Maydica 54 (2009): 485-499

DOUBLED HAPLOIDS IN HYBRID MAIZE BREEDING

H.H. Geiger[1], G.A. Gordillo[1,2]

[1] *University of Hohenheim, Institute of Plant Breeding, Seed Science, and Population Genetics, 70593 Stuttgart, Germany*

Received January 14, 2010

ABSTRACT - Use of doubled haploid (DH) lines produced by *in vivo* induction of maternal haploids are routinely used in maize (*Zea mays* L.) breeding. Major advantages of DH lines in hybrid breeding are (i) maximum genetic variance, (ii) complete homozygosity, (iii) short "time to market", (iv) simplified logistics, (v) reduced expenses, and (vi) optimal aptitude for marker applications. The present paper briefly reviews the experimental basis of the haploid induction technology, explains alternative DH-line-based breeding schemes, describes the features of a new software for optimizing such schemes, and presents and discusses selected optimization results. Modern inducer genotypes display induction rates of 8 to 10% on average. Various morphological and physiological markers warrant a fast and cheap identification of haploid kernels and/or seedlings. Artificial chromosome doubling procedures have successfully been adapted to large-scale commercial applications. Most likely, haploid embryogenesis is caused by defective sperm cells. After fusion with the egg cell, the chromosomes of the sperm cell degenerate and are stepwise eliminated from the primordial cells. The induction rate is under polygenic control. One cycle of DH-line development with two stages of testcross evaluation takes only four years if off-season nurseries are available. Cycle length can be shortened to three years if the first three breeding steps (recombination, haploid induction, and DH-plant production) are completed in a single year. Genome-wide marker-assisted selection can effectively be incorporated into DH-line based breeding schemes. To maintain selection response in the long run, the loss of genetic variation needs to be minimized by setting lower limits to the effective population size (N_e). Recurrent selection and line development may be combined to a single integrated breeding scheme. A new software MBP (Version 1.0) maximizes the expected annual genetic gain subject to budget and N_e restrictions. Input variables include estimated variance and covariance components, type of tester, haploid induction parameters, and costs of the individual breeding activities. For calculating N_e, the software considers genetic drift caused by both sampling and selection. Optimization results demonstrate that (i) schemes with only one stage of testcross evaluation provide faster breeding progress than those with two or more stages, (ii) genetic interlinking between staggered breeding programs is more efficient than a closed-population approach. Combined phenotypic and genome-wide selection holds great promise in accelerating future breeding progress.

KEY WORDS: *Zea mays* L.; *in vivo* haploid induction; DH-line development; Optimization of breeding schemes; Maintenance of genetic diversity.

INTRODUCTION

Doubled Haploid (DH) lines are routinely applied in many commercial hybrid maize (*Zea mays*) breeding programs. Major advantages of DH lines compared to selfed lines include (i) maximum genetic variance between lines for *per se* and testcross performance from the first generation, (ii) reduced breeding cycle length, (iii) perfect fulfillment of DUS (distinctness, uniformity, stability) criteria for variety protection, (iv) reduced expenses for selfing and maintenance breeding, (v) simplified logistics, and (vi) increased efficiency in marker-assisted selection, gene introgression, and stacking genes in lines. To our knowledge, all present commercial DH-line breeding programs are based on *in vivo* induction of maternal haploids (SEITZ, 2005; BARRET *et al.*, 2008; ROTARENKO *et al.*, 2009). Other techniques have proven to be less effective or too genotype specific.

Dedicated to Dr. Ronald L. Phillips, Regents Professor, University of Minnesota, on the occasion of his 70th birthday and his retirement.

[2] Present address: AgReliant Genetics Llc, Ivesdale, IL 61851, USA

* For correspondence (e.mail: geigerhh@uni-hohenheim.de).

Because of the genetic, methodological, and logistic advantages, further progress in maize breeding is expected to increase considerably with the development of DH lines. Yet, the success of employing DH lines depends on a robust and efficient haploid induction technology as well as on breeding strategies that make optimum use of the breeder's genetic, technical, and monetary resources (GORDILLO and GEIGER, 2008a-c).

An important aspect of the DH approach is to limit the loss of genetic variance due to random drift and selection by maintaining a minimum effective population size (N_e). To comply with this requirement, considerable numbers of DH parent lines need to be intercrossed for starting a new selection cycle (GORDILLO and GEIGER, 2008b).

In the present paper we want to (i) briefly review the experimental basis of the *in vivo* haploid induction technology, (ii) explain alternative DH-line-based breeding schemes, (iii) describe the features of MBP (version 1.0), a new computer software for optimizing DH-line breeding schemes, (iv) present selected optimization results, and (v) discuss the relative merits of alternative breeding schemes.

IN VIVO INDUCTION OF MATERNAL HAPLOIDS

To induce maternal haploids, the donor plant is pollinated by a specific maize stock (line, single cross, or population) called inducer. Beside regular F_1 kernels, the pollination results in a certain proportion of kernels with a haploid maternal embryo and a regular triploid endosperm. Such kernels display a normal germination rate and lead to viable haploid seedlings (RÖBER *et al.*, 2005; GEIGER, 2009). After artificial chromosome doubling the successfully treated seedlings are selfed leading to completely homozygous and homogeneous progenies (DH lines). Contrary to *in vitro* induction techniques such as anther or microspore culture, no tissue culture step is involved.

Induction rate

CHASE (1952) observed spontaneous haploids in US Corn-Belt germplasm at a rate about 0.1%. This value was too low for any commercial application. Later COE (1959) found an inbred line called Stock6 with an induction rate of 1 to 2%. This line became the ancestor of all subsequently developed inducer lines. Considerable improvement in the induction rate was achieved by groups working in India (SARKAR *et al.*, 1994), the former Soviet Union (TYRNOV and ZAVALISHINA, 1984; CHALYK, 1994; SHATSKAYA *et al.*, 1994b), France (LASHERMES and BECKERT, 1988; BORDES *et al.,* 1997), and Germany (DEIMLING *et al.*, 1997; RÖBER *et al.*, 2005).

One of the presently most effective inducers is line RWS developed at the University of Hohenheim (RÖBER *et al.*, 2005). It was derived from a cross between an inbred line originating from the Russian inducer synthetic KEMS (SHATSKAYA *et al.*, 1994) and the French inducer line WS14 (LASHMERMES and BECKERT, 1988) and is adapted to the temperate climate of Central Europe but is also effective in tropical environments (RÖBER *et al.*, 2005). Averaged across a wide range of donors and environments, it has an induction rate of about 8%. A sister line of RWS, RWK-76, developed from the reciprocal cross (WS14 × KEMS) provided an average induction rate of 9 to 10%. The same rate was observed for the cross RWS × RWK-76 (unpubl. data). To avoid an overestimation of the induction rate it is important to use donor genotypes (females) being homozygous for recessive markers such as the mutants *liguleless* or *glossy*.

Significant induction rate differences were observed between donor genotypes (RÖBER *et al.*, 2005). However, the range of variation was small compared with that reported for *in vitro* culture techniques (PETOLINO and THOMPSON, 1987; COWEN *et al.*, 1992; MURIGNEUX, 1994; BÜTER, 1997; SPITKÓ *et al.*, 2006). Considerable variation in induction rate may also be caused by environmental factors (RÖBER *et al.*, 2005). Moreover, the induction rate may depend on the method and time of pollination. ROTARENKO *et al.* (2009) obtained the best results by manual pollination (compared with open pollination in an isolated plot) three days after silk appearance. Yet experienced breeders get high induction rates under open-pollination as well (personal communications).

Haploid identification

A key issue to apply the *in vivo* haploid induction approach on a commercial scale is an efficient screening system allowing the breeder to differentiate between kernels or seedlings generated by haploid induction and those resulting from regular fertilization. The most efficient haploid identification marker is the 'red crown' or 'navajo' kernel trait encoded by the dominant mutant allele *R1-nj* of the 'red color' gene *R1*. In the presence of the domi-

nant pigmentation genes *A1* or *A2* and *C2*, *R1-nj* conditions deep pigmentation of the aleurone layer (endosperm tissue) in the crown (top) region of the kernel. In addition, it causes pigmentation of the scutellum (embryo tissue). Pigmentation varies in extent and intensity depending on the genetic background of the particular donor genotype (for details see MaizeGDB at http://maizegdb.org/).

