Theor Appl Genet (1994) 87:1016–1020

Linkage disequilibrium in crosses between Illinois maize strains divergently selected for protein percentage

J. W. Dudley

Department of Agronomy, University of Illinois, 1102 S. Goodwin Ave., Urbana, IL 61801, USA

Recived: 4 May 1993 / Accepted: 2 August 1993

Abstract. The objectives of this study were two fold: (1) to determine whether divergent selection for kernel protein concentration, which produced the Illinois high protein (IHP), Illinois low protein (ILP), reverse low protein (RLP), and reverse high protein (RHP) maize (*Zea mays* L.) strains, had generated coupling-phase linkages among genes controlling protein concentration or other traits and (2) to measure the effectiveness of random mating in reducing linkage disequilibrium in segregating generations from crosses between the strains. To achieve these objectives, design III progenies from the F_2 and F_6 (produced by random mating the F_2) from the crosses of IHP × ILP, IHP × RHP, ILP × RLP, and RHP × RLP were evaluated. Estimates of additive variance for percent protein in the crosses of IHP × ILP and ILP × RLP were significantly less in the F_6 than in the F_2 indicating the presence of coupling-phase linkages in the parents and their breakup by random mating. In addition, a significant reduction in dominance variance for grain yield from the F_2 to the F_6 in IHP × ILP suggested the presence of repulsion-phase linkages. No other evidence of coupling- or repulsion-phase linkages was found for any of the traits measured. These results demonstrate the effectiveness of long-term divergent selection in the development of coupling-phase linkages and of random mating to dissipate linkage disequilibrium.

Key words: Maize – Long-term selection – Quantitative genetics – Corn breeding – Protein

Communicated by A. R. Hallauer
Research supported by the Illinois Agricultural Experiment Station

Introduction

Recent work on the identification of associations between molecular marker loci and genes controlling quantitative traits (for reviews see Stuber 1992; Dudley 1993) has increased interest in the measurement and extent of linkage disequilibrium. Recent theoretical work (Dudley 1992, 1993) suggests random mating to reduce linkage disequilibrium at loosely-linked loci as a mechanism for identifing molecular markers associated with smaller chromosomal regions controlling quantitative traits. One measure of the extent of linkage disequilibrium in F_2 populations is the change in different types of genetic variance with random mating. Additive genetic variance estimates from F_2 populations are biased upward by linkage disequilibrium if coupling-phase linkages are predominant and downward if repulsion-phase linkages are predominant (Comstock and Robinson 1952). Dominance genetic variance estimates from F_2 populations are always biased upward regardless of the linkage phase. Because these biases result from linkage disequilibrium, they will decrease with random mating. Experimental estimates for maize grain yield from F_2 and advanced random mating generations demonstrated the bias in estimates of the degree of dominance caused by the presence of repulsion-phase linkages (Gardner 1963; Moll et al. 1964; Hallauer and Miranda 1988).

Because long-term divergent selection should create coupling-phase linkages, F_2 generations from crosses between divergently-selected populations should be in linkage disequilibrium caused by an excess of coupling-phase gametes. In such populations, estimates of additive genetic variance from random-mated generations following the F_2 should be lower than those from the F_2 (Comstock and Robinson 1952).

Long-term selection for oil and protein concentration in corn grain has created populations with divergent means for oil and protein (Dudley 1976; Dudley and Lambert 1992). Moreno-Gonzalez et al. (1975), in the cross between the 68th generation of the Illinois high oil (IHO) and Illinois low oil (ILO) strains, found significant reductions in estimates of additive genetic variance for percent oil from the F_2 to the F_6, demonstrating the presence of coupling-phase linkages.

The objectives of the present study were to determine: (1) whether divergent selection for kernel protein concentration, which produced the Illinois high protein (IHP) and Illinois low protein (ILP) strains, had generated coupling-phase linkages among genes controlling protein concentration or unselected traits, (2) whether 22 generations of reverse selection, which produced the reverse low protein (RLP) and reverse high protein (RHP) strains, had generated coupling-phase linkages differentiating RLP from ILP and RHP from IHP, (3) whether coupling-phase linkages still differentiated RHP and RLP, and (4) to measure the effectiveness of random mating in reducing linkage disequilibrium.

Materials and methods

The IHP, ILP, RHP, and RLP strains of corn were used as parental materials. IHP resulted from 70 generations of selection for high kernel protein concentration and ILP from 70 generations of selection for low kernel protein concentration in the open-pollinated cultivar Burr's White (Dudley and Lambert 1992). RHP resulted from 22 generations of selection for low protein concentration, starting with generation 48 of IHP, while RLP resulted from 22 generations of selection for high protein, starting with generation 48 of ILP. Details of the selection procedures have been published elsewhere (Dudley and Lambert 1992).

