Theor Appl Genet (1993) 87:217–224

Quantitative trait loci influencing protein and starch concentration in the Illinois Long Term Selection maize strains

I. L. Goldman *, T. R. Rocheford, J. W. Dudley

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

Received: 11 November 1992 / Accepted: 11 February 1993

Abstract. A study was initiated to determine the number, chromosomal location, and magnitude of effect of QTL (quantitative trait loci or locus depending on context) controlling protein and starch concentration in the maize (*Zea mays* L.) kernel. Restriction fragment length polymorphism (RFLP) analysis was performed on 100 F_3 families derived from a cross of two strains, Illinois High Protein (IHP), X Illinois Low Protein (ILP), which had been divergently selected for protein concentration for 76 generations as part of the Illinois Long Term Selection Experiment. These families were analyzed for kernel protein and starch in replicated field trials during 1990 and 1991. A series of 90 genomic and cDNA clones distributed throughout the maize genome were chosen for their ability to detect RFLP between IHP and ILP. These clones were hybridized with DNA extracted from the 100 F_3 families, revealing 100 polymorphic loci. Single factor analysis of variance revealed significant QTL associations of many loci with both protein and starch concentration ($P < 0.05$ level). Twenty-two loci distributed on 10 chromosome arms were significantly associated with protein concentration, 19 loci on 9 chromosome arms were significantly associated with starch concentration. Sixteen of these loci were significant for both protein and starch concentration. Clusters of 3 or more significant loci were detected on chromosome arms 3L, 5S, and 7L for protein concentration, suggesting the presence of QTL with large effects at these locations. A QTL with large additive effects on protein and starch concentration was detected on chromosome arm 3L. RFLP alleles at this QTL were found to be linked with RFLP alleles at the *Shrunken-2* (*Sh2*) locus, a structural gene encoding the major subunit of the starch synthetic enzyme ADP-glucose pyrophosphorylase. A multiple linear regression model consisting of 6 significant RFLP loci on different chromosomes explained over 64% of the total variation for kernel protein concentration. Similar results were detected for starch concentration. Thus, several chromosomal regions with large effects may be responsible for a significant portion of the changes in kernel protein and starch concentration in the Illinois Long Term Selection Experiment.

Key words: Restriction fragment length polymorphism (RFLP) – Mapping – Illinois Long Term Selection Experiment – Quantitative trait loci (QTL) – Protein – Starch – *Zea mays* L.

Communicated by A. R. Hallauer

* *Present address:* Department of Horticulture, University of Wisconsin-Madison, 1575 Linden Drive, Madison, WI 53706, USA

Correspondence to: T. R. Rocheford

Introduction

Maize (*Zea mays* L.) has long been utilized as an important human and animal food source due to the high levels of starch and protein in its kernel. Starch, which is primarily stored in endosperm tissue, and protein, which is found in both the endosperm and embryo, occupy approximately 85% of the kernel on a dry-weight basis. Both starch and protein are economically important products in wet and dry milling of corn grain (Glover 1988). Variation in protein and starch concentration in the maize kernel is quantitatively inherited (Dudley 1977), with as many as 173 effective factors estimated to control the protein response in a mass-selected maize population (Dudley and Lambert 1992). Traditionally, quantitative traits, or those characteristics conditioned by the action of many genes and displaying continuous variation in phenotype, have been investigated via statis-

tical models. In recent years, techniques developed by molecular geneticists, such as the identification of restriction fragment length polymorphism (RFLP), have begun to play an important role in the manner in which quantitative genetic variation is analyzed (Lander and Botstein 1989; Paterson et al. 1988; Edwards et al. 1992). The advent of high-density DNA marker linkage maps in many plant species has provided the opportunity to identify quantitative trait loci (QTL) by their association with segregating RFLP loci, thus allowing the examination of the genetic basis of quantitative trait variation at the molecular level.

During recent years, much time and effort has been invested in an effort to identify DNA markers associated with important agronomic traits in plants, such as yield, plant height, soluble solids, and disease resistance (Edwards et al. 1987; Paterson et al. 1988; Young et al. 1988; Osborn et al. 1987; Edwards et al. 1992). Limited research, however, has been attempted in the search for DNA markers associated with kernel chemical composition traits. Evidence gathered in recent years has suggested that breeding efforts designed to produce elevated levels of kernel protein and starch may result in higher value products in the corn processing industry (Glover 1988). Phenotypic extremes for kernel protein and starch concentration exist among maize strains involved in the illinois Long Term Selection Experiment (ILTSE). This experiment was initiated in 1896 by C. Hopkins in an effort to investigate changes in maize kernel chemical composition resulting from mass selection for protein and oil concentration (Hopkins 1899). Due to the large negative correlation between kernel protein and starch concentration, divergent selection has also resulted in large phenotypic extremes for kernel starch concentration. The ILTSE strains exhibit considerable RFLP (Sughroue and Rocheford 1993), making these genetic materials well suited for studies designed to analyze quantitative variation in kernel composition at the molecular level. Our objective was to use RFLP analysis to determine the number, magnitude of effect, and location of QTL controlling protein and starch concentration of the maize kernel in a segregating population derived from the cross IHP × ILP.

