Theor Appl Genet (1993) 87:33–37

Quantitative genetic studies of A/B zeins using a new model to test non-allelic interactions

T. P. Bogyo [1], J. W. Paulis [2], J. A. Bietz [2], J. S. C. Smith [3]

[1] Washington State University, Pullman, WA 99164-3144, USA
[2] Food Physical Chemistry Research, National Center for Agricultural Utilization Research, USDA-ARS, Peoria, IL 61604-3999, USA
[3] Department of Technical Support, Pioneer Hi-Bred International, Inc, 7250 NW 62nd Avenue, Johnston, IA 50131, USA

Received: 12 February 1993 / Accepted: 1 March 1993

Abstract. A genetic model developed by Bogyo et al. (1988) for quantitatively inherited triploid endosperm characters (an extension of the well-known Mather-Jinks model) is not well-suited for estimating epistatic interaction effects because it requires the assumption that, in segregating loci, all alleles positively affecting a particular character are in one of the inbred parental lines. To better explain zein inheritance in maize, a new model was developed not relying on this assumption. This model was tested by quantitative analysis of A/B zeins, the predominant prolamin storage proteins of maize, using reversed-phase high-performance liquid chromatography of two inbred lines, their reciprocal F_1 crosses, the F_2 generation, backcrosses, and reciprocal backcrosses to both parent lines. The model required epistatic components to be included for an excellent fit for most protein peaks.

Key words: Zein – *Zea mays* L. – Corn – Genetic linkages – RP-HPLC – Proteins – Prolamins – Inheritance – Quantitation – Model

Introduction

Reversed-phase high-performance liquid chromatography (RP-HPLC) has resolved prolamins of near-isogenic single and double maize mutants (Paulis et al. 1990, 1992) into many components. For example, A/B zeins resolve into 20–24 peaks. Amounts of proteins in 17 of these peaks were significantly affected by individual mutant genes. When two mutant genes were present in a single genotype, many peaks showed significant quantitative epistatic interactions. Since there are at least twice as many A/B zeins known as we can resolve into peaks by RP-HPLC, it is assumed that each peak is affected by several different loci. A quantitative genetic analysis is, therefore, appropriate for the further investigation of the inheritance of A/B zeins. A genetic model for quantitatively inherited endosperm characters was proposed by Bogyo et al. (1988). This model extended the diploid model of Mather (1949) and Mather and Jinks (1971) to accommodate endosperm characters. Unfortunately, the model, which includes non-allelic interaction (epistasis) parameters, can only be used to estimate these parameters under the rather unrealistic assumption that in segregating loci all alleles positively affecting a particular character are in one of the two inbred parent genotypes. A new model, not dependent on this unrealistic assumption, was needed. Such a model is described below.

Main effect model

In triploid tissue, such as the endosperm of members of the grass family, one parent (in grasses, the female) contributes two fused gametes, while a single gamete comes from the other parent. Consequently, in the crosses (including reciprocals) of two inbred lines with different genes at a locus, there are four genotypic values, G's, to be considered, G_{AAA}, G_{AAa}, G_{aaA}, and G_{aaa}. The mean of these values is

$$\mu = \frac{G_{AAA} + G_{AAa} + G_{aaA} + G_{aaa}}{4}.$$

Scientific Paper No. 3 Program in Statistics, Washington State University, Pullman WA 99164. The mention of firm names or trade products does not imply that they are endorsed or recommended by the US Department of Agriculture over other firms or similar products not mentioned

Communicated by A. R. Hallauer

Correspondence to: T. P. Bogyo

The mean of the double dose of A averaged over the single doses of alleles is

$$G_{AA\_} = (G_{AAA} + G_{AAa})/2$$

with a double dose of effect of A of

$$A_A^* = G_{AA\_} - \mu = (G_{AAA} + G_{AAa} - G_{aaA} - G_{aaa})/4$$
$$= -(G_{aa\_} - \mu) = -A_a^* .$$

