Theor Appl Genet (1989) 77: 501–504

Optimization of recurrent selection on the phenotypic value of doubled haploid lines

A. Gallais

CNRS-INRA-UPS, Ferme du Moulon, F-91190 Gif sur Yvette, France
INA-PG, 16 Rue Claude Bernard, F-75231 Paris Cedex 05, France

Received July 21, 1988; Accepted December 6, 1988
Communicated by A. R. Hallauer

Summary. Three methods of recurrent selection using doubled haploids (DH) are compared with the same resources: individual line selection (with one DH line per plant), selection between and within families, and selection index of family and line within family values. It seems that for low heritabilities at the individual level, the selection index method is the most efficient method, followed by the selection between and within families. Because of the difficulty in optimizing selection intensities, the selection between and within families can be less efficient than individual line selection. For very low heritabilities, it will be better to use selection index with a relatively low number of lines per haploidized plant, to minimize the rate of decrease of the effective population size. For a large range of situations, recurrent selection with only one DH line per plant seems one of the more efficient and the simpler methods.

Key words: Haplodiploidization – Line development

Introduction

Haplodiploidization can be defined as a system of mating which allows the derivation of all possible lines from a given genotype (Gallais 1978, 1979, 1986, 1988). Its use in recurrent selection was discussed by Gallais (1986, 1988). It was shown that one of the most efficient methods using doubled haploids (DH) would be a recurrent selection on the phenotype of the DH lines. We have shown that recurrent selection with only one DH line per haploidized plant in the population will be very efficient and simple in comparison to a method using haplodiploidization to develop a progeny test on line value, with intercrossing of the mother plant. However, we have not discussed the interest of a combined selection which can be developed when there are several lines from each haploidized mother plant. In this case two strategies are possible. The simpler is to select the best families of lines and, within these best families, select the best lines. However, from the selection index theory (Lush 1943), we can predict that with such a family structure, the most efficient selection will be that based upon an index of the family value and the value of the lines within a family. The objective of this paper is to study these two methods in comparison with the single DH method.

Efficiency of the combined selection

Description of the experimental design

We assume that we have l lines from each of the p haploidized plants. To test all the lines in a plot design, families are grouped in sets of the same size, with b repetitions within the set. We assume that the harvest is at the plot level. An analysis of variance can be performed (Table 1), with a family effect and a line/family effect, according to the following model for the phenotypic value P_{ijk} of the line j from plant i in the repetition k, within a set of families:

$P_{ijk} = \mu + M_i + L_{ij} + b_k + E_{ijk}$, where

M_i is the effect of the plant i with a variance σ_M^2, $1 \leq i \leq p$; L_{ij} is the effect of line j within plant i with a variance $\sigma_{L/M}^2$, $1 \leq j \leq l$; b_k is the effect of the repetition k, and is considered as a fixed effect, $1 \leq k \leq b$; E_{ijk} is the residual effect at the plot level, with a variance σ_e^2 (in the absence of interaction between block and other factors, E_{ijk} is due only to environment).

Table 1. Analysis of variance for a design with l DH lines per haploidized plant. $\sigma_M^2 = \sigma_{L/M}^2 = \sigma_{A_L}^2$ (Gallais 1986)

Source of variation	Mean square	Expected mean square
Plant/set	CM_M	$\sigma_e^2 + b\,\sigma_{L/M}^2 + b\,l\,\sigma_M^2$
Line/plant/set	$CM_{L/M}$	$\sigma_e^2 + b\,\sigma_{L/M}^2$
Residual	CM_e	σ_e^2

Using the phenotypic plant effect $\tilde{M}_i = P_{i..} - P_{...}$, to predict the line value A_{ij} of the progeny after intercrossing of the line ij we have:

$$\hat{A}_{ij} = \frac{\operatorname{cov} A_{ij}(P_{i..} - P_{...})}{\operatorname{var}(P_{i..} - P_{...})}\,(P_{i..} - P_{...})$$

$$\operatorname{cov} A_{ij}(P_{i..} - P_{...}) = \frac{p-1}{p}\left[\frac{1}{l}\operatorname{cov} A_{ij}P_{ij} + \frac{l-1}{l}\operatorname{cov} A_{ij}P_{ij'}\right]$$

but $\operatorname{cov} A_{ij}P_{ij} = \sigma_{A_L}^2$ (Gallais 1978, 1988)

and $\operatorname{cov} A_{ij}P_{ij'} = 1/2\,\sigma_{A_L}^2$

so $\operatorname{cov} A_{ij}(P_{i..} - P_{...}) = \frac{p-1}{p}\,\frac{l+1}{l}\,1/2\,\sigma_{A_L}^2$

$$\operatorname{var}(P_{i..} - P_{...}) = E(P_{i..} - P_{...})^2 = \frac{p-1}{p}\,(CM_M/b\,l)$$

The genetic advance due to family selection will be:

$$\Delta G_M = 2\,i_B\sqrt{\operatorname{var}\hat{A}} = i_B c_B\,\frac{\sigma_{A_L}^2}{\sqrt{\operatorname{var} M}}$$

with $c_B = \frac{l+1}{l}\sqrt{\frac{p-1}{p}} \sim \frac{l+1}{l}$ because p is great.