Materials and methods

Genetic materials and chemical analyses

Details of specific selection procedures, chemical analyses, and statistical evaluations of the genetic strains developed from the ILTSE have been reported (Dudley and Lambert 1992). A population was developed from a cross of five to seven IHP plants × five to seven ILP plants from cycle 76 of the ILTSE. Mean protein concentrations for cycle 76 were 25% for IHP and 4% for ILP. Mean starch concentrations for cycle 76 were 42% for IHP and 74% for ILP. F_1 plants were selfed to obtain an F_2 bulk population, and F_2 plants were self-pollinated to produce 100 F_3 families. Each F_3 family was represented by a single row 5 m long with 0.76 m between rows. The experiment was grown in 1990 and 1991. Five to seven plants per plot were self-pollinated, and equal numbers of seeds from each selfed ear within a plot were bulked for chemical analysis. Protein and starch concentration in the kernel was measured on a sample of the bulked seed on a Dickey-john GAC III near-infrared analyzer (Hymowitz et al. 1974; Dudley and Lambert 1992).

Table 1. Chromosomal location and relative position of RFLP loci used in this study

Chromosome	Locus
1	P 200689[a], NPI 234, NPI 286, NPI 262, NPI 401, NPI 272, UMC 67, UMC 23, NPI 447, UMC 72B
2	NPI 239, UMC 6, NPI 269, UMC 134, NPI 456, UMC 5, NPI 123, UMC 125, UMC 137, UMC 36
3	P 2000905, NPI 249, UMC 10, P 200511, BNL 10.24, UMC 39, UMC 17, UMC 16, P 100080, *Sh2*, NPI 432, UMC 63, UMC 96, NPI 457, NPI 420
4	UMC 123, *Bt2*, NPI 270, NPI 410, P 100025, UMC 66, UMC 104A, NPI 593, NPI 451
5	NPI 409, UMC 72A, UMC 90, UMC 27, UMC 43, UMC 1, NPI 213, NPI 237, P 100017, UMC 104B
6	UMC 85, NPI 377, UMC 21, NPI 252, BNL 5.47, UMC 38, UMC 62, NPI 280
7	NPI 277, *o2*, UMC 116, P 200746, NPI 455, BNL 8.32, BNL 8.39, NPI 433, BNL 16.06, BNL 8.44
8	NPI 110, NPI 276, NPI 260, BNL 9.08, UMC 93, UMC 30
9	NPI 253, UMC 113, *Sh1*, NPI 266, BNL 5.10, NPI 222, UMC 20, NPI 427, NPI 443, UMC 94
10	BNL 10.17, NPI 285, NPI 264, NPI 232

[a] Loci listed in relative chromosomal order beginning with distal end of short arm based upon map position described by Coe (1992)

Restriction fragment length polymorphism analysis

Genomic DNA clones used as probes were selected from collections of mapped maize clones provided by University of Missouri-Columbia (UMC), Brookhaven National Laboratory (BNL), and Pioneer Hi-Bred International (P). Clones obtained from P also included the Native Plants, Inc (NPI) clone set. DNA isolation, restriction enzyme digestion with *Eco*RI, electrophoresis, Southern blot transfer, probe preparation, and hybridizations were essentially performed according to established procedures (Feinberg and Vogelstein 1983; Hoisington 1991; Lee 1991; Saghai-Maroof et al. 1984; Sambrook et al. 1989).

QTL mapping

A set of 90 previously screened genomic clones spaced approximately every 20 cM throughout the maize genome were hybridized to DNA from F_3 families and subsequently revealed a total of 100 polymorphic loci (Table 1). Ten clones (UMC 17, UMC 16, UMC 72, NPI 593, UMC 104, BNL 5.47, BNL 8.39,

NPI 455, NPI 253, and NPI 266) revealed 2 polymorphic loci. Two of these clones (UMC 72 and UMC 104) have revealed multiple RFLP loci in another maize cross, and each locus has been placed on an RFLP map (Coe 1992). Chromosomal locations of the remaining 8 additional loci are unknown and are listed as only a single locus in Table 1. All chromosomal locations of RFLP loci were taken from Coe (1992) and unpublished information from NPI. RFLP alleles at a locus were assigned letter designations 'A' [highest molecular weight (MW) fragment] and 'B' (next highest MW fragment). Additional RFLP alleles at a locus (when present) were designated 'C', 'D', etc. by descending MW. Genotypes were assigned accordingly.