NANDA and CHASE (1966) and GREENBLATT and BOCK (1967) first used the red crown mutant as selectable marker in haploid induction experiments. To be effective, the donor has to have colorless seeds and the inducer needs to be homozygous for *R1-nj* and the aforementioned dominant pigmentation genes. A kernel resulting from haploid induction then has a red crown (regular triploid endosperm) and a non-pigmented scutellum, whereas a regular F_1 kernel displays pigmentation of both the aleurone and scutellum (GEIGER, 2009). If only the egg cell but not the central cell is fertilized, the kernel has a pigmented (diploid) embryo and a non-pigmented, diploid maternal endosperm and aborts during early kernel development. Kernels resulting from (unintended) selfing or outcrossing with other colorless donors do not show any pigmentation. The red crown marker does not work if the donor genome is homozygous for *R1* or for dominant anthocyanin inhibitor genes such as *C1-I*. These genes occur quite frequently in European flint or tropical materials (BELICUAS *et al*., 2007). In that case, haploid identification is possible in the early seedling stage if the inducer is homozygous for the genes *B1* and *Pl1* which in conjunction condition light-independent pigmentation of the coleoptile and root of the F_1 seedlings.

Another cheap and fast haploid identification method was suggested by ROTARENKO *et al.* (2007). The authors observed that kernels with a haploid embryo have a significantly lower oil concentration than those with a diploid F_1 embryo. This is due to the reduced size of haploid embryos when compared with diploid embryos. Inducers with an above-average oil concentration should be best suited for this approach. High-protein inducers might work analogously.

Properties of haploids

Haploid plants are smaller and less vigorous than corresponding diploid homozygous lines (CHASE, 1952; AUGER *et al.*, 2004). Most haploids display a certain level of female fertility (CHALYK, 1994; GEIGER *et al*., 2006) but in general haploids lack male fertility (CHASE, 1952; CHALYK, 1994). However, certain donor genotypes were detected that provided haploid plants producing traces of pollen which could successfully be used for selfing (CHALYK, 1994). ZABIROVA *et al.* (1993) identified a donor genotype from which one third of the induced haploids were male fertile. The donor resulted from four cycles of selection for that trait.

Artificial chromosome doubling

Chromosome doubling used to be a serious constraint in producing doubled haploids on a commercial scale. Spontaneous doubling was observed only in very few germplasm sources (CHASE, 1964; SHATSKAYA *et al.*, 1994a). A breakthrough was accomplished by GAYEN *et al.* (1994) who cut off the tip of the haploid coleoptiles and immersed the seedlings into a 0.06% colchicine solution plus 0.5% DSMO (dimethyl sulfoxid) for 12 hours at 18°C. DEIMLING *et al.* (1997) further increased the efficacy of the method by reducing the roots to 20 to 30 mm and placing the immersed seedlings in the dark. After the colchicine treatment, the seedlings are carefully washed in water and subsequently grown in the greenhouse to the 5- to 6-leaf stage (during the first days under high humidity). Thereafter, the treated plants are transferred to the field. EDER and CHALYK (2002) applied the method to a broad range of donor genotypes and achieved an average doubling rate of 49%. About 50 to 60% of the successfully treated plants shed pollen and could be selfed. Thus, about one out of three colchicinized seedlings produced seeds. The number of viable seeds per ear varies from less than 5 to more than 20 (unpublished data). Since colchicine is highly toxic to humans, most breeding companies meanwhile are applying less hazardous proprietary substances for chromosome doubling. Generally, the latter are sprayed onto the haploid seedlings in the 3- to 5-leave stage while the plants are still in a greenhouse (personal communications from various breeders).

Biological mechanism

Principally, two mechanisms leading to maternal haploids have been hypothesized: (1) One of the two sperm cells provided by the inducer is defective but yet able to fuse with the egg cell. During subsequent cell divisions, the inducer chromosomes degenerate and are stepwise eliminated from the primordial cells. The second sperm cell fuses with the central cell and leads to a regular triploid endosperm. (2) One of the two sperm cells is not able

to fuse with the egg cell but can trigger haploid embryogenesis. The second cell fuses with the central cell as under the first hypothesis.

Experimental support for the first hypothesis comes from studies of WEDZONY *et al.* (2002). The authors fixed ovaries of selfed RWS plants at regular intervals during the first 20 days after pollination. In accordance with the induction rate of RWS, 18 out of 203 embryos contained micronuclei in every cell of the shoot primordia. Micronuclei varied in number and diameter displaying the typical characteristics of metabolically inactive chromatin. In some equatorial plates also chromosome fragments were observed. Micronuclei elimination started during the globular state of embryogenesis. These observations are indirectly corroborated by results of FISCHER (2004) and LI *et al.* (2009). In both studies small fractions of the inducer genome were detected in the haploids by molecular marker techniques. In Fischer´s experiment, using inducer line RWS, paternally transmitted DNA was detected in 1.4% of the haploids whereas in Li´s *et al.* experiment, using the Chinese inducer line CAUHOI, the proportion was 43%. However, on average only 1.8% of the inducer genome was transmitted. Generally the transmitted segments had replaced the homologous maternal segments. Since CAUHOI is a high-oil genotype, paternal transmission was also detected by an increased oil concentration in some of the haploids carrying paternal chromosome segments.

ROTARENKO *et al.* (2009) pollinated two donor inbred lines with two inducer lines and their F_1. Representative samples of field-grown haploid progenies revealed significant differences between inducers for plant height, leaf length, leaf width, ear length, and number of kernels in 9 out of 25 cases (trait/donor/inducer combinations). This again indicates the possibility of paternal transmission and it seems that it is a common phenomenon of some inducers.

In the second hypothesis, there is, thus far, no experimental supportive evidence. However, it is conceivable that inducers with reproductive abnormalities such as aneuploid microsporocytes (CHALYK *et al.*, 2003) or increased heterofertilization rate (ROTARENKO and EDER, 2003) might be able to induce haploid embryogenesis without penetration of the sperm cell into the egg cell.

Genetics of haploid induction

Several inheritance studies concordantly show that the *in vivo* induction of maternal haploids is under polygenic control (LASHERMES and BECKERT 1988; DEIMLING *et al.*, 1997; RÖBER *et al.,* 2005). Therefore, transferring the haploid inducing ability into other, e.g. tropical, genetic backgrounds is a time- and labor-consuming exercise. Marker-based approaches would considerably alleviate this task. However, no genome-wide QTL analysis of modern inducers has so far been reported in the literature.

Using various dent inbred lines as donor parents and the French inducer line PK6 as pollinator, BARRET *et al.* (2008) detected and fine-mapped a locus on chromosome 1 influencing *in vivo* haploid induction. The PK6 allele significantly increased the induction rate in many of the above combinations. In segregating populations it showed gametophytic expression with incomplete penetrance and was correlated with segregation distortion against the inducer. Segregation distortion was also observed by DEIMLING *et al.* (1997) at several marker loci in various DH populations.

In an experiment with the French inducer line WS14 and a line derived from the Russian inducer synthetic KEMS, RÖBER *et al.* (2005) obtained close agreement between the F_1 and the mid-parent value for the induction rate. In a more recent experiment with the two German inducer lines RWS and RWK-76, the F_1 even reached the same induction rate as the better parent (GEIGER, 2009). Using an F_1 inducer facilitates large-scale haploid induction programs since they are more vigorous and stress tolerant than inbred inducers.

BREEDING METHODOLOGY

Line development schemes

In this paper, four DH line-based breeding schemes are described. The Standard Scheme (Fig. 1, left hand side) consists of following steps:

- Creating new variation by intercrossing selected lines for starting a new breeding cycle.
- *In vivo* haploid induction in generation F_1 (=S_0).
- Identifying haploids, chromosome doubling, and selfing of resulting DH plants for production of DH lines.
- Visual evaluation of DH lines *per se* in the field and, in parallel, seed multiplication of the lines.
- Production of testcross seed of selected lines with one or more testers from an unrelated ('opposite') gene pool.
- Evaluation of the testcrosses in multi-environment yield trials.