The index B is for between family selection.

Using the line/plant phenotypic effect $(\tilde{L}_{ij} = P_{ij} - P_{i..})$ to predict the line value of progeny after intercrossing of line ij, we have:

$$\hat{A}_{ij} = \frac{\operatorname{cov} A_{ij}(P_{ij.} - P_{i..})}{\operatorname{var}(P_{ij.} - P_{i..})}\,(P_{ij.} - P_{i..})$$

$$\operatorname{cov} A_{ij}(P_{ij.} - P_{i..}) = \operatorname{cov} A_{ij}P_{ij} - \frac{1}{l}\operatorname{cov} A_{ij}P_{ij} - \frac{l-1}{l}\operatorname{cov} A_{ij}P_{ij'} = \frac{l-1}{l}\,\frac{1}{2}\,\sigma_{A_L}^2$$

$$\operatorname{var}(P_{ij.} - P_{i..}) = \frac{l-1}{l}\,(CM_{L/M}/b)$$

So the genetic advance due to within family selection will be:

$$\Delta G_W = 2\,i_W\sqrt{\operatorname{var}\hat{A}} = i_W c_W\,\frac{\sigma_{A_L}^2}{\sqrt{\operatorname{var} L/M}}$$

with $c_W = \sqrt{(l-1)/l}$.

The total genetic advance due to combined selection will be:

$$\Delta G_C = i_B c_B\,\frac{\sigma_{A_L}^2}{\sqrt{\frac{l+1}{l}\sigma_{A_L}^2 + \sigma_e^2/b\,l}} + i_w c_w\,\frac{\sigma_{A_L}^2}{\sqrt{\sigma_{A_L}^2 + \sigma_e^2/b}}$$

For individual line selection, the genetic advance will be:

$$\Delta G_L = i\,\frac{2\,\sigma_{A_L}^2}{\sqrt{2\,\sigma_{A_L}^2 + \sigma_e^2/b}} = i\,h\sqrt{2\,\sigma_{A_L}^2}\ \text{with}\ h^2 = \frac{2\,\sigma_{A_L}^2}{2\,\sigma_{A_L}^2 + \sigma_e^2/b}$$

The ratio $\Delta G_C/\Delta G_L$ is:

$$\Delta G_C/\Delta G_L = \sqrt{2\,\sigma_{A_L}^2 + \sigma_e^2/b}\left[i_B c_B\Big/\sqrt{\frac{l+1}{l}\sigma_{A_L}^2 + \sigma_e^2/b\,l} + i_W c_W\Big/\sqrt{\sigma_{A_L}^2 + \sigma_e^2/b}\right]\Big/2\,i$$

Introducing the parameters h^2 we get the following expression:

$$\frac{\Delta G_C}{\Delta G_L} = \left[i_B\frac{l+1}{l}\Big/\sqrt{[h^2(l-1)+2]/l} + i_W\sqrt{\frac{l-1}{l}}\Big/\sqrt{2-h^2}\right]\Big/(i\sqrt{2})$$

Hence, the ratio depends only on h^2, l and selection intensities i_B, i_W, i. Numerical calculation of the ratio can be developed for a given set of values of the parameters. The change in b is equivalent to the change of heritability. h^2 will vary between 0.10 and 0.60. To determine the selection intensities, assume that 240 lines are studied with $l=4$ or $l=2$, and that 30 lines are intercrossed the rate of selection, is 12.5%. Then we have the following table:

	Rate of selection			
	Situation	Between	Within	Total
$l=4$	1	50%	25%	12.5%
	2	25%	50%	12.5%
	3	16.6%	75%	12.5%
$l=2$	4	25%	50%	12.5%
	5	12.5%	100%	12.5%

The results are given in Table 2 for each situation and each value of the heritability.

It appears that when only one line is selected within the best family, the combined individual family selection is less efficient than individual line selection. As expected when the between family selection intensity increases, the relative efficiency of combined selection is greater than 1

Table 2. Relative efficiency of combined between-within family selection to individual line selection

Heritability*	Situation				
	$l=4$			$l=2$	
	1	2	3	4	5
0.10	0.90	1.11	1.17	0.97	1.03
0.20	0.88	1.07	1.11	0.96	1.01
0.40	0.85	1.00	1.02	0.94	0.97
0.60	0.82	0.91	0.92	0.92	0.93

* defined at the level of means of b repetitions

Table 3. Relative efficiency of an index of family and line/family values

h^2 l	0.10	0.20	0.40	0.60
4	1.25	1.19	1.10	1.04
2	1.10	1.08	1.04	1.02

mainly for lower heritabilities. However we must note that heritability of 0.10 at the level of mean performance of the lines ($h^2_{\bar{x}}$) corresponds to a lower heritability at the plot level h^2_p and to a very low heritability at the individual level. Example: with three repetitions, a value of $h^2_{\bar{x}} = 0.10$ corresponds to $h^2_p = 0.03$, and $h^2_{\bar{x}} = 0.20$ corresponds to $h^2_p = 0.07$. A heritability of 0.10 at the individual level corresponds to a heritability of mean values of 0.4–0.5 with three repetitions, ten individuals per plot, and with an environmental correlation between individuals of the same plot of 0.2–0.4. The conclusion is that proposed combined selection will not be as efficient as individual line selection for a large range of realistic heritabilities.