Individual marker loci were tested for linkage to QTL by an analysis of variance of mean performance of marker classes essentially according to procedures used by Edwards et al. (1987). Significant differences in marker class means were interpreted as indicating the linkage of marker loci to QTL controlling protein or starch concentration. Significance was set at $P<0.05$ since the purpose of this experiment was to maximize the identification of chromosomal regions that may be associated with the control of protein and starch concentration. At this probability level we recognize that some Type 1 errors may occur. Linear and quadratic contrasts were performed for each locus with two alleles exhibiting a significant F-test ($P<0.05$). Linear effects contrasted mean performance of homozygous marker classes (additive gene action); quadratic effects contrasted the mean of homozygous marker classes with the heterozygous marker class (dominant gene action). Since the F_3 families used in this study were generated from a multiple-parent cross of highly inbred strains, multiple RFLP alleles were detected at some loci. Single factor analysis of variance was performed on these loci, but no linear or quadratic contrasts were performed. Additionally, accurate linkage relationships among RFLP loci could not be calculated due to the unknown RFLP allelic constitution of the multiple IHP and ILP parents.

To further characterize variation in protein and starch concentration, multiple linear regression models for these traits were constructed by adding significant marker loci as main effects and in pairwise combinations into a multiple linear regression model in a stepwise fashion. Because of the confounding effect of serial correlation between linked loci, a single significant marker locus displaying a significant effect on protein or starch concentration was chosen from each chromosome and added to the multiple regression model. Only those loci or pairs of loci which were significant at $P<0.05$ remained in the model.

Results and discussion

Phenotypic evaluation of protein and starch concentration

The analysis of variance combined over years revealed highly significant differences ($P<0.01$) among F_3 families for protein and starch concentration (data not shown). The year X family interaction was highly significant for both protein and starch concentration, but this interaction was due primarily to changes in the magnitude of means. In 1990 mean F_3 family concentrations were 12.1% for protein and 67.4% for starch. Mean F_3 family concentrations were 13.8% for protein and 59.5% for starch in 1991. Data were thus combined over years for mapping analyses (Fig. 1). Sufficient moisture was available for plant growth and development during the 1990 growing season, whereas moisture was limiting during the 1991 growing season. Differences in moisture availability were primarily responsible for changes in the magnitude of means between these 2 years.

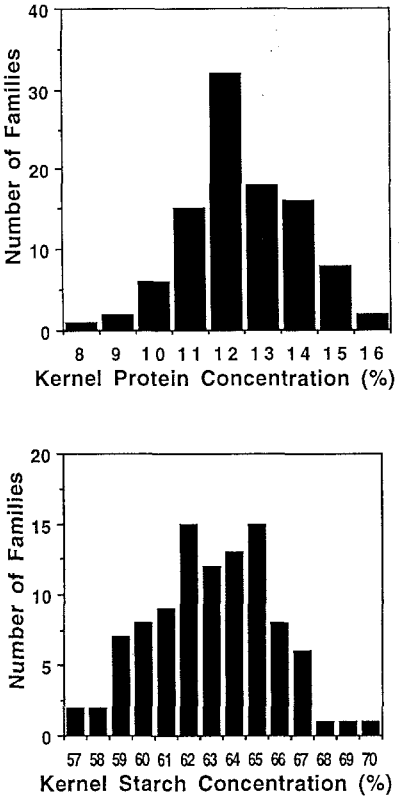

Fig. 1. Distribution of kernel protein (*top*) and starch (*bottom*) concentration among 100 F_3 families grown in replicated trials in 1990 and 1991. Mean protein concentration for cycle 76 IHP and ILP was 25% and 4%, respectively. Mean starch concentration for cycle 76 IHP and ILP was 42% and 74%, respectively

Identification of QTL associated with protein and starch concentration

Among the 100 F_3 families, 22 loci distributed on 10 chromosome arms were determined to be significantly associated with protein concentration, and 19 loci on 9 chromosome arms were significantly associated with starch concentration (Table 2 and Fig. 2). Sixteen loci were significant for protein and starch concentration. Clusters of 3 or more significant loci were detected on chromosomes arms 3L, 5S, and 7L for protein concentration, suggesting the presence of either a QTL with large effects or multiple QTL at these locations. Significant associations between marker loci surrounding QTL on 5S and 7L were detected for 3 adjacent loci for protein and starch concentration, and the QTL association on 3L was detected by 7 adjacent loci. The significant association on 3L spanned a longer approximate map distance than significant loci on 5S and 7L. Marker loci flanking significant loci on 5S and 7L were tested and determined

not to be significantly associated with protein and starch concentration. This suggests that 1 or more QTL in the 3L region may be of a larger magnitude than those located on 5S and 7L period. A cluster of 2 marker loci significant for protein concentration was detected on 5L. Marker class means were compared at significant marker loci for protein and starch concentration (Table 3). Homozygous marker classes differed by up to 1.7% in protein concentration and 3% in starch concentration.

