Theor Appl Genet (2007) 115:289–297
DOI 10.1007/s00122-007-0544-y

ORIGINAL PAPER

Linkage disequilibrium in two European F_2 flint maize populations under modified recurrent full-sib selection

K. C. Falke · H. P. Maurer · A. E. Melchinger · H.-P. Piepho · C. Flachenecker · M. Frisch

Received: 18 December 2006 / Accepted: 27 March 2007 / Published online: 28 April 2007

Abstract According to quantitative genetic theory, linkage disequilibrium (LD) can hamper the short- and long-term selection response in recurrent selection (RS) programs. We analyzed LD in two European flint maize populations, KW1265 × D146 (A × B) and D145 × KW1292 (C × D), under modified recurrent full-sib selection. Our objectives were to investigate (1) the decay of initial parental LD present in F_2 populations by three generations of intermating, (2) the generation of new LD in four (A × B) and seven (C × D) selection cycles, and (3) the relationship between LD changes and estimates of the additive genetic variance. We analyzed the F_2 and the intermated populations as well as all selection cycles with 104 (A × B) and 101 (C × D) simple sequence repeat (SSR) markers with a uniform coverage of the entire maize genome. The LD coefficient D and the composite LD measure Δ were estimated and significance tests for LD were performed. LD was reduced by intermating as expected from theory. A directional generation of negative LD between favorable alleles could not be observed during the selection cycles. However, considerable undirectional changes in D were observed, which we attributed to genetic sampling due to the finite population size used for recombination. Consequently, a long-term reduction of the additive genetic variance due to negative LD was not observed. Our experimental results support the hypothesis that in practical RS programs with maize, LD generated by selection is not a limiting factor for obtaining a high selection response.

Communicated by M. Xu.

K.C. Falke and H.P. Maurer are contributed equally to the manuscript.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-007-0544-y) contains supplementary material, which is available to authorized users.

K. C. Falke · H. P. Maurer · A. E. Melchinger (✉) · C. Flachenecker · M. Frisch
Institute of Plant Breeding, Seed Science, and Population Genetics, University of Hohenheim, 70593 Stuttgart, Germany
e-mail: melchinger@uni-hohenheim.de

H.-P.Piepho
Institute for Crop Production and Grassland Research, University of Hohenheim, 70593 Stuttgart, Germany

Introduction

Recurrent selection (RS) is a cyclical strategy designed to ensure long-term selection response by increasing the frequency of favorable alleles while maintaining the genetic variance in populations (Hallauer 1985). Selection generates linkage disequilibrium (LD) between alleles whose frequencies were increased by selection (Nei 1963). This newly generated LD is negative, which means by definition that the covariance between favorable alleles at different loci is negative. Negative LD reduces the additive genetic variance (σ_A^2) of the traits under selection (Bulmer 1971), and thus may result in a decline of long-term selection response.

In RS procedures in maize, open-pollinated-varieties or synthetics have mostly been employed as base populations (Hallauer and Miranda 1988; Bernardo 2002). In contrast, F_2 base populations have been used only in a few studies but mostly with remarkable success (Russell et al. 1973; Genter 1982; Moll 1991; Landi and Frascaroli 1993). In F_2 populations, there is a positive covariance between alleles originating from the same parental line at linked loci. To reduce the extent of this parental LD and its negative

effects on the selection gain, Johnson (1982) suggested three generations of random intermating before starting with RS. Although parental LD in base populations of RS programs may severely limit the selection response, to our knowledge, no previous results have been published on the reduction in parental LD through intermating.

Allele frequency changes due to selection were analyzed in several experimental studies (Labate et al. 1999; Pinto et al. 2003; Coque and Gallais 2006; Falke et al. 2007). However, the extent of LD between alleles whose frequencies were increased by selection, and the effect of this LD on σ_A^2 and the selection response were only investigated theoretically and with computer simulations (Hospital and Chevalet 1996). No experimental results have been published yet.

As a part of a maize breeding project, long-term recurrent full-sib selection programs with two European F_2 flint maize populations, previously employed in QTL studies (Schön et al. 1994; Mihaljevic et al. 2004, 2005a), were initiated for analyzing the selection response. After completing three generations of intermating, four cycles of RS were conducted in cross KW1265 × D146 (A × B) and seven cycles in cross D145 × KW1292 (C × D) to determine changes in the population structure at the phenotypic and molecular level. In several companion studies, we investigated the selection response at the phenotypic level (Flachenecker et al. 2006a, b, c) and determined allele frequency changes during RS at the molecular level (Falke et al. 2007).

The main goal of the present study was to analyze the LD in the RS programs of these two F_2 flint maize populations. In particular, our objectives were to (1) investigate the genetic distance among the base and intermated populations as well as the subsequent selection cycles, (2) examine the decay of parental LD between linked loci after three generations of intermating, (3) determine the generation of new LD between alleles whose frequencies were increased by selection during the RS procedure, and (4) relate changes in LD to estimates of σ_A^2.

