Theor Appl Genet (1989) 77:495–500

Multivariate analyses of polypeptide synthesis in developing maize embryos

J. G. Boothe and D. B. Walden

Department of Plant Sciences, University of Western Ontario, London, Ontario, Canada N6A 5B7

Received September 12, 1988; Accepted January 4, 1989
Communicated by P. L. Pfahler

Summary. Variation in polypeptide synthesis was examined in developing maize embryos of two inbred and two hybrid genotypes. Multivariate analyses were used to evaluate the variation among two-dimensional, electrophoretic separations of polypeptides. Several features of the data set were revealed. Similar developmental patterns were exhibited by all genotypes and no evidence was obtained for differential rates of development for inbreds and hybrids. The differential synthesis of two subsets of polypeptides during embryo development was observed. The multivariate methods employed in this study were a valuable aid in interpreting the results from a large and complex data set.

Key words: Multivariate analyses – 2D electrophoresis – Maize – Embryo development – Variation among polypeptides

Introduction

The electrophoretic separation of polypeptides has proved to be a valuable technique for the analysis of gene expression. Numerous studies in plant systems have employed this technique to examine variation occurring as a function of development (Dure et al. 1981; Zivy et al. 1984; Misra and Bewley 1985; Sawney et al. 1985; Sanchez-Martinez et al. 1986) and genotype (Zivy et al. 1983, 1984; Dunbar et al. 1985; Bahrman and Thiellement 1987; Damerval et al. 1987). In both cases, the results consist of qualitative and quantitative changes associated with a number of polypeptide species that may vary either independently or in groups. Tracking changes in individual polypeptides is often arbitrary, and the large number of variables involved hampers the detection of any global patterns which may be present in the data. This is particularly true when two-dimensional (2D) electrophoresis is employed enabling the resolution of hundreds to more than a thousand different polypeptides. The problem is compounded further in studies where large numbers of gels are involved, as is often the case when several developmental stages or genotypes are considered. It would be useful if a method were available which at once summarized the data, identified major sources of variation and revealed any inherent patterns.

Multivariate techniques have been applied recently to investigate genetic affinities among different varieties of rice, based on densitometric tracings of proteins separated on one-dimensional gels (Aliaga-Morell et al. 1987). These techniques have also been used to determine taxonomic relationships between various lines of maize through isozyme allele frequencies (Smith 1984; Bretting et al. 1987; Smith and Smith 1987) and chromatographically and electrophoretically separated proteins (Damerval et al. 1987; Smith and Smith 1987). To assess variability among proteins separated on 2D gels, Anderson et al. (1984) described methods for using densitometric measurement of intensity and software estimation of spot position in conjunction with multivariate analysis. In the present study, we employ a technique developed by Fewster and Walden (1987), which does not require densitometric measurement of intensity nor spot identification, yet captures sufficient information from the distribution of polypeptides on a 2D gel to permit the data to be analyzed by principal coordinate (PCoA) and concentration analyses. Our objective was to identify patterns of variation in polypeptide synthesis among developing maize embryos of different genotypes.

Materials and methods

Plant material

Two inbred cultivars of maize (OH43 and M14) and their reciprocal hybrids (OH43/M14 and M14/OH43) were selected for this study. All plants were grown under field conditions in our maize nursery in London, Ontario, Canada in 1985 and 1986. Controlled pollinations were made on the same date for each genotype in order to exclude differences in development due to environmental effects. Ears were collected at approximately 5-day intervals between 15 and 52 days after pollination (DAP).

Radio-labelling, extraction and electrophoresis

Ears were cut in the central region to expose a ring of kernels. Embryos were labelled in situ as described by Kriz (1982). Individual kernels were injected with 3 µl of ^{35}S methionine (NEN; specific activity > 1,000 Ci/m*M*) using a Hamilton syringe; the cut ends of the ears were covered with moist filter paper and the ears were incubated in the dark for 2 h at 27 °C.

Following the 2 h radio-labelling period, the embryos were dissected; three to five embryos were combined to form each sample and homogenized in a buffer containing 200 m*M* TRIS(hydroxymethyl)aminomethane-HCl (pH 7.5), 5% sodium dodecyl sulphate, 7.5% 2-mercaptoethanol and 1 m*M* phenylmethylsulphonyl fluoride (PMSF) (Hughes et al. 1981). Radioactive incorporation was determined as described by Mans and Novelli (1960).

