Plant Breeding **128**, 213—216 (2009)

doi:10.1111/j.1439-0523.2008.01570.x

Short Communication

Optimum allocation of test resources and relative efficiency of alternative procedures of within-family selection in hybrid breeding

C. F. H. Longin[1], H. P. Maurer[1,2], A. E. Melchinger[1,4] and M. Frisch[1,3]

[1]Institute of Plant Breeding, Seed Science and Population Genetics, University of Hohenheim, Fruwirthstr 21, D-70593 Stuttgart, Germany; [2]Present address: State Plant Breeding Institute, University of Hohenheim, Fruwirthstr 21, D-70593 Stuttgart, Germany; [3]Present address: Institute of Agronomy and Plant Breeding II, Justus-Liebig-University Giessen Heinrich-Buff-Ring 26-38, D-35392 Giessen, Germany; [4]Corresponding author, E-mail: melchinger@uni-hohenheim.de

With 1 figure and 1 table

Received April 10, 2008/Accepted August 3, 2008
Communicated by W. E. Weber

Abstract

Selection within families can be conducted as family deviation (FDS) or strict within-family selection (WFS). Our objectives were to (i) investigate two breeding schemes combining selection among families with FDS or WFS and (ii) compare the optimum allocation of test resources for these breeding schemes. We focused on selection among S_1 families and doubled haploid (DH) lines within S_1 families and used Monte Carlo simulations to determine the selection gain ($\Delta\hat{G}$), its standard deviation ($SD_{\Delta\hat{G}}$), and the average coefficient of coancestry among the selected DH lines ($\overline{\Theta}$). For breeding schemes focusing only on within-family selection, as employed in animal breeding, the maximum $\Delta\hat{G}$, its $SD_{\Delta\hat{G}}$, and $\overline{\Theta}$ were larger for FDS than for WFS. However, for breeding schemes combining among- and within-family selection, as employed commonly in plant breeding programmes, the maximum $\Delta\hat{G}$, $SD_{\Delta\hat{G}}$, and $\overline{\Theta}$ were almost equal for FDS and WFS. Furthermore, the optimum allocation of test resources was similar for FDS and WFS. We conclude that FDS and WFS are equally suited for short- and long-term success in breeding schemes where among-family selection is followed by within-family selection.

Key words: selection gain — Monte Carlo simulation — S_1 family — DH line

Several approaches for selection among and within families have been discussed in the literature (cf. Wricke and Weber 1986, Falconer and Mackay 1996). Selection between families is usually based on the family mean. For selection within families, two alternative procedures were reported (Dempfle 1990, Hill et al. 1996). In strict within-family selection (WFS), the best lines are selected from each family based on their rank within the family. Under family deviation selection (FDS), the deviation of each line from its respective family mean is determined in absolute units and those lines with largest deviations are selected from the entire population of lines. However, in the classic textbooks of quantitative genetics (Wricke and Weber 1986, Falconer and Mackay 1996), either no clear distinction is made between FDS and WFS or a wrong formula for the selection intensity within families is used for FDS (Hill et al. 1996). In animal breeding, a 5–8% larger selection gain and an up to 50% smaller effective population size were reported for FDS compared with WFS (Hill et al. 1996). However, this study compared FDS with WFS performing selection only among lines without prior selection among families, which is unrealistic for plant breeding. Consequently, our objectives were to (i) investigate the efficiency of breeding schemes comprising a first selection step among S_1 families and a second selection step with FDS or WFS among doubled haploid (DH) lines and (ii) compare the optimum allocation of test resources under these schemes.

Monte Carlo simulations were used to compare the maximum selection gain of two hybrid breeding schemes. Both breeding schemes focused on one-stage testcross selection among and within families originating from one cross of two homozygous lines. Regarding a promising breeding scheme identified by Longin et al. (2007), S_1 families and DH lines within S_1 families were considered as test candidates. The base population comprised N_F S_1 families, each consisting of N_L DH lines, which were evaluated in P test locations with one replication. Breeding scheme 1, described by Hill et al. (1996), focused only on selection within families and was performed either as FDS (BS1-FDS) or as WFS (BS1-WFS). Thus, selection is carried out only within and not among S_1 families. In breeding scheme 2, selection was carried out in two steps. In the first step, the best S_1 families were selected according to their family mean averaged across all N_L DH lines within an S_1 family. In the second step, the best DH lines were selected from the preselected S_1 families with FDS (BS2-FDS) or WFS (BS2-WFS). The allocation of test resources N_F, N_L, and P was considered optimum if it maximized the selection gain for the investigated scenario.

