Theor Appl Genet (2006) 112: 903–912
DOI 10.1007/s00122-005-0192-z

ORIGINAL PAPER

C. Friedrich H. Longin · H. Friedrich Utz
Jochen C. Reif · Wolfgang Schipprack
Albrecht E. Melchinger

Hybrid maize breeding with doubled haploids: I. One-stage versus two-stage selection for testcross performance

Received: 29 July 2005 / Accepted: 2 December 2005 / Published online: 25 January 2006

Abstract Optimum allocation of resources is of fundamental importance for the efficiency of breeding programs. The objectives of our study were to (1) determine the optimum allocation for the number of lines and test locations in hybrid maize breeding with doubled haploids (DHs) regarding two optimization criteria, the selection gain ΔG_k and the probability P_k of identifying superior genotypes, (2) compare both optimization criteria including their standard deviations (SDs), and (3) investigate the influence of production costs of DHs on the optimum allocation. For different budgets, number of finally selected lines, ratios of variance components, and production costs of DHs, the optimum allocation of test resources under one- and two-stage selection for testcross performance with a given tester was determined by using Monte Carlo simulations. In one-stage selection, lines are tested in field trials in a single year. In two-stage selection, optimum allocation of resources involves evaluation of (1) a large number of lines in a small number of test locations in the first year and (2) a small number of the selected superior lines in a large number of test locations in the second year, thereby maximizing both optimization criteria. Furthermore, to have a realistic chance of identifying a superior genotype, the probability P_k of identifying superior genotypes should be greater than 75%. For budgets between 200 and 5,000 field plot equivalents, $P_k > 75\%$ was reached only for genotypes belonging to the best 5% of the population. As the optimum allocation for $P_k(5\%)$ was similar to that for ΔG_k, the choice of the optimization criterion was not crucial. The production costs of DHs had only a minor effect on the optimum number of locations and on values of the optimization criteria.

Keywords Optimum allocation · Selection gain · Probability · Superior genotype · Monte Carlo simulation

Communicated by H. Becker

C. Friedrich H. Longin and H. Friedrich Utz contributed equally to this work.

C. F. H. Longin · H. F. Utz · J. C. Reif · W. Schipprack
A. E. Melchinger (✉)
Institute of Plant Breeding, Seed Science, and Population Genetics, University of Hohenheim, 70593 Stuttgart, Germany
E-mail: melchinger@uni-hohenheim.de
Fax: +49-711-4592343

Introduction

Optimum allocation of financial and breeding resources is of fundamental importance for the efficiency of breeding programs and selection strategies. Advances in the production of doubled haploids (DHs) by in vivo haploid induction (Bordes et al. 1997; Röber 1999) offer a promising alternative to recurrent selfing for rapid inbred line development in hybrid maize breeding. Currently, DHs are adopted as a routine method in commercial maize breeding programs in North America (Seitz 2005) and Europe (Schmidt 2004). Their efficient use requires an optimization of the entire breeding scheme in order to maximize progress from selection.

A selection strategy may involve one or several stages of selection. In the latter case, the initial population of lines is evaluated in 1 year and a superior subset is selected for further evaluation and selection in subsequent year(s). To quantify the progress from k selection stages, various criteria have been used such as (1) the selection gain (ΔG_k) (Cochran 1951; Utz 1969) and (2) the probability of identifying superior genotypes (P_k) (Keuls and Sieben 1955; Robson et al. 1967; Johnson 1989; Knapp 1998). In recurrent selection, ΔG_k represents the most widely used criterion to compare different methods and optimize the selection progress in population improvement (cf. Choo and Kannenberg 1988; Gallais 1991). For a given population, ΔG_k is a function of the heritability (h^2) and selection intensity (i_α), and increases with larger values for both parameters (Bernardo 2002). Heritability increases with an increasing number of test locations, years, and replications in performance trials, whereas i_α depends on the selected fraction (α) and the

probability distribution of the lines. With a fixed number of finally selected lines (N_f), i_α increases with a larger number of initial lines. Hence, a plant breeder with a fixed budget has to find a compromise between (a) the number of initial lines and (b) the intensity of their testing as determined by the number of test locations, years, and replications. This requires an optimization of the test resources for each breeding scenario.

For ΔG_k the optimum allocation of test resources for a fixed budget was investigated with numerical integration, assuming an infinite population size (Utz 1969), and with stochastic simulations, assuming a finite population size (Finney 1966; Young 1976). However, production costs of DHs have so far not been taken into account. In addition, most studies on the optimum allocation of resources were conducted more than 30 years ago and the limited computing power available at that time restricted the number of scenarios considered.

In addition to medium- and long-term germplasm improvement, plant breeders are forced to focus on rapid development of competitive varieties. For the latter purpose, P_k represents a suitable criterion (Johnson 1989). For a given population and a fixed number of N_f, P_k is increased by (1) an increasing number of lines in order to have at least the number of desired superior genotypes in the initial sample and (2) an increasing h^2 to warrant a high probability of detecting them. For one-stage selection, Robson et al. (1967) and Johnson (1989) investigated the impact of h^2, the initial sample size, and N_f on the probability (P_1) that all N_f have genotypic values exceeding a given threshold. Knapp (1998) extended this approach to marker-assisted selection. Nevertheless, these studies investigated P_k only for given values of h^2 and α, disregarding the optimum allocation of resources. Furthermore, ΔG_k and P_k have not yet been compared for one- and two-stage selection.

In this study, we optimized the allocation of test resources in hybrid maize breeding with DHs under one- and two-stage selection for testcross performance with a given tester by using Monte Carlo simulations. For different assumptions regarding the budget, ratio of variance components, and value of N_f, we (1) determined the optimum allocation of the number of lines and test locations for ΔG_k and P_k, (2) compared both optimization criteria including their standard deviations, and (3) investigated the influence of production costs of DHs on the optimum allocation of test resources.

