ORIGINAL PAPER

Investigating the efficiency of the single backcrossing breeding strategy through computer simulation

Jiankang Wang · Ravi P. Singh ·
Hans-Joachim Braun ·
Wolfgang H. Pfeiffer

Abstract A strategy combining single backcrossing with selected bulk breeding has been successfully used in wheat improvement at CIMMYT to introgress rust resistant genes from donor parents to elite adapted cultivars. In this research, the efficiency of this breeding strategy was compared to other crossing and selection strategies through computer simulation. Results indicated this breeding strategy has advantages in retaining or improving the adaptation of the recurrent parents, and at the same time transferring most of the desired donor genes in a wide range of scenarios. Two rounds of backcrossing have advantages when the adaptation of donor parents is much poorer than that of the adapted parents, but the advantage of three rounds of backcrossing over two rounds is minimal. We recommend using the single backcrossing breeding strategy (SBBS) when three conditions are met: (1) multiple genes govern the phenotypic traits to be transferred from donor parents to adapted parents, (2) the donor parents have some favorable genes that may contribute to the improvement of adaptation in the recipient parents, and (3) conventional phenotypic selection is being applied, or individual genotypes cannot be precisely identified. We envisage that all three conditions commonly exist in modern breeding programs, and therefore believe that SBBS could be applied widely. However, we do not exclude the use of repeated backcrossing if the transferred genes can be precisely identified by closely linked molecular markers, and the donor parents have extremely poor adaptation.

J. Wang (✉)
Institute of Crop Science, CIMMYT China and The National
Key Facility for Crop Gene Resources and Genetic
Improvement, Chinese Academy of Agricultural Sciences,
No. 12 Zhongguancun South Street, 100081 Beijing, China
e-mail: wangjk@caas.net.cn; jkwang@cgiar.org

R. P. Singh · H.-J. Braun
International Maize and Wheat Improvement Center
(CIMMYT), Apdo. Postal 6-641, 06600 Mexico D.F., Mexico

W. H. Pfeiffer
HarvestPlus Challenge Programme and International Center
for Tropical Agriculture (CIAT-HarvestPlus), A.A., 6713 Cali,
Colombia

Introduction

The main elements of the wheat improvement program of the International Maize and Wheat Improvement Center (also known as CIMMYT, the Spanish acronym for Centro Internacional de Mejoramiento de Maiz y Trigo) have been shuttle breeding at two contrasting locations in Mexico, wide adaptation, durable resistance to rust and Septoria diseases, international multi-environment testing, and the appropriate use of genetic variation to enhance yield gains (Rajaram 1999; van Ginkel et al. 2002; Ortiz et al. 2007). Each year, CIMMYT grows two wheat cycles in Mexico, its host country: one from November to April in Ciudad Obregón (27°N, 39 m above sea level), and the other from May to October in Toluca (19°N, 2,640 m above sea level). The development of high yielding, widely adapted, stable wheat germplasm with durable disease resistance and acceptable end-use qualities is the main objective of CIMMYT's wheat breeding efforts. To achieve this objective, CIMMYT wheat breeders have applied various breeding strategies.

Pedigree selection was used primarily from 1944 to 1985. From 1985 until the second half of the 1990s, the main selection method was a modified pedigree/bulk method (MODPED) (van Ginkel et al. 2002; Wang et al.

2003), which successfully produced many of the widely adapted wheat cultivars now being grown in the developing world. This method was replaced in the late 1990s by the selected bulk method (SELBLK) (Singh et al. 1998b; van Ginkel et al. 2002) in an attempt to improve resource-use efficiency. The MODPED method begins with pedigree selection of individual plants in the F_2 generation, followed by three generations (F_3–F_5) of bulk selection, and pedigree selection in the F_6. In the SELBLK method, spikes of selected F_2 plants within each cross are harvested in bulk and threshed together, resulting in one F_3 seed lot per cross. Selected bulk selection is also used from F_3 to F_5, while pedigree selection is used only in the F_6 (Wang et al. 2003, 2004). The major advantage of SELBLK over MODPED is that fewer seed lots need to be harvested, threshed, and visually selected for seed appearance. In addition, with SELBLK there are significant savings in time, labor, and costs associated with nursery preparation, planting, and plot labeling, and potential sources of error are avoided (van Ginkel et al. 2002). Simulations indicated the savings in resources did not result in a penalty in genetic gains on yield (Wang et al. 2003). In addition, the number of crosses remaining from SELBLK is always higher than that from MODPED, which means that delaying pedigree selection favors diversity.

Regarding the crossing strategies, top (or three-way) crosses and double (or four-way) crosses were employed to increase the genetic variability of breeding populations in the early 1970s. By the late 1970s, double crosses were dropped due to their poor results relative to single crosses, top crosses, and limited backcrosses. From the 1980s onwards, all crosses onto selected F_1s were either single crosses, backcrosses, or top crosses (van Ginkel et al. 2002). Single and top (or three-way) crosses are commonly used among adapted parental lines, while backcrosses are preferred for transferring a few useful genes from donor parents to adapted lines (Allard 1960; Stoskopf et al. 1993). In CIMMYT, the single backcrossing approach (one backcross to the adapted parent) was initially aimed at incorporating resistance to rust diseases based on multiple additive genes (Singh and Huerta-Espino 2004). However, it soon became apparent that the single backcross approach also favored selection of genotypes with higher yield potential. The reason why single backcrossing shifts the progeny mean toward the higher side is that it favors the retention of most of the desired major additive genes from the recurrent, while simultaneously allowing the incorporation and selection of additional useful small-effect genes from the donor parents.

Reddy and Comstock (1976) studied the effects of heritability and gene number on fixation of desired alleles from a simplified backcross breeding method. The efficiency of using backcrossing to introgress one or two major genes that can be uniquely identified by genetic markers has been investigated by various authors (Frisch and Melchinger 2001; Ribaut et al. 2002; Bregitzer et al. 2008). In their studies, repeated backcrossing is useful to recover the recipient genome to the greatest possible without worrying about losing of the introgressed genes. However, the single-gene inherited traits are rare in breeding, as is a donor parent with only one or two useful genes. Some genes may have major effects but are without completely linked molecular markers that breeders can use. For most breeding traits, such as durable rust resistances and grain quality, conventional phenotypic selection is still valid and widely applied. Therefore, repeated backcrossing may not be desirable when breeding for durable polygenic resistance, as it was devised to incorporate a single, or a few, major gene(s) while causing the least disturbance to the genetic make-up of the recurrent parent. Coincidentally, Bregitzer et al. (2008) recently reported the efficiency of single backcrossing for eliminating agronomic and quality alterations caused by somaclonal variation in transgenic barley.