We assumed selection for a quantitative trait like yield and, hence, employed normally distributed genotypic and phenotypic values. The phenotypic value of a line was:

$$Y = f + m_f + l + m_l,$$

with

$$f \sim N(0, \sigma_F^2),$$

$$l \sim N(0, \sigma_L^2),$$

$$m_f \sim N(0, \sigma_{F\times y}^2 + \sigma_{F\times p}^2 \big/ P + \sigma_{F\times y\times p}^2 \big/ P),$$

and

$$m_l \sim N(0, \sigma_{L\times y}^2 + \sigma_{L\times p}^2 \big/ P + \sigma_{L\times y\times p}^2 \big/ P + \sigma_e^2/(PR)),$$

where f and l are the effects of the families and lines, respectively, and m_f and m_l the effects masking the former effects. σ_F^2 is the genotypic variance among S_1 families and σ_L^2 the genotypic variance among DH lines within S_1 families. $\sigma^2_{F\times y}$, $\sigma^2_{F\times p}$, $\sigma^2_{F\times y\times p}$ are the interaction variances of S_1 families with years, locations, and years × locations, $\sigma^2_{L\times y}$, $\sigma^2_{L\times p}$, $\sigma^2_{L\times y\times p}$ the interaction variances of DH lines with years, locations, and years × locations, and σ^2_e is the plot error variance. In the absence of epistasis, the variances for testcross performance are $\sigma^2_F = \sigma^2_L = \sigma^2_g/2$ (Melchinger 1988).

Selection gain ($\Delta\hat{G}$) was determined as the difference between the mean genotypic value of the selected fraction and the mean genotypic value of the base population. The average coefficient of coancestry among the selected DH lines ($\overline{\Theta}$) was estimated by averaging the coefficient of coancestry of all possible pairwise crosses among the selected DH lines avoiding reciprocals. We used 25 000 simulation runs per investigated scenario to warrant a high numerical accuracy of the simulation results. We assumed a budget of $B = N_F \times N_L \times P$ = 200, 1000, and 5000 testcross plot equivalents, the selection of the best N_s = 4 or 25 DH lines, and variance components of σ^2_F; $\sigma^2_{F\times y}$; $\sigma^2_{F\times p}$; $\sigma^2_{F\times y\times p}$; σ^2_L; $\sigma^2_{L\times y}$; $\sigma^2_{L\times p}$; $\sigma^2_{L\times y\times p}$; σ^2_e = 0.5; 0.1; 0.1; 0.2; 0.5; 0.1; 0.1; 0.2; 1 (VC1), = 0.5; 0.25; 0.25; 0.5; 0.5; 0.25; 0.25; 0.5; 2 (VC2), and = 0.5; 0.625; 0.625; 1.25; 0.5; 0.625; 0.625; 1.25; 5 (VC3), respectively. To assess the optimum composition of finally selected test candidates (N_s), we investigated all possible integer combinations of test candidates for N_s = 4 and 25, i.e. the number of selected S_1 families × selected DH lines within S_1 families.

Results and Discussion

For the final selection of one S_1 family with several DH lines, the selection procedures FDS and WFS were by definition equal (cf. Fig. 1a). However, for the final selection of the best DH lines from several S_1 families, $\Delta\hat{G}$ was different for FDS and WFS as well as for BS1 and BS2. In BS1, maximum $\Delta\hat{G}$ was up to 6%, its $SD_{\Delta\hat{G}}$ up to 40%, and $\overline{\Theta}$ up to 10% larger for FDS compared with WFS (Table 1). These findings were in accordance with the results for selection within full-sib families (Hill et al. 1996). Thus, in BS1 the use of FDS maximizes short-term success, while the smaller $\overline{\Theta}$ for WFS compared to FDS suggests an advantage of WFS for long-term success. Breeding scheme BS1 may be of interest in animal breeding, but not in plant breeding. Therefore, we focus on breeding scheme BS2 in the further discussion.