Materials and methods

Selection strategies

In a standard maize breeding scheme (Fig. 1), a total of N_1 DH lines generated from one or several F_1 crosses via in vivo haploid induction are available at the beginning of the evaluation and selection process. A certain

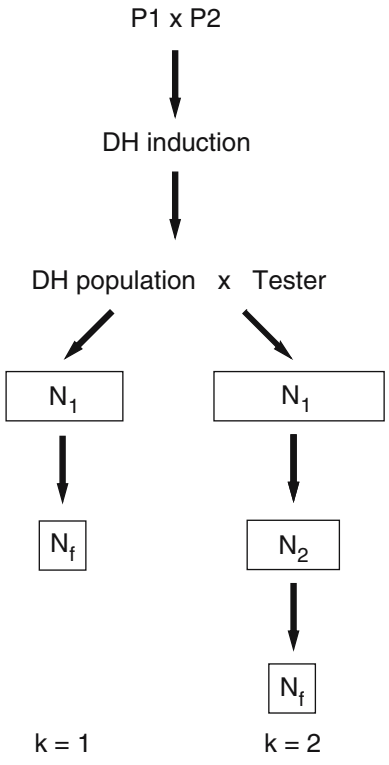

Fig. 1 Hybrid maize breeding scheme with production of doubled haploid (*DH*) lines, their testcross progenies and testcross evaluation in several test locations with one-stage ($k=1$) or two-stage selection ($k=2$). (N_1 = number of initial lines; N_2 = subset of superior lines selected after the first stage of two-stage selection; N_f = number of finally selected lines)

number N_f of phenotypically best DH lines are selected. We compared $N_f=1$ and $N_f=5$. The target variable Y is the genotypic value of testcross performance with a given tester T for a certain trait or index of traits. The tester can be any population with an arbitrary structure such as an inbred line, single cross, or random mating population. With one-stage selection, selection is based on field tests in a single year. With two-stage selection, field tests are conducted in 2 years with a subset of the most superior lines N_2 selected after the first year being evaluated in the second year. At stage j ($j=1, 2$), selection among N_j DH lines is based on variable X_j, the phenotypic mean of testcross performance at this stage with tester T evaluated in L_j locations with R_j replications. At stage $j=2$, the selection among lines could alternatively be based on an index of their performance in the first and second year. However, this would affect the optimum allocation and the selection gain only marginally (Utz 1969; Young 1976). Without an upper limit on L_j, $R_j=1$ is optimal regarding ΔG_k (Sprague and Federer 1951; Utz 1969). Thus, we set $R_j=1$.

Economic frame and quantitative-genetic parameters

We investigated three assumptions ($C=0$, 0.5, 1) concerning the production cost of one DH line relative to

the cost of one field plot for evaluating testcross progenies. For instance, $C = 0.5$ means that the production cost of one DH line is equal to half the cost of one field plot. $C = 0.5$ corresponds to the actual costs of DH production in breeding companies most advanced in the DH technique (Seitz, personal communication). $C = 1$ is a realistic assumption at the beginning of establishing the DH technique in a breeding program. With further improvements in the DH technique, the costs of DH production may become negligible in the future ($C = 0$).

A fixed total budget B for (1) producing the DH lines and (2) evaluating their testcross progenies in k selection stages was defined in terms of testcross plot equivalents as

$$B = N_1 C + \sum_{j=1}^{k} N_j L_j R_j \quad (1)$$

assuming equal plot sizes in all selection stages. We compared three budgets with $B = 200$, 1,000, 5,000 plot equivalents. An overview of the notation used throughout this treatise is given in Table 1.

Three ratios of variance components (σ_g^2:σ_{gl}^2:σ_{gy}^2:σ_{gly}^2:σ_e^2) were considered, where σ_g^2 refers to the genotypic variance, σ_{gl}^2 to the variance of genotype × location interactions, σ_{gy}^2 to the variance of genotype × year interactions, σ_{gly}^2 to the variance of genotype × location × year interactions, and σ_e^2 to the error variance. We set VC1 = 1:0.25:0.25:0.5:1, VC2 = 1:0.5:0.5:1:2, and VC3 = 1:1:1:2:4, resulting in a heritability on a plot basis of 0.33, 0.20, and 0.11, respectively. These ratios were chosen based on combined analyses of variance of grain yield in (1) recent official maize variety performance tests in Germany (VC1, Laidig, personal communication), (2) DH populations of commercial breeding programs (VC2, Gordillo and Geiger 2004), and (3) official maize variety performance tests in Southwest Germany (VC3, P. Herrmann, unpublished data).

Table 1 Notation used in this treatise

h^2	Heritability
i_α	Selection intensity for a certain selected fraction $\alpha = N_f / N_1$
j	Selection stage
B	Fixed total budget in field plot equivalents
C	Production costs of one DH line relative to the costs of one field plot for evaluating testcross progenies
DH	Doubled haploid
ΔG_k	Selection gain after k stages of selection
$\Delta\hat{G}_k$	ΔG_k estimated by Monte Carlo simulations
$\Delta\hat{G}_k^*$	Value of $\Delta\hat{G}_k$ at the corresponding optimum allocation (N_j^*, L_j^*)
N_f	Number of finally selected lines
N_j, L_j, R_j	Number of lines, locations, or replications at stage j in performance trials
N_j^*, L_j^*	Optimum number of lines and locations maximizing the optimization criterion in the set of admissible allocations
$P_k(q)$	Probability of identifying lines with genotypic values exceeding a fixed (100−q)% quantile of the corresponding normal distribution $N(0, \sigma_g^2)$ after k stages of selection
$\hat{P}_k(q)$	$P_k(q)$ estimated by Monte Carlo simulations
$\hat{P}_k^*(q)$	Value of $\hat{P}_k(q)$ at the corresponding optimum allocation (N_j^*, L_j^*)
VC	Ratio of variance components σ_g^2: σ_{gl}^2: σ_{gy}^2: σ_{gly}^2: σ_e^2

Simulation model

Genotypic and phenotypic values were generated separately for each combination of the above factors. Genotypic values were sampled from a normal distribution $N(0, \sigma_g^2)$. Non-genetic values were sampled from a normal distribution $N(0, \sigma_{m_j}^2)$, with

$$\sigma_{m_j}^2 = \sigma_{gy}^2 + \frac{\sigma_{gl}^2}{L_j} + \frac{\sigma_{gly}^2}{L_j} + \frac{\sigma_e^2}{L_j R_j} \quad (2)$$

representing the non-genetic variance. Phenotypic values were then generated by adding non-genetic values to the genotypic values. For two-stage selection, genotypic and phenotypic values were sampled out of a multivariate normal distribution MVN(μ, V) with $\mu^T = (0, 0, 0)$ and

$$V = \begin{pmatrix} \sigma_g^2 & \sigma_g^2 & \sigma_g^2 \\ \sigma_g^2 & \sigma_{x_1}^2 & \mathrm{cov}_{x_1x_2} \\ \sigma_g^2 & \mathrm{cov}_{x_1x_2} & \sigma_{x_2}^2 \end{pmatrix}. \quad (3)$$

The covariance between the phenotypic values at stage $j = 1$ and $j = 2$ was determined as $\mathrm{cov}_{x_1x_2} = \sigma_g^2 + (L_c\sigma_{gl}^2)/(L_1L_2)$, with L_c representing the number of locations common to both selection stages (Utz 1969). We assumed $L_c = L_1$. The two optimization criteria and their SDs were then calculated and stored. This procedure was repeated for each factor combination and choice of N_j and L_j, with a new set of realizations of random variables (further referred to as runs). The number of runs required to warrant an accuracy of 0.01 for the optimization criterion was calculated based on the standard error of the arithmetic mean as $(3\mathrm{SD}/0.01)^2$ (Berry and Lindgren 1996). Between 7,000 and 70,000 simulation runs were required for the different scenarios.