The following crosses were made: IHP × ILP, IHP × RHP, RHP × RLP, and ILP × RLP. Starting with the F_2, each cross was random mated for four generations by pollinating about 200 ears with bulk pollen each generation. A design III mating design (Comstock and Robinson 1952), in which individual plants of a given generation were crossed, as males, to both parental populations, was produced in 1979 for the F_2 and the F_6 of each cross. For IHP × ILP, IHP × RHP, and ILP × RLP 64 plants from each generation were crossed to each parent whereas 80 plants were sampled from each generation for RHP × RLP. Eighty plants were used for RHP × RLP because of the small differences in mean percent protein between the two strains. For each cross, the design III progenies for the F_2 and F_6 generations were grown in separate experiments. Each experiment consisted of a replications-in-blocks design with two replications of 16 entries (eight males with two crosses per male organized in a factorial arrangement) in each block. Blocks of the F_2 experiment from a given cross were alternated with blocks of the F_6 experiment of the same cross in the field. Thus the F_2 and F_6 progenies from a given cross sampled the same land area. There were ten blocks for the RHP × RLP cross and eight blocks for each of the other crosses. All experiments were grown in both 1980 and 1981 on the Agronomy South farm at Urbana, Illinois. However, because of low yields in 1980, data from IHP × RHP progenies were obtained only in 1981. Because of missing data in 1980, only 55 F_6 plants, instead of 64, were used in the analysis of the IHP × ILP F_6 progenies. Plots were single rows, 0.76 m apart and 5.3 m long. They were machine planted and thinned to 16 plants per row. Plots were machine harvested and a sample of grain was collected for chemical analysis. Because there is little effect of the pollen parent on protein percentage (Woodworth and Jugenheimer 1948) and primary interest was in protein and yield, controlled pollination to provide seed for chemical analysis was not considered necessary.

Protein and oil percentages (dry matter basis) were measured using a Dickey-john near infra-red analyzer (Hymowitz et al. 1974). Grain weights were converted to t ha^{-1} at 15.5% moisture. Plant height was measured as cm to the flag leaf node. Plant height data were collected in both 1989 and 1981 for IHP × ILP but only in 1981 for the other crosses.

Statistical analysis

The Design III analysis of variance for the F_2 and F_6 generations followed that given by Comstock and Robinson (1952) and Moreno-Gonzalez et al. (1975). For F-tests and estimation of variance components, males were considered random, the parents fixed, and years random. As suggested by Snedecor and Cochran (1967), F-tests were used to test the significance of components of variance within each generation. Differences between corresponding components of variance in the two generations were considered significant if 90% confidence intervals, calculated as suggested by Snedecor and Cochran (1967, pp 244–245), did not overlap. Because means for protein percentage were linearly associated with the standard deviations, protein data were transformed using a logarithm to the base 10 ($\log_{10}$ transformation).

Genetic interpretation of the Design III analysis has been discussed extensively (Comstock and Robinson 1952; Gardner et al. 1953). Moreno-Gonzalez et al. (1975) extended the theory to show that the effect of having non-homozygous parents was negligible if gene frequencies in the two parents were highly divergent. The critical components of variance are σ^2_m, the component of variance associated with the mean square for variation among male parents, and σ^2_{ml}, the component of variance associated with the interaction between male plants and parent lines. As shown by Comstock and Robinson (1952) the expectations of these components are:

$$\sigma^2_m = 1/2 \sum_j q_j(1-q_j)u_j^2 + \sum_{j,k}(pt-rs)_{j,k}u_j u_k$$

and

$$\sigma^2_{ml} = \sum_j q_j(1-q_j)a_j^2 u_j^2 + 2\sum_{j,k}^{r}(pt-rs)_{j,k}a_j u_j a_k u_k + 2\sum_{j,k}^{c}(rs-pt)_{j,k}a_j u_j a_k u_k$$

where p, r, s, and t are the frequencies of B_jB_k, B_jb_k, b_jB_k, and b_jb_k gametes, respectively; B_j, b_j, B_k, and b_k are the alleles at loci j and k, respectively. $\sum^{c}_{j,k}$ and $\sum^{r}_{j,k}$ are the sums over the jth and kth pairs of loci when the initial linkage phases are coupling and repulsion, respectively. These expressions account for the covariance components and are the source of linkage bias. In terms of genetic variance components, $\sigma^2_m = 1/4\,\sigma^2_g$ and $\sigma^2_{ml} = \sigma^2_d$, where σ^2_g and σ^2_d are the additive and dominance components respectively, provided that gene frequencies at segregating loci are 0.5.

