Quantitative Trait Loci Underlying Gene Product Variation: A Novel Perspective for Analyzing Regulation of Genome Expression

C. Damerval, A. Maurice, J. M. Josse and D. de Vienne

Station de Génétique Végétale, INRA/UPS/CNRS-URA 1492, La Ferme du Moulon, 91190 Gif-sur-Yvette, France

Manuscript received September 8, 1993
Accepted for publication January 28, 1994

ABSTRACT

A methodology to dissect the genetic architecture of quantitative variation of numerous gene products simultaneously is proposed. For each individual of a segregating progeny, proteins extracted from a given organ are separated using two-dimensional electrophoresis, and their amounts are estimated with a computer-assisted system for spot quantification. Provided a complete genetic map is available, statistical procedures allow determination of the number, effects and chromosomal locations of factors controlling the amounts of individual proteins. This approach was applied to anonymous proteins of etiolated coleoptiles of maize, in an F_2 progeny between two distant lines. The genetic map included both restriction fragment length polymorphism and protein markers. Minimum estimates of one to five unlinked regulatory factors were found for 42 of the 72 proteins analyzed, with a large diversity of effects. Dominance and epistasis interactions were involved in the control of 38% and 14% of the 72 proteins, respectively. Such a methodology might help understanding the architecture of regulatory networks and the possible adaptive or phenotypic significance of the polymorphism of the genes involved.

POLYMORPHISM in gene regulation has long been thought to be an important basis for morphological and adaptative changes (WALLACE 1963; WILSON *et al.* 1977; PAIGEN 1989). Data from several organisms, *e.g.*, mice, maize, fishes and especially Drosophila, show evidence for a large diversity within and between species for genes controlling the activity/amount of various enzymes and the timing and the tissue specificity of enzyme expression (DICKINSON 1980; ALLENDORF *et al.* 1982; BENNETZEN *et al.* 1984; SKADSEN *et al.* 1990; LAURIE *et al.* 1990; FANG and BRENNAN 1992). Within species, there may be several polymorphic loci distributed throughout the genome that affect the expression of a given structural gene (POWELL and LICHTENFELS 1979; LAURIE-AHLBERG *et al.* 1980; WILTON *et al.* 1982; LAURIE-AHLBERG and BEWLEY 1983; COLAS DES FRANCS and THIELLEMENT 1985; BUSH and PAIGEN 1992). In some cases, adaptative significance of these polymorphisms has been suggested on the basis of relationships between levels of enzyme activity/amount and components of fitness (AYALA and MCDONALD 1980; ALLENDORF *et al.* 1983; MATSUO and YAMAZAKI 1984; KLARENBERG *et al.* 1987). In maize, divergence between inbred lines for individual protein amounts appeared to be related to the hybrid performances for agronomical traits (LEONARDI *et al.* 1991).

However, it generally appeared somewhat difficult to reconcile the high levels of regulatory polymorphism observed with its possible selective and phenotypic value (BUSH and PAIGEN 1992). A rapidly expanding body of evidence bearing on the molecular mechanisms of gene regulation shows that several *trans*-acting factors are needed for efficient transcription of any given gene and that multiple interactions can produce composite regulators with distinct activities depending on the mix of active factors in a given cell (*e.g.*, RUPPERT *et al.* 1990, MINER and YAMAMOTO 1991). Relying on such combinatorial principles, DICKINSON (1988) demonstrated that a high degree of polymorphism is to be expected as a consequence of interactions in the network, independently from the action of natural selection. Other regulatory models involving cascades (PINNEY *et al.* 1988) may be expected to have similar consequences.

Prior to analyzing the polymorphism in regulatory networks, it was important to develop a methodology to dissect the genetic architecture of quantitative variations of numerous gene products simultaneously. Considering the amounts of various proteins as quantitative characters, various fundamental questions about the control of gene expression can be addressed. (i) Is it possible to give an estimate of the number of factors controlling the amount of any given protein? (ii) Are intralocus (dominance) and/or interlocus (epistasis) interactions common phenomena in gene regulation? (iii) Is it possible to give a genetic basis to coexpression of proteins? Inferences concerning the number of genes that affect a quantitative trait have been attempted using various statistical methods (CASTLE 1921; PANSE 1940; WRIGHT 1968; JINKS and TOWEY 1976; COMSTOCK and ENFIELD 1981; MATHER and JINKS 1982; ZENG *et al.* 1990). However, these procedures are sensitive to deviations from strong hypotheses, such as equality of gene effects, absence of linkage, and all increasing alleles in one parent. Dense

TABLE 1

Map symbols, function and sources of the cDNAs used as RFLP probes

Symbol	Function	Source
SC1	B anthocyanin regulator	CHANDLER *et al.* (1989)
SC15	Pollen-specific cDNA	HANSON *et al.* (1989)
SC19	Knotted	VOLLBRECHT *et al.* (1991)
SC39	C1	PAZ-ARES *et al.* (1987)
SC53	Ribosomal protein S11	LEBRUN and FREYSSINET (1991)
SC56	α-Tubuline-1	MONTOLIU *et al.* (1989)
SC54B45	Ribosomal protein L7	LEBRUN and FREYSSINET (1991)
SC61B13	ADPG pyrophosphorylase (leaf)	PRIOUL *et al.* (1994)
SC66SS1	Sh1 sucrose synthase	SHELDON *et al.* (1983)

genetic maps obtained using restriction fragment length polymorphism (RFLP) and other markers make it possible to examine this question in a more direct way. Taking advantage of linkage disequilibrium with marker loci, mendelian factors underlying quantitative variation (quantitative trait loci, QTLs) can be mapped throughout the genome. Their effects can be estimated in terms of proportion of the variance of the trait they account for, or in terms of allelic substitution effects (BECKMAN and SOLLER 1983; PATERSON *et al.* 1988; KEIM *et al.* 1990; EDWARDS *et al.* 1992). The dominance/recessivity and epistatic interactions can also be studied (DOEBLEY *et al.* 1990; BEAVIS *et al.* 1991; DE VICENTE and TANKSLEY 1993). Results obtained in various species support a model for quantitative variation wherein effects of individual factors range from essentially qualitative down to vanishingly small. Moreover, the increasing alleles are often distributed among parents.

In this study, we used the procedures for QTL mapping to analyze the genetic determinism of the quantitative variation of 72 anonymous proteins separated by high resolution two-dimensional polyacrylamide gel electrophoresis (2-D PAGE, O'FARRELL 1975) in an F_2 population of maize. The amounts of individual proteins were assessed through the integrated optical density of spots. A genetic linkage map constructed with both RFLP and protein markers provided by the 2-D PAGE was used to locate regulatory factors that we called protein quantity loci (PQLs). For 35% of the proteins analyzed, at least 2 PQLs were found, including epistatic factors, and up to 12 PQLs could be observed for a single protein. Since no functional criterion was imposed for the choice of the gene products under study, it appeared that the multifactorial control of protein variation is a general feature of genome expression.

MATERIAL AND METHODS

Plant material: Two genetically distant maize lines were used: *F2*, a flint line from the Institut National de la Recherche Agronomique (France), subsequently referred to as *Lc* (to avoid possible confusion with the F_2 progeny), and *Io*, an American dent line from the Iodent group. Eighty-five kernels of the F_2 progeny harvested on three F_1 plants (female *Io* × male *Lc*) and kernels of each parental line were allowed to germinate in the dark at 24° for 8 days. For protein analysis, the etiolated coleoptiles were taken off and immediately frozen in liquid nitrogen. The F_2 and parental seedlings were then planted out and allowed to resume growth in a greenhouse before being transferred in the field at Saint Martin-de-Hinx (France). Pieces of leaves of adult plants were harvested for RFLP analysis.

Nucleic acid manipulation: Genomic DNA was extracted following BURR *et al.* (1988). A 7.5-μg sample of DNA was digested by *Eco*RI or *Hin*dIII, loaded onto a 0.8% (w/v) agarose gel with 1 × TPE buffer and electrophoresed overnight. After shaking the gel for 10 min in 0.25 M HCl, the DNA was transferred overnight onto a nylon membrane (Positive filters, Appligene) in 0.4 M NaOH. Filters were then neutralized with 2 × SSC for at least 20 min. Probe labeling was done according to FEINBERG and VOGELSTEIN's (1983) procedure, using 50 μCi [^{32}P]dCTP, and hybridizations were done as described by CHURCH and GILBERT (1984). The probes used were chosen from the core map (GARDINER *et al.* 1993) or corresponded to cDNAs (Table 1), and were amplified using PCR.

Protein electrophoresis and quantification: Denaturing protein extraction and two-dimensional electrophoresis were as in DAMERVAL *et al.* (1986) and LEONARDI *et al.* (1991). The silver staining was slightly modified from DAMERVAL *et al.* (1987), following BURSTIN *et al.* (1993). The Kepler software package was used to analyze the 2-D gels (Large Scale Biology Corp., Rockville, Maryland). The various steps of the quantification procedure were as in DAMERVAL and DE VIENNE (1993). Briefly, the gels were digitized using 256 gray levels with a spatial resolution of 100 μM using an Eikonix scanner, which generates 2048 × 2048 pixel images. The optical density range was 0.9. For each protein spot in each 2-D gel, intensity was expressed as a "volume," computed by fitting the integrated OD of the spot to a double gaussian model.