Materials and methods

Plant materials

Four early maturing homozygous European flint lines, KW1265, D146, D145, and KW1292, subsequently referred to as A, B, C, and D, respectively, were used as parents to produce 380 $F_{2:3}$ lines of cross A × B and 140 $F_{3:4}$ lines of cross C × D. Lines A and D are proprietary elite inbreds developed by KWS SAAT AG (Einbeck, Germany), lines B and C are public elite inbreds developed by the University of Hohenheim (Stuttgart, Germany). The F_2 populations were intermated for three generations by applying a chain-crossing scheme with 240 F_2 plants to produce the F_2Syn3 populations. The selection procedure in each selection cycle was described in detail by Flachenecker et al. (2006a, b). Briefly, four (A × B) and seven (C × D) cycles of modified recurrent full-sib selection were completed for recombination of superior genotypes, using a pseudo-factorial mating scheme based on the suggestion of Cockerham and Burrows (1980). Full-sib families were selected on the basis of a selection index. For calculating the selection index, (1) grain yield and dry matter content were expressed in percent of the mean of F_2 check entries and (2) relative trait values received a weight of 1 for grain yield and 2 for dry matter content [i.e., the weight vector was $\mathbf{b}' = (1,2)$]. Evaluation of the full-sib families in each selection cycle was conducted in field trials at three locations in South Germany. The experimental design in each environment was an α-lattice (10 × 15) with three replications.

SSR analyses

A total of 104 (A × B) and 101 (C × D) codominant SSR markers polymorphic between the parental lines and warranting a uniform coverage over the entire maize genome was employed for genotyping. For SSR marker analyses, we used random subsets of 146 $F_{2:3}$ lines out of the 380 $F_{2:3}$ lines in A × B and 110 $F_{3:4}$ lines out of the 140 $F_{3:4}$ lines in C × D as well as 148 F_2Syn3 plants and the parents of the 36 families with the highest selection index in each selection cycle of both RS programs. DNA extraction as well as SSR amplification and detection were described in detail by Falke et al. (2006, 2007).

Principal coordinate analysis

Modified Rogers' distances (MRD, Wright 1978, p. 91) were estimated between parental lines (A, B and C, D), the population of $F_{2:3}$ (A × B) or $F_{3:4}$ (C × D) lines, the intermated populations F_2Syn3, and the various selection cycles of both RS programs, A × B and C × D, using 104 (A × B) and 101 (C × D) SSR marker loci. Based on MRD estimates, principal coordinate analyses (PCoA) (Gower 1966) were carried out to reveal associations among the base and intermated populations and the various selection cycles of the RS programs.

Assessment of parental LD in the intermated populations

Parental LD was assessed for all marker pairs in population of $F_{2:3}$ lines (A × B) and $F_{3:4}$ lines (C × D) and their corresponding F_2Syn3 populations with the linkage disequilibrium coefficient D (Weir 1996, p. 113)

$$D_{xy} = p_{xy} - p_x p_y. \quad (1)$$

Here, p_x and p_y are the allele frequencies of the alleles x and y at two loci originating from the same parental line, and p_{xy} is the frequency of gametes carrying both alleles x and y. The gametic frequencies p_{xy} were estimated with a maximum likelihood approach (Weir 1996, pp. 73–76), which assumes Hardy–Weinberg equilibrium (HWE) at both loci. The LD coefficient D was plotted as a function of recombination frequencies, which were calculated with the inverse of Haldane's (1919) mapping function from the map distances estimated in $F_{2:3}$ (A × B) and $F_{3:4}$ (C × D) by Falke et al. (2007).

No theoretical results exist on the expected decay of parental LD with increasing recombination frequencies for intermated populations produced by the chain-crossing method. Therefore, we simulated the intermating procedure (assuming no interference in crossover formation) with 500 replications to determine the expected LD decay. Observed values were compared with these expectations.

All loci pairs in the base and intermated populations were tested for significant parental LD with a Monte Carlo approximation of Fisher's exact test (Zaykin et al. 1995) at a significance level of $\alpha = 0.05$. For the underlying Monte Carlo method, 17,000 replications were used (Guo and Thompson 1992).

Assessment of LD in selection cycles

In the intermated populations and in the selection cycles of both RS programs, we investigated the LD between marker alleles tightly linked to QTL alleles with a positive effect on the selection index. Subsequently, we refer to these alleles as ''favorable alleles''. To determine the set of favorable alleles, we employed the following strategy. We chose all marker loci whose significant allele frequency changes were attributed to selection, as detected with Waples' test (1989) in a companion study (Falke et al. 2007). For each marker, the allele with the largest positive allele frequency change Δp from F_2Syn3 to the final selection cycle C4 (A × B) and C7 (C × D) was assigned to the set of favorable alleles. This strategy was necessary because non-parental alleles were observed since the initial selection cycles in both populations due to a contamination with foreign pollen (A × B: C1; C × D: C2) (Falke et al. 2007). The sets of favorable alleles consisted of 27 (A × B) and 14 (C × D) marker alleles

LD between the favorable alleles was assessed by (1) the linkage disequilibrium coefficient D (Eq. 1) and (Eq. 2) the composite linkage disequilibrium measure Δ (Weir 1996). The LD coefficient D was chosen because it is in direct relationship with σ_A^2 (Lynch and Walsh 1998, p. 102). The measure Δ was chosen, because it does not require the ML estimation of gamete frequencies, and therefore provides a means to assess the robustness of D with respect to deviations from HWE. Significance of D and Δ was tested with χ^2 tests. All analyses were carried out for populations F_2Syn3 and the selection cycles of both RS programs.

For the analysis of the relation between LD and σ_A^2 we employed the LD coefficient D and restricted maximum likelihood (REML) estimates of the variance components determined by Flachenecker et al. (2006a, b).

All computations and simulations were performed with software PLABSOFT (Maurer et al. 2004), which is implemented as an extension of the statistical software R (R Development Core Team 2004).