Optimum allocation and optimization criteria

An admissible allocation of test resources refers to tuples (N_j, L_j) for all stages j, such that Eq. 1 is satisfied. An element (N_j^*, L_j^*) is denoted as an optimum allocation if it maximizes the optimization criterion in the set of admissible allocations. For each run, the mean genotypic value of the N_f selected lines was calculated and the selection gain was estimated by averaging over all Monte Carlo runs for the allocation considered $(\Delta\hat{G}_k)$. The variance among these runs was used to calculate the

corresponding SD $\left(\mathrm{SD}_{\Delta\hat{G}_k}\right)$. In addition, the number of selected lines with genotypic values exceeding a fixed $(100-q)\%$ quantile of the corresponding normal distribution $N(0, \sigma_g^2)$ was determined for each run and divided by N_f. The probability of identifying superior genotypes was estimated by averaging these values over all Monte Carlo runs for the allocation considered $\left(\hat{P}_k(q)\right)$. The variance among these runs was used to calculate the corresponding SD $\left(\mathrm{SD}_{\hat{P}_k(q)}\right)$. We examined q values of 25, 5, 1, and 0.1%, with corresponding standardized genotypic thresholds of 0.67449, 1.64485, 2.32635, and 3.09023, respectively.

The optimum allocation of test resources for each scenario was obtained by a grid search in $\mathbf{Z}$, the space of admissible resource allocations. For instance, for $B=200$, $N_f=1$, VC1, $C=0$, and one-stage selection, the optimum choice of N_1 was determined by varying the selected fraction α_1 between 0.01 and 0.30 for each L_1 between one and a number that allowed a clear identification of the optimum of L_1. Thus, at least 100 calculations were performed to identify the optimum allocation for each scenario. Let OC represent the optimization criterion ΔG_k or $P_k(q)$. Let (N_j^o, L_j^o) be the allocation, where OC assumes its numerical maximum value in the simulations

$$\mathrm{OC}\left(N_j^o, L_j^o\right) = \max_{(N_j,L_j)\epsilon\mathbf{Z}} \mathrm{OC}\left(N_j, L_j\right). \quad (4)$$

Since OC is only estimated with a precision of 0.01, the optimum allocation (N_j^*, L_j^*) was determined following Utz (1969) such that the number of locations was minimum among all allocations within 0.01 drop-off of $\mathrm{OC}(N_j^o, L_j^o)$, i.e.,

$$L_j^* = \min_{L_j}\left\{\left(N_j, L_j\right)\epsilon\mathbf{Z} | \mathrm{OC}\left(N_j^o, L_j^o\right) - \mathrm{OC}\left(N_j, L_j\right) < 0.01\right\}. \quad (5)$$

The reason being that breeders prefer for technical reasons tests in fewer locations if this affects the OC only marginally.

The values of each optimization criterion at its corresponding optimum allocation (N_j^*, L_j^*) were denoted as $\Delta\hat{G}_k^*$ and $\hat{P}_k^*(q)$. Simulation programs were written in C and implemented in the statistical software R (R Development Core Team 2004).

Results

The optimization criteria were similarly affected by deviations from the optimum allocation of test resources for one- and two-stage selection, $N_f=1$ or 5, and production costs of DHs. Thus, only response curves for $\Delta\hat{G}_1$ and $\hat{P}_1(1\%)$ as a function of L_1 were presented for varying budgets and ratio of variance components assuming one-stage selection, $N_f=1$, and $C=0.5$ (Fig. 2). With increasing L_1, the optimization criteria $\Delta\hat{G}_1$ and $\hat{P}_1(1\%)$ increased up to an optimum and decreased slightly thereafter. Both response curves were flat in the vicinity of the maximum. The increase in $\Delta\hat{G}_1$ and $\hat{P}_1(1\%)$ was largest between $L_1=1$ and $L_1=4$. Curves for $\mathrm{SD}_{\hat{P}_1(1\%)}$ displayed similar trends as those for $\hat{P}_1(1\%)$, with a maximum at the optimum allocation of $\hat{P}_1(1\%)$. In contrast, curves for $\mathrm{SD}_{\Delta\hat{G}_k}$ decreased with increasing L_1.

The consequences of one-stage versus two-stage selection, varying N_f, and budgets on the optimum allocation of test resources and optimization criteria were hardly affected by the ratio of variance components and production costs of DHs (data not shown). Hence, the results on the influence of the former group of factors were presented exemplarily for intermediate values VC2 and $C=0.5$ (Table 2). The optimum number of initial lines N_1^* and test locations for two-stage selection was about twice as large as for one-stage selection. This was due to the optimum allocation of two-stage selection, which comprised a large number of initial lines N_1^* tested in a small number of test locations L_1^* at the first stage, and a small number of selected lines N_2^* tested in a large number of test locations L_2^* at the second stage. Furthermore, under the same allocation of resources $\Delta\hat{G}_k^*$, and values of $\hat{P}_k(5\%)$, $\hat{P}_k(1\%)$, and $\hat{P}_k(0.1\%)$ were on average 20, 30, 50, and 80%, respectively, higher than for one-stage selection. Reducing N_f from five to one resulted in (1) smaller values of N_j^* but larger values of L_j^* in the last selection stage, and (2) an increase in $\Delta\hat{G}_k^*$ and corresponding values for $\hat{P}_k(5\%)$, $\hat{P}_k(1\%)$, and $\hat{P}_k(0.1\%)$ of 20, 30, 60, and 110%, respectively. However, SD of these estimates were also increased by more than 60% on average. For one-stage selection, increasing the budget from $B=200$ to $B=5{,}000$ resulted in a more than 10-fold increase in N_1^* and a twofold increase in L_1^*. For two-stage selection, N_1^* and N_2^* increased more than 15- and 5-fold, whereas L_1^* and L_2^* increased twofold and threefold, respectively. In addition, the average increase in $\Delta\hat{G}_k^*$ and corresponding values for $\hat{P}_k(5\%)$, $\hat{P}_k(1\%)$, and $\hat{P}_k(0.1\%)$ was 65, 125, 300, and 650%, respectively.