For each spot, the relationship between volume and total protein amount was examined in the two parental lines, using a range of five loadings (5, 15, 25, 35 and 45 μl of extract corresponding to about 15 to 135 μg of total proteins). Two protein extracts, each obtained from a mix of about 100 coleoptiles, were made for each line and were twice run for each of the five loadings.

For each of 60 F_2 individuals, two to four 2-D gels were run that constituted replicates. Replicates taken from the loading range analysis were used to characterize the parental lines. About 40 μg of protein were loaded for each 2-D gel.

Genetic map: Two sources of markers were used to construct the map from 85 F_2 individuals: (i) 76 RFLP probes selected from a set of 116 probes for their codominant inheritance and (ii) monogenic position shift of proteins, as revealed by 2-D PAGE: two spots, with similar aspect and close to each other, are mutually exclusive in the parental lines, are both present in F_1 hybrids, and segregate codominantly in the

F_2s as allelic products (all individuals display at least one spot). Forty-two such protein loci were found.

Pairwise and multipoint linkage analyses were performed using Mapmaker 2.0 (LANDER *et al.* 1987). Linkage groups were constructed with a lodscore threshold of 4.0 and a maximum recombination value of 0.3. The map distances were calculated using KOSAMBI's (1944) function.

The equality of allelic frequencies was tested with a χ^2 test (1 d.f.), and genotypic frequencies were compared to $p^2{:}2pq{:}q^2$ frequencies (p from the previous analysis), using a χ^2 test with 1 d.f.

Scaling procedure for quantitative protein data: Experimental variations made it necessary to scale the 2-D gel data. First, for each loading of the extracts of the two genotypes *Lc* and *Io*, the spot volumes were scaled using principal component analysis to minimize variations between gels (BURSTIN *et al.* 1993). For each spot, the scaled volume was then plotted against the protein loading. The linearity of the response curves was tested using classical F statistics, and the parameters of the curve (intercept and slope values) could be estimated for 113 spots. Twenty-six of these spots, distributed all over the gel surface, were chosen as references because (i) they did not differ significantly in intensity between *Lc*, *Io* and their F_1 over the loading range and (ii) the linear model explained at least 60% of the variation in spot volume. Using the population of reference spots, an effective loading q_j of each 2-D gel j in the experiment was estimated as follows:

$$q_j = \sum_i [(v_{ij} - b_i)/a_i]/26$$

where v_{ij} was the volume of the ith reference spot in gel j, and a_i and b_i the slope and intercept, respectively, of the curve for the ith reference spot. Then the volume of any spot k in gel j could be calculated as:

$$vc_{kj} = a_k(q - q_j) + v_{kj}$$

where vc_{kj} and v_{kj} were the corrected and raw volume, respectively, of spot k in gel j, a_k the slope of the response curve of the kth spot (calculated as the common estimate of the slopes obtained for spot k in each parental line), and q an arbitrary protein loading value, identical for all gels. These corrected volumes are therefore measures of protein amounts expressed in arbitrary units and corrected for gel-to-gel variation.

Quantitative analysis: Fifty-seven common spots and 15 pairs of protein position shift were selected for quantitative analysis in the parental lines and in 60 F_2 individuals (identity of spot position was checked using coelectrophoresis (1:1, v:v) of parental extracts). The occurrence of quantitative variation between parents was assessed using one-way ANOVA with the genotype (two levels, *Lc* and *Io*) as the factor. One-way ANOVAs with the F_2 genotype as the factor and the spot volume as the dependent variable showed significant genotype effect ($P \leq 0.05$) for 88% of the spots ($P \leq 0.001$ for 55% of them), indicating small values of residual variance. The mean volumes were computed for each spot in every F_2 genotype, and were used for the subsequent analyses.

Detection of PQLs: PQLs were searched for the 72 proteins quantitatively analyzed in the progeny. For the 15 proteins displaying a position shift, the two members of a pair were considered as allelic forms of the same protein, and the volumes of the two spots present in heterozygous genotypes were added (see DISCUSSION).

The volumes were normally distributed for 56 spots, and a $\log_{10}$ transformation normalized the distribution for 10 additional spots. Mapmaker/QTL software (version 2.0) was used to map the PQLs in intervals between markers (LANDER and BOTSTEIN 1989) for these proteins. A PQL was retained when it was detected with a lodscore threshold of 2.4. The lodscore value indicates how much more probable (in a decimal logarithm scale) the data are to have arisen assuming the presence of a PQL than assuming its absence. The threshold was chosen from theoretical considerations of LANDER and BOTSTEIN (1989), given that maize has 10 chromosomes, a genome length between 1500 and 2000 cM, and that the mean spacing between markers was about 18 cM in our map (expressed in Haldane distance, which is the metric used by Mapmaker/QTL). For a given trait, this threshold corresponded to a global type I error of 0.05.

One-way ANOVAs [PROC GLM routine in SAS (SAS Institute 1988)] with genotype at the marker locus as the factor (three levels) and mean volume as the dependent variable were used to detect PQLs for spots with non-normally distributed volumes, because ANOVA is robust regarding deviation from normality. The type I error was 0.0005. Assuming all markers independent, the probability of false positive occurrence would be about 0.05 for one trait.

ANOVAs were also performed for the 66 proteins with normally distributed volumes, and the significance was examined as related to lodscore values.

PQL effects: For several proteins, PQLs were detected by clusters of linked markers when using ANOVAs. We considered that these markers detected the same PQL, and the most significant association of the cluster was used to calculate the genetic parameter values.

The proportion of the total variation in spot volume explained by each PQL was recorded as an R^2 value (R^2 = ratio of the sum of squares explained by the marker locus to the total sum of squares). When two or more PQLs were detected for a polypeptide, the cumulative R^2 was computed using Mapmaker/QTL, or the arithmetic sum of the R^2s from ANOVAs.

For each PQL, the linearity test between the spot volume and the genotype at the marker locus was used to test dominance ($P \leq 0.05$). Differences between the mean volumes of the three genotypes at the marker locus were used to test overdominance (t-test, $P \leq 0.02$). The genetic models available in Mapmaker/QTL were also used to test the genetic parameters. When a discrepancy occurred between the two methods (six cases), the ANOVA results were preferred except when the PQL was located distantly from the flanking markers (two cases).

Effect of the allelic substitution on the protein amount was calculated in standard deviation units, as

$$(H - L)/SD$$

where H (respectively L) was the most (respectively least) intense homozygote mean volume, and SD the standard deviation of the spot volume in the F_2 progeny.

Epistasis was tested using two-way ANOVAs with interaction, with markers as the factors with three levels each. Due to the high number of analyses performed (5886 for each protein), a type I error of 10^{-5} was chosen. No statistical test was attempted to compare the nine two-locus genotypes due to the small size of some genotypic classes. The surfaces defined by the spot volume and the genotype at the two marker loci were visually inspected.

RESULTS

Genetic map: High levels of polymorphism were found between the two lines *Lc* and *Io*. One hundred and forty-three RFLP probes were tested with the restriction enzymes *Eco*RI and *Hin*dIII, 116 of which (81%) revealed polymorphism with at least one enzyme.

Forty-two pairs of monogenic protein position shift were found.

The 118 markers with codominant inheritance (76 RFLP and 42 protein markers) were used to construct the genetic map. One hundred and nine markers were arranged by Mapmaker in 10 linkage groups that were assigned to the 10 chromosomes of maize using markers previously located in the core map. Six RFLP markers remained unmapped. Thirty-nine of the 42 protein position shifts were mapped at chromosomal positions hereafter referred to as PSL for position shift loci, that were interspersed with the RFLP loci (Figure 1). As compared to the maize core map (GARDINER *et al.* 1993), at least three PSLs (PSL 44 on chromosome *1*, PSL29 on chromosome *6* and PSL23 on chromosome 7) marked regions without RFLP markers. More interestingly, PSL 48 allowed chromosome *10* to be continued about 30 cM beyond UMC44A, the most distal probe of the core map, and 20 cM beyond SC53, a ribosomal protein probe. The number of deviations from the expected 1:1 allelic frequency was only three, which is in the range expected by chance at the 5% level. Deviations from the expected p^2:$2pq$:q^2 distribution were observed for 14.9% of the markers (10% if the linked markers were not considered), that corresponded to excess (chromosomes *1*, *3*, *5*, *6* and *8*) or deficit (chromosomes *2*, *4* and *7*) of the heterozygous class. The meaning of these deviations was unclear, because no significant residual heterozygosity for the markers involved was found in a recombinant inbred line population derived from the same cross (data not shown).

The length of the map was 1490 cM using the Kosambi distance (1810 cM in Haldane distance), which is in the range of estimated genome lengths for maize (BEAVIS and GRANT 1991). The marker locations were consistent with the previously published maps, except for UMC63, located on chromosome *2* in the core map but on the long arm of chromosome *5* in our map, and UMC23 and UMC14, mapped on chromosome *1* in the core map but on chromosome *2* in our map. This result could be explained by duplicated sequences in chromosomes *1* and *2* (HELENTJARIS *et al.* 1988) and by lack of polymorphism of one of the duplicated loci. The pairs UMC139-UMC55 (chromosome *2*), UMC102-UMC10 (chromosome *3*), UMC166-UMC27 (chromosome *5*) appeared in reverse order as compared to the core map.