If we can consider only between family selection, we have:

$$\Delta G_B = i_B \frac{\frac{l+1}{l}\sigma^2_{A_L}}{\frac{l+1}{l}\sigma^2_{A_L} + \sigma^2_e/b\,l}$$

If $l = 1$, this gives "individual" line selection (with one line per haploidized plant). The relative efficiency of this method with $l > 1$, assuming the same selection intensity, will be greater than 1 for low heritabilities. However, with the same amount of resources (and the same number of intercrossed lines) when there are l selected ("sister") lines per selected family, the genetic base of the breeding population will decrease more rapidly with the family selection.

Efficiency of a selection index with family and line/family values

In this situation the line value A_{ij} of progeny, after intercrossing of a line ij, can be written:

$$A_{ij} = \beta_1 (P_{i..} - P_{...}) + \beta_2 (P_{ij.} - P_{i..})$$

With a balanced experimental design, the two predictors $(P_{i..} - P_{...})$ and $(P_{ij.} - P_{i..})$ have a zero covariance. β_1 and β_2, therefore, are simple regression coefficients and have been computed previously.

$$\beta_1 = \frac{l+1}{l}\, 1/2\, \sigma^2_{A_L} \bigg/ \left(\frac{l+1}{l}\sigma^2_{A_L} + \sigma^2_e/b\right)$$

and

$$\beta_2 = 1/2\, \sigma^2_{A_L} / (\sigma^2_{A_L} + \sigma^2_e/l)\,.$$

The genetic advance will be, assuming $\frac{p-1}{p} \sim 1$:

$$\Delta G_I = 2\,i\,\sqrt{\mathrm{var}\,\hat{A}}$$

$$= i\,\sigma^2_{A_L} \sqrt{\frac{[(l+1)/l]^2}{\frac{l+1}{l}\sigma^2_{A_L} + \sigma^2_e/b\,l} + \frac{(l-1)/l}{\sigma^2_{A_L} + \sigma^2_e/b}}$$

In terms of h^2:

$$\Delta G_I = i\,h\,\sigma_{A_L} \sqrt{\frac{[(l+1)/l]^2}{[h^2(l-1)+2]/l} + \frac{(l-1)/l}{(2-h^2)}}$$

So:

$$\Delta G_I / \Delta G_L = i \sqrt{\frac{[(l+1)/l]^2}{[h^2(l-1)+2]/l} + \frac{(l-1)/l}{(2-h^2)}} \bigg/ \sqrt{2}$$

For these situations, the selection index, as expected, appears more efficient for low heritabilities (Table 3). However, as we have shown previously, heritability at the level of means ($h^2_{\bar{x}}$) of 0.40 corresponds to a heritability of 0.10 at the individual level, with three repetitions and ten individuals per plot and an environmental correlation between individuals of the same plot of 0.2–0.4. For $h^2 > 0.4$. and $l < 4$, the index method will not be greatly superior to the method with $l = 1$. Furthermore the index method, as the previous method, will lead to a quicker decrease in the genetic base of the breeding population (due to the selection of "sister" lines).

Conclusion

As expected for low heritabilities we have:

Index family-individual/line family > selection between and within families > individual selection.

But this relation appears only for very low heritability at the individual level. For realistic heritabilities at the level of means it does not appear to be a clear advantage of the index method over individual line selection. For the combined selection between and within family, due to the difficulty of optimizing between and within selection intensities, the genetic advance can be lower than with individual line selection, particularly for higher heritabilities. Both index and combined selection lead to a quicker decrease of the effective size of the breeding population. From this point of view, recurrent selection with only one DH line per plant is preferred. This method appears as one of the best in a large range of realistic situations. However with very low heritability (e.g. h^2 at the plot level less than 0.06), it will be better to have three or four lines per plant and use selection index.

References

Gallais A (1978) Le concept de valeur en lignées et son utilisation possible en amélioration des plantes. Ann Amélior Plant 29:1–22

Gallais A (1979) The concept of varietal ability in plant breeding. Euphytica 28:811–823

Gallais A (1986) Place de l'haplodiploidisation dans les schémas de sélection. Le Sélectionneur français 36:47–58

Gallais A (1988) A method of line development using doubled haploids: the single double haploid descent recurrent selection. Theor Appl Genet 75:330–332

Lush JL (1943) Family merit and individual merit as bases for selection. Am Nat 81:241–261, 362–379