The influence of different ratios of variance components and production costs of DHs on the optimum allocation of test resources and optimization criteria was hardly affected by the number of selection stages, N_f, and budget. Therefore, representative results on the influence of both factors were given for two-stage selection, $B=1{,}000$, and $N_f=1$ (Table 3). An increase in the non-genetic variance from VC1 to VC3 resulted in a reduction in N_1^* and an increase in L_j^* for $\Delta\hat{G}_k^*$, $\hat{P}_k^*(5\%)$, and $\hat{P}_k^*(1\%)$. For $\hat{P}_k^*(0.1\%)$, N_1^* was also reduced with increasing non-genetic variance, but N_2^* increased and L_j^* was fairly stable. The optimum allocation of test resources based on the same VC but different optimization criteria differed largely for small values of q ($q=0.1\%$) and large non-genetic variance (VC3). For instance, for VC3 and $C=0.5$ the optimum number of lines N_j^* was approximately doubled and the optimum number of locations L_j^* was halved for

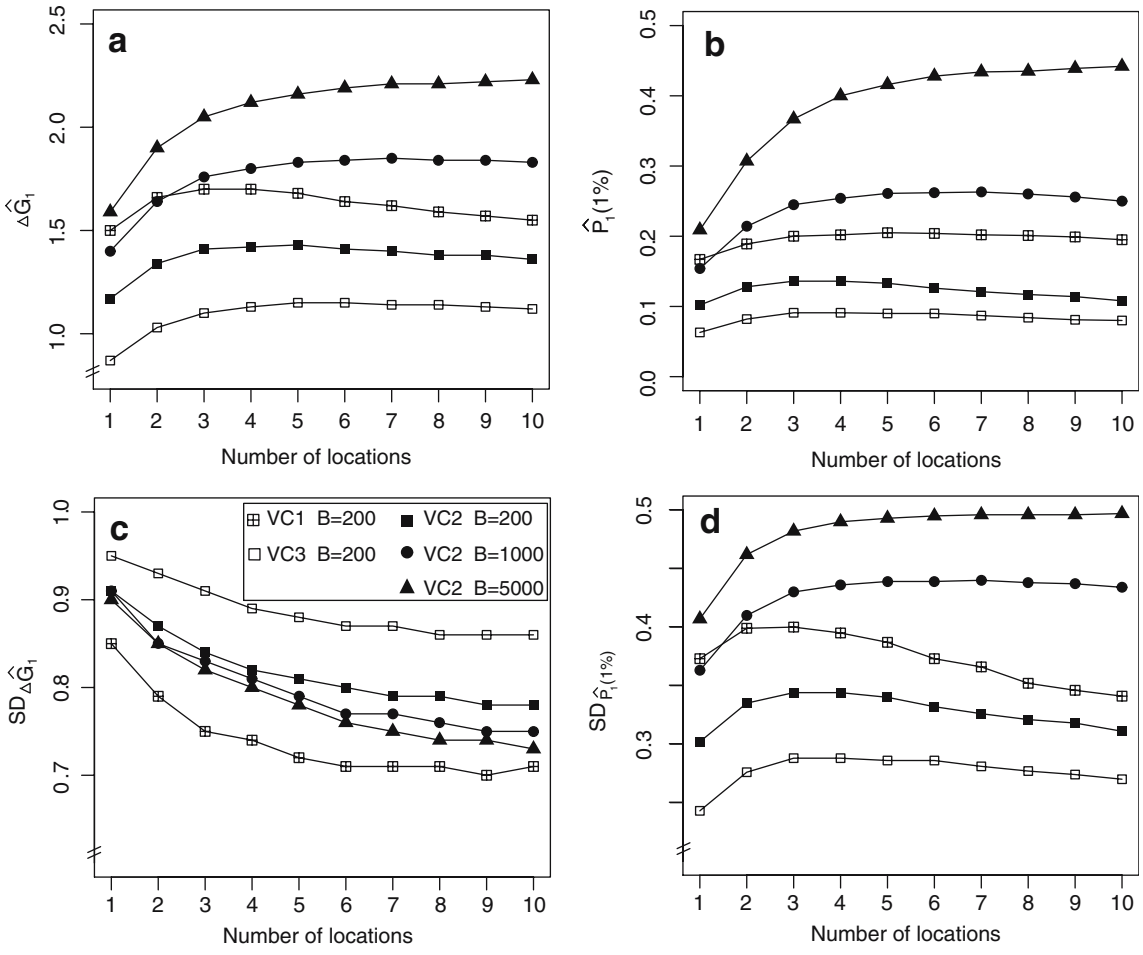

Fig. 2 **a** Selection gain $\Delta\hat{G}_1$, **b** probability $\hat{P}_1(1\%)$ of identifying one line with a genotypic value belonging to the 1% best genotypes of the population, **c**, **d** corresponding standard deviation $SD_{\Delta\hat{G}_1}$ and $SD_{\hat{P}_1(1\%)}$, respectively, as a function on the number of locations for one-stage selection assuming $C=0.5$, and $N_f=1$. For explanation of abbreviations, see Table 1

Table 2 Optimum allocation of test resources maximizing selection gain ($\Delta\hat{G}_k^*$), values of $\Delta\hat{G}_k^*$, and corresponding probabilities $\hat{P}_k(q)$ of identifying N_f lines with genotypic values belonging to the 5, 1, and 0.1% best genotypes of the population assuming $C=0.5$ and VC2. For explanation of abbreviations, see Table 1

Assumptions			Optimum allocation				Selection gain		Corresponding probabilities $\hat{P}_k(q)$					
k[a]	N_f	B	N_1^*	N_2^*	L_1^*	L_2^*	$\Delta\hat{G}_k^*$	SD[b]	$\hat{P}_k(5\%)$	SD[b]	$\hat{P}_k(1\%)$	SD[b]	$\hat{P}_k(0.1\%)$	SD[b]
1	1	200	44	–	4	–	1.42	0.81	0.39	0.49	0.13	0.34	0.02	0.14
1	1	1,000	133	–	7	–	1.85	0.76	0.60	0.49	0.27	0.44	0.05	0.22
1	1	5,000	588	–	8	–	2.22	0.74	0.78	0.42	0.44	0.50	0.12	0.32
1	5	200	57	–	3	–	1.08	0.36	0.25	0.20	0.07	0.11	0.01	0.04
1	5	1,000	222	–	4	–	1.52	0.36	0.43	0.23	0.16	0.16	0.03	0.07
1	5	5,000	769	–	6	–	1.92	0.35	0.64	0.22	0.30	0.21	0.06	0.11
2	1	200	93	10	1	6	1.68	0.78	0.52	0.50	0.21	0.40	0.04	0.19
2	1	1,000	298	17	2	15	2.20	0.70	0.79	0.41	0.42	0.49	0.10	0.30
2	1	5,000	1,560	50	2	22	2.64	0.67	0.94	0.06	0.68	0.22	0.25	0.19
2	5	200	90	16	1	4	1.25	0.37	0.31	0.21	0.09	0.13	0.01	0.05
2	5	1,000	461	44	1	7	1.80	0.35	0.58	0.23	0.24	0.20	0.04	0.09
2	5	5,000	1,502	83	2	15	2.30	0.32	0.84	0.17	0.48	0.23	0.12	0.15