Genetic determinism of protein amount variation: *Common polypeptides:* For the analysis of genetic determinism of protein amounts, 57 spots common to the two lines were retained on the following criteria: (i) response curve of the spot volume against the protein loading linear in most of the loading range and (ii) reliable quantification both in parents and progeny. Since a linear relationship existed between spot volume and protein amount, the volumes could be considered as measures of protein amounts.

Volume differences between the two lines were found significant at $P \leq 0.05$ for 49 of them, and were slightly below the threshold otherwise. Most of the spots displayed normally distributed volumes in the progeny, supporting the hypothesis of oligo- or polygenic control of gene product amount. For about half the proteins, the distribution was not centered on the mid-parental value, which was inconsistent with a purely additive control of the amount for these polypeptides. The shift in intensity relative to the mid-parental value was more often toward higher than to lower intensities (14 cases *vs.* 5, respectively).

Using the 109 mapped markers, 52 PQLs were identified that controlled the variation in amount of 33 common polypeptides.

Position shift variants: Among the 42 pairs of position shift variants selected for segregation analysis, 15 were also quantitatively analyzed (Figure 2). For the remaining pairs, one or both spots were in a crowded area or did not follow the criteria previously defined.

Between-line differences for the volumes of the allelic spots appeared in 10 pairs out of 15. In the F_2 progeny, two situations were observed. (i) Allele-specific expression (seven pairs): the spot volumes differed between the homozygous classes, and the heterozygotes had the two spots with unequal volumes, each one displaying a dose effect (Figure 3A) and (ii) no allele-specific expression (three pairs): the homozygous classes had similar mean spot volumes, that was only slightly lower than the most intense parental spot; the heterozygotes displayed the two spots with equal volumes, with a dose effect (Figure 3B). *Trans*-acting unlinked factor with dominant action could explained that observation. Since a dose effect appeared in the heterozygotes, the factor might be limitant. The absence of some genotypic classes (spots of low intensity in homozygous and heterozygous genotypes were expected according to the hypothesis) would be due to the small size of the progeny.

For the five pairs with no between-line difference, a dose effect was observed in the progeny (Figure 3C).

For PQL detection, the two proteins of each pair were considered as forms of the same protein, and 18 PQLs were found for 9 of these 15 proteins (Table 2). Six of the PQLs were coincident with the PSL, four of them possibly accounting for the observed allele-specific expression. No PQL was found that could support the hypothesis of a distant *trans*-acting regulator with dominant effect that was put forward for three proteins. PQLs were found for three of the five pairs which did not display difference for spot volume between the parental lines.

Genetic parameters of the PQLs: When considering together the 15 proteins displaying position shift and the 57 proteins common to both parental lines, 70 PQLs affecting 42 proteins were found. Forty of them were detected with a lodscore equal to or greater than 3.0,

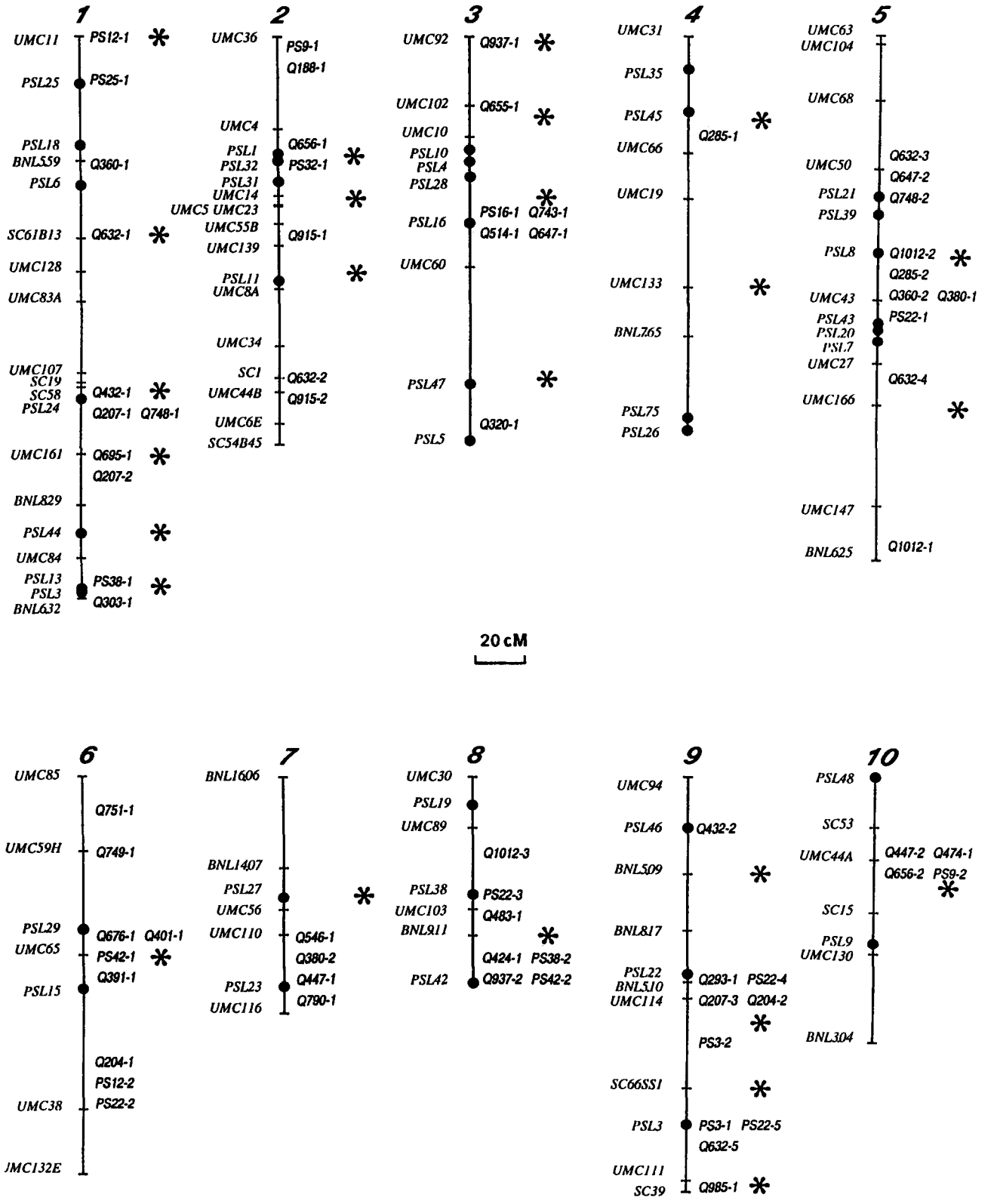

FIGURE 1.—Genetic linkage map derived from *Io* × *Lc* F_2 population of 85 plants. Symbols to left of chromosomes are RFLP and protein markers. RFLP markers are indicated with dashes whereas protein markers are indicated with black circles. Symbols to right of chromosomes represent approximate positions of significant PQLs detected for 9 proteins displaying position shift (PS) and 37 common proteins (Q) in the F_2 population (see Table 2). Asterisks point to regions involved in epistatic controls.