Maximum $\Delta\hat{G}$ as well as its $SD_{\Delta\hat{G}}$ were almost equal for BS2-FDS and BS2-WFS, regardless of whether the finally selected DH lines originated from the same or from different S_1 families (Table 1, Fig. 1a). For instance, for $N_s = 4 \times 1$, VC2, and a budget of 1000 testcross plot equivalents, we determined $\Delta\hat{G}$ = 1.60 and $SD_{\Delta\hat{G}}$ = 0.40 for BS2-FDS and BS2-WFS. This can be explained by the large contribution of the among-family selection to the total $\Delta\hat{G}$. This contribution is similar for FDS and WFS and results from the higher heritability for among-family selection compared with within-family selection (cf. Wricke and Weber 1986). Consequently, for BS2, a large number of N_F at the expense of a low number of N_L was optimum. With decreasing N_L due to increasing N_F, however, the difference in $\Delta\hat{G}$ and $SD_{\Delta\hat{G}}$ for FDS and WFS was reduced (Fig. 1b).

Furthermore, values of $\overline{\Theta}$ were almost equal for BS2-FDS and BS2-WFS, which is in contrast to findings for BS1 (Table 1). For FDS, higher values of $\overline{\Theta}$ were expected compared with WFS, because of the distribution of finally selected DH lines across S_1 families. In WFS, an equal number of DH lines per S_1 family is selected. In contrast, in FDS, the DH lines with largest deviation from the family means are selected across all families. Thus, the S_1 families contribute different numbers of DH lines to the selected fraction increasing $\overline{\Theta}$ in comparison with WFS. The increase of $\overline{\Theta}$ in FDS compared to WFS increased with decreasing N_F and increasing N_L. However, the large optimum N_F and low optimum N_L in BS2 reduced the difference in $\overline{\Theta}$ for BS2-FDS and BS2-WFS. Thus, short- and long-term success of BS2-FDS and BS2-WFS are expected to be similar.

In all scenarios of BS2, the greatest $\Delta\hat{G}$, its smallest $SD_{\Delta\hat{G}}$, and the smallest $\overline{\Theta}$ were determined for the final selection of several S_1 families each with one DH line, i.e. $N_s = 4 \times 1$ or $N_s = 25 \times 1$ (Table 1). This is due to the fact that for these values of N_s, the optimum number of N_F was maximized in comparison with other values of N_s. Consequently, N_L = 1 was optimum for $N_s = 4 \times 1$ or $N_s = 25 \times 1$, and breeding schemes BS2-FDS and BS2-WFS represent poor selection among S_1 families. For $N_s = 2 \times 2$ or $N_s = 5 \times 5$, poor selection among but not within S_1 families was less efficient than BS2 due to the inability to use the gain from selection among DH lines within S_1 families (data not shown). Nevertheless, the optimum N_F was clearly larger than N_L also for $N_s = 2 \times 2$ or $N_s = 5 \times 5$ emphasizing the importance of family selection for the considered breeding situations.

Response curves of $\Delta\hat{G}$ as a function of the number of test locations or S_1 families and the optimum allocation of test