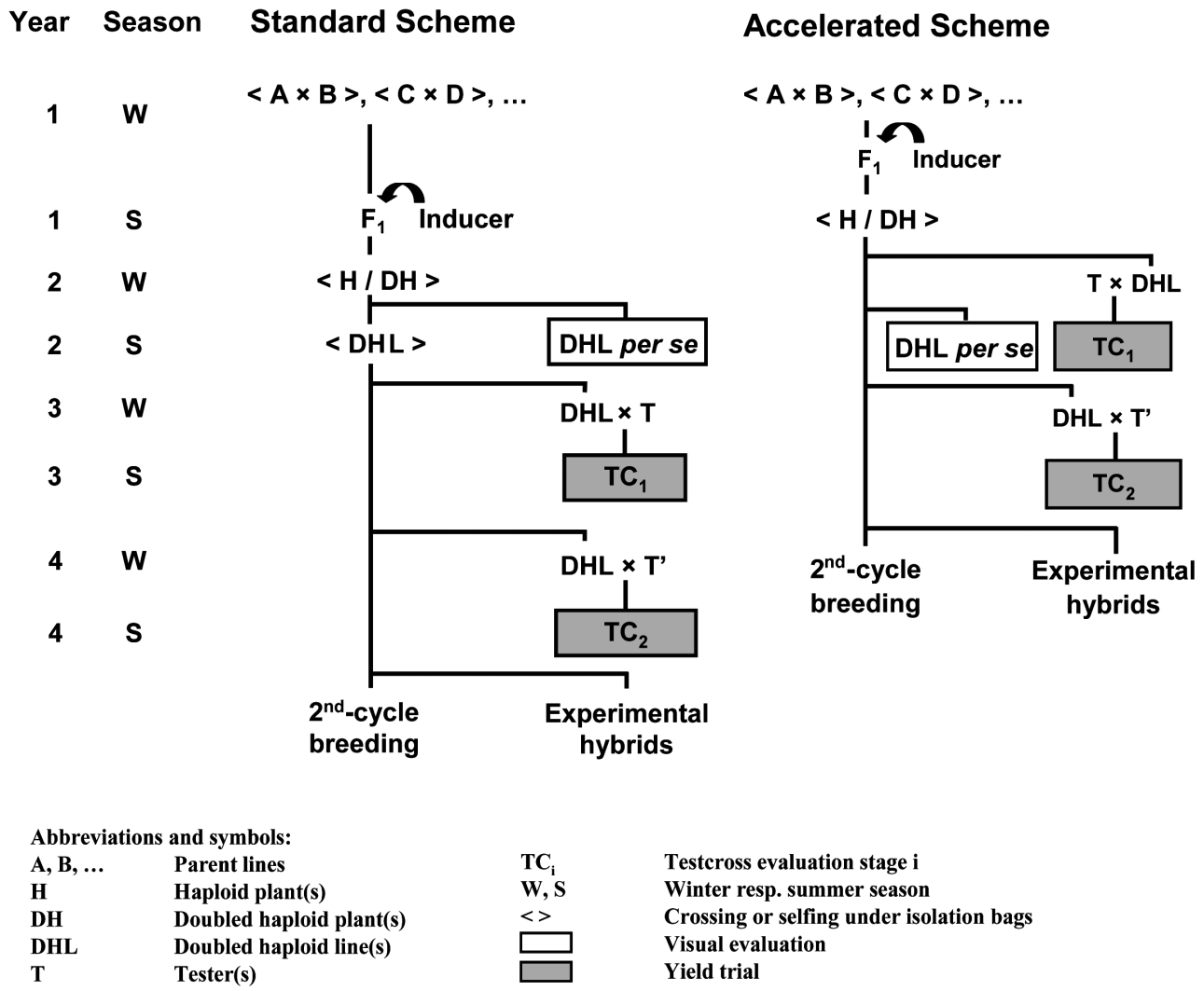

FIGURE 1 - Standard and Accelerated Scheme of DH-line development with two stages of testcross selection.

With one-stage testcross evaluation, selection is based on field tests in one year. With two-stage selection, tests are continued in a second year with the best lines selected in the first year. Analogously, with three-stage selection, the best lines selected in the second year are again tested in the third year. As the number of candidate lines is reduced over successive stages, the number of testers and location is increased. The Standard Scheme requires six, eight, or ten seasons with one, two, and three stages of selection, respectively. Using off-season nurseries, the cycle is completed after three, four and five years, respectively. Thereafter experimental hybrids are built up and evaluated for commercial use (not shown in Fig. 1).

The cycle length of the Standard Scheme can be shortened if a round-year nursery is available allowing completion of the first three breeding steps in a single year. In this Accelerated Scheme (Fig. 1, right hand side), the evaluation of DH lines in observation plots is not conducted before but parallel to the first testcross evaluation. Moreover, the seed for the first stage of testcross selection is produced by hand-pollination using the tester(s) as seed parent(s). The time saving compared to the Standard Scheme is one year per cycle.

If a breeding population requires strong selection against lack of adaptedness, it may be worth considering early selection in selfing generations S_1 or S_2 before starting with induction crosses. An example is given in Fig. 2. The scheme starts with the production of double crosses allowing the breeder

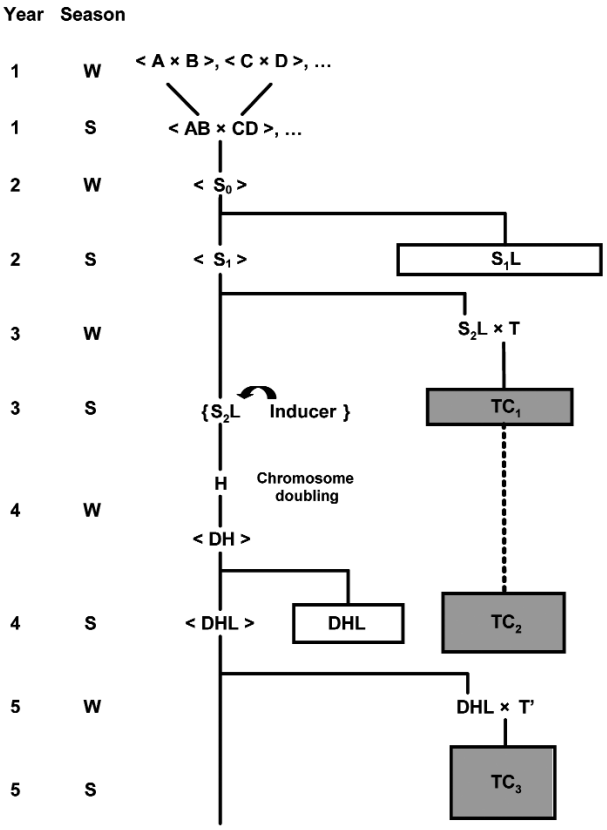

FIGURE 2 - Development of DH lines from preselected S_2 lines (for abbreviations and symbols see Fig. 1).

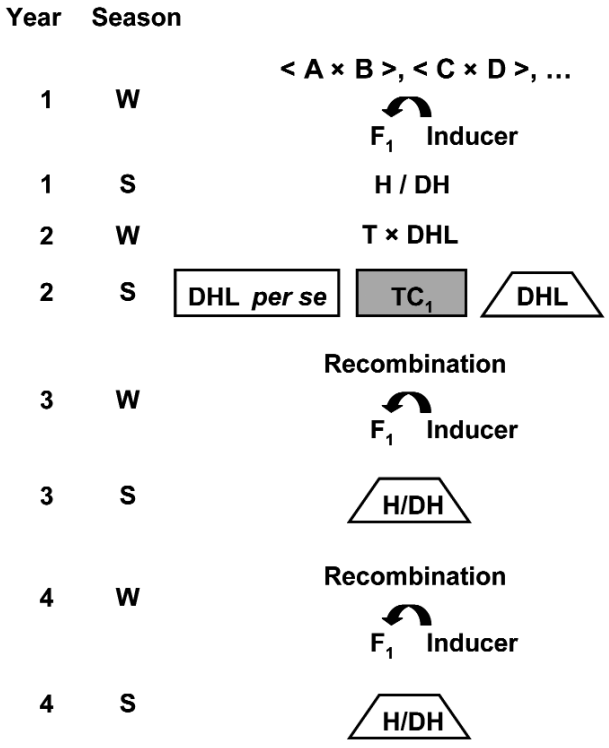

FIGURE 3 - DH-line development with integrated genome-wide selection (PS = phenotypic selection, GS = genome-wide selection, ⏢ = symbol for genotyping) (for other abbreviations and symbols see Fig. 1).

to vary the proportion of poorly adapted parents in a cross and to increase recombination between adapted and non-adapted genome segments. In year 2, S_1 lines are visually evaluated in multi-location observation plots and advanced by selfing. Testcrosses of S_2 lines derived from selected S_1 lines are then evaluated in two-stage yield trials over two years. After the first test year, haploid induction is started with selected S_2 lines as donors. A first visual scoring of DH lines is possible parallel to the second stage of S_2 line testcross evaluation. Finally, selected DH lines are tested for combining ability at a third stage of testcross evaluation. This scheme has the same cycle length as the three-stage Standard Scheme.