Use of structural genes in QTL mapping

An effort was made to utilize cloned maize genes with well-characterized effects on kernel starch and protein biosynthesis. cDNA clones of four structural genes revealed RFLP between IHP and ILP. cDNA clones of *Shrunken-2* (*Sh2*, Bhave et al. 1990), *Brittle-2* (*Bt2*, Bae et al. 1990), *opaque-2* (*o2*, Schmidt et al. 1990), and *Shrunken-1* (*Sh1*, Sheldon et al. 1983) were used to determine the presence or absence of QTL associations with loci that have qualitative mutants known to affect starch or protein biosynthesis. The *sh2* mutant affects the larger of the two subunits that comprise the enzyme ADP-glucose pyrophosphorylase (Tsai and Nelson 1966, Bhave et al. 1990) and results in gross modification of starch metabolism in the maize kernel. The *bt2* mutant is responsible for affecting the production of a smaller subunit of ADP-glucose pyrophosphorylase (Dickinson and Preiss 1969) and also results in shrunken, brittle seed. The *sh1* mutant is responsible for causing a lesion in the enzyme sucrose synthase (Chourey and Nelson 1976) and results in a severely shrunken kernel. The *o2* mutant reduces the zein protein fraction in the endosperm and results in an increase in the relative content of lysine (Mertz et al. 1964).

Highly significant effects on kernel protein and starch concentration were detected by hybridizing a cDNA clone of *Sh2* to DNA extracted from the F_3 families. Six other marker loci on chromosome arm 3L were significantly associated with protein and starch concentration as well as 1 marker on 3S and 1 marker in the chromosome 3 centromere region. These marker loci may be showing significance because they are linked to the QTL associated with *Sh2*. Two loci located near *Sh2*, NPI 432 and UMC 63 had R^2 values similar to *Sh2*, indicating that they explained similar proportions of the phenotypic variation. Map distances calculated by other workers (Coe 1992) show a 22 cM distance between NPI 432 and *Sh2*, while UMC 63 and *Sh2* map to the same position on 3L. These findings suggest that the *Sh2* locus may be directly responsible for alteration of the protein and starch concentration in the maize kernel. Alternatively, the high level of significance detected for *Sh2* may signal the presence of QTL with large effects on protein and starch concentration in close proximity to the *Sh2* locus on chromosome arm 3L. Other loci involved in starch and protein metabolism may therefore be present in this chromosomal region. No significant effects on kernel

Table 2. RFLP locus, chromosomal location, mean squares, and orthogonal contrasts of significant marker loci from the analysis of variance for protein and starch concentration

RFLP	Chromosome arm	Contrast[a]	*df*	Protein concentration	Starch concentration
NPI 262	1S		8	ns	27.5*
NPI 447	1L		5	10.8*	ns
NPI 269	2S	Linear	1	ns	ns
		Quadratic	1	70.4**	152.2**
UMC 125	2L		3	18.8**	44.0*
NPI 249	3S	Linear	1	ns	ns
		Quadratic	1	45.1**	121.8**
UMC 10	3CE	Linear	1	24.7*	70.2*
		Quadratic	1	ns	ns
UMC 17	3L	Linear	1	31.3**	130.3**
		Quadratic	1	ns	ns
UMC 16	3L	Linear	1	48.9**	171.1**
		Quadratic	1	ns	ns
NPI 432	3L	Linear	1	61.9**	202.3**
		Quadratic	1	ns	ns
P100080	3L	Linear	1	34.3**	ns
		Quadratic	1	ns	ns
Sh2	3L	Linear	1	46.5**	157.5**
		Quadratic	1	ns	ns
UMC 63	3L		4	16.4**	54.4**
UMC 96	3L	Linear	1	ns	51.2*
		Quadratic	1	ns	ns
UMC 72A	5S	Linear	1	33.1**	129.1**
		Quadratic	1	ns	ns
UMC 90	5S	Linear	1	15.1*	53.2*
		Quadratic	1	21.3*	ns
UMC 43	5S	Linear	1	30.1**	ns
		Quadratic	1	ns	ns
UMC 1	5S	Linear	1	ns	ns
		Quadratic	1	33.4**	153.8**
NPI 237	5L	Linear	1	ns	ns
		Quadratic	1	39.6**	107.2**
UMC 104B	5L	Linear	1	29.5**	ns
		Quadratic	1	ns	ns
UMC 21	6L	Linear	1	ns	ns
		Quadratic	1	ns	66.2*
P200746	7L	Linear	1	ns	84.3*
		Quadratic	1	23.6*	64.3*
NPI 455	7L	Linear	1	25.6**	ns
		Quadratic	1	ns	ns
BNL 8.39	7L	Linear	1	ns	ns
		Quadratic	1	40.9**	108.1**
NPI 443	9L	Linear	1	ns	ns
		Quadratic	1	33.7**	116.9**

* and ** denote significance based upon the *F*-test at the $P<0.05$ and $P<0.01$ levels, respectively. ns, not significant

[a] Orthogonal contrasts performed on all significant marker loci with 2 RFLP alleles. Linear effect contrasts mean difference between homozygous marker classes (additive gene action) and quadratic effect contrasts mean of homozygous marker classes with heterozygous marker class (dominant gene action)