Results

PCoA based on MRD estimates between population $F_{2:3}$ (A × B) and $F_{3:4}$ (C × D), F_2Syn3 and all selection cycles explained 98.3% (A × B) and 96.0% (C × D) of the molecular variance by the first two principal coordinates (PCs) (Fig. 1). In both RS programs, PC1 revealed a clear separation of the parental lines (A and B, C and D). All populations and selection cycles were in between the two parental lines with respect to PC1. In comparison with PC1, PC2 explained considerably less molecular variance (A × B: 6.6%; C × D: 10.4%). PC2 separated population $F_{2:3}$ (A × B) and $F_{3:4}$ (C × D), their corresponding F_2Syn3 and the selection cycles (A × B: C1 – C4; C × D:

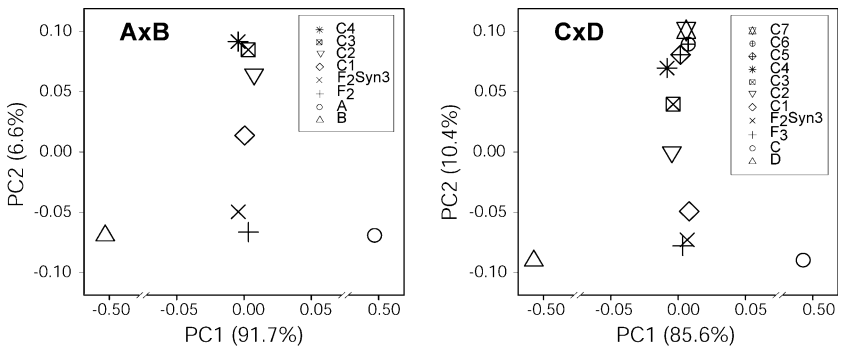

Fig. 1 Principal coordinate analysis of all populations and selection cycles from RS programs of A × B and C × D based on modified Rogers' distance calculated from SSR marker loci. PC1 and PC2 refer to the first and second principal coordinate, respectively. *Numbers in parentheses* indicate the proportion of molecular variance explained by the principal coordinates

C1 – C7) according to their chronological order. Final selection cycles (C4 and C7) were clearly separated from population $F_{2:3}$ (A × B) and $F_{3:4}$ (C × D). MRD estimates for all populations and selection cycles were presented in a supplementary table for both RS programs. Performing the PCoA between populations and selection cycles without parental lines, 85.1% (A × B) and 81.7% (C × D) of the molecular variance was explained by the first two PCs (data not shown). The proportion of the molecular variance, which was assessed by PC2 in the analysis with parents, was also captured by PC1 in the analysis without parents. The proportion of molecular variance from the populations and selection cycles was higher without parental lines (A × B: 73.6%; C × D: 70.1%) than with them (A × B: 6.6%; C × D: 10.4%). PC2 was approximately a quadratic function of PC1, so the ranking of the populations and selection cycles was unaltered compared to the analysis with parents.

The proportion of linked loci pairs in significant parental LD detected with Fisher's exact test decreased from 0.522 ($F_{2:3}$) to 0.357 (F_2Syn3) in A × B and from 0.488 ($F_{3:4}$) to 0.309 (F_2Syn3) in C × D. The extent of parental LD in population F_2Syn3 decreased with increasing recombination frequency, in accordance with the expected values obtained from simulations (Fig. 2).

Favorable alleles were mostly in linkage equilibrium in population F_2Syn3 of both RS programs (Fig. 3). In cycle C1, positive LD (indicated by red shading) was generated in both RS programs. Furthermore, the extent of negative LD (indicated by blue shading) was increased in C1 of C × D. In the final selection cycles (A × B: C4; C × D: C7), the extent of negative LD was increased in both RS programs, while the extent of positive LD was decreased in A × B and increased in C × D. Major deviations between the two LD measures D and Δ were neither observed in population F_2Syn3 nor in the first or final selection cycles. For several loci pairs, LD could not be determined (missing values, indicated by gray shading), because (1) non-parental alleles were absent in population F_2Syn3 but only present in later selection cycles or (2) the gamete frequency could not be determined with the ML procedure because of missing genotype classes.

In the RS program of A × B, the increase in the number of loci pairs with significantly positive LD from F_2Syn3 to cycle C1 detected by the χ^2 test was associated with the highest estimates of σ_A^2 for the selection index (Table 1, selection response and estimates of variance components were taken from Flachenecker et al. 2006a, b). In subsequent selection cycles, an increase in the number of loci pairs in negative LD was associated with a significant decrease in σ_A^2 for selection index. In C × D, the non-significant decrease in σ_A^2 for selection index was associated with a non-directional increase and decrease in loci pairs in positive and negative LD during the selection procedure.

Discussion

In previous studies, the selection response of modified full-sib recurrent selection programs in populations A × B and C × D was evaluated at the phenotypic level with classical quantitative genetic methods (Flachenecker et al. 2006a, b, c). At the molecular level, we investigated the effects of random genetic drift and selection on allele frequency changes in QTL regions by using SSR markers (Falke et al. 2007). We observed a comparatively high realized selection gain for selection index (A × B: 5.25%; C × D: 3.64% per selection cycle; Flachenecker et al. 2006a, b). Several QTL regions for selection index were detected in population $F_{2:3}$ (A × B) and $F_{3:4}$ (C × D) (Falke et al. 2007). At some of them, flanking markers showed significant changes in allele frequencies due to selection during the RS procedure (Falke et al. 2007). In the present study, we analyzed the development of the LD over several cycles of RS as well as the effects of changes in allele frequencies and LD on trends of σ_A^2.