Table 1. Means (from Dudley 1977) of parental strains for traits measured in Design III experiments

Trait	IHP	ILP	RHP	RLP	SE[a]
Grain yield (t ha^{-1})	2.04	3.63	3.83	5.26	0.02
Protein (%)	26.1	5.8	10.6	11.2	0.15
Oil (%)	5.4	4.2	4.8	4.4	0.11
Plant height (cm)[b]	155	179	185	204	4.8

[a] Standard error of a mean
[b] Measured in 1 year only. Other traits measured in 2 years

Results and discussion

Although the parental strains were not grown in this study, means for generation 70 were reported by Dudley et al. (1977) and are shown in Table 1. IHP and ILP differed significantly for grain yield, plant height, and oil percentage, as well as for protein percentage. For yield, plant height, and protein percentage, RHP and IHP differed significantly as did RLP and ILP. However, RHP and RLP differed significantly only for yield and plant height.

Because differences between overall means of the F_2 and F_6 progenies were small (Table 2), sampling during the random-mating generations was considered adequate. Estimates of σ_m^2 were significantly different from 0 for all traits in both generations except for grain yield in the F_2 and F_6 and moisture percentage in the F_6 of the IHP × ILP cross, the $\log_{10}$ protein percentage in the F_6 of ILP × RLP, the oil percentage in the F_6 of IHP × RHP, and the grain yield in the F_2 of RHP × RLP (Table 3). Significant reductions in estimates of σ_m^2 from the F_2 to the F_6 were found only for $\log_{10}$ protein concentration and only in the IHP × ILP and ILP × RLP crosses. The estimate of variance associated with linkage disequilibrium for $\log_{10}$ percent protein was larger for the IHP × ILP (8.5×10^{-4}) cross than for the ILP × RLP (5.0×10^{-4}) cross. Thus 70 generations of divergent selection had created coupling-phase linkages differentiating IHP and ILP. The 22 generations of reverse selection were adequate to develop coupling-phase linkages differentiating ILP

Table 3. Estimates of σ_m^2 and 90% Confidence limits (,) from the F_2 and F_6 generations of the design III experiments

Trait and cross	Generation	
	F_2	F_6
Yield (t ha^{-1} × 10^{-2})		
IHP × ILP	7.19 (−1.33, 16.89)	4.27 (−3.66, 13.47)
IHP × RHP	4.98 (1.17, 12.02)*	9.12 (4.10, 15.57)*
ILP × RLP	15.90 (4.44, 31.25)*	14.62 (3.35, 29.12)*
RHP × RLP	7.24 (−0.44, 16.96)	9.41 (1.07, 19.98)*
Protein $\log_{10}$ (%) × 10^4		
IHP × ILP	12.4 (7.80, 18.16)*	4.1 (1.15, 7.42)[a]
IHP × RHP	4.3 (2.15, 7.00)*	1.6 (0.22, 3.66)*
ILP × RLP	5.4 (2.67, 9.20)*	0.4 (−1.26, 2.47)[a]
RHP × RLP	2.5 (1.24, 4.09)*	2.9 (1.60, 4.73)*
Oil (%)		
IHP × ILP	0.04 (0.02, 0.06)*	0.03 (0.01, 0.05)*
IHP × RHP	0.02 (0.00, 0.05)*	0.01 (−0.00, 0.03)
ILP × RLP	0.01 (0.01, 0.02)*	0.01 (0.00, 0.01)*
RHP × RLP	0.01 (0.00, 0.02)*	0.02 (0.01, 0.03)*
Moisture (%)		
IHP × ILP	0.57 (0.32, 0.83)*	0.24 (−0.00, 0.52)
IHP × RHP	0.36 (0.22, 0.55)*	0.31 (0.15, 0.52)*
ILP × RLP	0.73 (0.53, 1.37)*	0.39 (0.12, 0.73)*
RHP × RLP	0.53 (0.33, 0.80)*	0.51 (0.31,0.78)*
Plant height (cm)		
IHP × ILP	21.6 (12.6, 35.0)*	28.0 (17.2, 45.0)*
IHP × RHP	26.8 (12.9, 48.1)*	25.2 (13.8, 42.7)*
ILP × RLP	25.1 (6.6, 50.2)*	19.2 (2.3, 42.1)*
RHP × RLP	41.2 (19.6. 70.3)*	21.6 (3.2, 45.0)*