QuLine (previously called QuCim, and freely available from http://www.uq.edu.au/lcafs/qugene/) is a QU-GENE (Podlich and Cooper 1998) application module that was specifically developed to simulate wheat breeding at CIMMYT (Wang et al. 2003, 2004), but has the potential to simulate most methodologies for developing inbred lines. It has been used to compare two selection strategies (Wang et al. 2003), to study the effects on selection of dominance and epistasis (Wang et al. 2004), to predict cross-performance using known gene information (Wang et al. 2005), to optimize marker assisted selection to efficiently pyramid multiple genes (Wang et al. 2007a), and to investigate the use of identified QTL-marker associations and design-led breeding approaches for improving rice quality (Wang et al. 2007b). Our objective in this study was to use QuLine to investigate the efficiency of a single backcrossing breeding strategy relative to other strategies involving repeated backcrosses, in a wide range of scenarios, where there are many genes to be introgressed through conventional phenotypic selection.

Materials and methods

The genotype by environment system used in simulation

The genotype by environment (GE) system defined in QU-GENE (Podlich and Cooper 1998; Wang et al. 2003, 2004) consists of all the required information on environment types, traits, and genes, as well as their effects on traits in different environment types, and trait heritabilities, among

the most important factors. One environment type (i.e., no gene by environment interaction was considered) and two general traits were defined in this study. One trait is so-called adaptation, and the other is the trait to be introduced from the donor parents, or introduced donor trait (DT). Adaptation is a combined index of various breeding target traits (except DT), such as maturity, plant height, yield and yield components, and quality. Traits that need to be introduced from the donor parents are not included in the definition of adaptation. Though defined as one single trait in the simulation study, DT can include two or more traits that are lacking in the adapted lines and need to be transferred from the less adapted parents.

We consider 200 additive genes contributing to the expression of adaptation, and ten additive genes affecting DT, which are distributed on the 21 chromosomes. Ten adaptation genes and one DT gene are evenly distributed on each of the first ten chromosomes, and the distance between two neighboring genes is set at 10 cM. On each of the other 11 chromosomes, ten adaptation genes are evenly distributed and the distance between two neighboring genes is also set at 10 cM. Assuming A and a are the two alleles at an adaptation locus, the adaptation values of the three genotypes AA, Aa, and aa are 0.5, 0.25, and 0, respectively. Therefore, the highest adaptation value is 100, which indicates all the favorable alleles are fixed, and the lowest value is 0, which means no favorable alleles are present. Assuming D and d are the two alleles at a locus affecting DT, the DT values of the three genotypes DD, Dd, and dd are 1, 0.5, and 0, respectively. Therefore, the highest DT value is 10 (all favorable alleles present), and the lowest value is 0 (no favorable alleles present). Broad-sense heritability at the individual plant level was set at 0.5 for both traits.

Parental lines used in the CIMMYT's Global Wheat Program

Breeding materials have been grouped according to mega-environments (MEs; Rajaram et al. 1994; Rajaram 1999). Within each ME, lines are sub-grouped based on country of origin or specific character expression, e.g., disease resistance, abiotic stress tolerance, and industrial quality. Parental lines used in breeding for each ME include (1) major cultivars released in different target countries (i.e., elite adapted lines, or EAL), (2) elite CIMMYT and other germplasm identified through international and national testing (i.e., adapted lines, or AL), and (3) advanced lines exhibiting desirable expression of one specific trait or group of traits, including those developed through inter-specific hybridization (Mujeeb-Kazi and Hettel 1995), and other germplasm within the Global Wheat Program (i.e., intermediate adapted lines, or IAL) (Table 1; Rajaram 1999; van Ginkel et al. 2002).

Considerable genetic diversity enters the breeding system in the form of introductions from most collaborating countries. In addition, inter-specific hybridization and breeding to host plant resistances provide unique gene combinations for inclusion in the breeding program (Table 1). Once introduced into the breeding programs, these materials are classified according to MEs with regard to disease resistance, agronomic type, and adaptability, and then considered for crossing. The adaptation of an individual is actually determined by the frequency of favorable genes and gene combinations, which, in QU-GENE and QuLine, allows the definition of different parental groups based on different gene frequencies. Table 1 gives a summary of CIMMYT's wheat germplasm, classified according to the percentage of favorable adaptation genes

Table 1 Estimated percentages of favorable alleles or gene combinations in different parental lines used in wheat breeding at CIMMYT

Category	Favorable genes[a] (%)	Example	Total parental lines (%)
Elite adapted lines (EAL)	80–85	Major released cultivars in targeted mega-environments (MEs) either developed by CIMMYT or by partners	10
Adapted lines (AL)	75–80	Elite advanced lines from CIMMYT's International Nursery and Yield Trials	60
Intermediate adapted lines (IAL)	65–75	Advanced lines from CIMMYT's Yield Trials in Ciudad Obregón and Toluca, Mexico	10
Un-adapted (or non-adapted) lines (UAL)	20–40	Land races	2
Second generation of re-synthesized wheat (SYNII)	40–60	Derived lines between the first generation of re-synthesized wheat derivatives and adapted lines	10
First generation of re-synthesized wheat (SYNI)	20–40	Derived lines between primary re-synthesized wheat and adapted lines	5
Primary re-synthesized wheat (SYN0)	0–30	Inter-specific crosses between *Triticum durum* and *Aegilops tauschii*	3

[a] Percentages of favorable alleles were estimated by CIMMYT wheat breeders

Table 2 Crosses between different parental groups used in wheat breeding at CIMMYT

Category[a]	Percentage of total crosses[b]
(EAL + AL) × (EAL + AL)	65
(EAL + AL) × IAL	10
(EAL + AL) × UAL	5
(EAL + AL) × SYNII	10
(EAL + AL) × SYNI	7
(EAL + AL) × SYN0	3

[a] See Table 1 for the meaning of each parental group

[b] Percentages of total crosses were estimated by CIMMYT wheat breeders

and gene combinations, estimated by CIMMYT wheat breeders.

About 2,000 crosses are made onto these introductions every year, and twice a year, around 15% of the parental stocks are replaced with outstanding introductions from within and outside CIMMYT. Each season about 65% of crosses are between adapted and elite adapted lines, i.e., (EAL + AL) × (EAL + AL), to achieve the short-term breeding objective of combining different favorable genes or gene combinations in one breeding cycle (Table 2; van Ginkel et al. 2002). Other crosses are targeted at mid- and long-term breeding objectives (Table 2).

Single backcrossing combined with the selected bulk selection strategy

To transfer multiple gene based resistance to a susceptible adapted cultivar or any other selected genotypes, CI-MMYT's breeders use a "single backcross-selected bulk" scheme (Singh and Huerta-Espino 2004; Singh and Tretho-wan 2007), where a cultivar or any other genotype is crossed with a group of 8–10 resistance donors; 20 spikes of the F_1

plants from each cross are then backcrossed to obtain 400–500 BC_1F_1 seeds (400 in Table 3). From the BC_1F_1 generation onward, selection for resistance and other agronomic features is conducted under high rust pressure. Because additive genes are partially dominant, BC_1F_1 plants carrying most of the genes show intermediate resistance and can be selected visually. About 1,200 plants per cross are space-grown in the BC_1F_2, whereas about 400 plants are maintained in the BC_1F_3 to BC_1F_5 populations (Table 3). Plants with desirable agronomic features and low to moderate terminal disease severity in early generations (BC_1F_1, BC_1F_2, and BC_1F_3) and plants with low terminal severity in later generations (BC_1F_4 and BC_1F_5) are retained. Bulk selection is used until the BC_1F_4 generation, and pedigree selection is used in the BC_1F_5 (Table 3).