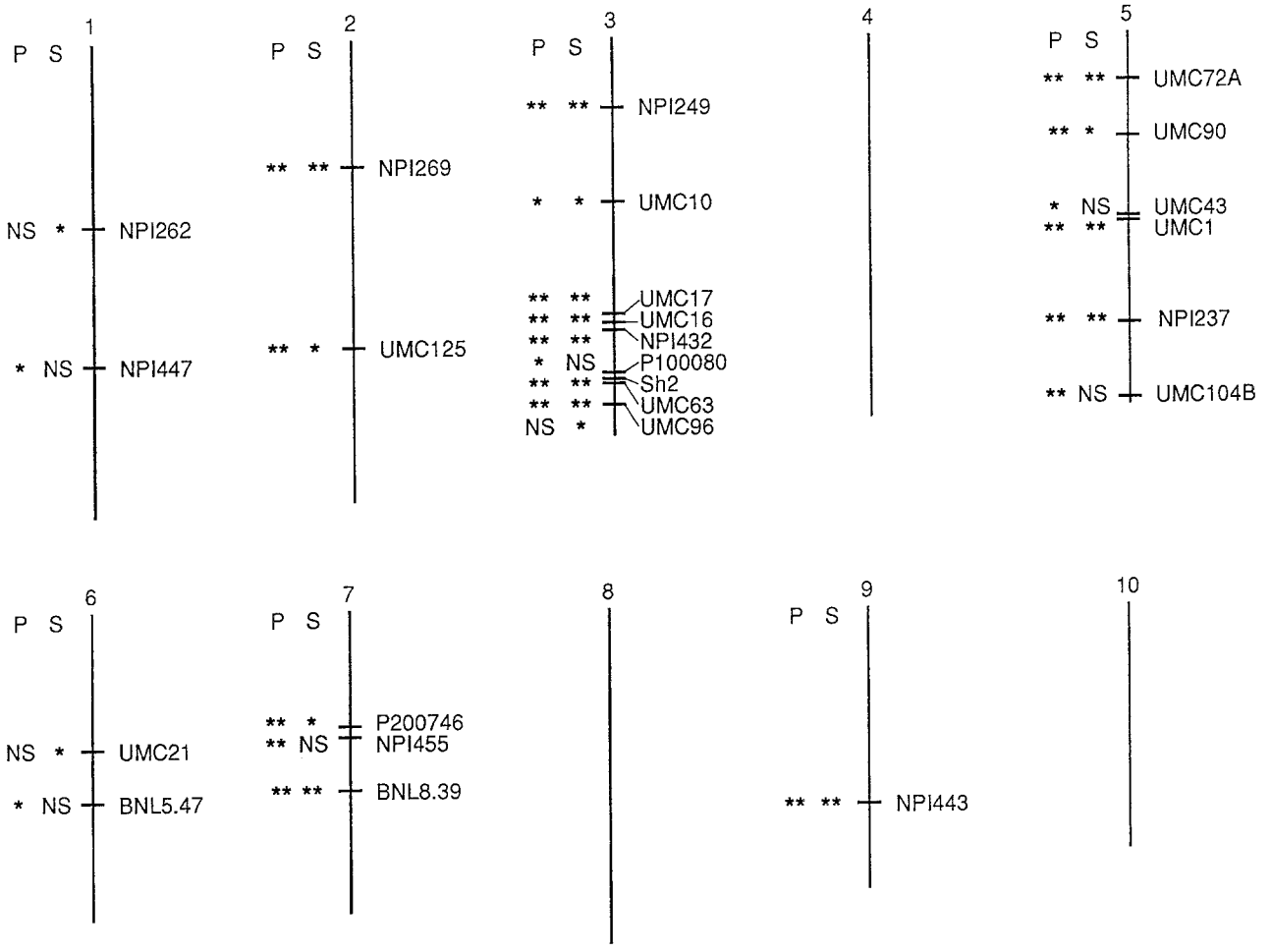

Fig. 2. Approximate chromosomal location of significant marker locus-QTL associations for protein and starch concentration detected in Illinois Long Term Selection Experiment strains. Significance levels for QTL associations are given under the headings *P* and *S*, corresponding to protein and starch. * and ** denote significance based upon *F*-test at the 0.05 and 0.01 levels, respectively. All map positions derived from Coe (1992) and unpublished data from Native Plants, Inc

protein or starch concentration were detected with the *Sh1*, *Bt2*, and *o2* probes.

The enzyme produced at the *Sh2* locus, (ATP: a *D*-glucose-1-phosphate adenyltransferase, EC 2.7.7.27) catalyzes the synthesis of ADP-glucose and pyrophosphate from ATP and glucose 1-phosphate, and may constitute one of the principal steps in starch biosynthesis in plants (Tsai and Nelson 1966, Preiss 1982). Mutant alleles at the *Sh2* locus in the homozygous recessive condition inhibit starch synthesis in the maize endosperm. The lack of significant association between the *Bt2* locus and kernel protein and starch concentration suggests that the subunit of ADP-glucose pyrophosphorylase controlled by the *Sh2* locus, and not the subunit produced by the *Bt2* locus, may be involved in the genetic control of the traits. The possible role of the *Sh2* locus is intriguing in that continuous progress in the ILTSE may have been the result of selection for favorable combinations of alleles at major loci, such as *Sh2*.