[a] $k=1$, one-stage selection; $k=2$, two-stage selection
[b] SD = standard deviation of estimates among runs

$\hat{P}_k^*(0.1\%)$ in comparison with $\Delta\hat{G}_k^*$. In addition, $\Delta\hat{G}_k^*$, $\hat{P}_k^*(5\%)$, $\hat{P}_k^*(1\%)$, and $\hat{P}_k^*(0.1\%)$ were reduced by approximately 25, 35, 65, and 70%, respectively, with increasing non-genetic variance. For $C=1$ compared with $C=0$, N_j^* decreased about 50%, whereas L_j^* changed only slightly. The reduction in $\Delta\hat{G}_k^*$ and $\hat{P}_k^*(q)$ for

Table 3 Optimum allocation of test resources maximizing selection gain $\Delta\hat{G}_k^*$ or probability $\hat{P}_k^*(q)$ of identifying one line ($N_f = 1$) with a genotypic value belonging to the 5, 1, and 0.1% best genotypes of the population for two-stage selection assuming $B = 1{,}000$. For explanation of abbreviations, see Table 1

Assumptions		Optimum allocation					
VC	C	N_1^*	N_2^*	L_1^*	L_2^*	OC[a]	SD[b]
$\Delta\hat{G}_k^*$							
1[c]	0	739	29	1	9	2.59	0.63
	0.5	498	28	1	9	2.51	0.63
	1	396	23	1	9	2.44	0.64
2[d]	0	660	34	1	10	2.26	0.73
	0.5	298	17	2	15	2.20	0.70
	1	251	19	2	13	2.16	0.70
3[e]	0	346	22	2	14	1.89	0.80
	0.5	224	12	3	18	1.85	0.78
	1	199	12	3	17	1.82	0.78
$\hat{P}_k^*(5\%)$							
1	0	760	30	1	8	0.93	0.25
	0.5	480	31	1	9	0.92	0.27
	1	350	30	1	10	0.90	0.30
2	0	620	38	1	10	0.80	0.40
	0.5	293	19	2	14	0.78	0.41
	1	257	19	2	12	0.76	0.43
3	0	580	35	1	12	0.61	0.49
	0.5	271	23	2	14	0.60	0.49
	1	254	17	2	14	0.58	0.49
$\hat{P}_k^*(1\%)$							
1	0	739	29	1	9	0.66	0.47
	0.5	480	31	1	9	0.61	0.49
	1	392	27	1	8	0.56	0.50
2	0	690	31	1	10	0.46	0.50
	0.5	312	20	2	11	0.42	0.50
	1	257	19	2	12	0.40	0.49
3	0	667	37	1	9	0.28	0.45
	0.5	288	28	2	10	0.27	0.44
	1	271	17	2	11	0.26	0.44
$\hat{P}_k^*(0.1\%)$							
1	0	832	24	1	7	0.21	0.41
	0.5	538	32	1	6	0.17	0.38
	1	427	29	1	5	0.15	0.35
2	0	760	40	1	6	0.12	0.32
	0.5	526	35	1	6	0.10	0.30
	1	412	35	1	5	0.09	0.28
3	0	730	45	1	6	0.06	0.24
	0.5	480	35	1	8	0.06	0.23
	1	395	42	1	5	0.05	0.21

[a] OC = optimization criterion
[b] SD = standard deviation of estimates among runs
[c] VC1 = 1:0.25:0.25:0.5:1
[d] VC2 = 1:0.5:0.5:1:2
[e] VC3 = 1:1:1:2:4

$C = 1$ versus $C = 0$ was small, and ranged from $5\%\left(\Delta\hat{G}_k^*\right)$ to $30\%\left(\hat{P}_k^*(0.1\%)\right)$.

Discussion

The selection gain (ΔG_k) is the most widely used criterion to optimize selection processes, but an infinite sample size was assumed in most studies (Cochran 1951; Hanson and Brim 1963; Utz 1969; Tomerius 2001; Grüneberg et al. 2004). As breeding populations normally are relatively small, we determined ΔG_k for finite sample sizes. However, both assumptions result in similar optimum allocations and marginally reduced gains for the finite sample case (Cochran 1951; Finney 1966; Utz 1969). We compared ΔG_k with an alternative optimization criterion, the probability $P_k(q)$ of identifying superior genotypes. Both ΔG_k and $P_k(q)$ were estimated by Monte Carlo simulations for one- and two-stage selection, assuming a Gaussian normal distribution of (1) genotypic and (2) phenotypic values. Experimental verification of the latter assumption requires a large population size in view of the low power of statistical tests for deviations from a Gaussian normal distribution. However, an extremely extensive QTL mapping experiment in maize (Schön et al. 2004) with testcross

progenies of 976 F_5 lines evaluated in 19 locations provided no evidence that phenotypic means for yield deviated from a Gaussian normal distribution. Likewise, the large number of detected QTL with small effects resulted in an approximative Gaussian normal distribution of genotypic values due to the Central Limit Theorem (Schön et al. 2004). Nevertheless, further research is needed to check the assumptions on probability distributions. DH populations should be an excellent tool for this purpose, because natural selection during inbreeding is minimized, if selection during in vivo haploid induction can be neglected.

We chose an accuracy of 0.01 for the optimization criteria to limit the number of simulation runs to a manageable number. Increasing the accuracy up to 0.0001 would require 5,000,000–600,000,000 simulation runs. However, the length of the resulting optimum allocation interval for an accuracy of 0.01 is only a minor problem for practical breeding purposes due to the extremely flat response curves.

Comparison of optimization criteria

In a first step, we compare the two optimization criteria under the assumption of no non-genetic variance ($h^2 = 1$) and one-stage selection (Fig. 3), because two-stage selection offers advantages only for $h^2 < 1$. Our simulation results for $\Delta\hat{G}_1$ and $SD_{\Delta\hat{G}_1}$ were in harmony with means and standard deviations of order statistics (Pearson and Hartley 1972). For $\hat{P}_1(q)$ the results were in agreement with those reported by Robson et al. (1967, Appendix 6 and Table 2). Thus, our Monte Carlo simulations were sufficiently accurate to estimate ΔG_k and $P_1(q)$. Furthermore, simulations can provide estimates for $SD_{\Delta G_2}$, $P_2(q)$, $SD_{P_1(q)}$, and $SD_{P_2(q)}$, which were not reported in previous studies.