The SBBS described above can be readily defined in QuLine (Wang et al. 2004). In this simulation study, we consider that selection for resistance is equivalent to selection for DT, and that selection for other agronomic traits is equivalent to selection for adaptation.

The design of simulation experiments

We assumed that the frequency of favorable adaptation genes in the adapted lines was 0.8 (Table 1), and the frequency of favorable DT genes in the donor lines was 1. We used "group" to represent a set of adapted parental lines or a set of donor parental lines with similar adaptation and DT defined by gene frequency. Different adapted parental groups were denoted by A0, A2, A4, A6, and A8, to represent the frequencies of favorable DT genes: 0, 0.2, 0.4, 0.6, and 0.8, respectively. Different donor parental groups were denoted by D0, D1, D2, D3, D4, D5, D6, and D7, to represent the frequencies of favorable adaptation genes : 0 (e.g., SYN0), 0.1 (e.g., SYN0), 0.2 (e.g., SYN0), 0.3 (e.g., UAL), 0.4 (e.g., UAL), 0.5 (e.g., SYNII), 0.6 (e.g., SYNII),

Table 3 The single backcrossing breeding strategy currently used in wheat breeding at CIMMYT

Generation	Seed propagation method[a]	No. of crosses or families grown	Individuals per cross or family	No. of selected crosses or families	No. of selected individuals in each cross or family
F_1	Hand pollination between adapted and donor lines	100	20	100	20
BC_1F_1	Backcrossing to the adapted parents	100	400	100	50
BC_1F_2	Selfing	100	1200	100	30
BC_1F_3	Selfing	100	400	100	10
BC_1F_4	Selfing	100	400	100	10
BC_1F_5	Selfing	100	400	100	10
BC_1F_6	Selfing	1,000	200	10	200
Final selected advanced lines		10			

[a] The seed propagation method specifies how the seed is formed for the next generation (Wang et al. 2004). Available methods in QuLine are clone, selfing, random mating, backcross, topcross, doubled haploids, and noselfing (random mating but self pollination avoided)

Table 4 Generation title and seed harvest method (or generation advance method) with four crossing strategies used in simulation

Crossing strategies[a]				Generation advance method[b]
B0	B1	B2	B3	
F_1	F_1	F_1	F_1	Bulk
F_2	BC_1F_1	BC_1F_1	BC_1F_1	Bulk
F_3	BC_1F_2	BC_2F_1	BC_2F_1	Bulk
F_4	BC_1F_3	BC_2F_2	BC_3F_1	Bulk
F_5	BC_1F_4	BC_2F_3	BC_3F_2	Bulk
F_6	BC_1F_5	BC_2F_4	BC_3F_3	Pedigree
F_7	BC_1F_6	BC_2F_5	BC_3F_4	Bulk

[a] No backcrossing is done in B0. B1 is the single backcrossing breeding currently used in wheat breeding at CIMMYT. Two rounds of backcrossing are done in B2, and three rounds of backcrossing are done in B3

[b] There are two options in QuLine for the generation advance method (Wang et al. 2004): pedigree and bulk. When pedigree is used, each selected individual will form a family in the next generation. When bulk is used, all selected individuals in a family will be harvested in bulk and form one family in the next generation

and 0.7 (e.g., IAL), respectively. A total of 40 potential parental groups for crossing, i.e., between (A0, A2, A4, A6, A8) and (D0, D1, D2, D3, D4, D5, D6, D7), are made in QuLine. In each group 100 crosses are made between 40 adapted lines and ten donor lines. The GE system is used to calculate the phenotypic values for defined traits of each individual genotype, and different adapted and donor parental groups are defined by the QU-GENE engine.

We use notation B1 to represent the SBBS defined in Table 3. Three other crossing strategies were also considered for comparison (Table 4): B0 for no backcrossing, B2 for two rounds of backcrossing, and B3 for three rounds of backcrossing (Table 4). For a reasonable comparison, the same segregating population size and selection intensity were used with all four crossing strategies (Table 4).

Besides the four crossing strategies, we also considered six selection schemes (Table 5), relevant to wheat breeding at CIMMYT. For example, breeders sometimes select for agronomic traits first, then select for host plant resistances. This strategy will be referred to as AD. When breeders select for host plant resistance first and then select for agronomic traits; this is similar to DA. At other times, breeders select for some agronomic traits first, then for host plant resistance, and finally for the remaining agronomic traits; this is similar to ADA. When breeders select for resistance at flowering, and again at maturity, this may be similar to DAD. ADAD and DADA indicate that adaptation and DT are selected for twice in a generation. These crossing and selection strategies can be readily defined in QuLine as well (Wang et al. 2003, 2004).

A total of 960 scenarios were simulated in this study: five groups of adapted lines having different numbers of DT genes (A0, A2, A4, A6, and A8), eight groups of donor lines having different adaptations (D0, D1, D2, D3, D4, D5, D6, and D7), four crossing strategies (B0, B1, B2, and B3), and six selection schemes (AD, DA, ADA, DAD, ADAD, and DADA). Each scenario was simulated 100 times, and the top ten advanced lines with highest adaptation and DT were finally selected and used for comparison.

Results

Comparison of the genetic gains from different selection schemes

Under a specified GE system, the response to selection depends on the selection intensity applied. In selection schemes AD and DA, the same intensity was applied for both traits (Table 5), resulting in similar genetic advances on both traits (Table 6). However, there is a trend toward a

Table 5 Selected proportions for individual traits in each generation using six selection schemes

Generation	Selection scheme[a]					
	AD	DA	ADA	DAD	ADAD	DADA
F_1	No selection applied					
BC_1F_1, BC_2F_1, and BC_3F_1	0.354	0.354	0.500	0.500	0.595	0.595
F_2, BC_1F_2, BC_2F_2, BC_3F_2	0.158	0.158	0.292	0.292	0.398	0.398
F_3 to F_6, BC_1F_3 to BC_1F_5, BC_2F_3, BC_2F_4, BC_3F_3	0.158	0.158	0.292	0.292	0.398	0.398
F_7, BC_1F_6, BC_2F_5, BC_3F_4	0.100	0.100	0.215	0.215	0.316	0.316