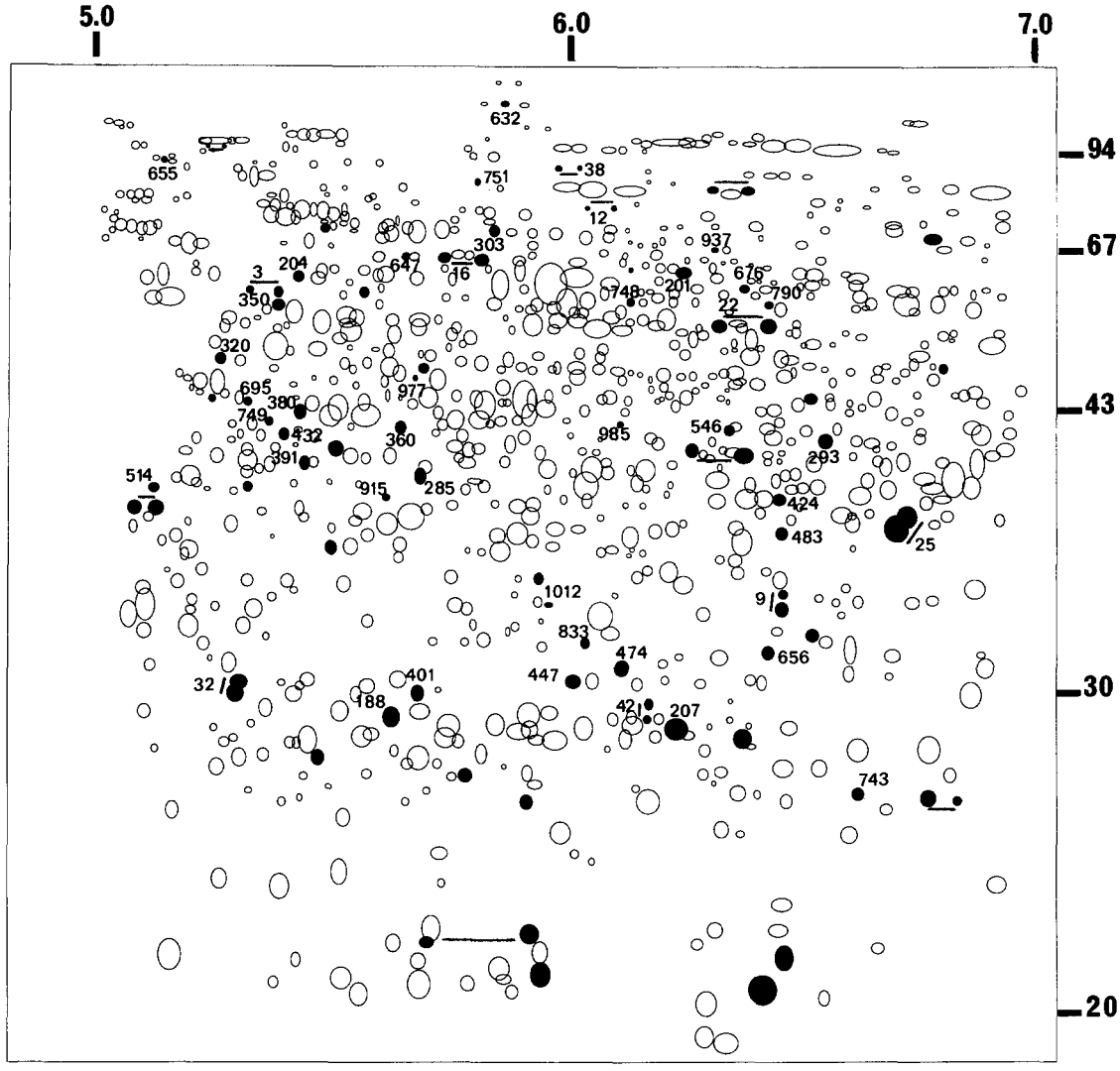

FIGURE 2.—Schematic 2-D protein pattern of the 966 spots scored in *Io* and *Lc* lines. Spots in black were quantitatively analyzed in the F_2 progeny. The two members of the 15 pairs of position shift variants are joined by a solid line. Figures indicate the position shift variants (PS in Table 2 and Figure 1) and the proteins common to both parental lines (Q in Table 2 and Figure 1) for which PQLs have been detected. Horizontal scale: isoelectric points; vertical scale: molecular masses.

and most of them were detected by ANOVA with $P \leq 0.01$ (Table 2). They were located throughout the 10 linkage groups and did not seem to be preferentially close to protein rather than to RFLP markers (Figure 1). Thirty polypeptides had no PQL detected, while 20 had at least two PQLs, the maximum number being five (two proteins) (Figure 4). Different chromosomes were involved in the control of most of these 20 proteins. For 11 of them, the alleles increasing the amount came from both parental lines, resulting in transgressive genotypes in the progeny.

Additivity was observed for 33 PQLs and dominance for 35 PQLs. Two PQLs with (pseudo)overdominant effects were detected that affected the expression of two proteins with position shift. Dominance for high amount was three times more frequent than dominance for low amount.

The effects of the PQLs, measured as the proportion of total variation they explained (R^2), ranged from 14% (lower statistical limit) to 67% (Figure 5A). Cumulative R^2 values ranged from 37% to 90%, with values above 60% for 13 proteins. The effect of the allelic substitution in standard deviation units ranged from 0.1 to 3.3, with the most frequent class around 1.5 (Figure 5B).

Epistasis: Two-way ANOVAs with markers as factors and a type I error of 10^{-5} allowed the detection of 38 significant interactions, while only 4 were expected by chance. These interactions occurred between unlinked markers, except in four cases. The 43 markers involved belong to 26 chromosomic regions, most of which already contained PQLs (Figure 1).

Ten polypeptides were subject to epistatic control, four of which had no previously detected PQL (Table 2). From 2 to 10 epistatic PQLs were detected per polypeptide that belonged to one to seven chromosomes. For most proteins, the epistatic PQLs did not map with the

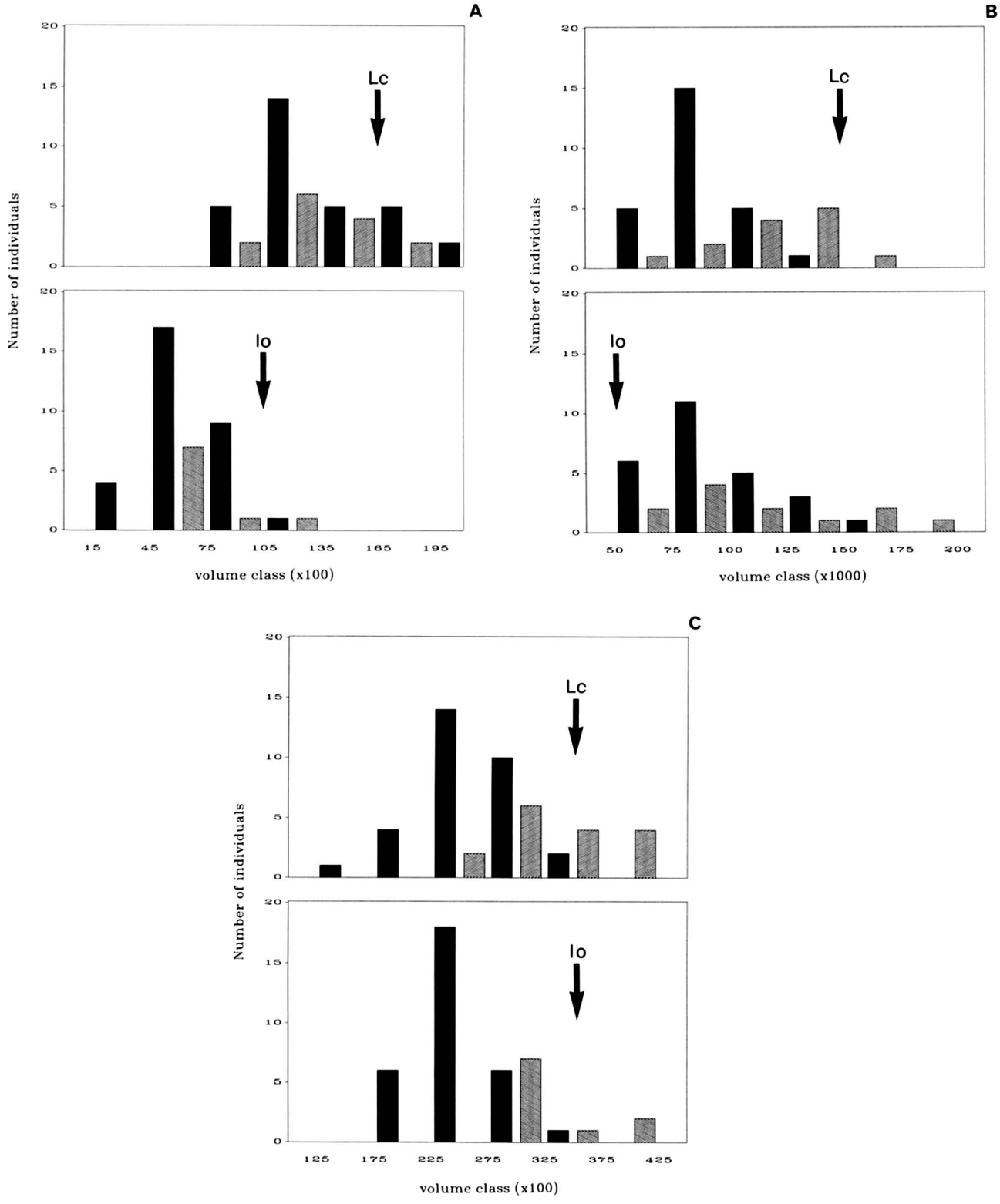

FIGURE 3.—Histogram of spot volumes for three examples of position shift variants. The histograms of the spots of a pair are represented one above the other. Homozygotes for *Io* (respectively *Lc*) alleles are left hatched (respectively right hatched), and heterozygotes are in black. The arrows point to the mean spot volume in *Io* or *Lc* line. (**A**) Allele-specific expression: the allele spots display significant volume difference between the parental lines. In the progeny, the homozygous genotypes have spot volume centered on the parental volume, and a dose effect appears in the heterozygotes. (**B**) No allele-specific expression: the allele spots display significant volume difference between the parental lines. In the progeny, both homozygous genotypes have volumes centered on the highest parental volume, and a dose effect appears in the heterozygotes. (**C**) No between line difference exists for the allele spot volumes. In the progeny, a dose effect appears in the heterozygous genotypes.