Marker-assisted selection can readily be incorporated into DH line-based breeding schemes. This may be exemplified for a genome-wide selection (GS) approach similar to the one recently proposed by BERNARDO and YU (2007). In this scheme (Fig. 3), the development and the first *per se* and testcross evaluation of DH lines in the field are conducted in two years just as in the Accelerated Scheme in Fig. 1. In addition to this, all DH lines under test are genotyped in parallel for a set of densely spaced markers distributed over the whole genome. Marker effects on the trait(s) of interest are estimated to predict genotypic values of the DH lines *per se* and their testcrosses by summing up all the marker effects for selection among candidates. A main feature of genome-wide selection is that it focuses on prediction of performance without identifying a subset of markers significantly associated with the trait(s) of interest (for alternative estimation procedures and computational details see MEUWISSEN *et al.*, 2001; BERNARDO and YU, 2007). Mean phenotypic and predicted genotypic values are combined to obtain overall breeding value estimates at the first stage of selection. The selected lines are then intercrossed and haploids are produced from the F_1s. The haploids are genotyped in the regular growing season of the third year and only those with the highest predicted genotypic superiority are intercrossed in the following winter season. This second

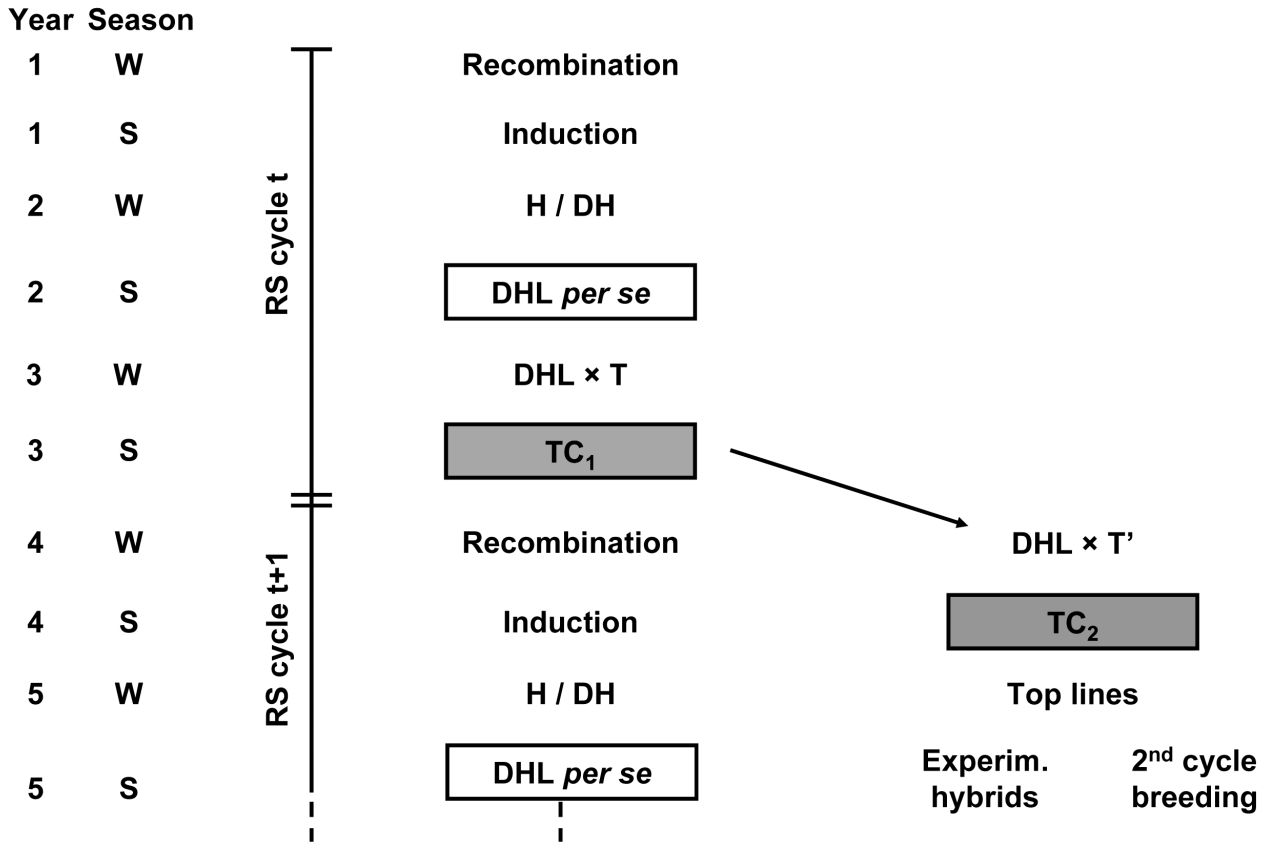

FIGURE 4 - Example of an integrated recurrent selection (RS) and line development scheme (for abbreviations and symbols see Fig. 1).

(marker-only) phase of the scheme may be repeated once more (as in Fig. 3). Marker effects need to be re-estimated periodically to adjust for a decay of linkage disequilibrium and for genotype × year interactions. The first phase of combined phenotypic and marker-based selection plus the two subsequent solely marker-based selection stages take four years like the three-stage Accelerated Scheme based on phenotypic selection alone.

Integrating recurrent selection into hybrid parent line development

The goal of recurrent selection (RS) is the cyclical genetic improvement of the breeding population for quantitatively inherited traits by increasing the frequency of favorable alleles without seriously reducing the genetic variability (HALLAUER and MIRANDA, 1981). In the context of hybrid breeding this means extensive testing for line *per se* and testcross performance and thereafter recombining the best candidates for starting a new RS cycle. It therefore is logical (and practiced by many hybrid breeders) to combine RS and Line Development (LD) in a single integrated breeding scheme. RS and LD may comprise the same or different numbers of selection stages. Frequently, however, the RS cycle is completed after the first round of testcross evaluation (Fig. 4).

The success in LD depends on the genetic potential of the respective breeding population and the superiority of the lines selected from that population. Generally, an extremely high selection intensity is practiced in LD, since no lower limit for N_e has to be met. As a consequence, a higher selection gain is achieved than in RS. However, recombining the small number of lines selected as hybrid parents would not suffice to maintain the genetic variation over many cycles. Hence, in the short term, the breeding success mainly depends on the genetic gain in LD, whereas in the medium and long term, progress is mainly determined by the cumulative response to RS (Fig. 5). From this point of view, the RS component represents the mainstream of a breeding program and the relative weights which

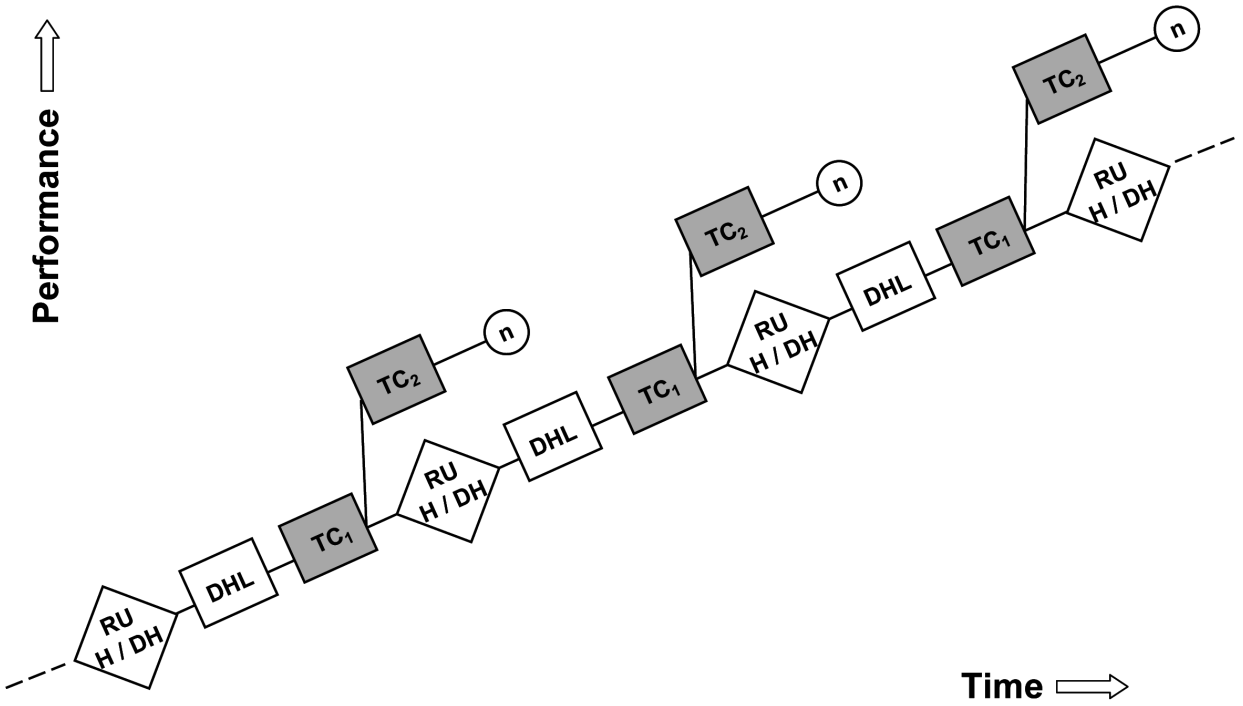

FIGURE 5 - Flow diagram of three cycles of an integrated recurrent selection (RS) and line development (LD) program.

should be given to the genetic gain in LD *versus* RS depend on the intended time span of that program.