Correspondingly, we arrive at the single-dose effects of these genes as

$$G_{\_\_A} = (G_{AAA} + G_{aaA})/2$$

$$A_A = (G_{\_\_A} - \mu) = (G_{AAA} + G_{aaA} - G_{AAa} - G_{aaa})/4 = -A_a .$$

The interaction, dominance, between double and single doses is

$$D_{AAA} = G_{AAA} - G_{AA\_} - G_{\_\_A} + \mu$$
$$= (G_{AAA} - G_{AAa} - G_{aaA} + G_{aaa})/4$$
$$= -D_{AAa} = -D_{aaA} = D_{aa}$$

Thus,

$$G_{AAA} = \mu + A_A^* + A_A + D_{AAA} \qquad G_{aaA} = \mu + A_a^* + A_A + D_{aaA}$$

$$G_{AAa} = \mu + A_A^* + A_a + D_{AAa} \qquad G_{aaa} = \mu + A_a^* + A_a + D_{aaa} .$$

Now, moving to a general notation for a cross of a (female) parent i with a (male) parent j, let the G's and effects denote the sums over loci and

$$G_{ij} = \mu + A_i^* + A_j + D_{iij} .$$

The reciprocal is $G_{ji} = \mu + A_j^* + A_i + D_{jji}$. The parents may be denoted as G_{ii} or G_{jj} with

$$G_{ii} = \mu + A_i^* + A_i + D_{iii} .$$

Using the same notation, the selfs, $G_{(ij)(ij)}$ or $G_{(ji)(ji)}$ are equivalent,

$$G_{(ij)(ij)} = \mu + \frac{A_i^*}{2} + \frac{A_j^*}{2} + \frac{A_i}{2} + \frac{A_j}{2} + \frac{D_{iii}}{4} + \frac{D_{jjj}}{4} + \frac{D_{iij}}{4} + \frac{D_{jji}}{4} .$$

For backcrosses and reciprocal backcrosses we have

$$G_{(ij)i} = \mu + \frac{A_i^*}{2} + \frac{A_j^*}{2} + A_i + \frac{D_{iii}}{2} + \frac{D_{jji}}{2}$$

and

$$G_{i(ij)} = \mu + A_i^* + \frac{A_i}{2} + \frac{A_j}{2} + \frac{D_{iii}}{2} + \frac{D_{iij}}{2} .$$

This notation can be extended to further backcrosses. For observed values of the various entries we use P's instead of G's with the same identification and $P_x = G_x + e_x$ where the e's denote experimental errors. Table 1 gives coefficients of expected parameters for triploid genotypes derived from homozygous lines.

Table 1. Coefficients of expected genetic parameters for triploids in generations derived from crosses from two homozygous lines

Generation	μ	A_1^*	A_2^*	A_1	A_2	D_{111}	D_{112}	D_{221}	D_{222}
P_{11}	1	1	0	1	0	1	0	0	0
P_{22}	1	0	1	0	1	0	0	0	1
P_{12}	1	1	0	0	1	0	1	0	0
P_{21}	1	0	1	1	0	0	0	1	0
$P_{(12)(12)}$	1	½	½	½	½	¼	¼	¼	¼
$P_{(12)1}$	1	½	½	1	0	½	0	½	0
$P_{1(12)}$	1	1	0	½	½	½	½	0	0
$P_{(12)2}$	1	½	½	0	1	0	½	0	½
$P_{2(12)}$	1	0	1	½	½	0	0	½	½

Table 2. Coefficients of expected reduced genetic parameters for triploids in generations derived from crosses from two homozygous lines

Generation	μ	A^*	A	D
P_{11}	1	1	1	1
P_{22}	1	−1	−1	1
P_{12}	1	1	−1	−1
P_{21}	1	−1	1	−1
$P_{(12)(12)}$	1	0	0	0
$P_{(12)1}$	1	0	1	0
$P_{1(12)}$	1	1	0	0
$P_{(12)2}$	1	0	−1	0
$P_{2(12)}$	1	−1	0	0

Because of the symmetries in the model, such as $A_i^* = -A_i^*$ and $D_{iii} = D_{jjj} = -D_{iij} = -D_{jji}$, it is possible to collapse Table 1 into one with only four main effect parameters, as in Table 2.