TABLE 2

Characteristics of the loci controlling the amount of 46 proteins in the F_2 progeny (*Io* × *Lc*)

Protein	PQL	Chromosome	Allele	Effect	LOD	*P*
PS3^{AS}	pql1*	*9*	*Lc*	d	3.30	$3\cdot10^{-3}$
	pql2	*9*	*Lc*	d	3.90	$2\cdot10^{-4}$
PS9^{AS}	pql1	*2*	*Io*	a	3.26	$9\cdot10^{-4}$
	pql2	*10*	*Lc*	d	2.73	$2\cdot10^{-3}$
PS12	pql1	*1*	*Io*	a	3.03	$4\cdot10^{-4}$
	pql2	*6*	*Io*	a	2.66	$8\cdot10^{-3}$
	pql3-5	*1*		i		
	pql6	*2*		i		
	pql7-8	*3*		i		
	pql9	*5*		i		
	pql10	*6*		i		
	pql11	*9*		i		
	pql12	*10*		i		
PS16	pql1*	*3*	*Lc*	o	6.16	$2\cdot10^{-6}$
PS22^{AS}	pql1	*5*	*Io*	d	3.49	$5\cdot10^{-4}$
	pql2	*6*	*Io*	a	2.99	$2\cdot10^{-3}$
	pql3	*8*	*Io*	a	2.51	$16\cdot10^{-3}$
	pql4*	*9*	*Lc*	d	5.56	$6\cdot10^{-6}$
	pql5	*9*	*Lc*	d	2.54	10^{-3}
PS25^{NAS}	pql1*	*1*	*Lc*	d	2.50	$5\cdot10^{-3}$
PS32^{AS}	pql1*	*2*	*Lc*	o	—	10^{-5}
	pql2	*2*		i		
	pql3-4	*3*		i		
PS38	pql1	*1*	*Io*	d	3.88	$3\cdot10^{-4}$
	pql2	*8*	*Io*	d	2.42	$9\cdot10^{-3}$
PS42^{AS}	pql1	*6*	*Lc*	d	2.63	$33\cdot10^{-3}$
	pql2*	*8*	*Lc*	a	3.22	10^{-3}
	pql3	*1*		i		
	pql4-5	*3*		i		
	pql6-7	*5*		i		
Q188	pql1	*2*	*Io*	d	7.31	$2\cdot10^{-4}$
Q201	pql1	*7*		i		
	pql2	*9*		i		
Q204	pql1	*6*	*Lc*	d	2.50	$11\cdot10^{-3}$
	pql2	*9*	*Io*	a	8.07	$<10^{-7}$
Q207	pql1	*1*	*Lc*	a	9.73	$4\cdot10^{-7}$
	pql2	*1*	*Lc*	a	8.99	$<10^{-7}$
	pql3	*9*	*Io*	d	2.92	$2\cdot10^{-3}$
Q285	pql1	*4*	*Io*	a	11.45	$<10^{-7}$
	pql2	*5*	*Io*	a	3.05	$6\cdot10^{-3}$
Q293	pql1	*9*	*Lc*	a	7.05	$3\cdot10^{-6}$
Q303	pql1	*1*	*Io*	a	2.56	$6\cdot10^{-4}$
Q320	pql1	*3*	*Io*	a	2.58	$4\cdot10^{-3}$
Q350	pql1	*2*		i		
	pql2	*9*		i		
Q360	pql1	*1*	*Lc*	a	3.04	$9\cdot10^{-3}$
	pql2	*5*	*Io*	a	2.50	$6\cdot10^{-3}$
Q380	pql1	*5*	*Io*	d	2.62	$11\cdot10^{-3}$
	pql2	*7*	*Io*	d	2.54	$11\cdot10^{-3}$
Q391	pql1	*6*	*Lc*	d	3.10	$15\cdot10^{-3}$
Q401	pql1	*6*	*Io*	d	2.44	$9\cdot10^{-3}$
Q424	pql1	*8*	*Io*	a	5.29	$6\cdot10^{-5}$
Q432	pql1	*1*	*Io*	d	2.60	$7\cdot10^{-3}$
	pql2	*9*	*Io*	d	2.71	$2\cdot10^{-3}$
Q447	pql1	*7*	*Io*	a	6.09	$2\cdot10^{-6}$
	pql2	*10*	*Lc*	d	3.22	$4\cdot10^{-4}$
Q474	pql1	*10*	*Lc*	d	2.50	$5\cdot10^{-3}$
Q483	pql1	*8*	*Lc*	a	2.63	$5\cdot10^{-3}$
Q514	pql1	*3*	*Io*	d	4.82	$3\cdot10^{-5}$
Q546	pql1	*7*	*Io*	a	2.65	$5\cdot10^{-3}$
Q632	pql1	*1*	*Io*	d	2.53	$4\cdot10^{-3}$
	pql2	*2*	*Lc*	a	3.83	$5\cdot10^{-4}$
	pql3	*5*	*Lc*	a	6.88	$3\cdot10^{-6}$
	pql4	*5*	*Lc*	d	3.17	$4\cdot10^{-3}$
	pql5	*9*	*Io*	d	2.75	$4\cdot10^{-3}$
Q647	pql1	*3*	*Lc*	a	7.52	$4\cdot10^{-7}$
	pql2	*5*	*Io*	d	2.89	$21\cdot10^{-3}$
Q655	pql1	*3*	*Lc*	a	3.06	$4\cdot10^{-3}$
Q656	pql1	*2*	*Io*	a	3.59	$7\cdot10^{-4}$
	pql2	*10*	*Lc*	d	3.20	$6\cdot10^{-4}$
Q676	pql1	*6*	*Lc*	a	4.09	$5\cdot10^{-4}$
Q695	pql1	*1*	*Io*	d	3.25	$7\cdot10^{-3}$
Q743	pql1	*3*	*Io*	a	3.03	10^{-3}
Q748	pql1	*1*	*Lc*	a	4.12	$5\cdot10^{-4}$
	pql2	*5*	*Io*	d	2.80	$2\cdot10^{-3}$
Q749	pql1	*6*	*Io*	a	3.49	$9\cdot10^{-4}$
Q751	pql1	*6*	*Io*	a	3.78	$9\cdot10^{-4}$
Q790	pql1	*7*	*Io*	a	3.20	$3\cdot10^{-3}$
Q833	pql1-3	*1*		i		
	pql4	*3*		i		
	pql5	*4*		i		
	pql6-7	*7*		i		
	pql8	*10*		i		
Q915	pql1	*2*	*Lc*	a	2.46	$25\cdot10^{-3}$
	pql2	*2*	*Lc*	d	2.91	$6\cdot10^{-3}$
	pql3	*2*		i		
	pql4-6	*4*		i		
Q937	pql1	*3*	*Io*	d	4.03	10^{-4}
	pql2	*8*	*Io*	d	2.69	$2\cdot10^{-3}$
	pql3-4	*9*		i		
Q977	pql1	*1*		i		
	pql2-4	*3*		i		
	pql5	*7*		i		
	pql6	*8*		i		
	pql7	*10*		i		
Q985	pql1	*9*	*Io*	d	—	$8\cdot10^{-5}$
Q1012	pql1	*5*	*Lc*	a	3.14	$8\cdot10^{-4}$
	pql2	*5*	*Io*	d	3.38	$8\cdot10^{-4}$
	pql3	*8*	*Io*	d	4.05	$8\cdot10^{-4}$
	pql4	*2*		i		
	pql5	*7*		i		

Chromosomal location, additivity (a), dominance (d), overdominance (o), epistasis (i), lodscore value (LOD) and type I error for the ANOVAs (*P*) are indicated. *Lc* or *Io* in the column 'Allele' indicate the parental origin of the allele responsible for the highest amount. For Position shift protein (PS) the asterisk indicates a PQL located close to the position shift locus; AS and NAS identify position shift variants with allele-specific and non-allele-specific expression, respectively (see text).

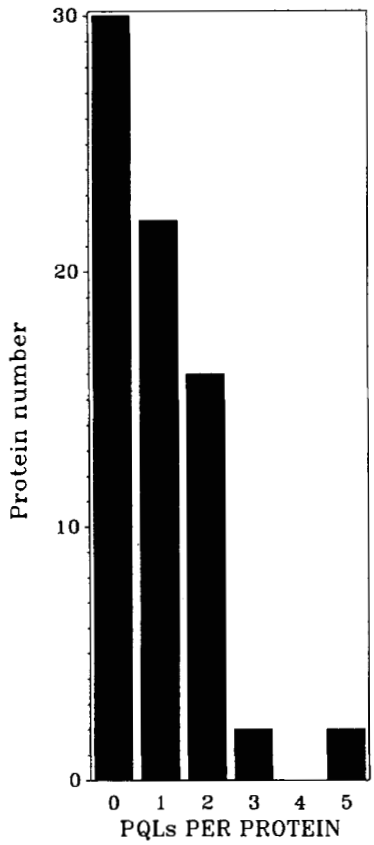

FIGURE 4.—Histogram of the number of PQLs detected per protein.

PQLs. The interactions can be only pairwise or more complex, involving sometimes cascades. The surfaces defined by the mean volumes of each two-locus genotype were of two kinds: (i) with one (more rarely two) genotype(s) strongly different from the other ones (15 surfaces); the outlying genotype had in most cases the highest volume (Figure 6) and (ii) surface of any form (four surfaces).

Coexpressed polypeptides: For the analysis of coexpression of proteins, we looked for PQLs located at similar map position, and we examined correlations between spot volumes.

Thirteen markers detected two to four PQLs that controlled different proteins. Fifty-five percent of the correlations computed between the volumes of the spots having such colocated PQLs were over 0.30 ($P \leq 0.05$) and 15% were over 0.50 ($P \leq 0.001$). As expected, the sign of these correlations was positive when the alleles responsible for the highest amount came from the same parent, negative in the reverse case, except in one case.

Correlation values over 0.50 involving 34 proteins were also found that could not be interpreted in terms of colocated PQLs. However, since for 21 of these proteins no PQL was detected, the hypothesis of undetected colocated PQLs could not be discarded.

DISCUSSION

Control of genetic expression: In maize, we found that the amounts of proteins, assessed through spot volumes, were continuously distributed in an F_2 progeny, and that it was possible to map Mendelian factors (PQLs) involved in their variation. Seventy PQLs were detected for 72 proteins, with a type I error of 0.05 for each protein. With this threshold, less than four false PQLs were expected by chance, *i.e.*, about 20 times less than what was found. If the type I error was more stringent, the power strongly decreased: for example, it fell from 0.74 to 0.49 when the type I error went from 0.05 to 0.01 for a PQL accounting for 15% of the variance of the trait (A. CHARCOSSET, personal communication).