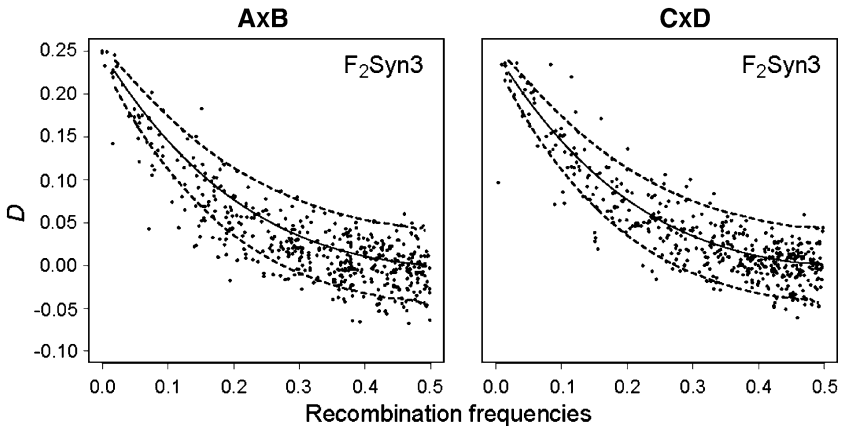

Fig. 2 Decay of parental linkage disequilibrium (LD) after three generations of intermating in populations A × B and C × D. Observed parental LD (*circles*) between linked pairs of SSR loci is plotted as a function of recombination frequencies. The mean and the respective 5% and 95% quantiles of the simulated LD decay are plotted as *solid and dashed lines*, respectively

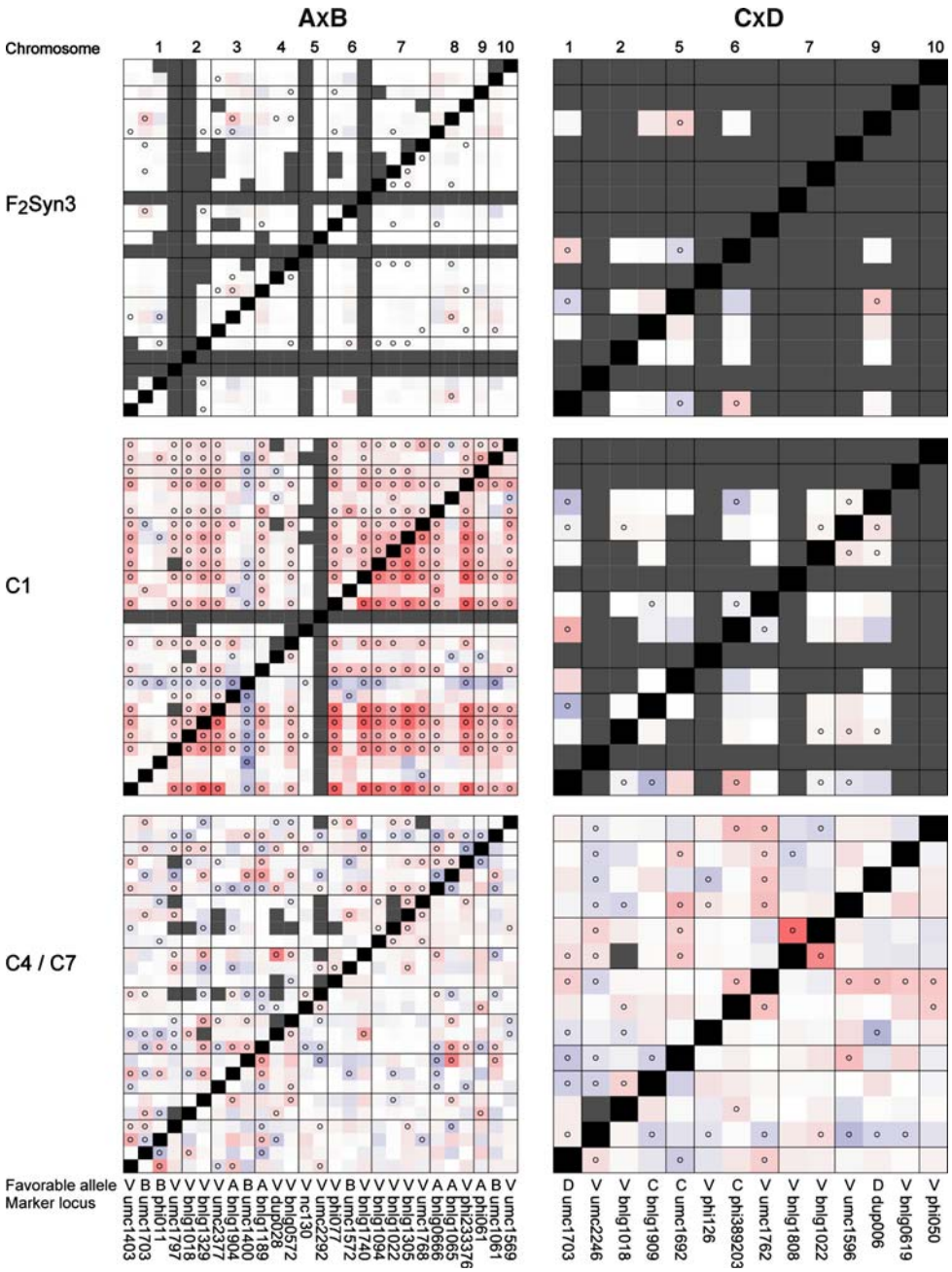

Fig. 3 Linkage disequilibrium (*LD*) between marker alleles whose frequencies were increased by selection measured as *D* (above diagonal) and Δ (below diagonal) in F_2Syn3 and in cycle C1 of both RS programs, in cycle C4 of A × B, and in cycle C7 of C × D. The respective favorable alleles are presented below the matrix (parental alleles A and B, C and D, and non-parental alleles V). Chromosomes are separated by *horizontal and vertical black lines*, *red coloring* indicates positive LD, *blue* negative LD, and *grey* missing values. Circles indicate significant LD detected with a χ^2 test