Interlinking staggered breeding cycles

In commercial hybrid breeding, a new breeding cycle is started in each gene pool (heterotic group) every year. Thus, several timely staggered breeding cycles are advanced in parallel. Such staggered breeding cycles may be genetically interlinked by composing the fraction selected for recombination not only of lines from the breeding cycle under consideration but also of preselected lines from the follow-up cycle (Fig. 6). This allows the breeder to increase selection intensity and yet comply with a predetermined effective population size (GORDILLO and GEIGER, 2008c).

Optimization software MBP (version 1.0)

GORDILLO and GEIGER (2008a) recently developed a computer software, MBP (version 1.0), for optimizing breeding plans based on DH lines. This software maximizes the annual expected gain in General Combining Ability (GCA) as a function of various quantitative genetic parameters and operational variables under the restrictions of a given annual breeding budget and a limited annual loss of genetic variance. It uses numerical methods for the calculation of normal integrals for the distribution of genotypic values under one-, two-, and three-stage selection. For computing the loss of genetic variance, MBP considers the combined effects of randomly and selectively caused allele frequency changes on N_e according to SANTIAGO and CABALLERO (1995). The upper limit for the decay of genetic variance is defined by $\Delta\sigma^2_g = 1/(2N_eY)$, where Y is the breeding cycle length in years.

A detailed description of the software along with a user's handbook is given in GORDILLO and GEIGER (2008a). Briefly, input variables comprise variance and covariance components, heritability coefficients,

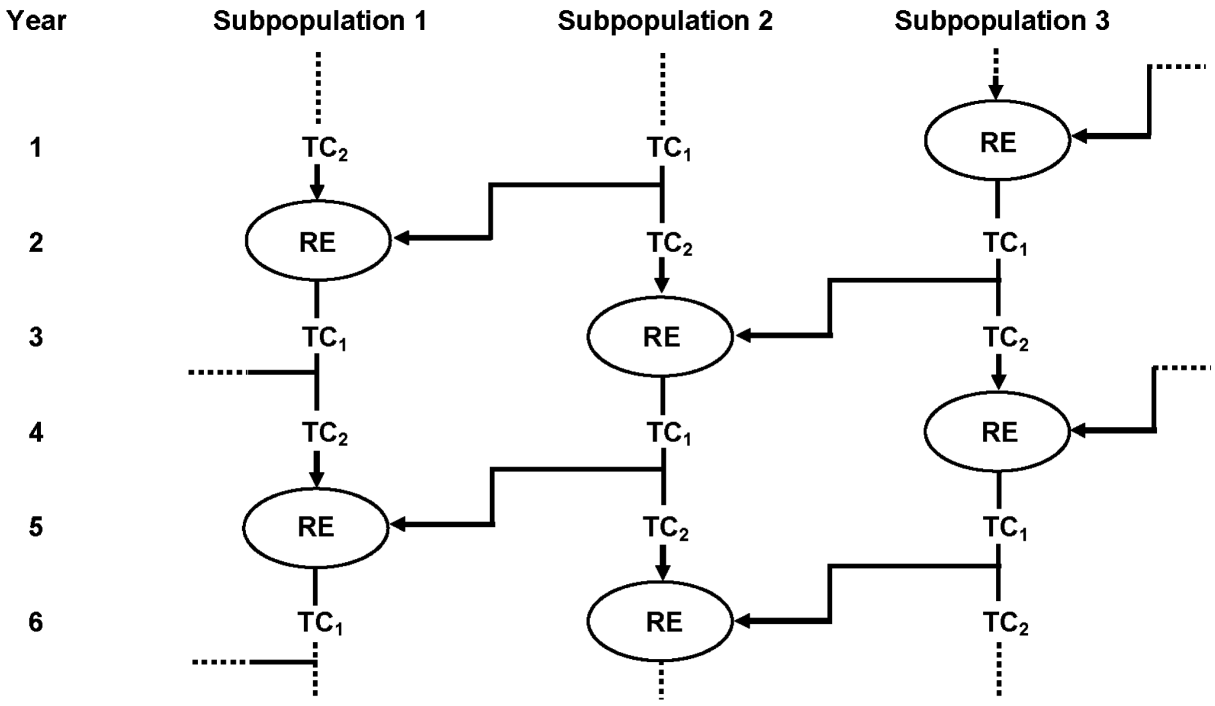

FIGURE 6 - Schematic representation of interlinking staggered breeding cycles within a given gene pool (reproduced from GORDILLO and GEIGER, 2008c, with kind permission of Crop Science).

haploid induction parameters, and monetary costs of the individual breeding steps. Parameter estimates were computed from field trial and haploid induction data, and labor cost quotations provided by collaborating breeders from Europe and North America. Inbred lines, single crosses, double crosses, and populations can be used as testers. All parameter settings can be varied by the user. The gain criterion is the GCA of the selected lines according to a base index (BRIM *et al.*, 1959; WILLIAMS, 1962) composed of the testcross performance for grain yield and grain dry matter content. When successive breeding cycles are interlinked (Fig. 6), the difference in the performance levels of the respective subpopulations is taken into account in estimating the selection response.

RESULTS OF OPTIMIZATION STUDIES

In this paper the following program specifications are used for all schemes considered:

- Annual budget: EUR 500,000.
- Proportion of lines visually pre-selected for *per se* performance: $\alpha = 0.5$.
- Single crosses as testers at all selection stages.
- Yield trials (at each selection stage): one year, multiple locations, non-replicated.
- Three finally selected lines per LD cycle.
- Gain criterion: I = GY + 2.5 DMC, where GY and DMC denote the GCA for grain yield (dt ha^{-1}, where dt = 10^{-1} t) and grain dry matter content (%), respectively.
- Annual loss of genetic diversity restricted to 2%.

Grain dry matter content is included in the gain criterion to counterbalance the generally negative genetic correlation between grain yield and grain dry matter in cool temperate climates. For simplicity, only the genetic gain for grain yield is reported in this paper.

Optimized LD schemes clearly show that multi-stage selection is more efficient than one-stage selection if judged by the genetic gain per cycle whereas the reverse is true regarding the gain per

TABLE 1 - *Optimum allocation and short-term genetic gain in DH-line development (LD) under one-, two-, and three-stage selection in the Standard, Accelerated, and S_2/DH Scheme assuming three finally selected lines.*

	Optimum allocation[1]									Genetic gain for yield (kg ha^{-1})	
Scheme	N_1	N_2	N_3	T_1	T_2	T_3	L_1	L_2	L_3	per cycle	per year
One-stage LD[3]											
Standard (3 ys)	1802	–	–	2	–	–	10	–	–	684	228
Accelerated (2 ys)	2318	–	–	2	–	–	8	–	–	646	323
Two-stage LD											
Standard (4 ys)	3573	57	–	1	7	–	5	20	–	812	203
Accelerated (3 ys)	5680	79	–	1	6	–	3	19	–	786	262
Three-stage LD											
Standard (5 ys)	4285	200	15	1	3	7	3	8	26	880	176
Accelerated (4 ys)	5758	146	11	1	3	10	3	12	29	852	213
Combined S_2/DH-line selection (5 ys)	9772	417	52	1	1	6	2	15	19	905	181

[1] N, T, L denote the optimum number of DH lines, testers, and locations, respectively; subscripts refer to selection stages.

TABLE 2 - *Optimum allocation and genetic gain from one- and two-stage recurrent selection (RS) in the Standard and Accelerated Scheme; annual loss of genetic variance restricted to 2%.*

		Optimum allocation[1]						Genetic gain for yield (kg ha^{-1})	
Scheme	N_{rec}	N_1	N_2	T_1	T_2	L_1	L_2	per cycle	per year
One-stage RS									
Standard (3 ys)	43	2481	-	2	-	6	-	504	168
Accelerated (2 ys)	57	3900	-	1	-	8	-	442	221
Two-stage RS									
Standard (4 ys)	32	3821	225	1	3	4	12	624	156
Accelerated (3 ys)	39	5574	274	1	3	3	12	576	192

[1] N, T, L denote the optimum number of DH lines, testers, and locations, respectively; subscripts refer to selection stages.
N_{rec} = number of recombined lines.