[a] Each letter in a selection scheme represents the trait [i.e., A is for adaptation, and D is for transferred donor traits (DT)] to be selected; the order of the letters represents the order in which they are selected. For simplicity, the same selection scheme is applied for all selected traits. Taking AD as an example, in generation BC_1F_1, 35.4% of individuals are first selected based on their adaptation, and then 35.4% of the retained individuals are selected based on DT, resulting in a total selected proportion of 12.5%. As the population size of BC_1F_1 is 400, 50 individuals are selected. Although a top selected proportion is given in this table, other options are also available in QuLine, such as selecting a fixed number of individuals and families, or selecting all individuals or families above or below a predefined phenotypic value (Wang et al. 2004)

Table 6 Mean genetic values for adaptation and transferred donor traits (DT) when the adapted parents will not have favorable genes introduced from the donor parents

Selection scheme	Crossing scheme	Adaptation								DT							
		D0	D1	D2	D3	D4	D5	D6	D7	D0	D1	D2	D3	D4	D5	D6	D7
AD	B0	59.1	62.6	66.0	69.1	72.5	75.7	79.5	83.1	8.7	8.8	8.9	8.9	9.1	9.2	9.3	9.4
	B1	75.2	76.4	77.7	78.8	80.3	81.6	83.2	85.0	6.1	6.2	6.3	6.5	6.7	6.9	7.1	7.3
	B2	80.8	81.3	81.7	82.3	82.9	83.5	84.5	85.3	4.3	4.4	4.6	4.7	4.9	5.0	5.3	5.5
	B3	82.9	83.0	83.3	83.6	84.0	84.3	84.7	85.2	2.8	2.9	3.1	3.2	3.3	3.4	3.7	3.8
DA	B0	56.8	60.4	64.0	67.6	71.4	74.6	78.8	82.5	9.3	9.2	9.2	9.3	9.4	9.4	9.5	9.6
	B1	73.3	74.6	76.3	77.6	79.3	80.6	82.5	84.1	6.8	6.8	6.9	7.0	7.1	7.2	7.4	7.7
	B2	79.8	80.2	80.9	81.5	82.4	83.1	84.0	84.6	4.7	4.8	4.9	5.0	5.2	5.3	5.5	5.7
	B3	82.3	82.7	82.9	83.2	83.7	84.0	84.5	84.5	3.1	3.2	3.3	3.4	3.5	3.7	3.7	4.0
ADA	B0	62.5	65.6	68.6	71.9	75.4	78.0	81.7	85.0	7.8	7.8	7.9	8.0	8.2	8.3	8.5	8.7
	B1	78.0	79.1	80.2	81.1	82.5	83.5	85.1	86.8	4.9	5.0	5.2	5.3	5.5	5.7	6.0	6.2
	B2	83.1	83.4	83.9	84.3	84.8	85.4	86.0	86.8	3.2	3.2	3.4	3.6	3.7	3.9	4.1	4.3
	B3	84.7	84.8	84.9	85.2	85.4	85.8	86.0	86.5	1.9	2.0	2.1	2.3	2.4	2.6	2.7	2.8
DAD	B0	52.9	56.8	60.9	64.6	68.9	72.3	76.9	80.8	9.9	9.9	9.9	9.9	9.9	9.9	9.9	10.0
	B1	70.2	71.8	73.3	75.4	77.0	78.8	80.9	82.6	7.9	8.0	8.1	8.2	8.3	8.3	8.5	8.7
	B2	77.4	78.0	78.8	79.5	80.5	81.3	82.5	83.4	5.9	6.1	6.2	6.2	6.3	6.4	6.6	6.7
	B3	80.3	80.7	81.1	81.6	81.9	82.5	83.1	83.5	4.3	4.4	4.5	4.5	4.6	4.7	4.8	4.9
ADAD	B0	58.8	62.4	65.7	69.0	73.0	75.8	79.8	83.3	9.1	9.2	9.2	9.3	9.3	9.4	9.6	9.6
	B1	75.1	76.5	77.5	79.0	80.4	81.8	83.7	85.4	6.5	6.5	6.7	6.8	7.0	7.2	7.5	7.5
	B2	81.0	81.6	81.8	82.7	83.1	83.8	84.8	85.6	4.5	4.6	4.7	4.9	5.1	5.2	5.5	5.6
	B3	83.0	83.3	83.5	83.8	84.2	84.4	84.9	85.3	3.1	3.2	3.3	3.5	3.6	3.8	3.9	4.0
DADA	B0	57.6	61.2	64.7	68.4	72.3	75.3	79.6	83.1	9.3	9.4	9.4	9.5	9.5	9.6	9.7	9.7
	B1	74.1	75.5	76.8	78.4	79.8	81.4	83.3	84.8	6.9	6.9	7.0	7.1	7.3	7.3	7.6	7.7
	B2	80.4	81.1	81.3	82.2	82.8	83.5	84.6	85.4	4.8	4.9	5.0	5.1	5.3	5.5	5.6	5.8
	B3	82.7	83.0	83.2	83.6	83.9	84.4	84.9	85.2	3.3	3.3	3.5	3.5	3.7	3.7	4.0	4.1

slightly faster genetic advance on the trait that was selected first (Table 6), indicating that the most important traits should be selected first in breeding. In ADAD and DADA, each trait was selected twice in a generation at different wheat development stages, e.g., host plant resistance to rust was selected at grain-filling and at maturity. But total selection intensity was the same as that for AD and DA. Therefore, genetic advances on adaptation and DT with ADAD and DADA are similar to those achieved with AD and DA (Table 6).

In scheme ADA, adaptation was selected for twice, and DT was selected once. The selection intensity for adaptation was higher than the intensity for DT. On the contrary, the selection intensity for DT was higher than the intensity for adaptation in scheme DAD. Thus, a faster genetic advance on adaptation was observed with ADA, while a faster advance on DT was observed with DAD (Table 6). In Table 6, the frequency of the introduced genes was 0 in the 40 adapted parents (A0). But similar results can be observed for adapted parents with other DT gene frequencies (A2, A4, A6, and A8; results not shown). The

little difference in genetic advance caused by the order of selected traits indicates that less expensive phenotyping traits can be selected first to minimize total breeding expenses.