Genetic diversity

For both RS programs, PC1 separated the parental lines with a MRD of 1.0, as expected from theory due to the employment of exclusively polymorphic SSR markers (Fig. 1). The population of $F_{2:3}$ lines (A × B) and $F_{3:4}$ lines (C × D) together with the intermated populations F_2Syn3 were in the center between respective parental lines. During the selection procedure, the selection cycles varied only slightly around this origin of PC1 and no directional shift in favor of one parental line was observed. This suggests that the parental lines of each selection program carried approximately similar numbers of favorable alleles, and the observed selection response was driven by recombination of alleles of different parental origin and selection of favorable allele combinations. This is in agreement with the small estimates of the sum of additive effects estimated in a generation mean analysis as well as the results from QTL mapping studies (Mihaljevic et al. 2005a, b).

PC2 separated the $F_{2:3}$ or $F_{3:4}$ from the intermated populations and the selection cycles according to their chronological order, with larger differences between sub-

Table 1 Selection response for selection index (±SE) and its components relative to the mean performance of six F_2 checks. Restricted maximum likelihood (REML) estimates of the additive genetic variance (σ_A^2 ± SE), their mean across selection cycles and the coefficient (*b*) of the linear regression across selection cycles of RS programs A × B and C × D. Proportion of loci pairs with favorable alleles in significant LD (P 0.05; χ^2 test statistic based on D) in population F_2Syn3 and the various selection cycles of both RS programs.

Selection cycle	Selection index (%)		Grain yield (% of F_2)		Grain moisture (% of F_2)		LD between favorable alleles		
	Selection response ± SE	σ_A^2 ± SE	Selection response ± SE	σ_A^2 ± SE	Selection response ± SE	σ_A^2 ± SE	No. of loci pairs	Proportion of loci pairs in significant LD	
								Negative	Positive
A × B									
F_2Syn3							223	0.067	0.040
C1	315.3 ± 2.3	207.5 ± 82.5*	120.2 ± 2.4	0.38 ± 0.16*	102.4 ± 0.6	92.5 ± 13.6**	314	0.067	0.475
C2	323.2 ± 4.7	198.9 ± 31.9**	129.1 ± 4.3	0.25 ± 0.05**	103.0 ± 1.0	111.8 ± 16.7**	270	0.085	0.437
C3	321.5 ± 4.9	174.5 ± 64.3**	124.2 ± 2.6	0.27 ± 0.13*	101.4 ± 1.7	167.2 ± 24.0**	303	0.172	0.155
C4	317.6 ± 4.9	140.3 ± 21.9**	123.1 ± 2.2	0.12 ± 0.07	102.8 ± 1.7	183.9 ± 24.4**	324	0.160	0.210
Mean		180.3 ± 49.9		0.26 ± 0.1		138.8 ± 197			
B	–0.31	–22.6*	0.10	–0.08	0.00	32.9*			
C × D									
F_2Syn3							12	0.167	0.167
C1	298.0 ± 1.8	49.4 ± 23.9	97.9 ± 1.9	0.21 ± 0.11	100.0 ± 0.2	32.8 ± 15.4*	30	0.200	0.133
C2	304.0 ± 5.1	274.1 ± 96.6**	105.4 ± 4.3	0.54 ± 0.22*	100.7 ± 0.7	162.1 ± 23.5**	52	0.173	0.250
C3	323.3 ± 5.3	226.2 ± 32.9**	119.6 ± 5.0	0.69 ± 0.10**	98.1 ± 0.7	188.7 ± 24.7**	57	0.228	0.474
C4	339.6 ± 4.9	25.3 ± 54.0	135.4 ± 4.5	0.00	97.9 ± 0.6	104.7 ± 15.7**	80	0.212	0.288
C5	339.6 ± 5.2	62.6 ± 45.7	132.4 ± 5.7	0.49 ± 0.25*	96.5 ± 1.0	226.7 ± 30.9**	84	0.167	0.143
C6	343.3 ± 5.8	70.0 ± 45.6	137.0 ± 6.0	0.50 ± 0.28	96.8 ± 1.1	229.4 ± 31.9**	81	0.173	0.148
C7	369.4 ± 7.1	123.4 ± 111.1	157.5 ± 8.5	0.17 ± 0.11	94.1 ± 1.7	475.6 ± 197.9*	89	0.180	0.213
Mean		118.7 ± 58.5		0.37 ± 0.15		202.9 ± 48.6			
B	11.2**	–3.6	9.1**	–0.02	–1.1**	37.2			

Selection response and variance components for selection index and its components were determined by Flachenecker (2006a, b)

*, ** Significant at the 0.05 and 0.01 probability level, respectively

sequent cycles in the earlier selection cycles than in the later ones. However, these differences do not match the differences in selection response for the selection index attained in the various selection cycles (Table 1). For example, in RS program C × D the largest selection response was realized in cycle C7, but the MRD between cycles C6 and C7 was very small (Fig. 1). In contrast, the differences between the selection cycles with respect to PC2 agreed well with migration effects due to contamination with foreign pollen observed since the initial selection cycles (Falke et al. 2007). Hence, the differences in PC2 are most probably caused by migration and, in consequence, migration had a considerable influence on the genetic structure of the population in both RS programs. Therefore, migration and subsequent selection of the migrated alleles might very well have been an important factor contributing to the relatively large observed selection response and the small reduction in σ_A^2 in the RS program of C × D.