Probably very few PQLs with high effects remained undetected given the density of the map. Since the QTL detection methodologies can locate only genes segregating for alleles with relatively large effects, the PQL numbers, from 1 to 5 for 42 proteins, were minimum estimates. The mean number of PQLs was similar for proteins with position shift polymorphism and for proteins with only quantitative variation. The number and relative effects of PQLs dramatically differed among the proteins. For example, the PQL detected for protein Q985 explained as much variation as the five PQLs of PS22 (*ca.* 70%). For the 30 proteins without PQLs in the present analysis, PQLs with minor effects could be found when using ANOVA with $P \leq 0.05$ (not shown).

When PQLs with interactive effects were taken into account, up to twelve chromosome regions could be involved in the control of a single protein. Twenty-five proteins had at least two PQLs and/or epistatic PQLs. At least in such cases a control by genetic elements distant from the structural gene can be concluded. The control of the amount of the proteins which display position variation could provide more precise information about the location of the PQLs relative to the structural gene, if the genetic basis of the shift could be elucidated. The two members of the pair could be (i) products of allelic structural genes, (ii) products resulting from the action of a polymorphic, modifier locus and (iii) products of two closely linked genes. The recombination fraction between two such genes should be below 3%, since no recombinants were detected in 85 individual progeny ($P \leq 0.05$). Actually in a collection of 21 genetically distant maize lines, both members of the pairs were never found together, which makes the latter hypothesis rather unlikely (J. BURSTIN, personal communication). The position shifts of proteins on 2-D gels are likely to have the same bases as the differences in electrophoretic mobility of enzymes, namely in many cases, differences in the structural gene (GELBART *et al.* 1976; LUSIS and PAIGEN 1978; THATCHER 1980; KELLY and FREELING 1980; LANGLEY *et al.* 1988). Molecular techniques (sequencing

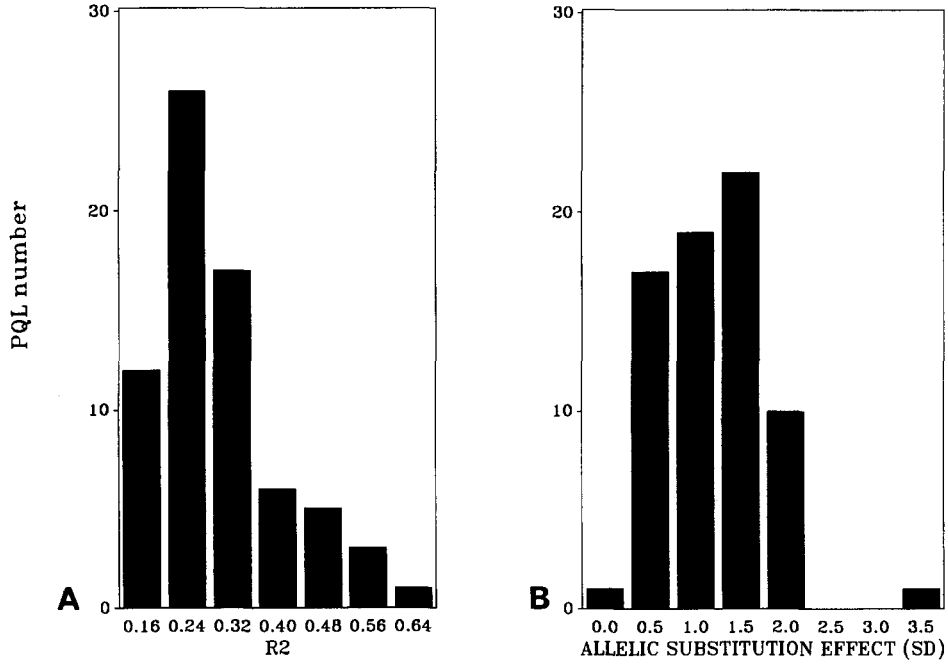

FIGURE 5.—Effects of PQLs. (**A**) Distribution of variation in spot volume explained by individual PQLs ($R2$). (**B**) Distribution of allele substitution effects in standard deviation units.

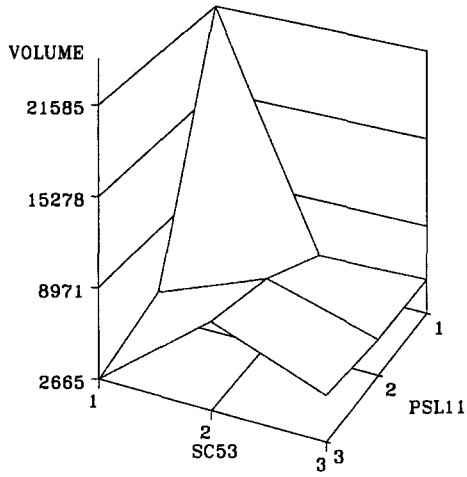

FIGURE 6.—Example of the most commonly observed surface of epistatic interaction between pairs of marker loci. The homozygous genotypes for each locus (x and y axes) are noted **1** and **3**, and the heterozygous genotypes are noted **2**. The spot volume is indicated on the z axis.

alleles, mapping probes of the structural genes) would indicate whether the gene responsible for the shift is indeed the structural gene.

Just as different levels of electrophoretic polymorphism of proteins seemed to exist as related to their structure and/or function (GILLESPIE and KOJIMA 1968; JOHNSON 1974; MCCONKEY 1982; SOMMER *et al.* 1992), categories of complexity in genetic control could exist, as related to the function of the protein or its level of action in the metabolism. Our results do not allow the identification of clear cut categories of complexity of protein amount determinism, but suggest, however, a great diversity of possible controls.

Intra- and interlocus interactions: About half of the PQLs had dominant or (pseudo)overdominant effects, and dominance was involved in the control of 64% of the 42 proteins with detected PQLs. Positive dominance appeared three times more frequently than negative dominance, which is consistent with the asymmetry of nonadditivity already observed in F_1 hybrids (DE VIENNE *et al.* 1988). Epistatic interactions between pairs of loci affecting protein amount concerned 14% of the proteins, and the most frequently observed situation resulted in one genotype with higher amount than the others. Thus, both dominance and epistasis could account for the shift of the amount distribution of some proteins toward a value higher than the midparental value.

Our results concerning epistasis contrast with those obtained from morphological or agronomical traits, where interactions are not commonly detected (EDWARDS *et al.* 1987; PATERSON *et al.* 1991; DE VICENTE and TANKSLEY 1993). This might be due to lower numbers of steps between gene expression and protein phenotypes than between gene expression and composite macroscopic phenotypes. Molecular interactions involving proteins, or proteins and nucleic acid sequences, play a major role in gene transcription (DYNAN and TJIAN 1985; MINER and YAMAMOTO 1991; JONES 1991; KAELIN *et al.* 1991; CONAWAY and CONAWAY 1993). Heteropolymer associating subunits encoded by different genes are documented not only for enzymes [see DE VIENNE and RODOLPHE (1985) for review], but also for transcriptional activators, some of which are known to act through heterodimerization (SMEAL *et al.* 1989; BENEZRA *et al.* 1990; LAMB and MCKNIGHT 1991; PYSH *et al.* 1993). Genetic variability affecting the interacting elements may generate epistasis. Transduction signal systems or regulatory hierarchies can also be sources of epistatic phenomena,

since disruption of one element can impair the functioning of others (ARTAVANIS-TSAKONAS 1988; AVERY and WASSERMAN 1992). In our study, the most frequently observed epistasis situation was highest protein amount in one (more rarely two) genotype. An interlocus heterodimer with only one combination of subunits being fully functional could explain such results. Other systems involving two components (*e.g.*, two different proteins cooperating for the same function, or a *cis*-acting sequence and a *trans*-acting factor) could produce a similar surface of interaction.

Coexpression of proteins: We found partial discrepancy between the set of proteins with colocated PQLs and the set of highly correlated proteins. When a high correlation was not supported by colocated PQLs, it could be that PQLs were undetected. In fact, when the PQLs detected using ANOVA with $P \leq 0.05$ were taken into account, it was possible to find at least one PQL accounting for most of the observed correlations. When no significant correlation was found in spite of colocated PQL(s), the simplest hypothesis was that the largest proportion of the amount variation was explained by unlinked, protein specific PQLs. When proteins shared more than one region with colocated PQLs, different allele effect distributions in the different regions might also account for lack of correlation.

The 70 PQLs were located in about 30 chromosomic regions that may contain several linked genes, or one pleiotropic gene. Analyzing other progenies or populations would help making a decision, since linkage disequilibrium would not be expected to persist across various backgrounds.

Conclusion: The proteins under study were chosen independently of any functional criteria, and therefore probably represented a broad spectrum of gene products. Thus, our results showed that numerous polymorphic elements controlling gene expression dispersed throughout the genome are a common feature of gene regulation, as was suggested from studies on enzyme genes. Numerous molecular processes are involved from gene transcription to protein maturation. Even if the level of action of the PQLs remains to be determined, it appeared that many proteins can tolerate polymorphism simultaneously at more than one step of their biosynthesis. The complexity of genetic control evidenced for 42 anonymous gene products illustrates the large diversity of possible protein phenotypes. A high flexibility of the genetic expression is thus theoretically available, that might be favorable in some developmental context, or in response to environmental changes. This large reservoir of regulatory variation can potentially serve to provide abundant phenotypic variation as the raw material for natural selection. However, probably not all the regulatory variations are relevant to adaptive and morphological variation, and additional physiological studies are required. Techniques are now available to identify and characterize proteins from various cellular compartments resolved by 2-D PAGE, from western blotting to microsequencing or determination of amino acid composition (BAUW *et al.* 1989; RASMUSSEN *et al.* 1991; SHAW 1993).