TABLE 3 - *Influence of the weights given to recurrent selection and line development (w_{RS} and w_{LD}, respectively) on the optimum allocation and annual genetic gain in the integrated RS/LD Standard Scheme; annual loss of genetic variance restricted to 2%; three lines selected as hybrid parents.*

	Optimum allocation[1]							Genetic gain for yield (kg ha^{-1}) per year	
w_{RS} : w_{LD}	N_{rec}	N_1	N_2	T_1	T_2	L_1	L_2	RS	LD
0 : 1	34	3573	57	1	7	5	20	143	204
0.50 : 0.50	32	3805	153	1	4	4	14	155	202
1 : 0	32	3831	225	1	3	4	12	156	199

[1] N, T, L denote the optimum number of DH lines, testers, and locations, respectively; subscripts refer to selection stages. N_{rec} = number of recombined lines.

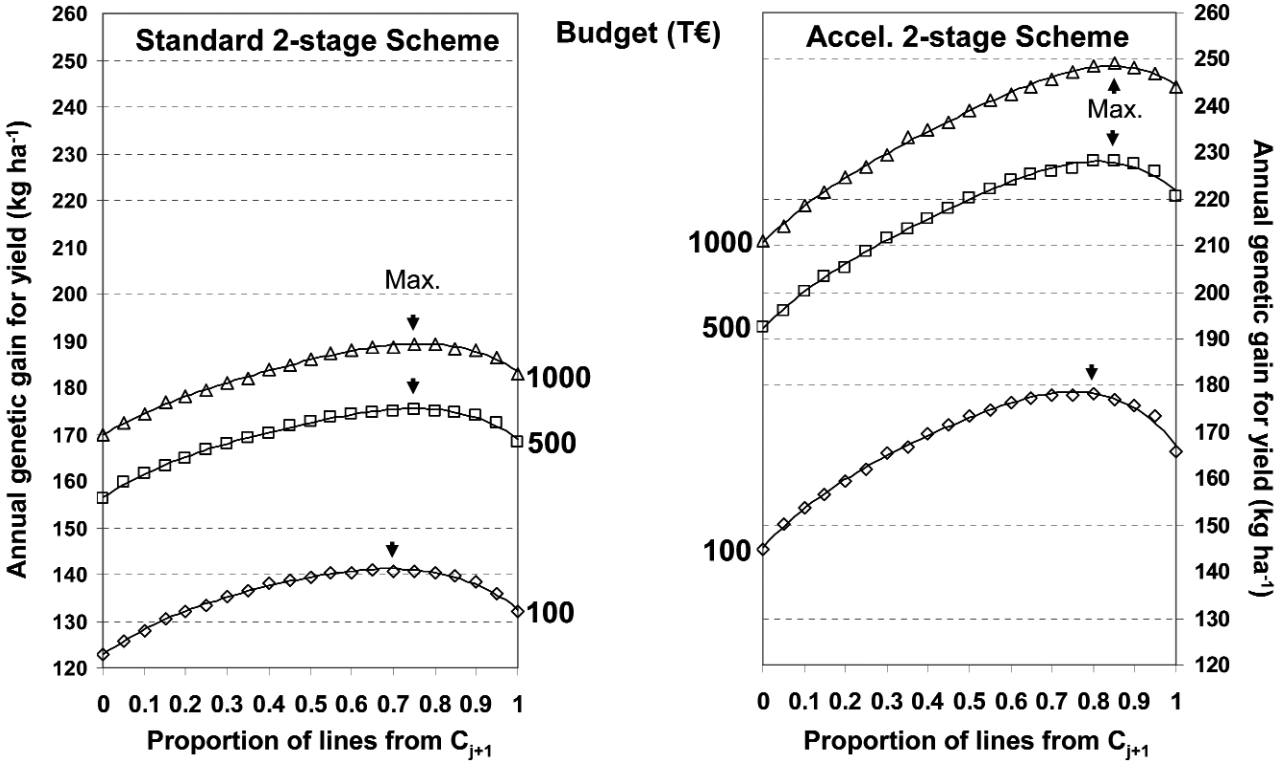

FIGURE 7 - Annual genetic gain in breeding cycle C_j as a function of the proportion of lines selected from interlinked cycle C_{j+1} (annual loss of genetic variance restricted to 2%).

year (Table 1). Similarly, the standard versions of the two schemes are superior over the accelerated ones if considered per cycle but inferior if considered on an annual basis. The highest annual genetic gain is expected under the one-stage Accelerated Scheme (323 kg ha^{-1}) and the lowest under the three-stage Standard Scheme (176 kg ha^{-1}). Combined S_2/DH-line selection (181 kg ha^{-1}) is slightly superior to the three-stage Standard Scheme but inferior to the Accelerated Scheme (213 kg ha^{-1}). Much larger populations can be used to start a new breeding cycle under multi-stage than under one-stage selection. Under multi-stage selection, the optimum number of testers and test sites considerably increases from stage to stage. Under one-stage selection two testers and 8 to 10 test sites are optimal.

The annual genetic gain from RS (Table 2) is significantly less than that in LD since the selection intensity must be reduced in RS to comply with the loss-of-genetic-diversity restriction. Under the most efficient RS scheme (accelerated one-stage selection), more than 50 selected DH lines need to be recombined for this purpose. As a consequence, the optimal number of testers and test sites also differs between RS and LD.

Integrating RS in LD allows the breeder to use the same breeding capacities for both purposes. Optimizing such an integrated program is possible for arbitrary weights given to the expected gains from RS and LD (GORDILLO and GEIGER, 2008a). Increasing the weight for RS leads to a considerable increase in the gain from RS, while it hardly impairs the gain in LD. It mainly requires a change in the number of entries and locations at the second stage of testcross selection (Table 3).

Interlinking staggered breeding cycles by recombining lines selected in a given breeding cycle started in year *j* and lines from another cycle started in the subsequent year *j+1* (Fig. 6) increases the expected genetic gain by about 10 to 15% under the Standard Scheme and 17 to 23% under the Accelerated Scheme (Fig. 7). The advantage of interlinking increases as the budget is increased. The optimal proportion of lines to be selected from breeding cycle *j+1* varies between 70 and 85%.

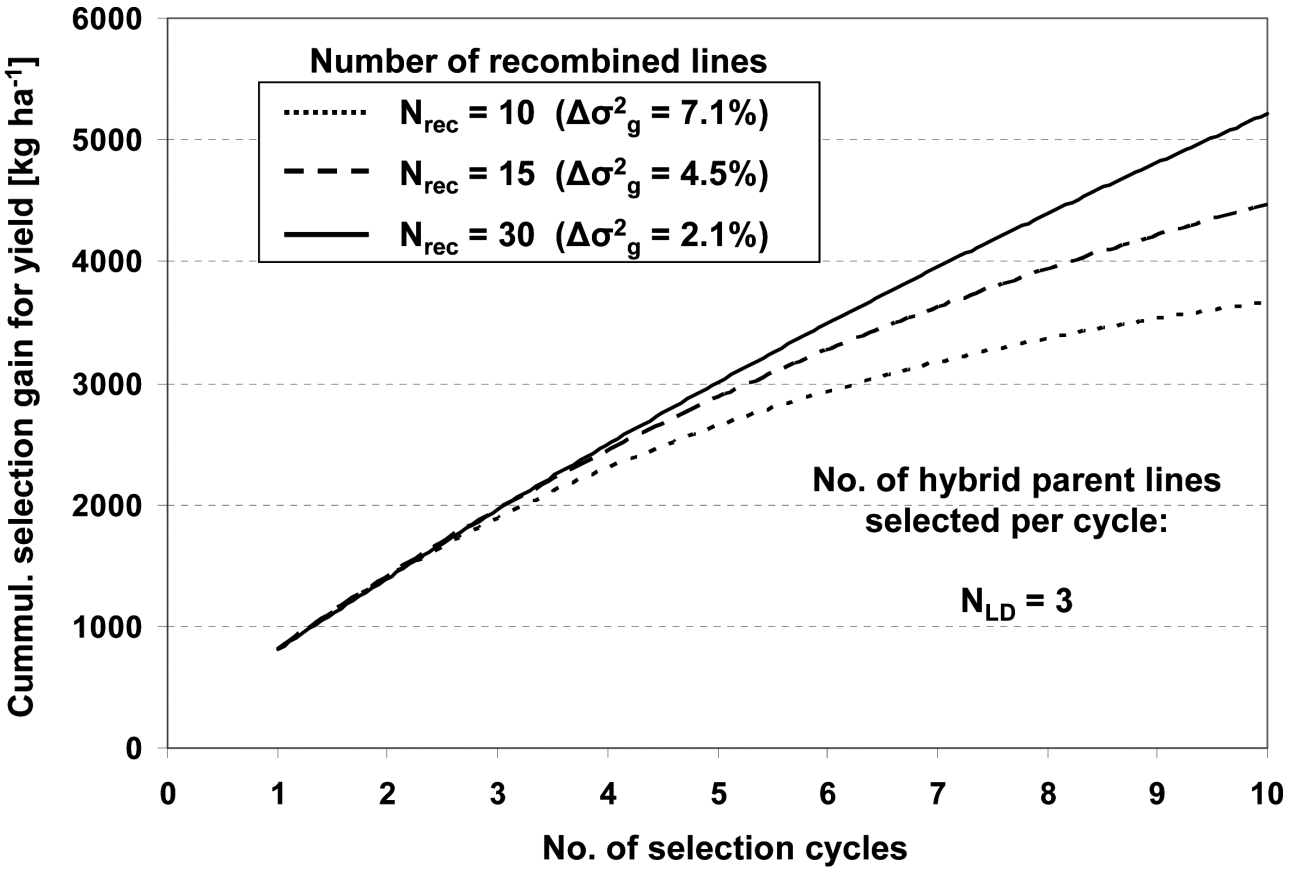

FIGURE 8 - Influence of the number of lines recombined for starting the next breeding cycle on the long-term genetic progress in line development ($\Delta\sigma^2_g$ = annual loss of genetic variance).