The response curves of both optimization criteria illustrated that the slopes decreased with an increasing number of lines (Fig. 3). This corroborates the well-known relationship that a linear increase in ΔG_1 requires an exponential increase in N_1 (Becker 1993). The choice

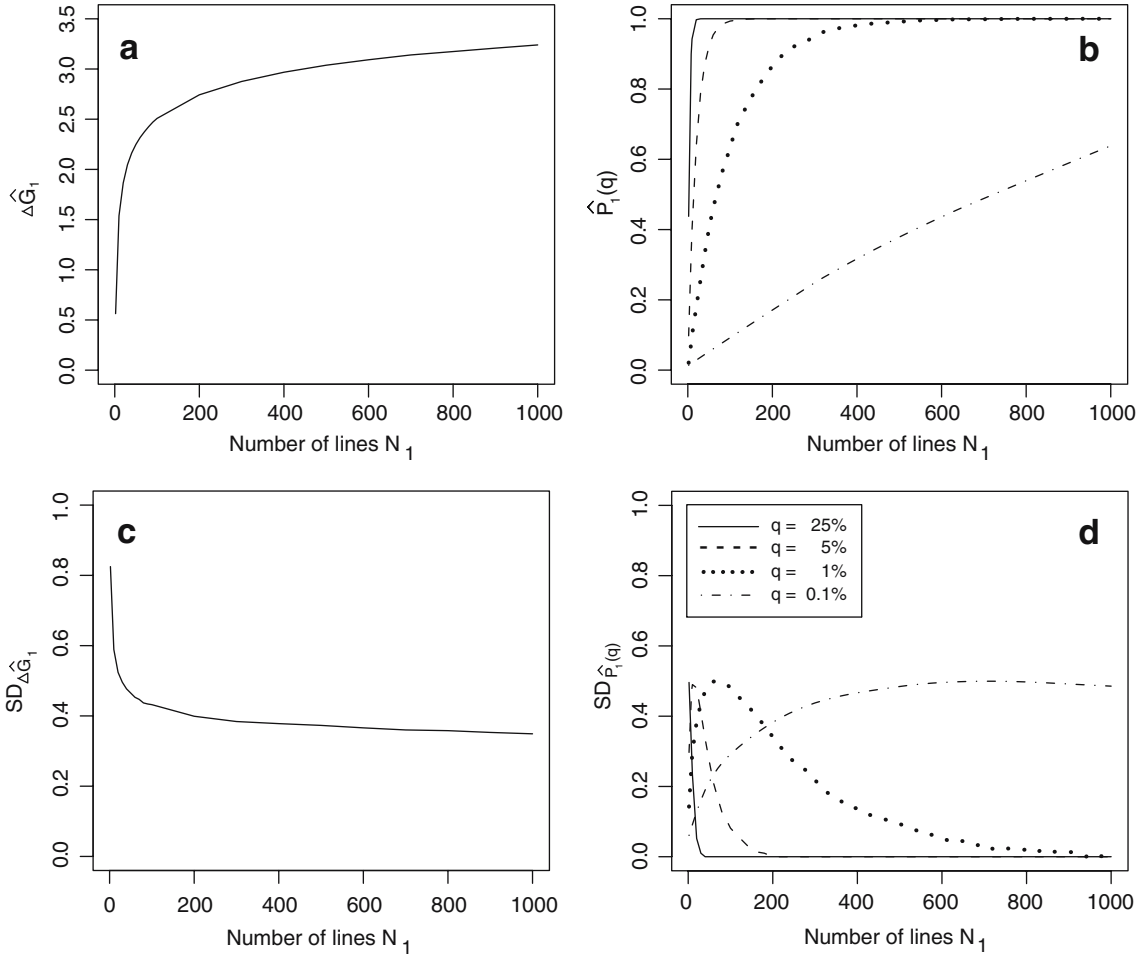

Fig. 3 **a** Selection gain $\Delta\hat{G}_1$, **b** probability $\hat{P}_1(q)$ of identifying one line with a genotypic value belonging to the q% best genotypes of the population, **c**, **d** corresponding standard deviation $SD_{\Delta\hat{G}_1}$ and $SD_{\hat{P}_1(q)}$, respectively, as a function of the number of lines assuming $h^2 = 1$. For explanation of abbreviations, see Table 1

of q had a strong influence on the curve of $\hat{P}_1(q)$, especially its slope. Response curves of $\hat{P}_1(5\%)$ and $\hat{P}_1(1\%)$ were similar in shape to the curve of $\Delta\hat{G}_1$. $\hat{P}_1(25\%)$ increased rapidly between 2 and 20 lines reaching for $N_1 > 20$ a 100% probability that the selected genotype belongs to the best 25% of the population. In contrast, the response curve of $\hat{P}_1(0.1\%)$ was almost linear with a low slope. Thus, for $N_1 = 1{,}000$ the probability that the selected genotype belongs to the best 0.1% of the population was still smaller than 65%. Consequently, for obtaining DHs with large genotypic values, a very large number N_1 of initial lines must be tested, which is in harmony with results of Robson et al. (1967), Johnson (1989), and Knapp (1998). In addition to rarely occurring positive recombinants, this result may explain that outstanding inbreds are identified only seldom in practice because the choice of N_1 is commonly much smaller than required. $\hat{P}_k(25\%)$ will be disregarded in our further discussion, because of being close to one.

In practice, selection is based on phenotypic and not on genotypic values and, thus, heritability is smaller than one. The influence of the different optimization criteria on the optimum allocation of test resources was hardly affected by the number of selection stages. Hence, only results for two-stage selection are discussed. Optimum allocation of test resources differed for the two optimization criteria and also for values of q, especially under large non-genetic variance (Table 3). The closest agreement between the optimum allocation of test resources maximizing $\hat{P}_k(q)$ and $\Delta\hat{G}_k$ was observed for $q = 5\%$. With decreasing values of q, an increased N_1^* and a decreased L_j^* were observed. Nevertheless, values of $\hat{P}_k^*(q)$ differed only slightly from $\hat{P}_k(q)$ at the optimum allocation of test resources with regard to $\Delta\hat{G}_k$. For instance, $\hat{P}_2^*(0.1\%) - \hat{P}_2(0.1\%)$ was below 0.01 for two-stage selection, $B = 1{,}000$, $N_f = 1$, VC2, and $C = 0.5$ (Tables 2, 3). This can be explained by the flat response curves of $\Delta\hat{G}_k$ and $\hat{P}_k(q)$ in the vicinity of the maximum (Fig. 2). For $\Delta\hat{G}_k$, it is attributable to the small slopes of the curves of h^2 for increased L_j, and i_α for decreased α (Becker 1993). For $\hat{P}_k(q)$, these findings are due to the small slopes of the curves of (1) h^2 and (2) the probability that genotypes belonging to the q% best genotypes of the population are among the lines for decreased α (Fig. 3).