Comparison of genetic gains from different crossing strategies

Genetic advances on adaptation and DT achieved with selection schemes AD, ADA, and DAD are shown in Fig. 1 for all scenarios. Other schemes (DA, ADAD, and DADA)

Fig. 1 Genetic means for adaptation and transferred donor traits ▶ (DT) in the final selected ten top lines. **a** Genetic mean for adaptation with selection scheme AD (in each generation adaptation is selected first, followed by selection for DT). **b** Genetic mean for DT with selection scheme AD. **c** Genetic mean for adaptation with selection scheme ADA (adaptation is selected first, followed by selection for DT, and adaptation is re-selected in each generation). **d** Genetic mean for DT with selection scheme ADA. **e** Genetic mean for adaptation with selection scheme DAD (DT is selected first, followed by selection for adaptation, and DT is re-selected in each generation). **f** Genetic mean for DT with selection scheme DAD
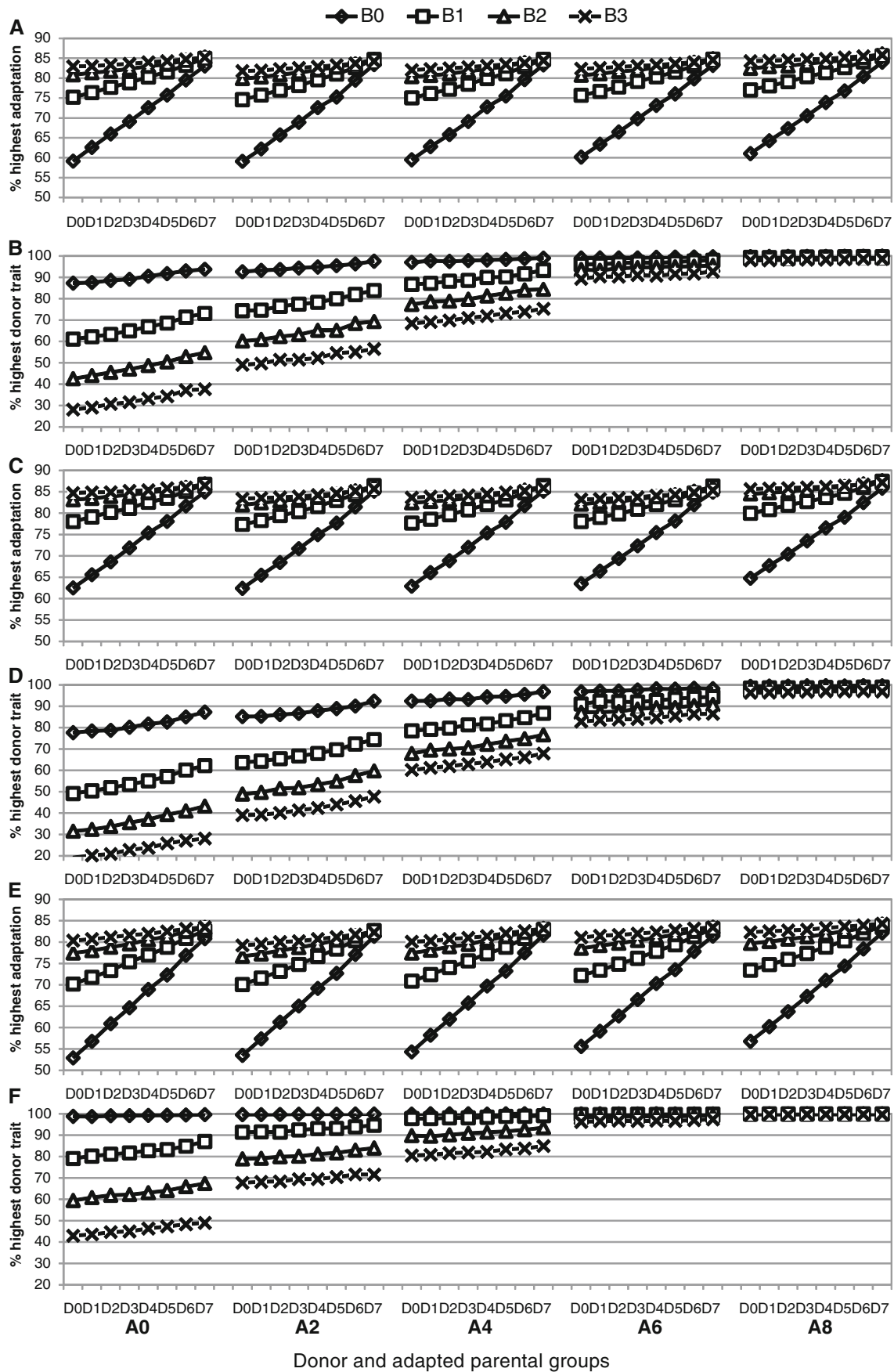
produced results similar to those from the AD scheme (Table 6) and therefore are not shown. As expected, backcrossing (B1, B2, or B3) appears to be more efficient in improving adaptation when the donor parents are less adapted (Fig. 1a, c, e). For example, when the adaptation of donor parents is around 20 (D2), the mean adaptations of the top ten lines derived from single crosses (B0) were 66.0, 68.6, and 60.9 for selection schemes AD, ADA, and DAD, respectively (Table 6). One round of backcrossing (B1) increases these values to 77.7, 80.2, and 73.3, respectively, whereas two rounds of backcrossing (B2) increase them to 81.7, 83.9, and 78.8, respectively, and three rounds of backcrossing (B3) increase them to 83.3, 84.9, and 81.1, respectively.

Single backcrossing (B1) had significant advantages over no backcrossing (B0), unless the adaptation of the donor parents was similar to that of the adapted lines, i.e., D7 (Fig. 1a, c, e). Two rounds of backcrossing (B2) still showed an advantage when the adaptation of donor parents was lower than that of the adapted parents, i.e., D0, D1, or D2. The advantage of three rounds of backcrossing (B3) over two rounds (B2) is minimal, especially considering the additional cost of hand pollination used in backcrosses.

Backcrossing reduces DT performance in the final selected lines (Fig. 1b, d, f). Due to random effects associated with phenotyping, the individuals selected for top DT values may not have the highest genotypic values, and therefore some DT genes may be lost during backcrossing. When ten genes needed to be introduced (A0), 87–94% of genes were introduced in the final top lines selected (Table 6; Fig. 1b) under scheme AD. When lower selection intensity was applied on DT (ADA), 78–87% of genes were selected; when higher selection intensity was applied on DT (DAD), 99–100% of genes were selected. When ten genes need to be introduced (A0), each additional generation of backcross leads to the loss of approximate two genes (Table 6; Fig. 1b, d, f). In most cases, backcrossing resulted in the loss of desired DT genes, except for A8 and selection scheme DAD, where fewer genes needed to be introduced and the selection intensity for DT was higher than that for adaptation (Fig. 1b, d, f).

The objective of backcrossing breeding is to transfer a few desired traits or genes from the donor parents to the adapted parents, while at the same time maintaining the adaptation of the recurrent or adapted parents (Allard 1960; Stoskopf et al. 1993). Therefore both adaptation and DT have to be considered when comparing different crossing and selection strategies. While backcrossing increases the adaptation of the final selected lines, it also increases the probability of losing the genes to be introduced (Fig. 1). When the number of introduced genes is high, e.g., more than six (as per A0 and A2), single backcrossing is capable of introducing more than 60% of the genes while

maintaining adaptation at a level similar to that of the adapted parents. Additional backcrossing results in losing 2 DT genes when the number of introduced genes is ten and losing at least one when the number of introduced genes is eight (Fig. 1b, d, f). When fewer genes are to be introduced (e.g., A6 and A8), a second round of backcrossing is useful for further improving adaptation without losing a significant number of DT genes. Additional backcrossing (as per B3) will not significantly improve adaptation and therefore is not recommended in conventional breeding.