Interaction components – model with epistasis

Using coefficients as in Table 2, we are able to include nonallelic interaction (epistasis) parameters in the model. Nine such parameters are possible: A^*A^*, A^*A, A^*D, AA^*, AA, AD, DA^*, DA, and DD. However, with only nine generations available, one can test at most eight coefficients, since nine would always give a perfect fit. Thus, for purposes of the present study, only four epistatic parameters were chosen: A^*A^*, A^*D, AD, and DD. The coefficients are listed in Table 3.

The analysis

The eight parameters of the model can be estimated from the nine observed means, being means of all peak areas from a particular generation from replicate RP-HPLC analyses. It is customary to first test a model with only main effects, using only the first four columns of the X

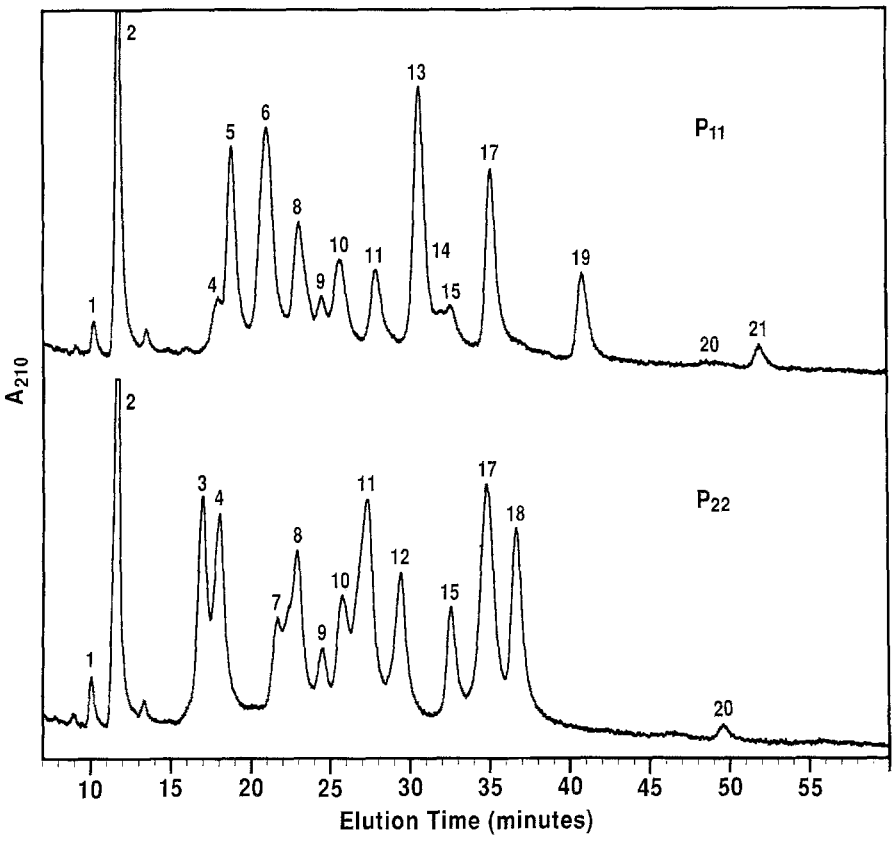

Fig. 1. RP-HPLC analyses of A/B zeins extracted with 70% ethanol + 0.5% NaOAc + 0.2% dithiothreitol from normal inbred maize lines PHG80 (P_{11}) and PHG50 (P_{22}). Peaks referred to in the text are numbered

Table 3. Coefficients of expected genetic parameters including epistasis for triploids in generations derived from crosses from two homozygous lines

Generation	μ	A^*	A	D	A^*A^*	A^*D	AD	DD
P_{11}	1	1	1	1	1	1	1	1
P_{22}	1	−1	−1	1	1	1	−1	1
P_{12}	1	1	−1	−1	1	−1	1	1
P_{21}	1	−1	1	−1	1	1	−1	1
$P_{(12)(12)}$	1	0	0	0	0	0	0	0
$P_{(12)1}$	1	0	1	0	0	0	0	0
$P_{1(12)}$	1	1	0	0	1	0	0	0
$P_{(12)2}$	1	0	−1	0	0	0	0	0
$P_{2(12)}$	1	−1	0	0	1	0	0	0