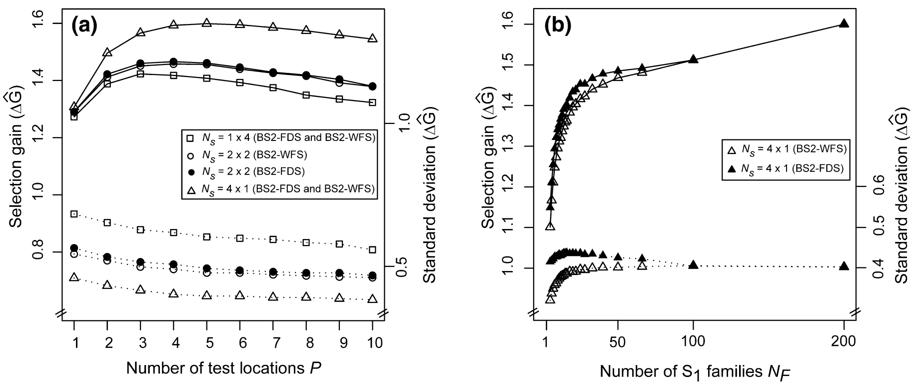

Fig. 1: Selection gain (solid line) and its standard deviation (dotted line) in breeding schemes BS2-FDS and BS2-WFS as a function of the number of (a) test locations and (b) S_1 families, assuming varying compositions of the finally selected number of S_1 families × DH lines within selected S_1 families (N_s). Note, BS2-FDS and BS2-WFS were identical for $N_s = 1 \times 4$ and $N_s = 4 \times 1$ in (a), and, thus, only one response curve was plotted. Similar trends were obtained for the response curves of $N_s = 25 \times 1$, $N_s = 5 \times 5$, and $N_s = 1 \times 25$ and, thus, these curves were not shown to make a clear illustration

Table 1: Optimum allocation of test resources maximizing selection gain ($\Delta\hat{G}$) in breeding schemes either (1) with family-deviation selection (BS1-FDS) or strict within-family selection (BS1-WFS) or (2) combining family selection with FDS (BS2-FDS) or WFS (BS2-WFS) (VC, set of variance components; P, N_F, N_L, optimum number of test locations, S_1 families, and DH lines within S_1 families, respectively; $SD_{\Delta\hat{G}}$, standard deviation of $\Delta\hat{G}$; $\overline{\Theta}$, average coefficient of coancestry among the selected DH lines)

	Assumptions			Optimum allocation					
Selection scheme	N_s	VC	Budget	P	N_F	N_L	$\Delta\hat{G}$	$SD_{\Delta\hat{G}}$	$\overline{\Theta}$
Breeding scheme 1									
BS1-FDS	5 × 5[1]	2	1000	4	5[2]	50	0.72	0.16	0.55
BS1-FDS	25 × 1	2	1000	3	25	13	0.70	0.17	0.51
BS1-FDS	4 × 1	2	1000	6	4	41	1.04	0.41	0.56
BS1-WFS	5 × 5	2	1000	4	5	50	0.71	0.12	0.54
BS1-WFS	25 × 1	2	1000	3	25	13	0.65	0.13	0.50
BS1-WFS	4 × 1	2	1000	6	4	41	0.99	0.29	0.50
Breeding scheme 2									
BS2-FDS	5 × 5[1]	2	1000	2	45	11	1.07	0.28	0.54
BS2-FDS	25 × 1	2	1000	3	333	1	1.15	0.17	0.50
BS2-FDS	2 × 2	2	1000	4	35	7	1.47	0.50	0.60
BS2-FDS	4 × 1	2	1000	5	200	1	1.60	0.40	0.50
BS2-WFS	5 × 5	2	1000	2	33	15	1.07	0.28	0.54
BS2-WFS	25 × 1	2	1000	3	333	1	1.15	0.17	0.50
BS2-WFS	2 × 2	2	1000	4	50	5	1.46	0.49	0.58
BS2-WFS	4 × 1	2	1000	5	200	1	1.60	0.40	0.50
BS2-FDS = BS2-WFS	4 × 1	1	1000	3	333	1	1.94	0.37	0.50
BS2-FDS = BS2-WFS	4 × 1	3	1000	8	125	1	1.20	0.44	0.50
BS2-FDS = BS2-WFS	4 × 1	2	200	3	66	1	1.18	0.42	0.50
BS2-FDS = BS2-WFS	4 × 1	2	5000	8	625	1	1.99	0.38	0.50

[1] Number of finally selected S_1 families × DH lines within selected S_1 families.

[2] In BS1, fixed values for N_F were used, because selection is carried out only within S_1 families.

resources were similar for FDS and WFS (Fig. 1, Table 1). For instance, for $N_s = 4 \times 1$, VC2, and a budget of 1000 testcross plot equivalents, the optimum allocation of test resources comprised $P = 5$, $N_F = 200$, and $N_L = 1$ in BS2-FDS and BS2-WFS. However, choice of $P = 4$ instead of the optimum $P = 5$ reduced $\Delta\hat{G}$ only to a small extent, if the number of N_F was increased in parallel. This is due to the flat response curves of $\Delta\hat{G}$ as a function of the number of test locations in the vicinity of their maximum (Fig. 1a). For the final selection of only one DH line per S_1 family, response curves of $\Delta\hat{G}$ increased continuously with increasing number of S_1 families (Fig. 1b). This can be explained by the requirement of large values of N_F to make maximum use of the higher heritability for among-family selection compared with within-family selection.