Samples were prepared for 2D polyacrylamide gel electrophoresis (PAGE) as described by Baszczynski et al. (1983) with slight modification. Urea was added to samples to give a concentration of 8 *M* and NP-40 to give a ratio of at least 8:1, NP-40:SDS. Ampholytes (LKB) in the range of 3.5–10 were added to give a concentration of 2% and the mixture was combined with an equal volume of a solution containing 9.5 *M* urea, 7.5% 2-mercaptoethanol, 1 m*M* PMSF and 8% NP-40. 2D-PAGE was performed as described by O'Farrell (1975). IEF gels were 10.5 cm in length and were prepared with a 4:1 mixture of 5–8 and 3.5–10 ampholytes. SDS gels were 10.5 cm in height with a linear acrylamide concentration gradient of 7.5%–17.5%. Fluorography was performed as described by Bonner and Laskey (1974) and the gels were exposed to preflashed (Laskey and Mills 1975) Kodak X-Omat XAR film.

Preparation of data for multivariate analysis

A representative fluorogram obtained from the 2D separation of maize embryo polypeptides is shown in Fig. 1. Superimposed on this fluorogram is the grid employed in coding the data for multivariate analysis (Fewster and Walden 1987). Two or more replicate fluorograms were compared for each sample. Over 300 spots could be resolved on all fluorograms between a range in pH of 5 and 8 and in molecular mass of 14 and 94 kD. At least 25 prominent spots exhibited qualitative and/or quantitative variation during the period between 15 DAP and 52 DAP. Of these, 2 spots showed genotypic specificity, one present in each inbred and both present in each of the hybrid cultivars. Since the majority of spots were common to all developmental stages and genotypes, the position of each grid line could be assigned accurately and the entire grid structure faithfully reproduced on all fluorograms. As expected, the actual physical area encompassed by individual grid units differed slightly among gels, particularly when electrophoresis was performed at different times. In such cases, comparison of replicates run at different times ensured that the functional areas of the grid units were equivalent for all fluorograms. Each grid unit represented an axis of variation upon which fluorograms from different individuals could be compared.

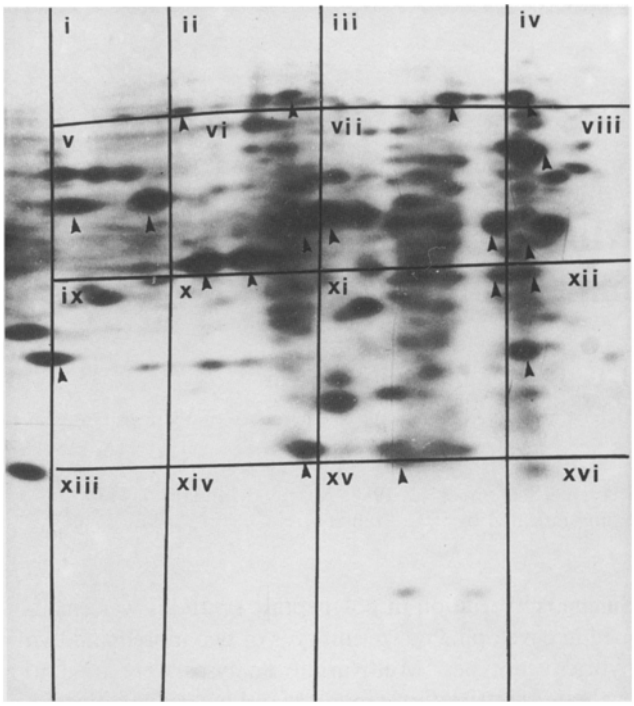

Fig. 1. Representative fluorogram obtained from the 2D separation of 35 DAP maize embryo polypeptides, overlaid with the grid used in coding the data for multivariate analyses. *Arrowheads* indicate marker spots used to determine the position of grid lines. Grid units are numbered i–xvi from left to right and top to bottom