DISCUSSION

DH lines display maximum genetic differentiation for *per se* and testcross performance from the first generation and allow the breeder to drastically reduce the 'time to market'. As a consequence, most internationally leading seed companies have converted their LD programs to the DH technology during the last years or have initiated this process. The technology has also found its way into research but much slower than in breeding because experienced staff and appropriate experimental facilities are needed to apply it successfully. Yet, the first large mapping population consisting of 720 DH lines was already established at the University of Hohenheim in the year 2000 (PRESTERL *et al.*, 2007).

Optimization results revealed that the one-stage LD and RS schemes provide the greatest annual gain from selection. This holds true for other settings (budget, genetic parameters, type of tester, labor cost etc.) of the MBP software as well (GORDILLO and GEIGER, 2008b,c). Similar results were obtained by LONGIN *et al.* (2006) and WEGENAST *et al.* (2008) using a different optimization software and other basic assumptions. However, in case of strong genotype × year interactions, one-stage selection may be considerably biased in specific years. This risk is much smaller under two-stage selection so that breeders generally prefer schemes with at least two years of testing despite the lower expected annual gain.

GALLAIS (2009) proposed to use full sibs between DH lines in reciprocal recurrent selection (DHFSRRS). Compared with the classical RRS scheme based on full sibs between S_0 plants, the expected annual genetic gain is considerably higher when using DH lines. However, DHFSRRS cannot easily be integrated into LD, since every line is crossed

with a different line from the opposite pool rather than with one or more common tester(s). Moreover, under full sib selection the variance between crosses is inflated by the variance of SCA effects, thus reducing the gain from selection. On the other hand, under the here presented Standard Scheme of RS maximal progress can only be achieved if the testers adequately represent the opposite gene pool.

With MBP the integrated RS/LD schemes can be optimized for RS or LD or any intermediate weighting of the two goals depending on the emphasis given to long-term or short-term selection response. When optimizing the scheme for LD alone, the expected short-term response is maximized but the gain significantly decays after a few breeding cycles if the loss of genetic variance is not limited by a minimum N_e. In contrast, if the scheme is optimized for RS alone, the progress in LD is only slightly reduced (Table 3). This may become plausible from the graphical illustration of three consecutive RS/LD cycles presented in Fig. 8. After the first cycle, the superiority of the finally selected hybrid parent lines results to about equal parts from the response to RS and the 'added' gain from the second stage of selection devoted to LD alone. But, whereas the response to RS is accumulating from cycle to cycle, the added gain in LD cannot be transferred to the next cycle. This means, in the long run the progress in LD asymptotically approaches that of RS. Thus, generally it appears advisable to weight RS higher than LD.

The DH technology is not only suited for improving elite hybrid materials but is also a highly effective tool for intra-population improvement of landraces and open-pollinated varieties (OPVs). WILDE *et al.* (2009) derived DH lines from three European flint maize landraces and assessed their combining ability to an elite dent single-cross tester. The mean testcross performance of the DH lines was similar to that of their parental landraces. Highly significant genetic variance existed in each of the three DH-line groups for grain yield, maturity, and tolerance to nitrogen deficiency. The *per se* performance was acceptable for most of the lines. From this the authors conclude that during the haploid stage the lines are effectively purged from recessive detrimental alleles occurring in the parental gene pool. Nevertheless the first-cycle success rate in generating DH lines is much lower in OPVs than in elite materials.

ROTARENKO *et al.* (2009) conducted four cycles of recurrent mass selection for vigor of induced maternal haploids in a synthetic population composed of four elite Corn-Belt dent lines. Selection response for grain yield was 13% per cycle raising the performance of the synthetic up to the level of the check hybrids. This indicates that 'Haploid Recurrent Selection' (CHALYK and ROTARENKO, 1999) does not only act against detrimental recessive mutants but also effectively raises the frequency of alleles enhancing agronomic performance.

Breeders occasionally observe haploids that display phenotypic characteristics or marker alleles of the inducer (various pers. comm.). FISCHER (2004) and LI *et al.* (2009) used molecular markers to analyze this phenomenon. Both studies revealed evidence for paternal DNA transmission. In Fischer's experiment 1.3% of the haploid plants were not truly maternal whereas in Li's *et al.* study this proportion amounted to 43%. The two studies used different marker densities, donor genotypes and inducers. Further research is needed to clarify whether the amount of paternal transmission varies among inducers and/or in breeding populations or whether the deviating results are essentially due to the different marker densities. In haploids showing any degree of male transmission, the average proportion of the inducer genome was 1.8%.

Obviously, genome-wide selection can readily be incorporated into DH-line-based breeding schemes (Fig. 3). Results of simulation studies by BERNARDO and YU (2007) indicate that part of the expenses of field tests can be saved if marker scores are included in the selection index and if one or two marker-only stages follow the first (combined) selection stage. Moreover, the disturbing effect of genotype × environment interaction on selection would diminish if consistent marker effects could be obtained from data assessed across multiple years and genetic backgrounds. This would allow the breeder to more effectively select for phenotypic stability even with only one stage of field-based testcross selection per breeding cycle. Yet, while theoretical studies show great potential of combined phenotypic and genome-wide selection, extensive research is still needed to validate those results in practice (HEFFNER *et al.,* 2009).

In conclusion, the *in vivo* haploid induction technology has provided an exciting avenue to increase rate of progress in hybrid maize breeding. Theoretical and experimental research results encourage the breeder to take advantage from this new genetic tool.

REFERENCES

AUGER D.L., T.S. REAM, J.A. BIRCHLER, 2004 A test for a metastable epigenetic component of heterosis using haploid induction in maize. Theor. Appl. Genet. **108:** 1017-1023.

BARRET P., M. BRINKMANN, M. BECKERT, 2008 A major locus expressed in the male gametophyte with incomplete penetrance is responsible for *in situ* gynogenesis in maize. Theor. Appl. Genet. **117:** 581-594.

BELICUAS P.R., C.T. GUIMARÃES, L.V. PAIVA, J.M. DUARTE, W.R. MALUF, E. PAIVA, 2007 Androgenetic haploids and SSR markers as tools for the development of tropical maize hybrids. Euphytica **156:** 95-102.

BERNARDO R., J. YU, 2007 Prospects for genomewide selection for quantitative traits in maize. Crop Sci. **47:** 1082-1090.

BORDES J., R.D. DE VAULX, A. LAPIERRE, M. POLLACSEK, 1997 Haplodiploidization of maize (*Zea mays* L.) through induced gynogenesis assisted by glossy markers and its use in breeding. Agronomie **17:** 291-297.

BRIM C.A., H.W. JOHNSON, C.C. COCKERHAM, 1959 Multiple selection criteria in soybeans. Agron J. **51:** 42-46.

BÜTER B., 1997 *In vitro* haploid production in maize. pp. 37-71. *In:* S.M. Jain, S.K. Sopory, R.E. Veilleux (Eds.), *In vitro* haploid production in higher plants. Kluwer Academic Publishers, Dordrecht, Boston, London.

CHALYK S.T., 1994 Properties of maternal haploid maize plants and potential application to maize breeding. Euphytica **79:** 13-18.

CHALYK S.T., V.A. ROTARENKO, 1999 Using maternal haploid plants in recurrent selection in maize. Maize Genet. Coop. Newsletter **73:** 56-57.

CHALYK S.T., A. BAUMANN, G. DANIEL, J. EDER, 2003 Aneuploidy as a possible cause of haploid-induction in maize. Maize Genet. Coop. Newsletter **77:** 29-30.