matrix (in Table 3). A significant departure from a good fit is measured by a $\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$, where the expected value is calculated by substituting observed values with expected values of the parameters. The vector of observations, y, can be represented by the linear model $y = X\beta + e$, where X is the matrix of known coefficients (as in Table 3), β the vector of parameters to be estimated, and e is a vector of random errors with known variance covariance matrix $V = \text{Var}(e)$. A weighted mean squares is used to estimate $\hat{\beta}$, which will minimize the weighted sum of squares $(y - X\beta)' V^{-1} (y - X\beta)$. The estimates are

$$\hat{\beta} = (X' V^{-1} X)^{-1} X' V^{-1} y .$$

The variance covariance matrix of these estimates is

$$\text{Var}(\hat{\beta}) = (X' V^{-1} X)^{-1} .$$

The sum of squares $y' V^{-1} y - \beta' X' V^{-1} y$ under normality assumptions has a chi-square distribution with n-p degrees of freedom, where n is the number of observations (generation means) and p the number of parameters estimated.

Materials and methods

Two inbred maize lines were chosen with very different RP-HPLC profiles (Fig. 1) for A/B zeins (see mean peak sizes in Table 4). Two homozygous inbreds (PHG80 and PHG50, both proprietary products of Pioneer Hi-Bred International Inc), were selected as parents P_1 and P_2 respectively, to make the nine seed populations needed to evaluate the theoretical model. The nine populations were made as follows (parentage as parentheses): (1) P_{11} (PHG80 selfed), (2) P_{22} (PHG50 selfed), (3) P_{12} (PHG80*PHG50), (4) P_{21} (PHG50*PHG80), (5) $P_{(12)(12)}$ (P_{12} selfed), (6) $P_{(12)1}$, (7) $P_{1(12)}$, (8) $P_{(12)2}$, and (9) $P_{2(12)}$. Each population was grown in adjacent nursery rows during 1987 at Johnston, Iowa. F_1 seed needed to make populations five to nine was produced previously from the same seed sources used to make P_{11}, P_{22}, P_{12}, and P_{21}. Since Smith and Smith (1986), studying alcohol-soluble maize inbred proteins from different