Decay of parental LD due to intermating

Parental LD in the base population of an RS program is expected to hamper the possible short- and long-term selection gain. Therefore, Johnson (1982) suggested three generations of random intermating before starting an RS program to reduce the parental LD and its negative effects on the selection gain. We adopted this idea and conducted three generations of intermating with our F_2 populations. The observed decay of LD was in good agreement with the theoretical expectations obtained from simulations (Fig. 2). However, whether the observed reduction of linked loci pairs in significant LD through intermating was an important cause of the large realized selection response cannot be definitely answered by our experimental setup. In particular, the breeding success depends on linkage and the linkage phase relationship between favorable alleles. Consequently, an interesting open question for further research is whether the initial time lag in the start of an RS

program due to intermating the base population pays off in terms of a greater realized selection response.

Another goal of intermating generations is to obtain an increased mapping accuracy of tightly linked loci. Intermated F_2 and intermated recombinant inbred line populations have been employed as mapping populations and the authors reported that the maps showed a map expansion (Liu et al. 1996; Lee et al. 2002; Winkler et al. 2003; Falque et al. 2005; Teuscher et al. 2005; Teuscher and Broman 2007). However, this is misleading (Falke et al. 2006; Martin and Hospital 2006), because since the introduction of map distances these were estimated with recombination frequencies, referring to only a single meiosis but not to recombination events accumulated in all intermating generations (Haldane 1919; Kosambi 1944; Stam 1993).

Increase in positive LD during the selection cycles

In both selection programs and at many loci pairs, we observed an increase in positive LD (i.e., a positive covariance) between favorable alleles (Fig. 3) during the selection cycles. This increase can be explained by the employed mating scheme of Cockerham and Burrows (1980) and the contamination with foreign pollen (Falke et al. 2007).

In the mating scheme of Cockerham and Burrows (1980), the best third of the selected plants (used as male parents in the pseudo factorial mating design) transmits their gametes with twice the dose as the remaining two thirds (used as female parents). This can result in positive LD between the alleles responsible for the superior performance of the male parents.

The contamination with foreign pollen (Falke et al. 2007) resulted in non-parental alleles with small allele frequency in the early cycles of the selection program. If these migrated alleles had a selective advantage, they showed a relatively large allele frequency change Δp, which was often detected by Waples' test (Falke et al. 2007). Therefore, many non-parental alleles were included in the sets of favorable alleles (Fig. 3) and, due to their linkage, contributed considerably to the positive LD observed in the final selection cycles C4 (A × B) and C7 (C × D) (Fig. 3).

Increase in negative LD due to selection

Selection is expected to increase the frequency of favorable alleles and simultaneously build up a negative LD between them (Bulmer 1971). Labate et al. (2000) observed slight increases in LD over 12 cycles of reciprocal RS. In this study, especially loci near fixation showed significant LD between each other. With simulations studies, Hospital and Chevalet (1996) found that the extent of negative LD increases at the beginning and decreases in later stages of the selection process, irrespective of the recombination frequency between linked loci. The time, when LD started decreasing, corresponded to the time when some loci reached fixation.

If the change in allele frequencies due to one cycle of selection at two loci is Δp_x and Δp_y, then the LD in the selection fraction is (Nei 1963)

$$D_{xy}^{(1)} = -\Delta p_x \Delta p_y. \quad (2)$$

The build-up of LD during selection is the net effect of two opposing forces: selection increases LD, while recombination reduces LD. However, this newly generated LD is reduced in every subsequent generation by recombination. As a first approximation for the LD reduction per recombination with our full-sib selection scheme, we use the expected LD reduction for random mating, which is $(1 - r)$, where r is the recombination frequency between the two loci (Falconer and Mackay 1996, p. 18). Assuming that in each generation (1) the newly generated LD due to selection equals D_{xy} (Eq. 1) and (2) the LD from the previous generation is reduced by a factor of $(1 - r)$, we have after n generations of selection and recombination (Nei 1963)

$$D_{xy}^{(n)} \leq {}^1\!/\!_r \left[1 - (1 - r)^n\right] D_{xy}^{(1)}. \quad (3)$$

We use two simple examples to illustrate the numerical magnitude of such expected allele frequency changes. (1) Very large allele frequency changes due to selection of $\Delta p_x = \Delta p_y = 0.2$ result in $D_{xy}^{(1)} = -0.04$. (2) $\Delta p_x = \Delta p_y = 0.1$, $r = 0.5$, and three generations of recombination result in $D_{xy}^{(3)} \leq -0.0175$. Very large populations would be necessary to detect such small changes with sufficient accuracy. We conclude that while the effects of selection could very well contribute to the observed increase in negative LD (Fig. 3), it seems hardly justified to attribute the considerable increase in negative LD exclusively to selection.

