at the first clonal year stage if normal selection at BB had been carried out. These results therefore support the conclusions of earlier work, that a high proportion of potentially desirable clones would be lost if individual clone selection was practised in the first clonal year of a potato breeding scheme.

To examine the use of cross prediction methods the following estimates were considered from each year and site; (1) the mean preference, averaged over breeders and clones within crosses; (2) the phenotypic, within progeny variance of mean preference; (3) the observed proportion of clones from each cross that transgressed a mean preference score of 5 and (4) the predicted proportion of clones that would be expected to transgress an overall preference score of 5. Calculation of (4) was based on (1) and (2) using the normal probability integral (Jinks and Pooni 1976). The correlations between environments for mean preference were all significant. The lowest coefficients were obtained when seedling mean preference was correlated with assessment in the field. However, even these coefficients ranged from 0.70 for seedlings and the first clonal year at BB, to 0.78, for seedlings and the second clonal year at MURR. Correlation coefficients between BB and MURR in the first clonal year were slightly lower than similar coefficients between the first and second clonal year. The highest correlation coefficient for mean preference was, in fact, between the two sites in the second clonal year.

Coefficients obtained by correlation between environments for the observed proportion selected (overall preference >5) were generally lower in magnitude than the predicted correlations (Table 2). However, caution should be exercised in that the pair-wise correlations given in Table 2 are not independent within each column. Nevertheless, the average coefficient for the observed proportion selected was 0.5 and that for the predicted proportion selected was 0.7. If the three methods of prediction are considered, then the highest correlation coefficients were obtained for mean preference followed by the proportion of selected clones,

Table 1 a and b. Number of clones that were **a** selected and rejected at BB in the first clonal year (BB-FCY) and at BB in the second clonal year (BB-SCY), and **b** number of clones that were selected and rejected at BB in the first clonal year (BB-FCY) and at MURR in the second clonal year (MURR-SCY); figures in brackets are the expected numbers, assuming the two rounds of selection to be independent

a BB-FCY v BB-SCY

		BB-FCY	
		Rejected	Selected
BB-SCY	Selected	112 (161.7)	69 (19.3)
	Rejected	1746 (1696.3)	153 (202.7)

b BB-FCY v MURR-SCY

		BB-FCY	
		Rejected	Selected
MURR-SCY	Selected	69 (84.0)	25 (10.0)
	Rejected	627 (611.9)	58 (73.1)

Table 2. Correlation coefficients (r) between the 52 progenies grown as seedlings in the glasshouse (G.H.) and at BB and MURR in the first (FCY) and second clonal years (SCY); coefficients are presented for mean preference (Mean), observed proportion of selected clones (mean preference >5, Obs.sel) and the predicted proportion of clones expected to transgress an overall preference score of 5, based on the overall preference and the within progeny variance (Pre.sel); levels of significance can be judged as: $r>0.28$ significant at $P<0.05$, $r>0.36$ significant at $P<0.01$, and $r>0.46$ significant $P<0.001$

		Mean	Obs.sel	Pre.sel
G. H.	v BB-FCY	0.70	0.38	0.61
G. H.	v MURR-FCY	0.74	0.50	0.65
G. H.	v BB-SCY	0.71	0.28	0.59
G. H.	v MURR-SCY	0.78	0.52	0.76
BB-FCY	v MURR-FCY	0.75	0.65	0.72
BB-FCY	v BB-SCY	0.83	0.49	0.75
BB-FCY	v MURR-SCY	0.80	0.57	0.63
MURR-FCY	v BB-SCY	0.77	0.59	0.67
MURR-FCY	v MURR-SCY	0.78	0.57	0.73
BB-SCY	v MURR-SCY	0.89	0.47	0.70

with the observed proportion of selected clones being least repeatable.

In general, plant breeders are more interested in the relative value of different crosses within a breeding programme (i.e., rankings) than absolute performance. In this context it may be of more interest to identify the best, second best and worst etc. than to compare actual scores. Rank correlations of mean preference between different environments were of a similar magnitude to the phenotypic correlations (Table 3). All pair-wise comparisons were highly significant ($P < 0.001$) and it should therefore be possible to identify the "better" crosses even when the progenies are grown in the most atypical environment (i.e., as seedlings in the glasshouse).

The aim of cross prediction in a practical breeding scheme would be to identify crosses that have a high probability of providing genetically superior clones. The most interesting comparison is, therefore, between predictions in the seedling and first clonal generation with the actual number of clones that were selected in the second clonal year, this being the best estimate of the true commercial potential of this material (Table 4). The correlation coefficients between mean performance in the first two clonal generations with the number of clones selected in the third generation were higher when MURR was the selection site than when BB was. All coefficients between the rank of mean preference and number selected in the second clonal year were negative (i.e., the better the rank (lower number) the higher the selection rate). At BB in the second clonal year, 181 clones were selected, from these 65 (36%) were from crosses ranked 1st to 10th for overall preference as seedlings (Table 5), while only 6 clones were selected from those crosses ranked 41st to 52nd in the glasshouse. If the predictions had been based on progeny evaluation of mean overall preference at BB in the first clonal year then 70 clones (39%) of second clonal year selections would have derived from crosses ranked 1st to 10th and only 10 clones from those crosses ranked 41st to 52nd. If the crosses had been assessed at MURR in the first clonal year a similar result would have been obtained in that the proportion of clones selected was directly related to the rank of crosses. Also, if the number of clones selected at MURR in the second clonal year was partitioned in this way an almost identical result was obtained.