The finding of a major QTL associated with a qualitative locus such as *Sh2* is consistent with the findings of Edwards et al. (1992) who reported significant associations between marker loci and QTL for plant height located in close proximity to genes with known effects of plant stature, such as dwarfing and brachytic genes. Results at the *Sh2* locus and at RFLP loci in close proximity to *Sh2* may lend support to the existence of clusters of linked functionally related loci (first proposed by Pontecorvo 1950), such as the *T* locus in mice (Bennett 1975). These data may reflect the presence of a cluster of loci involved in starch and/or protein biosynthesis. Alternatively, the large significant effect detected on 3L may be a reflection of the magnitude of significance of a single QTL located at *Sh2*, as proposed and discussed by Thompson (1975) and Robertson (1985). They postulated that minor wild-type allelic variants at major qualitative mutant loci may be responsible for substantial amounts of quantitative genetic variation. We evaluated

Table 3. Marker class means of 100 F_3 families at selected significant marker loci for protein and starch concentration

Marker class	Locus means					
	Marker locus					
	NPI 432	UMC 72A	NPI 237	UMC 1	P200746	NPI 443
Protein concentration	% Protein					
AA	11.9	12.1	13.6	13.1	13.7	12.4
AB	12.9	12.8	12.4	12.3	12.7	13.3
BB	13.6	13.5	13.1	13.4	12.7	12.5
LSD 0.05 AA/AB[a]	0.66	0.72	0.61	0.59	0.58	1.05
LSD 0.05 AA/BB[b]	0.73	0.77	0.69	0.59	0.75	0.26
LSD 0.05 AB/BB[b]	0.58	0.56	0.65	0.66	0.07	0.51
Starch concentration	% Starch					
AA	65.3	64.8	62.6	63.2	62.1	64.8
AB	63.5	63.8	64.3	64.8	63.9	62.7
BB	62.3	62.2	63.1	62.4	64.2	64.2
LSD 0.05 AA/AB	1.51	1.65	1.39	1.35	1.26	2.49
LSD 0.05 AA/BB	1.68	1.76	0.94	1.46	1.64	1.26
LSD 0.05 AB/BB	1.34	1.28	1.49	1.50	0.65	1.20

[a] LSD value for comparison of AA versus AB marker class means
[b] LSD value for comparison of AA versus BB marker class means
[c] LSD value for comparison of AB versus BB marker class means

structural genes at loci with well-characterized qualitative mutants known to affect kernel chemical and physical properties. The success of this approach in the case of *Sh2* suggests that wild-type alleles at loci with known qualitative mutants may exhibit quantitative effects. Future efforts to detect QTL in genetically well-characterized plant and animal species may focus on using a series of structural genes affecting closely related traits in an effort to further explore the possible relation of qualitative mutants and QTL.

Gene action present in QTL-marker locus associations

Significant additive and dominant gene effects were detected at marker loci associated with kernel protein and starch concentration (Table 2). Of the 18 marker loci with 2 RFLP alleles exhibiting a significant association with protein concentration, 9 showed only additive gene effects, while 7 showed only dominant gene effects. Two loci had both significant additive and significant dominant gene effects. Of the 16 marker loci exhibiting a significant association with starch concentrations 8 showed additive gene effects, while 7 showed only dominant gene effects. One locus showed significant effects of both types. Dominance effects at loci significant for protein and starch concentration were generally associated with high starch and low protein (or vice versa) for the heterozygous marker class as compared with the two homozygous classes (e.g. NPI 237 and UMC 1, Table 3). However, the dominant gene effect at P200746 (chromosome 7) was characterized by superior performance of the 'AA' marker class as comapred with the 'AB' and 'BB' marker classes.

The type of gene action and direction of dominance at markers in different genomic regions varied for both protein and starch concentration. Similar trends were observed by Edwards et al. (1992), who found that the type of gene action and direction of dominance varied for a single trait in different regions of the genome. The only consistent pattern of gene action occurred on 3L, where all significant marker loci with 2 RFLP alleles (UMC 17, P100080, NPI 432, *Sh2*, UMC 16, UMC 96) exhibited additive gene effects. Additive gene effects at many loci in the ILTSE strains has been suggested to be responsible for progress from selection through the accumulation of favorable alleles at many loci. Quantitative genetic analysis of cycle 70 from the ILTSE revealed that additive genetic variance was considerably larger than dominance genetic variance for protein concentration (J. W. Dudley, unpublished data). Loci linked to QTL with strong additive effects may be relatively more important in the genetic control of protein and starch concentration than loci with dominance effects.