* Significant at the 0.05 probability level based on the F test in the analysis of variance
[a] 90% confidence intervals of F_2 and F_6 estimates do not overlap

Table 2. Means (± standard error[a]) averaged across all crosses and years for traits measured in the design III studies

Trait	IHP × ILP	IHP × RHP	ILP × RLP	RHP × RLP
Yield(t ha^{-1}) F_2	3.51 ± 0.37	1.97 ± 0.38	3.80 ± 0.42	3.44 ± 0.35
F_6	3.36 ± 0.38	2.09 ± 0.34	3.79 ± 0.44	3.56 ± 0.35
Protein(%) F_2	12.8 ± 0.06	19.4 ± 0.10	7.8 ± 0.04	11.7 ± 0.03
F_6	13.2 ± 0.05	19.9 ± 0.10	8.3 ± 0.04	11.6 ± 0.03
Oil(%) F_2	5.2 ± 0.014	5.7 ± 0.025	4.5 ± 0.009	4.6 ± 0.009
F_6	5.0 ± 0.014	5.6 ± 0.022	4.4 ± 0.009	4.7 ± 0.009
Moisture (%) F_2	21.0 ± 0.04	20.4 ± 0.04	20.8 ± 0.06	21.3 ± 0.04
F_6	21.2 ± 0.05	20.1 ± 0.06	21.0 ± 0.06	21.0 ± 0.04
Plant height (cm) F_2	209 ± 0.34	209 ± 0.59	241 ± 0.48	235 ± 0.64
F_6	208 ± 0.33	211 ± 0.49	246 ± 0.48	231 ± 0.67

[a] standard error of the experiment mean

Table 4. Estimates of σ^2_{ml} and 90% confidence limits (,) from the F_2 and F_6 generation of the design III experiments

Trait and cross	Generation	
	F_2	F_6
Yield (t ha^{-1} × 10^{-2})		
IHP × ILP	58.5 (27.8, 92.1)*	8.9 (−5.9, 26.2)[a]
IHP × RHP	4.3 (−1.5, 16.9)	17.1(8.0, 30.7)*
ILP × RLP	10.9(−8.6, 35.2)	−3.7(−18.7, 14.5)
RHP × RLP	12.5 (1.0, 27.2)*	−0.9(−12.4, 12.6)
Log protein (%) × 10^4		
IHP × ILP	0.5 (−2.3, 4.0)	3.6 (0.6, 7.1)*
IHP × RHP	3.7 (1.0, 7.8)*	4.6 (1.5, 9.3)*
ILP × RLP	6.1 (1.5, 12.2)*	3.5 (0.3, 7.7)*
RHP × RLP	2.1 (1.1, 3.8)*	1.9 (0.4, 3.5)*
Oil (%)		
IHP × ILP	0.01 (−0.01, 0.3)	−0.01(−0.02,0.01)
IHP × RHP	0.04 (0.01, 0.10)	0.03 (0.01, 0.08)
ILP × RLP	0.01 (0.0, 0.01)*	0.01 (0.00, 0.01)*
RHP × RLP	0.01 (−0.0, 0.01)	0.01(−0.00,0.02)
Moisture (%)		
IHP × ILP	0.15 (−0.06, 0.40)	0.40 (0.02, 0.80)
IHP × RHP	0.36 (0.18, 0.61)*	0.27 (0.05, 0.58)*
ILP × RLP	0.09 (−0.16, 0.40)	0.38 (0.07, 0.77)*
RHP × RLP	0.38 (0.18, 0.64)*	0.42 (0.15, 0.75)*
Plant height (cm)		
IHP × ILP	28.4 (14.6, 48.1)*	9.5 (−2.9, 25.8)
IHP × RHP	6.4 (−10.0, 29.2)	44.9 (23.4, 77.6)*
ILP × RLP	16.1 (−7.4, 49.1)	22.3 (−2.6, 58.4)
RHP × RLP	30.4 (3.7, 66.9)*	44.2 (13.0, 89.0)*

* Significant at the 0.05 probability level based on the F test in the analysis of variance

[a] See Table 3

from RLP but not IHP from RHP. The mean protein percentages of RHP and RLP were similar and no evidence of linkage disequilibrium due to coupling-phase linkages was found. Thus, after 48 generations of selection, reverse selection may have dissipated enough of the coupling-phase linkages which differentiated IHP and ILP to prevent detection of a significant reduction in additive variance with four generations of random mating. Alternatively, RHP and RLP may have been at similar, intermediate gene frequencies and thus the assumptions necessary for detecting bias no longer hold. The lack of significant reductions in σ^2_m for traits other than protein, even though significant differences between the parental means were present for many of these traits, suggests that selection for protein did not cause the development of coupling-phase linkages for other traits.