The optimum allocation of test resources, $\Delta\hat{G}$ and its $SD_{\Delta\hat{G}}$, as well as $\overline{\Theta}$ were similar for FDS and WFS across a large range of budgets, variance components, and compositions of finally selected lines (Table 1). Variance components were chosen based on analyses of variance of grain yield in (i) testcross series of sugar beet of Central European breeding material (Borchardt and Geiger 1997) and recent official rye variety performance tests in Germany (VC1; F. Laidig, personal communications), (ii) DH populations in maize programmes of Central European breeding companies (VC2; Gordillo and Geiger 2004) and (iii) testcross series of rapeseed of Central European material (VC3; J. Möhring, personal communications). Thus, the similarity of BS2-FDS and BS2-WFS seems to be consistent across several crops and numerous breeding situations.

In conclusion, the relative superiority of FDS over WFS found by Hill et al. (1996), analysing only BS1, does not hold true for BS2. In applied plant breeding, selection among and within families is usually performed simultaneously, like in BS2. Hence, both procedures of within-family selection, FDS and WFS, can be used without affecting $\Delta\hat{G}$, its $SD_{\Delta\hat{G}}$, or the optimum allocation of test resources in short- and long-term selection. However, a careful choice of the number of finally selected families and lines within families as well as the weights for family- and within-family selection is necessary to maximize the progress from selection.

Acknowledgements

This research was conducted within the breeding and informatics (BRAIN) project of the Genome Analysis of the Plant Biological System (GABI) initiative (http://www.gabi.de). We thank Prof. Dr H. F. Utz, Dr F. Laidig, and J. Möhring for valuable suggestions. The authors appreciate the editorial work of Dr J. Muminovic and the helpful comments of three anonymous reviewers.

References

Borchardt, D. C., and H. H. Geiger, 1997: Theoretical studies on optimizing sugar beet breeding plans. In: P. Krajewski, and Z. Kaczmarek (eds), Proceedings Meeting of the EUCARPIA Section Biometrics in Plant Breeding, 61—66. Poznan, Poland.

Dempfle, L., 1990: Statistical aspects of design of animal breeding programs: a comparison among various selection strategies. In: D. Gianola, and K. Hammond (eds), Advances in Statistical Methods for Genetic Improvement of Livestock, 98—117. Springer-Verlag, Berlin.

Falconer, D. S., and T. F. C. Mackay, 1996: Introduction to Quantitative Genetics, 4th edn, 232—234. Longman Scientific & Technical Ltd, Essex.

Gordillo, A., and H. H. Geiger, 2004: Estimating quantitative-genetic parameters of European maize populations to optimize hybrid breeding methods by model calculations. In: Poster Abstract XVIIth EUCARPIA General Congress, Tulln, Austria.

Hill, W. G., A. Caballero, and L. Dempfle, 1996: Prediction of response to selection within families. Genet. Sel. Evol. **28,** 379—383.

Longin, C. F. H., H. F. Utz, J. C. Reif, T. Wegenast, W. Schipprack, and A. E. Melchinger, 2007: Hybrid maize breeding with doubled haploids: III. efficiency of early testing prior to doubled haploid production in two-stage selection for testcross performance. Theor. Appl. Genet. **115,** 519—527.

Melchinger, A. E., 1988: Means, variances, and covariances between relatives in hybrid populations with disequilibrium in the parent population. In: B. S. Weir, E. J. Eisen, M. M. Goodman, and G. Namkong (eds), Proc. Second Int. Conf. Quantit. Genet., 400—415. Sinauer, Sunderland, MA.

Wricke, G., and W. E. Weber, 1986: Quantitative Genetics and Selection in Plant Breeding, 214—216, 293—295. Walter de Gruyter, Berlin, New York.