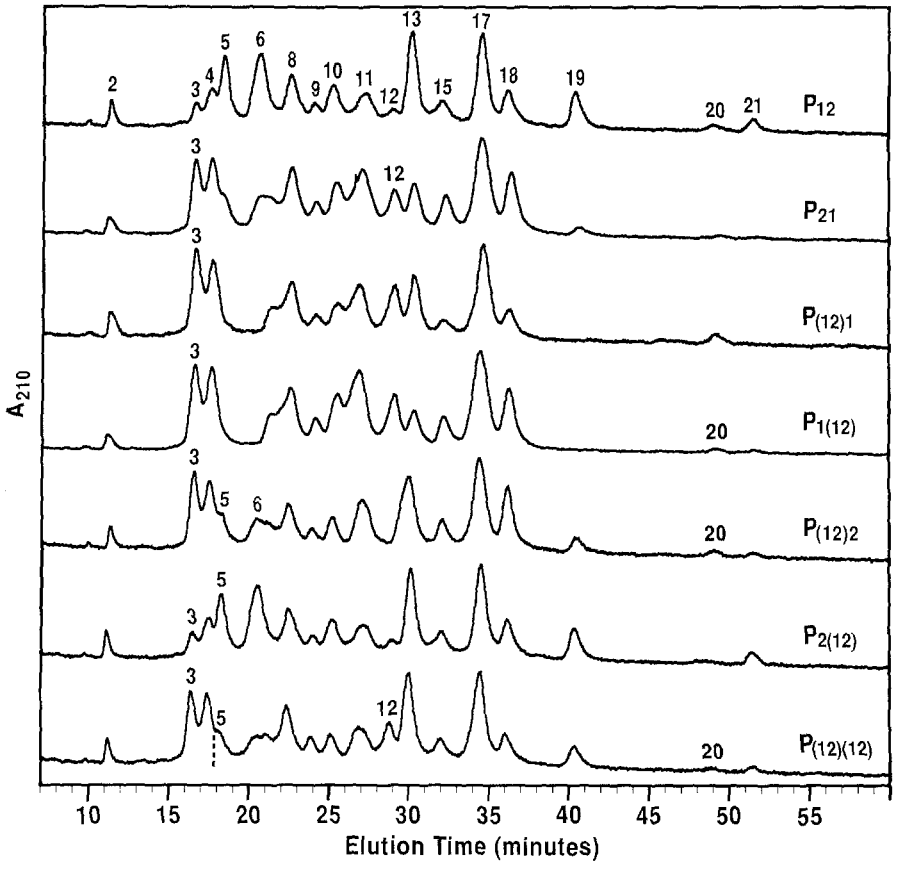

Fig. 2. RP-HPLC analyses of A/B zeins of lines derived from crossing maize inbreds PHG80 (P_{11}) and PHG50 (P_{22}). Pedigrees of lines shown (*top to bottom*, P_{12}, P_{21}, $P_{(12)1}$, $P_{1(12)}$, $P_{(12)2}$, $P_{2(12)}$, and $P_{(12)(12)}$ are described in the text

Table 4. Mean peak areas (× 1,000) of the two inbreds and the ratio of the larger area to the smaller one

Peak no.	P_1	P_2	Ratio
1	29.9	48.1	1.61
2	491.1	713.4	1.45
3	110.2	249.1	2.26
4	78.5	278.1	3.54
5	260.7	8.2	31.69
6	270.8	7.98	33.93
8	262.1	434.0	1.66
9	104.4	156.9	1.50
10	215.0	277.9	1.29
11	188.3	217.1	1.15
12	25.3	399.3	15.76
13	469.6	22.6	20.78
15	104.9	260.6	2.48
17	446.8	629.4	1.41
18	70.8	561.5	7.93
19	145.4	37.9	3.83
20	31.6	48.95	1.55

Table 5. Estimates of the genetic parameters (*in units of peak size* × 1,000) and their standard errors for Peaks 5, 6, and 13

Peak	$\hat{\mu}$	SE	$\hat{A}^*$	SE	$\hat{A}$	SE	$\hat{D}$	SE
5	183.7	5.4	80.2	5.9	56.4	6.2	−38.9	6.6
6	272.5	6.3	162.3	6.6	96.1	6.7	−6.2	6.9
13	278.9	7.3	113.9	8.5	102.2	8.0	−38.3	9.1

environments, have shown that environment has a minimal effect on RP-HPLC elution profiles, no genotype-environment interaction was expected in spite of the fact that not all of the generations were grown during the same year. From each population 10–25 ears were submitted for analysis. Individual seeds from each generation were ground and used for RP-HPLC (Fig. 2) (Paulis et al. 1990, 1992) except for the two parents for which, because of proprietary reasons, three replicates of pooled ground seeds were used.

Twenty kernels from the middle of three ears (6–8 kernels of each ear) from each population were cracked with pliers and ground separately in a Crescent Wig-L-Bug for approximately 1 min. A/B zeins were quantitatively extracted at room temperature by shaking 0.1 g meal of each kernel with a Buchler vortex evaporator for 2 h with 2 ml 70% (v/v) ethanol containing 0.5% (w/v) sodium acetate and 0.2% (v/v) dithiothreitol (Paulis and Bietz 1986). Extracts were stored at room temperature for 2 days prior to RP-HPLC.

For each parental line, RP-HPLC of A/B zeins was performed on three replicate extracts of each of three meals. For each cross, 20 individual kernels, as described above, were examined. Analyses were on a Waters system including a WISP 710B automatic sample injector, Model 660 solvent programmer, and Model 450 variable wavelength detector (210 nm, 0.1 absorbance units full scale). A Vydac phenyl column (250 × 4.6 mm) was used at 30 °C. The solvents were aqueous solutions of acetonitrile (ACN) containing 0.1% (v/v) trifluoroacetic acid (Paulis and Bietz 1986). Following injection of 20-µl samples, a linear 60-min gradient (47.5–57.9% ACN) was followed by 20 min isocratic elution at 57.9% ACN. The flow rate was 1 ml/min. This gradient uses a higher initial ACN concentration than the one previously described (Paulis et al. 1990) and resolves the A/B zeins