Genetic gains in adaptation and DT using SBBS

Positive genetic gains in DT can be observed in all scenarios (Table 7), indicating the efficiency of SBBS for transferring DT from the donor parents to the adapted parents. With selection scheme AD, transgressive segregation on adaptation was observed for donor groups D4–D7; with selection scheme ADA, transgressive segregation on adaptation was observed for donor groups D3–D7; and with selection scheme DAD, transgressive segregation on adaptation was observed for donor groups D6–D7, indicating SBBS can also select the favorable adaptation genes carried by the donor parents. This explains the 5–15% higher yield potential of lines with improved rust resistance selected by SBBS in CIMMYT's wheat breeding programs (Singh and Trethowan 2007).

Averaged across the eight donor parental groups and when the AD selection scheme is applied, SBBS can retain 66, 73, 82, 92, and 97% of the DT genes in adapted parental groups A0, A2, A4, A6, and A8, respectively. When the ADA selection scheme is applied, these numbers become 55, 60, 70, 82, and 94%, and when the DAD selection scheme is applied, these numbers become 83, 91, 97, 99, and 100%. Hence, when donor parents can also contribute to adaptation, i.e., they have some favorable adaptation genes such as in AL, IAL, and SYNII (Table 1), SBBS can retain these genes in the final selected advanced lines, which results in the transgressive advanced lines on adaptation.

Discussion

Backcrossing in plant breeding

Backcrossing is commonly used in plant breeding. In most cases, backcross breeding means that one or a few genes are transferred from a donor parent to an adapted line (Allard 1960; Stoskopf et al. 1993). Assuming DT can be precisely phenotyped, repeated backcrossing can efficiently recover the recurrent parent (Allard 1960). However, single-gene inherited traits are rare in breeding, as is a donor

Table 7 Genetic gains in adaptation and transferred donor traits (DT) using the single backcrossing breeding strategy

Selection scheme	Adapted parents	Trait	Donor parents							
			D0	D1	D2	D3	D4	D5	D6	D7
AD	A0	Adaptation	−4.79	−3.63	−2.33	−1.18	0.27[a]	1.57	3.22	4.97
		DT	6.11	6.22	6.33	6.50	6.69	6.86	7.13	7.31
	A2	Adaptation	−5.43	−4.28	−3.00	−1.87	−0.41	1.01	2.60	4.71
		DT	5.44	5.47	5.64	5.74	5.83	5.99	6.19	6.38
	A4	Adaptation	−4.97	−3.86	−2.80	−1.53	−0.10	1.13	2.86	4.71
		DT	4.67	4.72	4.82	4.86	5.00	5.02	5.16	5.33
	A6	Adaptation	−4.32	−3.35	−2.21	−0.85	0.32	1.55	3.04	4.76
		DT	3.59	3.61	3.66	3.70	3.67	3.72	3.76	3.79
	A8	Adaptation	−2.98	−1.99	−0.90	0.35	1.39	2.62	4.08	5.64
		DT	1.94	1.93	1.94	1.96	1.95	1.96	1.97	1.96
ADA	A0	Adaptation	−1.97	−0.89	0.21	1.11	2.53	3.49	5.08	6.76
		DT	4.91	5.04	5.19	5.34	5.51	5.71	6.02	6.22
	A2	Adaptation	−2.65	−1.74	−0.58	0.34	1.57	2.98	4.35	6.46
		DT	4.37	4.43	4.55	4.67	4.79	4.96	5.23	5.44
	A4	Adaptation	−2.32	−1.41	−0.34	0.79	1.99	3.08	4.76	6.43
		DT	3.85	3.92	3.97	4.13	4.17	4.33	4.47	4.67
	A6	Adaptation	−1.93	−0.98	−0.16	0.93	2.01	3.12	4.71	6.36
		DT	3.10	3.24	3.21	3.20	3.27	3.36	3.41	3.44
	A8	Adaptation	−0.03	0.85	1.88	2.67	3.62	4.62	5.99	7.48
		DT	1.86	1.86	1.86	1.86	1.88	1.88	1.90	1.90
DAD	A0	Adaptation	−9.82	−8.23	−6.69	−4.60	−3.04	−1.24	0.85	2.63
		DT	7.91	8.02	8.11	8.17	8.28	8.33	8.48	8.70
	A2	Adaptation	−9.93	−8.42	−6.82	−5.24	−3.35	−1.65	0.32	2.70
		DT	7.13	7.15	7.14	7.24	7.31	7.32	7.39	7.46
	A4	Adaptation	−9.13	−7.62	−5.97	−4.41	−2.74	−1.23	0.88	2.96
		DT	5.78	5.77	5.81	5.81	5.84	5.88	5.89	5.91
	A6	Adaptation	−7.80	−6.58	−5.16	−3.80	−2.16	−0.65	1.26	3.11
		DT	3.97	3.96	3.98	3.98	3.97	3.97	3.98	3.99
	A8	Adaptation	−6.60	−5.29	−4.04	−2.66	−1.17	0.27	2.05	3.83
		DT	2.00	2.00	2.00	1.99	2.00	2.00	2.00	2.00

[a] Italicized if positive, i.e., the mean adaptation of the top ten final selected lines is higher than that of the 40 adapted parents

parent with only one or two useful genes. When the number of genes to be transferred from the donor parents to adapted parents is relatively high (e.g., 5–10), it is less likely that all of the donor parents' favorable genes can be transferred to the adapted parents without reducing the adaptation of the recurrent parents in one breeding cycle. Taking durable resistance as an example, the combination of one major gene and 4–5 minor genes will produce satisfactory long-lasting resistance. In this case, the breeding objective is not to transfer all genes, but most of them. Through simulation, we found that the SBBS is capable of transferring more than 60% of favorable genes and, at the same time, improving the adaptation of the adapted parents.

In this study, rounds of backcrossing refer to the number of times backcrossing was done in the current breeding cycle, and backcrossing to derive the donor parents in previous breeding cycles was not counted. In CIMMYT, SYNI was normally derived from 3 to 4 rounds of repeated backcrossing, and SYNII from 2 to 3 rounds of repeated backcrossing. If a single backcross using a SYNII line as a donor is made, the backcrossing procedure from the primary donor, i.e., SYN0 or SYNI, is not considered.

Successful application of SBBS for improving durable rust resistance in CIMMYT

The genetic basis of slow rusting resistance in CIMMYT wheat lines started to become clear in the early 1990s (Singh et al. 1998a, 2000, 2005). High-yielding lines that combine four or five additive, minor genes for resistance to

both leaf and stripe rusts, and show near-immune levels of resistance were developed in the 1990s (Singh et al. 2000, 2005). Highly resistant lines are now being used systematically to transfer minor rust resistance genes to farmers' preferred cultivars that are well-adapted and grown across large areas but have become susceptible to rust races in Mexico.