The concept of ΔG_k is based on the superiority of the selected genotypes in comparison with their unselected base population. In contrast, $P_k(q)$ reflects the chance of developing competitive varieties that are better than the existing ones. To have a realistic chance of identifying a superior genotype, $P_k(q)$ should be greater than 75%, permitting only q values of about 5% for the budgets considered. The choice of the optimization criterion for these q values is not crucial, because the optimum allocation of test resources differed only slightly from those obtained by applying $\Delta\hat{G}_k$. For small values of q, different allocation optima were obtained for $\Delta\hat{G}_k$ and $\hat{P}_k(q)$, but probabilities $\hat{P}_k(q)$ were too low to be recommended as optimization criterion for the budgets investigated. Extending the formula of $P_1(q)$ given by Robson et al. (1967) to multi-stage selection could facilitate the optimum allocation of resources based on $P_k(q)$ due to a drastic reduction in computation time.

Standard deviations of optimization criteria

The choice of q had a large influence on the curves of $SD_{\hat{P}_1(q)}$ (Fig. 3). For instance, $SD_{\hat{P}_1(25\%)}$ decreased rapidly between 2 and 20 lines and reached zero for $N_1 = 40$, whereas $SD_{\hat{P}_1(0.1\%)}$ increased up to a maximum at $N_1 = 700$ and decreased slightly thereafter. These differences can be explained by the binomial nature of $P_k(q)$ with genotypes surpassing the defined threshold or not. Thus, $SD_{\hat{P}_k(q)}$ assumed its maximum for $\hat{P}_k(q) = 0.5$. In contrast, the response curve of $SD_{\Delta\hat{G}_1}$ decreased continuously with an increasing number of lines (Fig. 3) and test locations (Fig. 2). The small differences between values of $SD_{\Delta\hat{G}_1}$ for varying budgets (Fig. 2) can be explained by the small negative slope of $SD_{\Delta\hat{G}_1}$ for increasing values of N_1 (Fig. 3). As the curves of the optimization criteria were flat in the vicinity of the maximum (Fig. 2), their respective SD could serve as a secondary optimization criterion. However, curves of SD were also flat in the vicinity of the maximum of the optimization criteria, thus limiting their usefulness as additional optimization criterion.

Economic frame, quantitative-genetic parameters, and selection strategies

To assess their relative importance, the economic frame and quantitative-genetic parameters were varied in a range relevant for maize. Production costs C of DHs covered the entire range from recently established ($C = 1$) to further improved ($C = 0$) DH technology, with $C = 0.5$ corresponding to the actual costs in breeding companies advanced in the DH technique (Seitz, personal communication). The budget in our study can either refer to the resources available for evaluating the progenies of one cross ($B = 200 - 1{,}000$) or a complete breeding program ($B = 5{,}000$). For instance, considering the evaluation of 100 DH lines for each cross in two locations, 200 plots are required for 1 cross, and 5,000 plots for 25 crosses. The optimization of a complete breeding program would, however, require the assumption of equal means and segregation variances for progenies from different crosses. As these parameters usually differ among crosses (cf. Mihaljevic et al. 2004), optimization of breeding programs including these population parameters would be very promising but requires additional research.

The choice of N_f in this study reflects two situations. Commonly, numerous crosses are completely rejected before final evaluation and only few lines are selected in each of the remaining crosses. Thus, $N_f = 1$ represents a reasonable compromise for one specific cross. In contrast, in a complete breeding program, typically several

lines are finally selected. Consequently, $N_f = 1$ seems appropriate for $B = 200 - 1{,}000$ but $N_f = 5$ for $B = 5{,}000$.

For selection among genetically fixed lines, both optimization criteria depend on $\alpha = N_f/N_1$ and h^2. Variation in the budget or production costs of DHs mainly influenced α and, to a lesser extent, h^2. The budget was the major factor influencing values of both optimization criteria by its strong influence on α (Table 2). In contrast, production costs of DHs had only a minor effect on both optimization criteria. This can be explained by small changes in (1) α in comparison to changes of α for different budgets and (2) h^2 (Tables 2, 3). The slight trend towards larger values of L_j^* for $C = 1$ versus $C = 0$ reflects the fact that rejection of more expensive lines should be based on more reliable information.

The variance components were chosen according to recent estimates from large series of experiments within a broad sample of Central European maize breeding populations including DH populations (VC2, Gordillo and Geiger 2004), reflecting the typical situation for breeding programs with adapted maize populations. Variance components affect h^2 directly, and with VC3 the reduced h^2 could only partly be compensated by increased values of L_j^* with a parallel reduction in N_1^*. Altogether, we found a large reduction in values of $\hat{P}_k(1\%)$ and $\hat{P}_k(0.1\%)(> 50\%)$ with increased non-genetic variance. This is in accordance with previous studies (Keuls and Sieben 1955; Robson et al. 1967; Johnson 1989; Knapp 1998) analyzing the problem to identify superior genotypes under high non-genetic variance. Summarizing, our results underline the high impact of VC on the optimum allocation of resources with alternative breeding strategies.

Breeding is a continuous process and every year a new breeding cycle is initiated. Under this assumption, the annually available budget, for all cycles running in parallel is equal to the budget available for one entire cycle (Utz 1969). Consequently, comparisons between one- and two-stage selection can be made directly without dividing the optimization criteria by the years required in the selection strategy. Two-stage selection with optimum allocation of resources allows the evaluation of a large number of lines N_1 in a small number of test locations L_1. The N_2 lines selected in stage one are further evaluated in a large number of test locations L_2 to ascertain a high accuracy of the test results. This guarantees a low α and high h^2 and increases consequently both optimization criteria. In addition, response curves of the optimization criteria were flatter for two-stage selection than for one-stage selection, reducing the risk of choosing a non-optimal allocation. However, with one-stage selection breeders could exploit 1 year earlier the progress of selection by improved DH lines and hybrids developed from them.