Table 6. Estimates of the full model parameters (only significant estimates are listed)

Peak	μ	A^*	A	D	A^*A^*	A^*D	AD	DD
1	7.9***	−0.9		14.3***	2.0***	−3.4***	−4.4***	14.8***
2	50.6***			268***	17.1**	−52.5***	−59***	266.2***
3	181.2***	−177***			78.8***	50.6*	73.3***	−68.2***
4	217.1***	−212***	−37.3*		112.8***	59.3*	91.0***	−125***
8	249.5***	−139***	29.9***		151.3***	−17.0**		−55.6***
9	92.0***	−48.0***			46.4***		12.6*	
12	148.7***	−160***		13.4*	92.8***			−42.5***
15	135.6***	−60***			71.3***			
18	233***	−1,856			137.6***			
19	130.6***	25.7***		−34.2***	52.2***			−56.9***

* $P<0.05$, ** $P<0.01$, *** $P<0.001$
Unit measures × 1,000

better. Raw chromatographic data were stored in a ModComp computer for subsequent analyses (Paulis and Bietz 1986; Paulis et al. 1990).

Results and discussion

The main effects only model gave a very good fit for 3 of 17 peaks analyzed (peaks 5, 6, and 13). Amounts of proteins in these three peaks were very different in the inbred parents: for each, the average peak area of one parent was more than 20 times larger than that of the other (Fig. 1). Estimates of the four genetic parameters, with standard errors, are given in Table 5. All estimates are significantly different from zero except for $\hat{D}$ for Peak 6.

The full model was fitted to the remaining peaks (Figs. 1, 2). Of these, no fit could be obtained from peaks 10, 11, 17, and 20. The homozygous parents P_1 and P_2 showed similar area patterns for these 4 peaks. The remaining 10 peaks (1, 2, 3, 4, 8, 9, 12, 15, 18, and 19) gave a good fit when epistatic components were included in the model. Logarithmic transformation was needed for peaks 8 and 18 to achieve the fit. Estimates of parameters for these peaks are given in Table 6. An interesting pattern of parameter estimates is visible in this table. Significant epistatic parameters are more prevalent in early- than in late-eluting peaks. Of the main effects, the double-dose effect is present in most peaks, except peak 2, and all have negative signs except Peak 19. The epistatic interaction between double-dose additive effects (A^*A^*) is present in all peaks. The other epistatic components, A^*D, AD, and DD, are generally significant only in early peaks.

For future studies, more types of backcrosses will be necessary to estimate the five remaining epistatic parameters (A^*A, AA^*, AA, DA^*, and DA). Crosses such as $P_{((12)1)1}$, $P_{1((12)1)}$, etc. could further help elucidate the pattern of epistasis affecting different peaks. It would also be necessary to choose parents having RP-HPLC patterns that are more different for peaks 10, 11, 17, and 20.

Acknowledgements. We owe special thanks to Dr. C. C. Cockerham, who suggested the genetic model used in this study. We also thank E. Schmalzried and D. Ozuna for excellent technical assistance.

References

Bogyo TP, Lance RCM, Chevalier P, Nilan RA (1988) Genetic models for quantitatively inherited endosperm characters. Heredity 60:61–67

Mather K (1949) Biometrical genetics. Methuen, London

Mather K, Jinks JL (1971) Biometrical genetics. Cornell University Press, Ithaca, N.Y.

Paulis JW, Bietz JA (1986) Separation of alcohol-soluble maize proteins by reversed-phase high-performance liquid chromatography. J Cereal Sci 4:205–216

Paulis JW, Bietz JA, Bogyo TP, Darrah LL, Zuber MS (1990) Expression of alcohol-soluble endosperm proteins in maize single and double mutants. Theor Appl Genet 79:314–320

Paulis JW, Bietz JA, Bogyo TP, Nelsen TC, Darrah LL, Zuber MS (1992) Expression of A/B zeins in single and double maize endosperm mutants. Theor Appl Genet 85:407–414

Smith JSC, Smith OS (1986) Environmental effects on zein chromatograms of maize lines revealed by reverse-phase high-performance liquid chromatography. Theor Appl Genet 71:607–612