Multiple regression models predictive of protein and starch concentration

Significant marker loci were used to develop a predictive model for the genetic control of protein and starch concentration. To represent the effects of individual chromo-

Table 4. Multiple regression models and percentage of total phenotypic variation explained by marker loci for protein and starch concentration in the Illinois Long Term Selection strains

Model	Source of variation	*df*	Mean square	P>f
A. Protein: individual loci $R^2=0.647$				
	Sh2	2	4.4	0.0035
	NPI 269	2	15.1	0.0001
	NPI 447	4	11.1	0.0001
	NPI 237	2	2.4	0.0448
	BNL 8.39	2	11.6	0.0001
	NPI 443	2	4.2	0.0039
B. Protein: individual loci and epistatic interactions $R^2=0.838$				
	Sh2	2	2.7	0.0026
	NPI 269	2	19.7	0.0001
	NPI 447	4	4.8	0.0001
	NPI 237	2	3.5	0.0042
	BNL 8.39	2	4.3	0.0001
	NPI 443	2	7.8	0.0001
	Sh2*NPI 443	3	3.8	0.0001
	NPI 443*BNL 8.39a	4	2.9	0.0001
	Sh2*BNL 8.39a	4	4.5	0.0001
	Sh2*NPI 269	4	1.8	0.0040
	NPI 269*NPI 237	4	4.2	0.0001
C. Starch: individual loci $R^2=0.663$				
	Sh2	2	14.3	0.0061
	NPI 269	2	22.2	0.0005
	NPI 262	8	12.4	0.0001
	NPI 237	2	15.9	0.0036
	P 200746	2	19.2	0.0011
	NPI 443	2	29.0	0.0001
	UMC 21	2	8.5	0.0447
D. Starch: individual loci and epistatic interactions $R^2=0.780$				
	Sh2	2	9.2	0.0074
	NPI 269	2	45.7	0.0001
	NPI 262	8	22.0	0.0001
	NPI 237	2	25.6	0.0001
	NPI 443	2	50.7	0.0001
	UMC 21	2	17.9	0.0001
	NPI 269*NPI 237	4	30.1	0.0001

somes, a single representative marker locus was chosen from each of six chromosomes associated with significant effects for protein concentration and each of seven chromosomes associated with significant effects for starch concentration. Initially, these marker loci were entered into a multiple linear regression model in a stepwise fashion. Marker loci remained in the model if the locus was significant at $P<0.05$. Six individual marker loci on six chromosomes (1, 2, 3, 5, 7, 9) accounted for 64.7% (R^2) of the total variation for protein concentration in this experiment (Table 4). Following examination of this main effects model, two-way epistatic combinations between these marker loci were tested. Interactions were only performed between marker loci on different chromosomes. Pairwise comparisons between unlinked loci were used in an effort to eliminate confounding effects of co-segregating marker loci. Five significant two-way interactions were detected between these marker loci and added to the model. The resulting model containing both main effects and significant epistatic combinations accounted for 83.8% /R^2) of the variation for kernel protein concentration. Similarly, 7 individual loci explained 66.3% (R^2) of the variation in starch, and the inclusion of a single two-way interaction increased the R^2 for the model to 78.0%.

These R^2 values are higher than have been reported for similar multiple regression models describing QTL variation in maize for traits such as grain yield and plant height (Edwards et al. 1992) and tolerance to low-phosphorous stress (Reiter et al. 1991). Keim et al. (1990), however, found 5 independent RFLP marker loci and their interactions to account for 71% of the total phenotypic variation for hard seededness in soybean. These workers made use of an F_2 population developed from an interspecific cross in the genus *Glycine* between parents that differed widely for hard seededness.

Number of effective factors

Gain from selection for quantitative traits is usually considered to be due to the accumulation of favorable combinations of alleles at many loci. The numbers of effective factors separating IHP and ILP for protein at generation 76 were estimated to be 122 (Dudley 1977). An estimate of 28 effective factors was obtained using a modification of Wright's formula (Mather and Jinks 1977) applied to the F_3 family data obtained in this experiment. The F_3 family values are likely underestimates because of bias due to linkage disequlibrium. In an earlier study (J. W. Dudley, unpublished data), estimates of effective factors from a Design III (Comstock and Robinson 1948) study of the cross of generation 70 of IHP and ILP were 15 for the F_2 and 45 for the F_6 produced by random mating the F_2 generation four times. Estimates of effective factors from this study and from the unpublished data suggest that the estimates obtained by Dudley (1977) may have been overestimates. Our finding of three main clusters of markers significantly associated with QTL and the detection of 22 significant marker-QTL associations for protein is in closer agreement with the effective factor estimates from the design III study and from the F_3 families in this study than with Dudley's original estimates. However, comparisons of number of significant marker-QTL associations with numbers of effective factors need to be made with caution. Where clusters of significant marker-QTL associations exist, it is not possible to determine, without further experimentation, whether these associations are the result of the linkage of several markers with the same QTL or of the existence of a number of QTL located in the same chromosomal region. Random mating of the F_2 for some generations followed by re-estimation of marker-QTL associations should provide information on this question.

Given the rather large agronomic differences between 'Burr's White' and much of the modern Corn Belt germ plasm, the possibility exists that identification of some marker loci linked to genes controlling protein and starch concentration in this study may be unique to the ILTSE strains. Loci controlling the chemical and physical properties of the kernel may, however, be relatively constant across different *Zea mays* genetic material. Investigation of the marker locus-QTL associations reported here will be performed in other genetic backgrounds to provide information more relevant to marker-facilitated selection for protein and/or starch concentration in elite corn germ plasm.