Our methodology could be used to analyze the genetic control of the amount (or activity) variation of known-function proteins involved in a metabolic pathway. For example, we recently dissected the regulatory relationships between activities of various key enzymes of carbon metabolism in maize (M. CAUSSE, J. P. ROCHER, A. M. HENRY, A. CHARCOSSET, J. L. PRIOUL, and D. DE VIENNE, submitted for publication). Such analyses would help understanding the genetic architecture of composite networks and the possible adaptive significance of the polymorphism levels observed for their components.

We thank M. LE GUILLOUX for expert technical assistance for 2-D gel electrophoresis, and A. LEONARDI, M. CAUSSE and M. ZIVY for valuable comments on the manuscript.

LITERATURE CITED

ALLENDORF, F. W., K. L. KNUDSEN and S. R. PHELPS, 1982 Identification of a gene regulating the tissue expression of a phosphoglucomutase locus in rainbow trout. Genetics **102:** 259–268.

ALLENDORF, F. W., K. L. KNUDSEN and R. F. LEARY, 1983 Adaptive significance of differences in the tissue-specific expression of a phosphoglucomutase gene in rainbow trout. Proc. Natl. Acad. Sci. **80:** 1397–1400.

ARTAVANIS-TSAKONAS, S. 1988 The molecular biology of the *notch* locus and the fine tuning of differentiation in *Drosophila*. Trends Genet. **4:** 95–100.

AVERY, L., and S. WASSERMAN, 1992 Ordering gene function: the interpretation of epistasis in regulatory hierarchies. Trends Genet. **8:** 312–316.

AYALA, F. J., and J. F. MCDONALD, 1980 Continous variation: possible role of regulatory genes. Genetica **52/53:** 1–15.

BAUW G. J., J. VAN DAMME, M. PUYPE, J. VANDEKERKHOVE, B. GESSER *et al.*, 1989 Protein electroblotting and microsequencing strategies in generating protein data bases from two-dimensional gels. Proc. Natl. Acad. Sc. USA **86:** 7701–7705.

BEAVIS, W. D., and D. GRANT, 1991 A linkage map based on information from four F_2 populations of maize (*Zea mays* L.). Theor. Appl. Genet. **82:** 636–644.

BEAVIS, W. D., D. GRANT, M. ALBERTSEN and R. FINCHER, 1991 Quantitative trait loci for plant height in four maize populations and their association with qualitative genetic loci. Theor. Appl. Genet. **83:** 141–145.

BECKMANN J. S., and M. SOLLER, 1983 Restriction fragment length polymorphisms in genetic improvement: methodologies, mapping and costs. Theor. Appl. Genet. **67:** 35–43.

BENEZRA, R., R. L. DAVIS, D. LOCKSHON, D. L. TURNER and H. WEINTRAUB, 1990 The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. Cell **61:** 49–59.

BENNETZEN, J. L., J. SWANSON, W. C. TAYLOR and M. FREELING, 1984 DNA insertion in the first intron of maize *Adh1* affects message levels: cloning progenitor and mutant *Adh1* alleles. Proc. Natl. Acad. Sci. USA **81:** 4125–4128.

BURR, B., F. A. BURR, K. H. THOMPSON, M. C. ALBERTSON and C. W. STUBER, 1988 Gene mapping using recombinant inbreds in maize. Genetics **118:** 519–526.

BURSTIN, J., M. ZIVY, D. DE VIENNE and C. DAMERVAL, 1993 Analysis of scaling methods to minimize experimental variations in two-dimensional electrophoresis quantitative data. Applications to the comparison of maize inbred lines. Electrophoresis **14:** 1067–1673 .

BUSH R. M., and K. PAIGEN, 1992 Evolution of β-glucuronidase regulation in the genus *Mus*. Evolution **46:** 1–15.

CASTLE, W. E., 1921 An improved method of estimating the number of genetic factors concerned in cases of blending inheritance. Science **54:** 223.

CHANDLER, V. L., J. P. RADICELLA, T. P. ROBBINS, J. CHEN and D. TURKS, 1989 Two regulatory genes of the maize anthocyanin pathway are homologous: isolation of B utilizing R genomic sequences. Plant Cell **1:** 1175–1183.

CHURCH, M. G., and W. GILBERT, 1984 Genomic sequencing. Proc. Natl. Acad. Sc. USA **81:** 1991–1995.

COLAS DES FRANCS, C., and H. THIELLEMENT, 1985 Chromosomal localization of structural genes and regulators in wheat by 2D electrophoresis of ditelosomic lines. Theor. Appl. Genet. **71:** 31–38.

COMSTOCK, R. E., and F. D. ENFIELD, 1981 Gene number estimation when multiplicative genetic effects are assumed–growth in flour beetles and mice. Theor. Appl. Genet. **59:** 373–379.

CONAWAY, R. C., and J. W. CONAWAY, 1993 General initiation factors for RNA polymerase II. Annu. Rev. Biochem. **62:** 161–190.

DAMERVAL, C., and D. DE VIENNE, 1993 Quantification of dominance for proteins pleiotropically affected by *opaque*-2 in maize. Heredity **70:** 38–51.

DAMERVAL, C., D. DE VIENNE, M. ZIVY and H. THIELLEMENT, 1986 Technical improvements in two-dimensional electrophoresis increase the level of genetic variation detected in wheat seedling proteins. Electrophoresis **7:** 52–54.

DAMERVAL, C., M. LE GUILLOUX, J. BLAISONNEAU and D. DE VIENNE, 1987 A simplification of Heukeshoven and Dernick's silver staining of proteins. Electrophoresis **8:** 158–159.

DE VICENTE, W. C., and S. D. TANKSLEY, 1993 QTL analysis of transgressive segregation in an interspecific tomato cross. Genetics **134:** 585–596.

DE VIENNE, D., and F. RODOLPHE, 1985 Biochemical and genetic properties of oligomeric structures: a general approach. J. Theor. Biol. **116:** 527–568.

DE VIENNE, D., A. LEONARDI, C. DAMERVAL, 1988 Genetic aspects of variation of protein amounts in maize and pea. Electrophoresis **9:** 742–750.

DICKINSON, W. J., 1980 Evolution of patterns of gene expression in Hawaiian picture-winged *Drosophila*. J. Mol. Evol. **16:** 73–94.

DICKINSON, W. J., 1988 On the architecture of regulatory systems: evolutionary insights and implications. BioEssays **8:** 204–208.

DOEBLEY, J., A. STEC, J. WENDEL and M. EDWARDS 1990 Genetic and morphological analysis of a maize-teosinte F_2 population: implications for the origin of maize. Proc. Natl. Acad. Sci. USA **87:** 9888–9892.

DYNAN, W. S., and R. TJIAN, 1985 Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. Nature **316:** 774–778.

EDWARDS, M. D., C. W. STUBER and J. F. WENDEL, 1987 Molecular-marker-facilitated investigations of quantitative-trait loci in maize. I. Numbers, genomic distribution, and types of gene action. Genetics **116:** 113–125.

EDWARDS, M. D., T. HELENTJARIS, S. WRIGHT and C. W. STUBER, 1992 Molecular-marker-facilitated investigations 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.

FANG, X.-M., and M. D. BRENNAN, 1992 Multiple *cis*-acting sequences contribute to evolved regulatory variation of Drosphila *Adh* genes. Genetics **131:** 333–343.

FEINBERG, A., and B. VOGELSTEIN, 1983 A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132:** 6–13.

GARDINER, J. M., E. H. COE, S. MELIA-HANCOCK, D. A. HOISINGTON and S. CHAO, 1993 Development of a core RFLP map in maize using an F_2 immortalized population. Genetics **134:** 917–930.

GELBART, W. H., H. MCCARRON and A. CHOVNIK, 1976 Extension of the limits of the structural element in *Drosophila melanogaster*. Genetics **84:** 211–232.

GILLESPIE J. H., and K. KOJIMA, 1968 The degree of polymorphisms in enzymes involved in energy production compared to that in nonspecific enzymes in two *Drosophila ananassae* populations. Proc. Natl. Acad. Sci. USA **61:** 582–585.

HANSON, D. D., D. A. HAMILTON, J. L. TRAVIS, D. M. BASHE and J. P. MASCARENHAS, 1989 Characterization of a pollen-specific cDNA clone from *Zea mays* and its expression. Plant Cell **1:** 173–179.

HELENTJARIS, T., D. WEBER and S. WRIGHT, 1988 Identification of the genomic locations of duplicate nucleotide sequences in maize by analysis of restriction fragment length polymorphisms. Genetics **118:** 353–363.

JINKS, J. L., and P. M. TOWEY, 1976 Estimating the number of genes by genotype assay. Heredity **40:** 397–403.

JOHNSON G. B., 1974 Enzyme polymorphism and metabolism. Science **184:** 28–37.

JONES, N., 1991 Complex inhibitions. Curr. Biol. **1:** 224–226.