Bulking of selected plants, used in SBBS, imposes no restriction on the number of plants that can be selected in each generation because harvesting and threshing are quick and inexpensive, and the next generation is derived from a sample of the bulked seed. Because high rust resistance levels require the presence of four to five additive genes, the level of homozygosity from the F_4 generation onwards is usually sufficient to identify plants that combine adequate resistance with good agronomic traits. Moreover, selecting plants with low terminal disease severity under high disease pressure means that more additive genes may be present in those plants. Selection for seed characteristics is carried out on seeds obtained from individually harvested F_5 plants. Small plots of the F_6 lines are then evaluated for agronomic traits, homozygosity of resistance and other traits, before conducting yield trials (Singh and Huerta-Espino 2004).

Resistant derivatives of several cultivars and genotypes were recently bred using the above methodology. In each case derived lines were identified that not only carry high levels of resistance to leaf rust or yellow rust, or both, but also show 5–15% higher yield potential than the original cultivar. This was confirmed by the simulation results shown in Fig. 1, i.e., the top ten lines have higher adaptation than the original adapted parents in many scenarios. Therefore, we believe this approach to wheat improvement makes it possible to maintain the characteristics of the original cultivar while improving its yield potential and rust resistance.

Wider application of SBBS

Single backcrossing breeding strategy has proved to be efficient in CIMMYT's wheat breeding efforts aimed at transferring durable rust resistance to adapted lines. The extensive simulations conducted in this study confirmed SBBS's advantages over single crosses and repeated backcrossing. This strategy has allowed the simultaneous transfer not only of host plant resistance genes, but of other minor genes with small effects that increase yield potential or improve the grain quality of an adapted cultivar. As a result, the transgressive segregations on adaptation to the recurrent parents combined with durable resistance have been selected at CIMMYT in the past 20 years. Other simulations were conducted using a genetic model with dominance. Of the 200 adaptation genes used in this study,

74 were assumed to be purely additive (i.e., $d = 0$), 42 partially dominant (i.e., $d = 0.5a$), 42 completely dominant (i.e., $d = a$), and 42 over dominant (i.e., $d = 2a$). The results for selection scheme AD are shown in Fig. 2. Obviously, Fig. 2a and b are similar to Fig. 1a and b, respectively.

Though we used wheat in our simulation study, this breeding methodology can be used in other crops, such as maize. To give an example, breeding maize inbred lines combining the best nutritional and agronomic traits is a major objective for the HarvestPlus breeding program (Pfeiffer and McClafferty 2007). Since strong selection for Fe and Zn concentrations, and pro-vitamin A content was not applied in conventional maize inbred breeding, modern maize inbred lines normally have low levels of these traits. Donor parents with desired micronutrient content have been identified, but they are usually non-adapted. We used a GE system similar to maize (i.e., 100 adaptation and ten DT genes evenly distributed on ten chromosomes) to complete simulation experiments, and achieved similar results (Fig. 3). This suggests that SBBS could also be used in HarvestPlus maize breeding programs.

Multiple alleles, epistasis, and genotype by environment interactions can be defined in QU-GENE and then simulated in QuLine. The topic of this study is more relevant to biparental crosses, where multiple alleles can be ignored. More specifically, breeders may want to know whether one or two rounds of backcrossing are needed for a specific adapted parent and a specific donor. For this reason, the bi-allelic model used in this study may be more reasonable and suitable. When genes are interacting with each other, the favorable allele at each locus cannot be determined by the individual additive effects. Therefore, it is not easy to define the five recurrent parental groups (i.e., A0, A2, A4, A6 and A8) and the eight donor parental groups (i.e., D0–D7) from allele frequencies, if epistasis is present. Considering that additive variation may be more important, especially in self-pollinating crops, we doubt that the inclusion of a certain amount of epistatic variation (e.g., 30%) could significantly change the results and conclusions observed in this study.

In the last two decades, molecular markers have been widely used for studying the genetics of complex, quantitative traits in different crop species. Reviews have indicated that QTL mapping studies have typically detected 3–5 QTL for each trait (Kearsey and Farquhar 1998; Bernardo 2002). QTL having smaller effects are normally associated with non-significant test statistics and therefore not reported. The actual number of genes governing most breeding traits should therefore be greater than the number detected by QTL mapping. It seems unlikely that in the near future all genes for important breeding traits will be mapped, and that relevant closely linked molecular markers that breeders can use

Fig. 2 Genetic means for adaptation and transferred donor traits (DT) in the final selected ten top lines in a genetic model including pure additive, partial dominance, complete dominance, and overdominance for the 200 adaptation genes. **a** Genetic mean for adaptation with selection scheme AD (in each generation adaptation is selected first, followed by selection for DT). **b** Genetic mean for DT with selection scheme AD
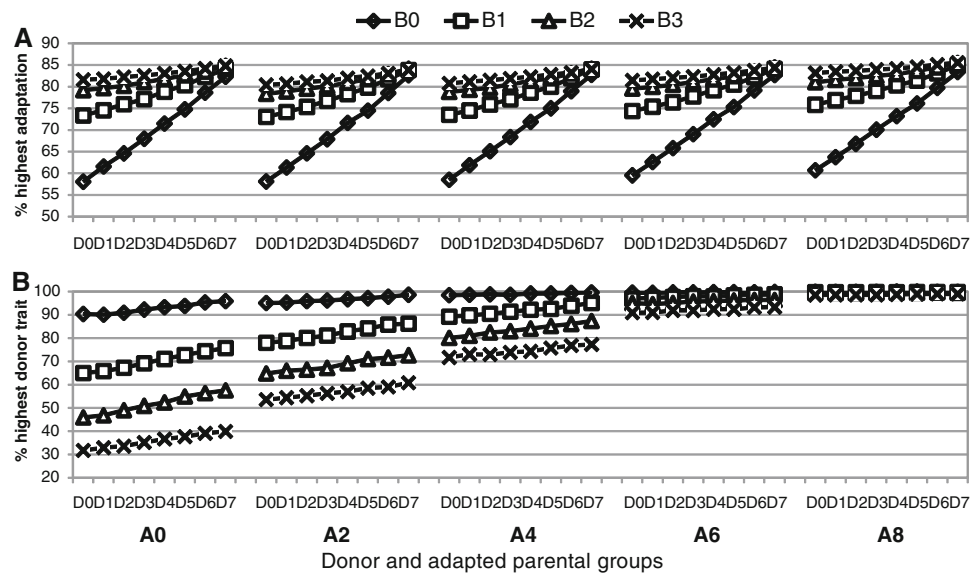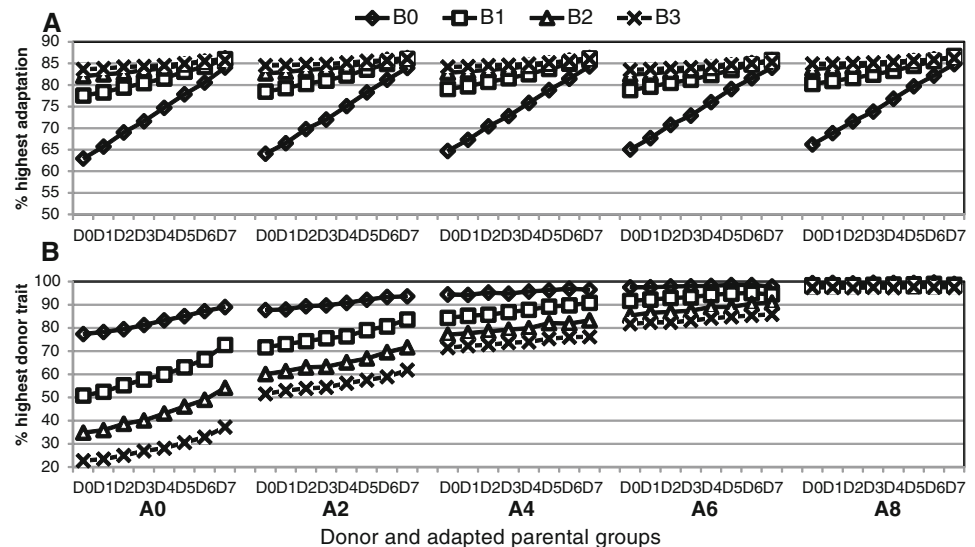