Discussion and conclusions

The results from this study agree with those given in earlier papers (Maris 1966; Tai 1975; Brown et al. 1987b) in demonstrating that selection of individual

Table 3. Rank correlations between mean preference scores on 52 progenies raised as seedlings in the glasshouse (G. H.) and at BB and MURR in the first (FCY) and second clonal years (SCY). All coefficients are significant ($P < 0.001$)

	G. H.	BB-FCY	MURR-FCY	BB-SCY
BB-FCY	0.66			
MURR-FCY	0.67	0.90		
BB-SCY	0.64	0.91	0.82	
MURR-SCY	0.68	0.70	0.87	0.78

Table 4. Correlation coefficients between mean preference as well as the rank of mean preference, recorded on seedlings grown in the glasshouse (G. H.) and at BB and MURR in the first clonal year (FCY), and the number of clones that were selected (overall preference > 5) at BB and MURR in the second clonal year (SCY). All coefficients are significant ($P < 0.001$)

		Number of clones selected	
		BB-SCY	MURR-SCY
G. H.	mean	0.59	0.68
G. H.	rank	–0.55	–0.55
BB-FCY	mean	0.63	0.68
BB-FCY	rank	–0.66	–0.67
MURR-FCY	mean	0.57	0.75
MURR-FCY	rank	–0.64	–0.75

Table 5. Number of clones selected at BB and MURR in the second clonal year (SCY) grouped according to progeny ranking of mean preference as seedlings grown in the glasshouse (G. H.) and at BB and MURR in the first clonal year (FCY)

	Number of clones selected	
	BB-SCY	MURR-SCY
Rank in G. H.		
1–10	65	33
11–20	37	22
21–30	43	19
31–40	30	17
41–52	6	3
Rank at BB-FCY		
1–10	70	32
11–20	46	23
21–30	33	16
31–40	22	16
41–52	10	7
Rank at MURR-FCY		
1–10	63	34
11–20	50	25
21–30	42	17
31–40	17	14
41–52	9	4

genotypes at the first clonal year stage of a potato breeding programme is ineffective and results in the loss of potentially valuable clones.

The present study suggests, however, that it is possible to identify superior progenies using cross prediction methods and these progenies are shown to yield a higher proportion of selected clones. The best predictor was found to be the mean preference score. The next best was the prediction based on the mean and within progeny variance of the preference score while the least repeatable predictor was the observed proportion of selected clones in the sample.

Brown et al. (1987 b), when examining the efficiency of selecting individual genotypes in the early generations, found that the correlation coefficients between seedling scores and clonal generations decreased with increasing years. For example, the largest correlation was between seedlings and first clonal year, followed by seedlings with the second clonal year. The lowest correlations were between seedlings and the third clonal year. They concluded that this was due to a "carry over" effect between clonal generations. In this study, however, the correlation coefficients between crosses did not appear to follow this trend. In fact, the coefficients showed a tendency to go in the opposite direction. This is the direction that might be expected in the absence of "carry over" effects between clonal generations when increasing generations are associated with increasing sample (i.e., plot) size.

The correlation coefficients between sites and years obtained in this study did however follow the trend noted by Caligari et al. (1986) and Brown et al. (1987 b), with higher correlation coefficients being obtained between common sites in different years than between different sites in the same year. The effect was not large but did imply a slight tendency for selection, as practised, to favour the particular locality at which it was practised.

Caligari and Brown (1986) found that the within progeny variance added little to the prediction where the target value was set at 5. They found, however, that the within progeny variance became increasingly important where higher target values were used. Where 5 is used as a target value, the lower limit for "selected" clones, then it may not be necessary to have a good estimate of the within progeny variance. The same effect could prove relevant to the present results in that, in the normal way, the population will be further reduced in later clonal generations until it simply consists of the one or two clones that will be submitted for statutory trials as potential commercial cultivars. If it was attempted to predict in which progenies these clones would occur, which would therefore be at very low frequency and, hopefully, associated with the higher scores, then the variance might become decisive and the prediction method might well appear more efficient again.

Although the 52 crosses were not a random sample, having been chosen deliberately to represent the range present in the initial crosses, when all 191 crosses were included the correlations between seedlings and BB in the first clonal year were 0.65 for mean preference, 0.37 for the observed proportion selected and 0.64 for the predicted proportion selected. There is, therefore no reason to suppose from these results that the conclusions would not hold if all 191 progenies had been included in the later generation.