Values of $\Delta\hat{G}_k$ and $\hat{P}_k(q)$ increased roughly to the same extent by (1) two-stage instead of one-stage selection, (2) a fivefold increase in the budget ($B = 200$ to $B = 1{,}000$), (3) a reduction in N_f from five to one, or (4) a quarter reduction in the non-genetic variance (VC1 instead of VC3). Except for the last factor, which is determined by the breeding material and target environments, all other factors can be chosen in favor of an increased selection response, but at the expense of a longer duration of the selection strategy (two-stage selection), higher costs (larger budget), and a higher risk of the final outcome (larger SD for $N_f = 1$). In particular, our results demonstrate that employing two-stage instead of one-stage selection represents a promising alternative to an increased budget.

Conclusions

The production costs of DHs had only a minor effect on the optimum allocation of breeding resources. Even if the current DH production using in vivo haploid induction is still relatively expensive, the compensation obtained through a reduced number of initial lines recommends their application. As DH costs are decreasing owing to expected improvements in the DH technique in the future, they will be only of secondary importance regarding the optimum allocation of resources.

For two-stage selection, a budget of approximately 1,000 field plot equivalents, and actual production costs of DHs, the allocation of test resources is roughly close to its optimum, if (1) the selected fraction $\alpha_1 = N_2/N_1$ is smaller than 0.10, (2) the number of test locations at the final selection stage exceeds at least six, and (3) about three quarters of the budget are invested in the first stage.

We attained a reasonable probability of success with continuous breeding for q values of about 5%. In these cases, the choice of the optimization criterion was relatively unimportant. However, for very large budgets the small probability of identifying outstanding genotypes is maximized if the number of lines is increased at the expense of the number of test locations. Optimization of complete breeding programs based on DHs is very promising, but selection theory must be extended for selection among and within crosses, consideration of different number and types of testers, and tests for line per se performance.

Acknowledgements This research was supported by funds from DFG, Grant No 1070/1, International Research Training Group "Sustainable Resource Use in North China" to C. F. H. Longin. The authors appreciate the editorial work of Dr. J. Muminović whose suggestions considerably improved the style of the manuscript. In addition, the authors thank Dr. F. Laidig, Bundessortenamt Hannover, Germany, and Dr. G. Seitz, AgReliant Genetics, Westfield, IN, USA for their valuable suggestions. We greatly appreciate the helpful comments and suggestions of two anonymous reviewers.

References

Becker H (1993) Pflanzenzüchtung (in German). Verlag Eugen Ulmer, Stuttgart, pp 129–132

Bernardo R (2002) Breeding for quantitative traits in plants. Stemma Press, Woodbury, p 205

Berry DA, Lindgren BW (1996) Statistics. Duxbury Press, New York, p 371

Bordes J, de Vaulx RD, Lapierre A, Pollacsek M (1997) Haplodiploidization of maize (*Zea mays* L.) through induced gynogenesis assisted by glossy markers and its use in breeding. Agronomie 17:291–297

Choo TM, Kannenberg LW (1988) Selection response and efficiency of doubled-haploid recurrent selection in a cross-fertilized species. Theor Appl Genet 75:410–414

Cochran WG (1951) Improvement by means of selection. In: Proc second Berkeley symp math stat prob, pp 449–470

Finney DJ (1966) An experimental study of certain screening processes. J Roy Stat Soc B 28:88–109

Gallais A (1991) A general approach for the study of a population of test-cross progenies and consequences for the recurrent selection. Theor Appl Genet 81:493–503

Gordillo GA, Geiger HH (2004) Estimating quantitative-genetic parameters of European maize populations to optimize hybrid breeding methods by model calculations. In: Poster abstract XVIIth EUCARPIA general congress 2004, Tulln, Austria

Grüneberg WJ, Abidin E, Ndolo P, Pereira CA, Hermann M (2004) Variance component estimations and allocation of resources for breeding sweetpotato under East African conditions. Plant Breed 123:311–315

Hanson WD, Brim CA (1963) Optimum allocation of test material for two-stage testing with an application to evaluation of soybean lines. Crop Sci 3:43–49

Johnson B (1989) The probability of selecting genetically superior S_2 lines from a maize population. Maydica 34:5–14

Keuls M, Sieben JW (1955) Two statistical problems in plant selection. Euphytica 4:34–44

Knapp SJ (1998) Marker-assisted selection as a strategy for increasing the probability of selecting superior genotypes. Crop Sci 38:1164–1174

Mihaljevic R, Utz HF, Melchinger AE (2004) Congruency of quantitative trait loci detected for agronomic traits in testcrosses of five populations of european maize. Crop Sci 44:114–124

Pearson ES, Hartley HO (1972) Biometrika tables for statisticians, 2nd edn. Cambridge University Press, London

R Development Core Team (2004) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-00-3, URL http://www.R-project.org

Röber FK (1999) Fortpflanzungsbiologische und genetische Untersuchungen mit RFLP-Markern zur in-vivo Haploideninduktion bei Mais. (in German). Ph.D. thesis, University of Hohenheim, Stuttgart, Germany

Robson DS, Powers L, Urquhart NS (1967) The proportion of genetic deviates in the tails of a normal population. Theor Appl Genet 37:205–216

Schmidt W (2004) Hybridmaiszüchtung bei der KWS SAAT AG (in German). In: Bericht über die 54. Tagung der Vereinigung der Pflanzenzüchter und Saatgutkaufleute Österreichs 2003, Gumpenstein, Austria, pp 1–6

Schön CC, Utz HF, Groh S, Truberg B, Openshaw S, Melchinger AE (2004) Quantitative trait locus mapping based on resampling in a vast maize testcross experiment and its relevance to quantitative genetics for complex traits. Genetics 167:485–498

Seitz G (2005) The use of doubled haploids in corn breeding. In: Proceedings of the forty first annual Illinois Corn Breeders' School 2005, Urbana—Champaign, Illinois, USA, pp 1–8

Sprague GF, Federer WT (1951) A comparison of variance components in corn yield trials: II. Error, year × variety, location × variety and variety components. Agron J 42:535–541

Tomerius AM (2001) Optimizing the development of seed-parent lines in hybrid rye breeding. Ph.D. thesis, University of Hohenheim, Stuttgart, Germany (http://opus-ho.uni-stuttgart.de/hop/volltexte/2001/10/pdf/tomerius.pdf)

Utz HF (1969) Mehrstufenselektion in der Pflanzenzüchtung (in German). Arbeiten der Universität Hohenheim, vol 49. Verlag Eugen Ulmer, Stuttgart, Germany

Young JC (1976) Varietal screening from finite normal populations. J Am Stat Assoc 71:87–92