CHASE S.S., 1952 Monoploids in maize. pp. 389-399. *In:* J.W. Gowen (Ed.), Heterosis. Iowa State College Press, Ames, IA, USA.

CHASE S.S., 1974 Utilization of haploids in plant breeding: Breeding diploid species. pp. 211-230. *In:* K.J. Kasha (Ed.), Proc. Intl. Symposium on Haploids in Higher Plants.

COE E.H., 1959 A line of maize with high haploid frequency. Am. Nat. **93:** 381-382.

COWEN N.M., C.D. JOHNSON, K. ARMSTRONG, M. MILLER, A. WOOSLEY, S. PESCITELLI, M. SKOKUT, S. BELMAR, J.F. PETOLINO, 1992 Mapping genes conditioning *in vitro* androgenesis in maize using RFLP analysis. Theor. Appl. Genet. **84:** 720-724.

DEIMLING S., F. RÖBER, H.H. GEIGER, 1997 Methodik und Genetik der *in-vivo*-Haploideninduktion bei Mais. Vortr Pflanzenzüchtg. **38:** 203-224.

EDER J., S.T. CHALYK, 2002 *In vivo* haploid induction in maize. Theor. Appl. Genet. **104:** 703-708.

FISCHER E., 2004 Molekulargenetische Untersuchungen zum Vorkommen paternaler DNA-Übertragung bei der *in-vivo*-Haploideninduktion bei Mais (*Zea mays* L.). Ph.D. thesis, University of Hohenheim, Stuttgart, Germany.

GALLAIS A., 2009 Full-sib reciprocal recurrent selection with the use of doubled haploids. Crop Sci. **49:** 150-152.

GAYEN P., J.K. MADAN, R. KUMAR, K.R. SARKAR, 1994 Chromosome doubling in haploids through colchicine. Maize Genet. Coop. Newsletter **68:** 65.

GEIGER H.H., 2009 Doubled haploids. pp. 641-659. *In:* J.L. Bennetzen, S. Hake (Eds.), Maize Handbook. Vol. II: Genetics and Genomics. Springer Verlag, Heidelberg, New York.

GEIGER H.H., M.D. BRAUN, G.A. GORDILLO, S. KOCH, J. JESSE, B.A.E. KRÜTZFELDT, 2006 Variation for female fertility among haploid maize lines. Maize Genet Coop. Newsletter **80:** 28-29.

GORDILLO G.A., H.H. GEIGER, 2008a MBP (version 1.0): A software package to optimize maize breeding procedures based on doubled haploid lines. J. Hered. **99:** 227-231.

GORDILLO G.A., H.H. GEIGER, 2008b Optimization of DH-line based recurrent selection procedures in maize under a restricted annual loss of genetic variance. Euphytica **161:** 141-154.

GORDILLO G.A., H.H. GEIGER, 2008c Alternative recurrent selection strategies using doubled haploid lines in hybrid maize breeding. Crop Sci. **48:** 911-922.

GREENBLATT I.M., M. BOCK, 1967 A commercially desirable procedure for detection of monoploids in maize. J. Hered. **58:** 9-13.

HALLAUER A.R., J.B. MIRANDA, 1981 Quantitative genetics in maize breeding. Iowa State Univ. Press, Ames, IA, USA.

HEFFNER E.L., M.E. SORRELLS, J.-L. JANNINK, 2009 Genomic selection for crop improvement. Crop Sci **49:** 1-12.

LASHERMES P., M. BECKERT, 1988 A genetic control of maternal haploidy in maize (*Zea mays* L.) and selection of haploid inducing lines. Theor. Appl. Genet. **76:** 405-410.

LI L., X. XU, W. JIN, S. CHEN, 2009 Morphological and molecular evidences for DNA introgression in a haploid induction via a high oil inducer CAUHOI in maize. Planta **230:** 367-376.

LONGIN C.F.H., H.F. UTZ, J.C. REIF, W. SCHIPPRACK, A.E. MELCHINGER, 2006 Hybrid maize breeding with doubled haploids: I. One-stage versus two-stage selection for testcross performance. Theor. Appl. Genet. **112:** 903-912.

MEUWISSEN T.H.E., B.J. HAYES, M.E. GODDARD, 2001 Prediction of total genetic value using genome-wide dense marker maps. Genetics **157:** 1819-1829.

MURIGNEUX A., 1994 Genotypic variation of quantitative trait loci controlling *in vitro* androgenesis in maize. Genome **37:** 970-976.

NANDA D.K., S.S. CHASE, 1966 An embryo marker for detecting monoploids of maize (*Zea mays* L.). Crop Sci. **6:** 213-215.

PETOLINO J.F., S.A. THOMPSON, 1987 Genetic analysis of anther culture response in maize. Theor. Appl. Genet. **74:** 284-286.

PRESTERL T., M. OUZUNOVA, W. SCHMIDT, E.M. MÖLLER, F.K. RÖBER, C. KNAAK, K. ERNST, P. WESTHOFF, H.H. GEIGER, 2007 Quantitative trait loci for early plant vigour of maize grown in chilly environments. Theor. Appl. Genet. **114:** 1059-1070.

RÖBER F.K., G.A. GORDILLO, H.H. GEIGER, 2005 *In vivo* haploid induction in maize - Performance of new inducers and significance of doubled haploid lines in hybrid breeding. Maydica **50:** 275-283.

ROTARENCO V.A., J. EDER, 2003 Possible effects of heterofertilization on the induction of maternal haploids in maize. Maize Genet. Coop. Newsletter **77:** 30.

ROTARENCO V.A., I.H. KIRTOCA, A.G. JACOTA, 2007 Possibility to identify kernels with haploid embryo by oil content. Maize Genet. Coop. Newsletter **81:** 11.

ROTARENCO V.A., G. DICU, M. SARMANIUC, 2009 Induction of maternal haploids in maize. Maize Genet. Coop. Newsletter **83** (http://www.agron.missouri.edu/mnl/83/46rotarenco.htm).

SANTIAGO E., A. CABALLERO, 1995 Effective size of populations under selection. Genetics **139:** 1013-1030.

SARKAR K.R., A. PANDEY, P. GAYEN, J.K. MANDAN, R. KUMAR, J.K.S. SACHAN, 1994 Stabilization of high haploid inducer lines. Maize Genet. Coop. Newsletter **68:** 64-65.

SEITZ G., 2005 The use of doubled haploids in corn breeding. pp. 1-7. *In:* Proc. 41th Annual Illinois Corn Breeders' School 2005. Urbana-Champaign, IL, USA.

SHATSKAYA O.A., E.R. ZABIROVA, V.S. SHCHERBAK, 1994a Autodiploid lines as sources of haploid spontaneous diploidization in corn. Maize Genet. Coop. Newsletter **68:** 51-52.

SHATSKAYA O.A., E.R. ZABIROVA, V.S. SHCHERBAK, M.V. CHUMAK, 1994b Mass induction of maternal haploids in corn. Maize Genet. Coop. Newsletter **68:** 51.

SPITKÓ T., L. SÁGI, J. PINTÉR, L.C. MARTON, B. BARNABÁS, 2006 Haploid regeneration aptitude of maize (*Zea mays* L.) – Lines of various origin and of their hybrids. Maydica **51:** 537-542.

TYRNOV V.S., A.N. ZAVALISHINA, 1984 Inducing high frequency of matroclinal haploids in maize. Dokl. Akad. Nauk. SSSR **276:** 735-738.

WEDZONY M., F. RÖBER, H.H. GEIGER, 2002 Chromosome elimination observed in selfed progenies of maize inducer line RWS. pp. 173. *In:* 7th Intl. Congress on Sexual Plant Reproduction. Maria Curie-Sklodowska University Press, Lublin, Poland.

WEGENAST T., C.F.H. LONGIN, H.F. UTZ, A.E. MELCHINGER, H.P. MAURER, J.C. REIF, 2008 Hybrid maize breeding with doubled haploids. IV. Number versus size of crosses and importance of parental selection in two-stage selection for testcross performance. Theor. Appl. Genet. **117:** 251-260.

WILDE K., H. BURGER, V. PRIGGE, T. PRESTERL, W. SCHMIDT, M. OUZUNOVA, H.H. GEIGER, 2009 Testcross performance of doubled-haploid lines developed from European flint maize landraces. Plant Breed. doi:10.1111/j.1439-0523.2009.01677.x

WILLIAMS J.S., 1962 The evaluation of a selection index. Biometrics **18:** 375-393.

ZABIROVA E.R., O.A. SHATSKAYA, V.S. SHCHERBAK, 1993 Line 613/2 as a source of a high frequency of spontaneous diploidization in corn. Maize Genet. Coop. Newsletter **67:** 67.