Further causes of an increase in negative LD can be sampling effects due to small population sizes. In contrast to selection, which is expected to build up LD in a directional process, sampling effects are expected to result in erratic changes in LD. To investigate the causes of the build-up of LD, we analyzed the five loci pairs showing the largest absolute positive or negative D values in the last selection cycle of both RS programs. However, D values did not show a directional, but rather an erratic change (Fig. 4). Thus, these changes in LD are attributable to sampling effects rather than selection. We therefore conclude that sampling effects due to small and finite number of selected plants ($N = 72$) were presumably an important

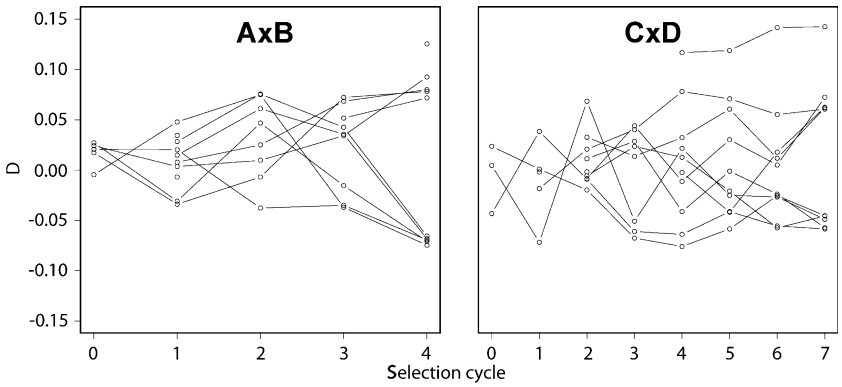

Fig. 4 Development of LD between alleles whose frequencies were increased by selection, measured by D over four cycles of RS in population A × B and seven cycles of RS in population C × D. LD was measured only for the five loci pairs with the highest and lowest LD values in the final selection cycles (A × B: C4 and C × D: C7)

factor contributing to the increase in LD during the selection cycles (Fig. 3).

Association between LD and additive genetic variance

In two companion studies, Flachenecker et al. (2006a, b) observed a decrease in σ_A^2 for selection index in both RS programs, which was only significant in A × B. In theory, the build-up of negative LD due to selection results in a reduction in σ_A^2. For the biallelic case and n loci, σ_A^2 and D_{xy} are related by (Lynch and Walsh 1998)

$$\sigma_A^2 = 2\sum_{x=1}^{n} \alpha_x^2 p_x(1-p_x) + 2\sum_{x=1}^{n}\sum_{y\neq x}^{n} \alpha_x\, \alpha_y\, D_{xy}, \qquad (4)$$

where n is the number of loci and, p_x is the frequency of allele x, D_{xy} the LD between the xth and yth locus, and α_x and α_y are the average effects of allele substitution at the xth and yth locus. However, a long-term reduction in σ_A^2 is only expected if uniformly negative D values were observed at many loci (Eq. 4). This was not the case in our experiment (Fig. 3) and therefore the LD generated by selection was hardly a factor in reducing the selection response by a reduction in σ_A^2.

Summarizing, we attribute the comparatively high selection response per cycle compared with other RS studies not only to the migration effects observed during the selection procedure but also to the applied mating scheme of Cockerham and Burrows (1980). Thus, the applied mating scheme offers an alternative for successful maize breeding by means of RS. Moreover, our experimental results support the hypothesis that the LD generated by selection is not a limiting factor for achieving high selection response in RS programs, in particular if an efficient recombination procedure is employed, which reduces negative LD between favorable alleles. This may be an explanation for of continued selection response in other long-term selection program, e.g., the Illinois long-term selection program (*cf.* Dudley and Lambert 2004). In particular, this hypothesis is supported by (1) the large phenotypic selection response in our study, and (2) the fact that a clear trend towards negative LD between favorable alleles was not observed at the molecular level.

Acknowledgments The present study was supported by a grant from the Deutsche Forschungsgemeinschaft, Grant No. FR 1615/1. The authors thank Dr. D. Klein and the staff of the Plant Breeding Research Station at Eckartsweier for conducting the field trials. We thank KWS SAAT AG and Pioneer Hi-Bred Northern Europe for their support with the SSR screening. We gratefully acknowledge Christine Beuter and Elisabeth Kokai-Kota for the technical assistance with the SSR analyses. The authors appreciate the editorial work of Dr. Jasmina Muminović, whose suggestions considerably improved the style of the manuscript.

References

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

Bulmer MG (1971) The effect of selection on genetic variability. Am Natl 105:201–211

Cockerham CC, Burrows PM (1980) Selection limits and strategies. Proc Natl Acad Sci USA 77:546–549

Coque M, Gallais A (2006) Genomic regions involved in response to grain yield selection at high and low nitrogen fertilization in maize. Theor Appl Genet 112:1205–1220

Dudley JW, Lambert RJ (2004) 100 generations of selection for oil and protein in corn. In: Janick J (ed) Plant breeding reviews, Part 1: Long-term selection: maize. Wiley, New York, pp 79–110

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics, 4th edn. Longman Sci and Tech, Harlow

Falke KC, Melchinger AE, Flachenecker C, Kusterer B, Frisch M (2006) Comparison of linkage maps from F_2 and three times intermated generations in two populations of European flint maize (*Zea mays* L.). Theor Appl Genet 113:857–866

Falke KC, Flachenecker C, Melchinger AE, Piepho H-P, Maurer HP, Frisch M (2007) Temporal changes in allele frequencies in two European F_2 flint maize populations under modified recurrent full-sib selection. Theor Appl Genet 114:765–776

Falque M, Decousset L, Dervins D, Jacob AM, Joets J, Martinant JP, Raffoux X, Ribiere N, Ridel C, Samson D, Charcosset A, Murigneux A (2005) Linkage mapping of 1454 new maize candidate gene loci. Genetics 170:1957–1966