KAELIN, W. G., JR., D. C. PALLAS, J. A. DECAPRIO, F. J. KAYE and D. M. LIVINGSTON, 1991 Identification of cellular proteins that can interact specifically with the T/E1A-binding region of the retinoblastoma gene product. Cell **64:** 521–532.

KEIM, P., B. W. DIERS, T. C. OLSON and R. C. SHOEMAKER, 1990 RFLP mapping in soybean: association between marker loci and variation in quantitative traits. Genetics **126:** 735–742.

KELLY, J., and S. FREELING, 1980 Purification of maize alcohol dehydrogenase-1 allozymes and comparison of their tryptic peptides. Biochim. Biophys. Acta **624:** 102–110.

KLARENBERG, A. J., K. SIKKEMA and W. SCARLOO, 1987 Functional significance of regulatory *map* and structural *Amy* variants in *Drosophila melanogaster*. Heredity **58:** 383–389.

KOSAMBI, D. D., 1944 The estimation of map distances from recombination values. Ann. Eugen. **12:** 172–175.

LAMB, P., and S. L. MCKNIGHT, 1991 Diversity and specificity in transcriptional regulation: the benefits of heterotypic dimerization. Trends Biochem. Sci. **16:** 417–422.

LANDER, E. S., and D. BOTSTEIN, 1989 Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics **121:** 185–199.

LANDER, E. S., P. GREEN, J. ABRAHAMSON, A. BARLOW, M. J. DALY *et al.*, 1987 Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics **1:** 174–181.

LANGLEY, C. H., A. E. SHRIMPTON, T. YAMAZAKI, M. MIYASHITA, Y. MATSUO *et al.*, 1988 Naturally occurring variation in the restriction map of the *Amy* region of *Drosophila melanogaster*. Genetics **119:** 619–629.

LAURIE, C. C., E. M. HEATH, J. W. JACOBSON and M. S. THOMPSON, 1990 Genetic basis of the difference in alcohol dehydrogenase expression between *Drosophila melanogaster* and *Drosophila simulans*. Proc Natl. Acad. Sci. USA **87:** 9674–9678.

LAURIE-AHLBERG, C. C., and G. C. BEWLEY, 1983 Naturally occurring genetic variation affecting the expression of *sn*-glycerol-3-phosphate dehydrogenase in *Drosophila melanogaster*. Biochem. Genet. **21:** 943–961.

LAURIE-AHLBERG, C. C., G. MARONI, G. C. BEWLEY, J. C. LUCCHESI and B. S. WEIR, 1980 Quantitative genetic variation of enzyme activities in natural populations of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **77:** 1073–1077.

LEBRUN, M., and G. FREYSSINET, 1991 Nucleotide sequence and characterization of a maize cytoplasmic ribosomal protein S11 cDNA. Plant Mol. Biol. **17:** 265–268.

LEONARDI, A., C. DAMERVAL, Y. HEBERT, A. GALLAIS and D. DE VIENNE, 1991 Association of protein amount polymorphism (PAP) among maize lines with performances of their hybrids. Theor. Appl. Genet. **82:** 552–560.

LUSIS, A. J., and K. PAIGEN, 1978 The large scale isolation of mouse β-glucuronidase and comparison of allozymes. J. Biol. Chem. **253:** 7336–7345.

MATHER, K., and J. L. JINKS, 1982 *Biometrical Genetics*, Ed. 3. Chapman & Hall, London.

MATSUO, Y., and T. YAMAZAKI, 1984 Genetic analysis of natural populations of *Drosophila melanogaster* in Japan. IV. Natural selection on the inducibility, but not on the structural genes, of amylase loci. Genetics **108:** 879–896.

MCCONKEY, E. H., 1982 Molecular evolution, intracellular organization, and the quinary structure of proteins. Proc. Natl. Acad. Sci. USA **79:** 3236–3240.

MINER, J. N., and K. R. YAMAMOTO, 1991 Regulatory crosstalk at composite response elements. Trends Biochem. Sci. **16:** 423–426

MONTOLIU, L., J. RIGAU and P. PUIGDOMENECH, 1989 A tandem of α-tubulin genes preferentially expressed in radicular tissues from *Zea mays*. Plant Mol. Biol. **14:** 1–15.

O'FARRELL, P. H., 1975 High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem. **250:** 4007–4021.

PAIGEN,K., 1989 Experimental approaches to the study of regulatory evolution. Am. Nat. **134:** 440–458.

PANSE, V. G., 1940 Application of genetics to plant breeding. II. The inheritance of quantitative characters and plant breeding. J. Genet. **40:** 283–302.

PATERSON, A. H., E. S. LANDER, J. D. HEWITT, S. PETERSON, S. E. LINCOLN *et al.*, 1988 Resolution of quantitative traits into mendelian factors, using a complete linkage map of restriction fragment length polymorphisms. Nature **335:** 721–726.

PATERSON, A. H., S. DAMON, J. D. HEWITT, D. ZAMIR, H. RABINOWITCH *et al.*, 1991 Mendelian factors underlying quantitative traits in tomato: comparison across species, generations, and environments. Genetics **127:** 181–197.

PAZ-ARES, J., D. GHOSAL, U. WIENAND, P. A. PETERSON and H. SAEDLER, 1987 The regulatory C1 locus of *Zea mays* encodes a protein with homology to myb proto-oncogenes products and with structural similarities to transcriptional activators. EMBO J. **6:** 3553–3558.

PINNEY, D. F., S. H. PEARSON-WHITE, S. F. KONIECZNY, K. E. LATHAM and C. P. EMERSON, JR., 1988 Myogenic lineage determination and differentiation: evidence for a regulatory gene pathway. Cell **53:** 781–793.

POWELL, J. R., and J. M. LICHTENFELS, 1979 Population genetics of *Drosophila* amylase. I. Genetic control of tissue-specific expression in *D. pseudoobscura*. Genetics **92:** 603–612.

PRIOUL, J.-L., E. JEANNETTE, A. REYSS, N. GREGORY, M. GIROUX *et al.*, 1994 Expression of ADPglucose pyrophosphorylase in maize grain and source leaf during grain filling. Plant Physiol. **104:** 179–187.

PYSH, L. D., M. J. AUKERMAN and R. J. SCHMIDT, 1993 OHP1: a maize basic domain/leucine zipper protein that interacts with Opaque2. Plant Cell **5:** 227–236.

RASMUSSEN H. H., J. VAN DAMME, M. PUYPE, B. GESSER, J. E. CELIS *et al.*, 1991 Microsequencing of proteins recorded in human two-dimensional gel protein data bases. Electrophoresis **12:** 873–882.

RUPPERT, S., M. BOSHART, F. X. BOSCH, W. SCHMID, R. E. K. FOURNIER *et al.*, 1990 Two genetically defined *trans*-acting loci coordinately regulate overlapping sets of liver-specific genes. Cell **61:** 895–904.

SAS INSTITUTE, INC., 1988 *SAS Users Guide:Statistics.* SAS Institute, Cary, N.C.

SHAW, G., 1993 Rapid identification of proteins. Proc. Natl. Acad. Sci. USA **90:** 5138–5142.

SKADSEN, R. W., S. M. RUZSA and J. G. SCANDALIOS, 1990 Separate molecular mechanisms regulate *CAT2* gene expression before and after glyoxysome maturation in two genetic lines of maize. Dev. Genet. **11:** 280–288.

SHELDON, E., R. FERL, N. FEDOROFF and L. C. HANNAH, 1983 Isolation and analysis of genomic clone encoding sucrose synthase in maize: evidence for two introns in *Sh1*. Mol. Gen. Genet. **190:** 421–426.

SMEAL, T., P. ANGEL, J. MEEK and M. KARIN, 1989 Different requirements for formation of Jun:Jun and Jun:Fos complexes. Genes Dev. **3:** 2091–2100.

SOMMER, R. J., M. RETZLAFF, K. GOERLICH, K. SANDER and D. TAUTZ, 1992 Evolutionary conservation pattern of zinc-finger domains of *Drosophila* segmentation genes. Proc. Natl. Acad. Sci. USA **89:** 10782–10786.

THATCHER, D. R., 1980 The complete amino acid sequence of alcohol dehydrogenase alleloenzymes (Adh^{N-11}, Adh^{S} and Adh^{UF}) from fruitfly *Drosophila melanogaster*. Biochem. J. **187:** 875–886.

VOLLBRECHT, E., B. VEIT, N. SINHA and S. HAKE, 1991 The developmental gene *knotted-1* is a member of a maize homeobox gene family. Nature, **350:** 241–243.

WALLACE, B., 1963 Genetic diversity, genetic uniformity and heterosis. Can. J. Genet. Cytol. **5:** 239–253.

WILSON, A. C., S. S. CARLSON and T. J. WHITE, 1977 Biochemical evolution. Annu. Rev. Biochem. **46:** 573–639.

WILTON, A. N., C. C. LAURIE-AHLBERG, T. H. EMIGH and J. W. CURTSINGER, 1982 Naturally occurring enzyme activity variation in *Drosophila melanogaster*. II. Relationships among enzymes. Genetics **102:** 207–221.

WRIGHT, S., 1968 *Evolution and the Genetics of Populations, Vol. 1. Genetics and Biometrical Foundations.* University of Chicago Press, Chicago.

ZENG, Z.-B., D., HOULE and C. C. COCKERHAM, 1990 How informative is Wright's estimator of the number of genes affecting a quantitative character? Genetics **126:** 235–247.

Communicating editor: W. G. HILL