Fig. 3 Genetic means for adaptation and transferred donor traits (DT) in the final selected ten top lines using the maize genome. **a** Genetic mean for adaptation with selection scheme AD (in each generation adaptation is selected first, followed by selection for DT). **b** Genetic mean for DT with selection scheme AD

will be developed. Thus for most traits, breeders will have to continue to rely on phenotypic selection. On the other hand, after 100 years of modern breeding, most donor parents have also been improved. Parents with intermediate adaptation but possessing particular elite traits that are lacking in elite adapted parents are common. Therefore, in the future, SBBS may be more and more widely applied in plant breeding.

References

Allard RW (1960) Principles of plant breeding. Wiley, New York

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

Bregitzer P, Dahleen LS, Neate S, Schwarz P, Manoharan M (2008) A single backcross effectively eliminates agronomic and quality alterations caused by somaclonal variation in transgenic barley. Crop Sci 48:471–479

Frisch M, Melchinger AE (2001) Marker-assisted backcrossing for simultaneous introgression of two genes. Crop Sci 41:1716–1725

Kearsey MJ, Farquhar AGL (1998) QTL analysis; where are we now? Heredity 80:137–142

Mujeeb-Kazi A, Hettel GP (1995) Utilizing wild grass biodiversity in wheat improvement: 15 years of wide cross research at CIMMYT. CIMMYT Research Report 2. CIMMYT, D.F., Mexico

Ortiz R, Trethowan R, Ortiz Ferrara G, Iwanaga M, Dodds JH, Crouch JH, Crossa J, Braun H-J (2007) High yield potential, shuttle breeding and new international wheat improvement strategy. Euphytica 157:365–384

Pfeiffer WH, McClafferty B (2007) HarvestPlus: breeding crops for better nutrition. Crop Sci 47(S3):S88–S105

Podlich DW, Cooper M (1998) QU-GENE: a platform for quantitative analysis of genetic models. Bioinformatics 14:632–653

Rajaram S (1999) Historical aspects and future challenges of an international wheat program. In: van Ginkel M, McNab A, Krupinsky J (eds) *Septoria* and *Stagonospora* diseases of cereals: a compilation of global research. CIMMYT, D.F, Mexico, pp 1–17

Rajaram S, van Ginkel M, Fischer RA (1994) CIMMYT's wheat breeding mega-environments (ME). In: Proceedings of the 8th international wheat genetics symposium. China Agricultural Scientech. Beijing, pp 1101–1106

Reddy BVS, Comstock RE (1976) Simulation of the backcross breeding method. I. Effects of heritability and gene number on fixation of desired alleles. Crop Sci 16:825–830

Ribaut J-M, Jiang C, Hoisington D (2002) Simulation experiments on efficiencies of gene introgression by backcrossing. Crop Sci 42:557–565

Singh RP, Huerta-Espino J (2004) The use of "single-backcross selected-bulk" breeding approach for transferring minor genes based rust resistance into adapted cultivars. In: Black CK, Panozzo JF, Rebetzke GJ (eds) Proceedings of 54th Australian cereal chemistry conference and 11th wheat breeders assembly, 21–24 September, 2004, Canberra. Australia. Cereal Chemistry Division, Royal Australian Chemical Institute, North Melbourne, pp 48–51

Singh RP, Trethowan R (2007) Breeding spring bread wheat for irrigated and rainfed production systems of the developing world. In: Kang MS, Priyadarshan PM (eds) Breeding major food staples. Blackwell Publishing, Ames, pp 109–140

Singh RP, Mujeeb-Kazi A, Huerta-Espino J (1998a) *Lr46*: a gene conferring slow-rusting resistance to leaf rust in wheat. Phytopathology 88:890–894

Singh RP, Rajaram S, Miranda A, Huerta-Espino J, Autrique E (1998b) Comparison of two crossing and four selection schemes for yield, yield traits, and slow rusting resistance to leaf rust in wheat. Euphytica 100:35–43

Singh RP, Huerta-Espino J, Rajaram S (2000) Achieving near-immunity to leaf and stripe rusts in wheat by combining slow rusting resistance genes. Acta Phytopathologica Hungarica 35:133–139

Singh RP, Huerta-Espino J, William HM (2005) Genetics and breeding for durable resistance to leaf and stripe rusts in wheat. Turk J Agric For 29:121–127

Stoskopf NC, Tomes DT, Christie BR (1993) Plant breeding: theory and practice. Westview Press, Boulder

van Ginkel M, Trethowan R, Ammar K, Wang J, Lillemo M (2002) Guide to bread wheat breeding at CIMMYT. Wheat Special Report 5. CIMMYT, D.F, Mexico

Wang J, van Ginkel M, Podlich D, Ye G, Trethowan R, Pfeiffer W, DeLacy IH, Cooper M, Rajaram S (2003) Comparison of two breeding strategies by computer simulation. Crop Sci 43:1764–1773

Wang J, van Ginkel M, Trethowan R, Ye G, DeLacy I, Podlich D, Cooper M (2004) Simulating the effects of dominance and epistasis on selection response in the CIMMYT Wheat Breeding Program using QuCim. Crop Sci 44:2006–2018

Wang J, Eagles HA, Trethowan R, van Ginkel M (2005) Using computer simulation of the selection process and known gene information to assist in parental selection in wheat quality breeding. Aust J Agric Res 56:465–473

Wang J, Chapman SC, Bonnett DB, Rebetzke GJ, Crouch J (2007a) Application of population genetic theory and simulation models to efficiently pyramid multiple genes via marker-assisted selection. Crop Sci 47:580–588

Wang J, Wan X, Li H, Pfeiffer W, Crouch J, Wan J (2007b) Application of the identified QTL-marker associations in rice quality improvement through a design breeding approach. Theor Appl Genet 115:87–100