Cross prediction has several advantages over the more traditional methods of potato breeding. First, it allows efficient selection to be carried out among glasshouse grown seedlings. This is the stage where most clones can be grown easily, in relatively small amounts of space and makes least demands on labour. The present study shows that even in this atypical environment, progeny assessment, as opposed to individual clone selection, would be effective. Thus the selection pressure could be increased by screening large numbers at this stage. Secondly, if the progeny sample that is assessed is representative of the whole progeny then, effectively, the breeder is assessing the whole potential of the cross i.e. as though he had grown a progeny of infinite size from the cross. Thirdly, the technique promises to provide the basis for rapidly assessing the success of parental genotypes in producing desirable progenies. It therefore would allow selection to be practised among the multitude of parental genotypes that are available and potentially increase the power of selection that the breeder could apply. Fourthly, because of the simplicity of the trials that are needed and the fact that only a sample of clones is required, the material can be tested in many different environments. Finally, it is easy to extend the approach to encompass more than one variable. Indeed, at SCRI cross prediction has not only been successfully carried out over the last three years for mean preference but also crosses have been assessed for resistance to tuber blight and foliage blight (Caligari et al. 1983) as well as potato cyst nematode *G. pallida* Pa2/Pa3 (Phillips 1981). Investigations are presently underway to extend the use of such cross prediction techniques to cover resistance to common scab, gangrene, tobacco rattle virus, potato leafroll virus and to assess cooking quality characters such as after cooking blackening and fry colour.

References

Anderson JAD, Howard HW (1981) Effectiveness of selection in the early stages of potato breeding. Potato Res 24: 289–299

Brown J, Caligari PDS (1988) The use of multivariate cross prediction methods in the breeding of a clonally reproduced crop (*Solanum tuberosum*). Heredity 60: 147–153

Brown J, Caligari PDS, Mackay GR, Swan GEL (1984) The efficiency of seedling selection by visual preference in a potato breeding programme. J Agric Sci 103:339–346

Brown J, Caligari PDS, Mackay GR (1987a) The repeatability of progeny means in the early generations of a potato breeding programme. Ann Appl Biol 110:365–370

Brown J, Caligari PDS, Mackay GR, Swan GEL (1987b). The efficiency of visual selection in the early generations of a potato breeding programme. Ann Appl Biol 110:257–263

Caligari PDS, Brown J (1986) The use of univariate cross prediction methods in the breeding of a clonally reproduced crop (*Solanum tuberosum*). Heredity 57:395–401

Caligari PDS, Mackay GR, Stewart HE, Waistie RL (1983) A seedling progeny test for resistance to potato foliage blight (*Phytophthora infestans* (Mont.) de Bary). Potato Res 27: 43–50

Caligari PDS, Powell W, Jinks JL (1985) The use of doubled haploids in barley breeding. 2. An assessment of univariate cross prediction methods. Heredity 53:353–358

Caligari PDS, Brown J, Abbott RJ (1986) Selection for yield and yield components in the early generations of a potato breeding programme. Theor Appl Genet 73:218–222

Fischen HJ (1984) Züchtung neuer Kartoffelsorten. Kartoffelbau 35:52–55

Jinks JL, Pooni HS (1976) Predicting the properties of recombinant lines derived by single seed decent. Heredity 45:305–312

Jinks JL, Pooni HS (1986) Description and illustration of the practical applications of biometrical genetics to plant breeding. In: Proc 6th Meet Eucarpia Sect. Biometrics in plant breeding, Birmingham, pp 1–20

Maris B (1966) The modifiability of characters important in potato breeding. Euphytica 15:18–31

Pfeffer C (1963) Vergleichende Untersuchungen über Auslesemöglichkeiten von im Freiland und in Töpfen kultivierten Kartoffelsämlingen. Züchter 33:6–11

Phillips MS (1981) A method of assessing potato seedling progenies for resistance to the white potato cyst nematode. Potato Res 24:101–103

Pooni HS, Jinks JL (1978) Predicting the properties of recombinant inbred lines derived from single seed descent for two or more characters simultaneously. Heredity 40: 349–361

Powell W, Caligari PDS, McNicol JW, Jinks JL (1985) The use of doubled haploids in barley breeding. 3. An assessment of multivariate cross prediction methods. Heredity 55: 249–254

Scholz M (1986) General strategy of potato breeding in the German Democratic Republic. In: Proc. EAPR/EUCARPIA breeding and variety assessment meeting, 1985, Cambridge/UK (in press)

Snape JW (1982) Predicting the frequency of transgressive segregants for yield and yield components in wheat. Theor Appl Genet 62:127–134

Tai GCC (1975) Effectiveness of visual selection for early clonal generation seedlings of potato. Crop Sci 15:15–18

Tai GCC, Young DA (1984) Early generation selection for important agronomic characteristics in a potato breeding programme. Am Potato J 61:419–434