For $\log_{10}$ protein percentage, significant dominance variance estimates were found for all crosses and generations except for the F_2 of IHP × ILP (Table 4). However, no significant shifts in estimates of dominance variance from F_2 to F_6 were found. Such a result could occur if genes showing dominance for protein percentage are scattered throughout the genome. Except for the significant reduction in dominance variance from the F_2 to the F_6 for grain yield in the IHP × ILP cross, there is little evidence that selection for protein concentration caused a build up of either coupling- or repulsion-phase linkages for unselected traits.

The results of this study, in agreement with those of Moreno-Gonzalez et al. (1975), demonstrate the effectiveness of divergent selection in building up coupling-phase linkages. In agreement with Gardner et al. (1953), Moll et al. (1964), and Moreno-Gonzalez et al. (1975), the result demonstrate the effectiveness of random-mating in breaking up linkage disequilibrium. Except for grain yield in the IHP × ILP cross, no evidence for the development of either coupling- or repulsion-phase linkages for traits other than the selected trait were found. These results have implications for work using molecular markers to identify associations with genes controling quantitative traits. Based on the apparent rapid dissipation of linkage disequilibrium in this study, measurement of molecular marker-QTL associations in the F_2 and again in subsequent random-mated generations, as previously proposed (Dudley 1993), should help identify smaller chromosome regions tightly linked to molecular markers.

References

Comstock RE, Robinsion HF (1952) Estimates of average dominance of genes. In: Gowen JW (ed) Heterosis. Iowa State College Press, Ames, Iowa, pp 494–516

Dudley JW (1976) 76 generations of selection for oil and protein percentage in maize. In: Pollak E, Kempthorne O, Bailey TB Jr (eds) Proc Int Conf Quantit Genet. Aug 16–21, 1976. Iowa State Univ. Press, Ames, Iowa pp 459–474

Dudley JW (1992) Theory for identification of marker locus-QTL association in populations by line crosses. Theor Appl Genet 85:101–104

Dudley JW (1993) Molecular markers in plant improvement: manipulation of genes affecting quantitative traits. Crop Sci 33:660–668

Dudley JW, Lambert RJ (1992) Ninety generations of selection for oil and protein in maize. Maydica 37:1–7

Dudley JW, Lambert RJ, de la Roche IA (1977) Genetic analysis of corn strains divergently selected for percent oil and protein. Crop Sci 17:111–117

Gardner CO (1963) Estimates of genetic parameters in cross-fertilizing plants and their implications in plant breeding. In: Hanson WD (ed) Statistical genetics and plant breeding. Natl Acad Sci, Natl Research Council Publ 982, Washington, DC, pp 225–252

Gardner CO, Harvey PH, Comstock RE, Robinson HF (1953) Dominance of genes controlling quantitative characters in maize. Agron J 45:186–191

Hallauer AR, Miranda, Fo JB (1988) Quantitative genetics in maize breeding. Iowa State Univ Press, Ames, Iowa

Hymowitz T, Dudley JW, Collins FI, Brown CM (1974) Estimation of protein and oil concentration in corn, soybean, and oat seed by near-infra red reflectance. Crop Sci 14:713–715

Moll RH, Lindsey, MF, Robinson HF (1964) Estimates of genetic variances and level of domainance in maize. Genetics 49:411–423

Moreno-Gonzalez J, Dudley JW, Lambert RJ (1975) A design III study of linkage disequilibrium for percent oil in maize. Crop Sci 15:840–843

Snedecor GW, Cochran WG (1967) Statistical methods (6th edn) Iowa State Univ Press, Ames, Iowa Stuber CW (1992) Biochemical and molecular markers in plant breeding. In: Janick J (ed) Plant breeding reviews. John Wiley and sons, Inc, New York, pp 37–61

Woodworth CM, Jugenheimer RW (1948) Breeding and genetics of high-protein corn. In: What's new in the production, storage and utilization of hybrid corn. A report of Hybrid Seed Corn Division's 3rd Ann Industry-Research Conference. Chicago, Illinois 1–2 Dec. ASTA Chicago, Illinois, pp 75–83