Acknowledgements. The authors would like to thank Yaakov Tadmor for enthusiastic technical assistance and William Bridges for stimulating discussions on statistical analysis. This investigation was supported by the Illinois Agricultural Experiment Station and a grant from the Midwest Plant Biotechnology Consortium, of which Pioneer Hi-Bred International and Northrup King are corporate sponsors.

References

Bae JM, Giroux M, Hannah LC (1990) Cloning and characterization of the *Brittle-2* gene of maize. Maydica 35:317–322

Bennett D (1975) The T-locus of the mouse. Cell 6:441–454

Bhave MR, Lawrence S, Barton C, Hannah LC (1990) Identification and moelcular characterization of *Shrunken-2* cDNA clones of maize. Plant Cell 2:581–588

Chourey PS, Nelson OE (1976) The enzymatic deficiency conditioned by the *shrunken-1* mutations in maize. Biochem Genet 14:1041–1055

Coe E (1992) Gene list and working maps. Maize Genet Coop Newsl 66:127–159

Comstock RE, Robinson HF (1948) Estimates of average dominance of genes. p. 494–516. In: Gowen JW (ed) Heterosis. Iowa State College Press, Ames, Iowa

Dickinson DB, Preiss J (1969) ADP-glucose pyrophosphorylase from maize endosperm. Arch Biochem Biophys 130:119–128

Dudley JW (1977) Seventy-six generations of selection for oil and protein percentage in maize. In: Pollack E, Kempthorne O, Bailey TB Jr (eds) Proc Int Conf Quant Genet. Iowa State University Press, Ames, Iowa, pp 459–474

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

Edwards MD, Stuber CW, Wendel JF (1987) Molecular-marker-facilitated investigations of quantitative trait loci in maize. 1. Numbers, genomic distribution and types of gene action. Genetics 116:113–125

Edwards MD, Helentjaris T, Wright S, Stuber CW (1992) Molecular-marker-facilitated investigation of quantitative trait loci in maize. 4. Analysis based on genome saturation with isozyme and restriction fragment length polymorphism markers. Theor Appl Genet 83:765–774

Feinberg A, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Glover DV (1988) Corn protein and starch-genetics, breeding, and value in foods and feeds. Am Seed Trade Assoc Proc 43:106–130

Hoisington D (1991) University of Missouri RFLP procedures manual. University of Missouri, Columbia, Mo.

Hopkins CG (1899) Improvement in the chemical composition of the corn kernel. Ill Agric Exp Stn Bull 55:205–240

Hymowitz T, Dudley JW, Collins FI, Brown CM (1974) Estimations of protein and oil concentration in corn, soybean, and oat seed by near-infrared light reflectance. Crop Sci 14:167–170

Keim P, Diers BW, Shoemaker RC (1990) Genetic analysis of soybean hard seededness with molecular markers. Theor Appl Genet 79:465–469

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199

Lee M (1991) Iowa State RFLP procedures manual. Iowa State University Press, Ames, Iowa

Mather K, Jinks J (1977) Introduction to biometrical genetics. Cornell University press, New York

Mertz ET, Bates LS, Nelson OE (1964) Mutant gene that changes protein composition and increases lysine content in maize endosperm. Science 145:279–280

Osborn TC, Alexander DC, Fobes JF (1987) Identification of restriction fragment length polymorphisms linked to genes controlling soluble solids content in tomato fruit. Theor Appl Genet 73:350–356

Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors, using a complete linkage map of restriction fragment length polymorphisms. Nature 355:721–726

Pontecorvo G (1950) New fields in the biochemical genetics of micro-organisms. Biochem Soc Symp Vol 4. Cambridge, England, pp 40–50

Preiss J (1982) Regulation of the biosynthesis and degradation of starch. Annu Rev Plant Physiol 33:431–454

Reiter RS, Coors JG, Sussman MR, Gabelman WH (1991) Genetic analysis of tolerance to low-phosphorous stress in maize using restriction fragment length polymorphisms. Theor Appl Genet 82:561–568

Robertson DS (1985) A possible technique for isolating genic DNA for quantitative traits in plants. J Theor Biol 117:1–10

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA 81:8014–8018

Schmidt RJ, Burr FA, Auckermann MJ, Burr B (1990) Maize regulatory gene *opaque-2* encodes a protein with a "leucine-zipper" motif that binds to zein DNA. Proc Natl Acad Sci 87:46–50

Sheldon E, Ferl R, Federoff N, Hannah LC (1983) Isolation and analysis of a genomic clone encoding sucrose synthetase in maize: evidence for two introns in *sh*. Mol Gen Genet 190:421–426

Sughroue and Rocheford (1993) Restriction fragment length polymorphism among Illinois long term selection strains. Theor Appl Genet (in press)

Thompson JN (1975) Quantitative variation and gene number. Nature 258:665–668

Tsai CY, Nelson OE (1966) Starch deficient maize mutants lacking adenosine diphosphate glucose pyrophosphorylase activity. Science 151:341–343

Young ND, Zamir D, Ganal MW, Tanksley SD (1988) Use of isogenic lines and simultaneous probing to identify DNA markers tightly linked to the *Tm-2a* gene in tomato. Genetics 120:579–585