Flachenecker C, Frisch M, Falke KC, Melchinger AE (2006a) Trends in population parameters and best linear unbiased prediction of

progeny performance in a European F_2 maize population under modified recurrent full-sib selection. Theor Appl Genet 112:483–491

Flachenecker C, Frisch M, Muminovic J, Falke KC, Melchinger AE (2006b) Modified full-sib selection and best linear unbiased prediction of progeny performance in a European F_2 maize population. Plant Breeding 125:248–253

Flachenecker C, Frisch M, Falke KC, Melchinger AE (2006c) Genetic drift and selection effects from recurrent selection programs in two F_2 populations of European flint maize. Theor Appl 113:1113–1120

Genter CF (1982) Recurrent selection for high inbred yields from the F_2 of maize single cross. Proc Ann Corn Sorghum Ind Res Conf 37:67–76

Gower JC (1966) Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53:325–338

Guo S, Thompson EA (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. Biometrics 48:361–372

Haldane JBS (1919) The combination of linkage values, and the calculation of distance between the loci of linked factors. J Genet 8:299–309

Hallauer AR (1985) Compendium on recurrent selection methods and their application. Crit Rev Plant Sci 3:1–33

Hallauer AR, Miranda JBF (1988) Quantitative genetics in maize breeding, 2nd edn. Iowa State University Press, Ames

Hospital F, Chevalet C (1996) Interactions of selection, linkage and drift in the dynamics of polygenic characters. Genet Res Camb 67:77–87

Johnson GR (1982) Two-locus theory in recurrent selection for general combining ability in maize. Theor Appl Genet 61:279–283

Kosambi DD (1944) The estimation of map distance from recombination values. Ann Eugen 12:172–175

Labate JA, Lamkey KR, Lee M, Woodman WL (1999) Temporal changes in allele frequencies in two reciprocally selected maize populations. Theor Appl Genet 99:1166–1178

Labate JA, Lamkey KR, Lee M, Woodman W (2000) Hardy–Weinberg and linkage equilibrium estimates in the BSSS and BSCB1 random mated populations. Maydica 45:243–255

Landi P, Frascaroli E (1993) Responses to four cycles of full-sib family recurrent selection in an F2 maize population. Maydica 38:31–37

Lee M, Sharopova N, Beavis WD, Grant D, Katt M, Blair D, Hallauer A (2002) Expanding the genetic map of maize with the intermated B73 × Mo17 (IBM) population. Plant Mol Biol 48:453–461

Liu SC, Kowalski SP, Lan TH, Feldmann KA, Paterson AH (1996) Genome-wide high-resolution mapping by recurrent intermating using *Arabidopsis thaliana* as a model. Genetics 142:247–258

Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland

Martin OC, Hospital F (2006) Two- and three-locus tests for linkage analysis using recombinant inbred lines. Genetics 173:451–459

Maurer, HP, Melchinger AE, Frisch M (2004) PLABSOFT: Software for simulation and data analysis in plant breeding. In: 17th EUCARPIA Gen Congr 2004 (post abstr). Tulln, Austria

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

Mihaljevic R, Schön CC, Utz HF, Melchinger AE (2005a) Correlations and QTL correspondence between line per se and testcross performance for agronomic traits in four populations of European maize. Crop Sci 45:114–122

Mihaljevic R, Utz HF, Melchinger AE (2005b) No evidence for epistasis in hybrid and per se performance of elite European flint maize inbreds from generation means and QTL analyses. Crop Sci 45:2605–2613

Moll RH (1991) Sixteen cycles of recurrent full-sib family selection for grain weight in two maize populations. Crop Sci 31:959–964

Nei M (1963) Effect of selection on the components of genetic variance. pp. 501–515. In:Hanson WD and Robinson HF (eds) Statistical genetics and plant breeding. National Academy of Sciences—National Research Council, Washington

Pinto LR, Carneiro Vieira ML, de Souza CL Jr, de Souza AP (2003) Reciprocal recurrent selection effects on the genetic structure of tropical maize populations assessed at microsatellite loci. Genet Mol Biol 26:355–364

R Development Core Team (2004) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Russell WA, Eberhard UA, Vega Q (1973) Recurrent selection for specific combining ability for yield in two maize populations. Crop Sci 13:257–261

Schön CC, Melchinger AE, Boppenmaier J, Brunklaus-Jung E, Herrmann RG, Seitzer JF (1994) RFLP mapping in maize: quantitative trait loci affecting testcross performance of elite European flint lines. Crop Sci 34(2):378–389

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: joinMap. Plant J 3:739–744

Teuscher F, Broman KW (2007) Haplotype probabilities for multiple-strain recombinant inbred lines. Genetics 175:1267–1274

Teuscher F, Guiard V, Rudolph PE, Brockmann GA (2005) The map expansion obtained with recombinant inbred strains and intermated recombinant inbred populations for finite generation designs. Genetics 170:875–879

Waples RS (1989) Temporal variation in allele frequencies: testing the right hypothesis. Evolution 43:1236–1251

Weir BS (1996) Genetic data analysis II. Sinauer Associates, Sunderland

Winkler CR, Jensen NM, Cooper M, Podlich DW, Smith OS (2003) On the determination of recombination rates in intermated recombinant inbred populations. Genetics 164:741–745

Wright S (1978) Evolution and genetics of populations. The University of Chicago Press, Chicago

Zaykin DV, Zhivotovsky L, Weir BS (1995) Exact tests for association between alleles at arbitrary numbers of loci. Genetica 96:169–178