The grid structure was drawn on acetate sheets overlaid onto each fluorogram using the marker spots as a guide. Each grid unit was assigned a score from 0–7 (Fewster and Walden 1987) based on the number and intensity of spots within as follows: score 0, no spots; scores 1 and 2, 1–3 and > 3 light spots respectively; scores 3, 4, 5, 6 and 7, 1, 2–4, 5–10, 11–15 and > 15 dark spots, respectively. A blind scoring procedure was adopted and the integrity of the method was checked by subsequently comparing the values assigned to replicate fluorograms. Normally, identical values were obtained for equivalent grid units on replicate gels of comparable resolution. One replicate from each pair was selected so that all fluorograms included in the analyses were of approximately equal density. The data set consisted of 32 fluorograms: one for each genotype at eight different sampling ages (15, 20, 25, 31, 35, 40, 45, 52 DAP). The coded data were analysed using two forms of eigenanalysis: principal coordinate (Gower 1966) and concentration analyses (Feoli and Orlóci 1979, 1985; Orlóci 1981).

Results

Principal coordinate analysis

A PCoA was conducted on the data obtained from fluorograms from each of the four genotypes at eight sampling ages. The vector diagram presented in Fig. 2 shows

the relative contribution of the 16 original variables to the first two axes of this analysis, which together comprised 59% of the total variation. The three vectors of greatest magnitude and, hence, those which contribute most to the variation represented in the first plane of the analysis are 3, 5 and 7, corresponding to the grid units of the same numbers in Fig. 1. Most variation attributable to variables 3 and 7 is manifested in the first or horizontal axis while that of variable 5, the vector of greatest magnitude, is represented in both the first and second axes. Polarity is observed along the first axis for variable 5 in relation to variables 3 and 7.

An ordered distribution of the data emerges across the plane described by the first two axes from the PCoA

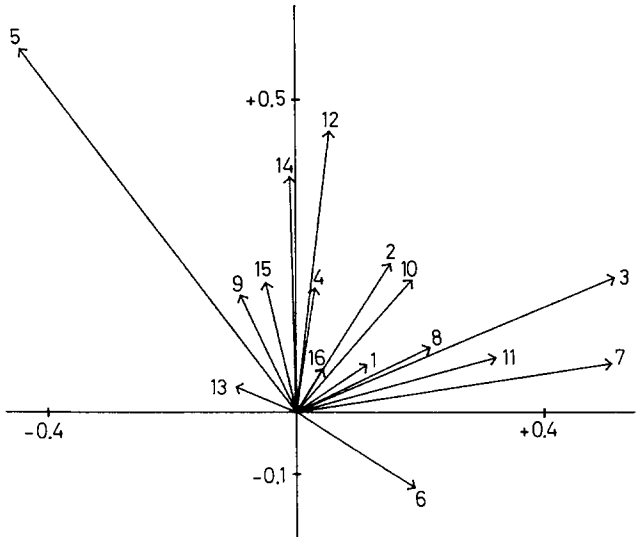

Fig. 2. Vector diagram showing the relative contribution (length and direction) of original variables to the first two component axes from a PCoA of coded fluorographic data. Numbers for each vector correspond with the grid units in Fig. 1

(Fig. 3). By utilizing the results presented in Fig. 2, it is observed that increasing embryo age corresponds to an increase in the value of variable 5 concomitant with a decrease in the values of variables 3 and 7. Development in all four genotypes appears to follow the same pattern.

Concentration analysis

The results obtained following concentration analysis of the data are presented in Fig. 4. Concentration analysis partitions the variation into a number of linear components based on the decomposition of Chi-square. For each component, the variation is presented in the form of a lattice of deviations from random expectation. The technique differs from PCoA in that it enables an assessment of the dissimilarity among individuals for each grid unit. The first lattice of deviations accounted for 36% of the total variation. 'Profiles' were determined for each cultivar/sampling age: deviations from random expectation along the vertical axis reflect the changes occurring in the values of grid units (the horizontal axis) and the cumulative value of the entire fluorogram compared to all others in the data set. A trend is evident in the shift from a positive to a negative slope occurring with increasing embryo age. Changes in the values of variables 3, 5 and 7 follow a pattern similar to that revealed by PCoA. Deviations associated with variable 5 shift from negative to positive during the course of development while the reverse is observed for variables 3 and 7. Shifts from negative to positive deviations for some variables such as 9, which exhibit little or no change in value during development, reflect a slight decrease in the cumulative value of all 16 variables with increasing age. No strong indication of genotype-specific trends is observed. Although, in some cases, a shift in deviation may appear to occur earlier or later in a particular genotype, there is not a consistent pattern through all sampling ages.

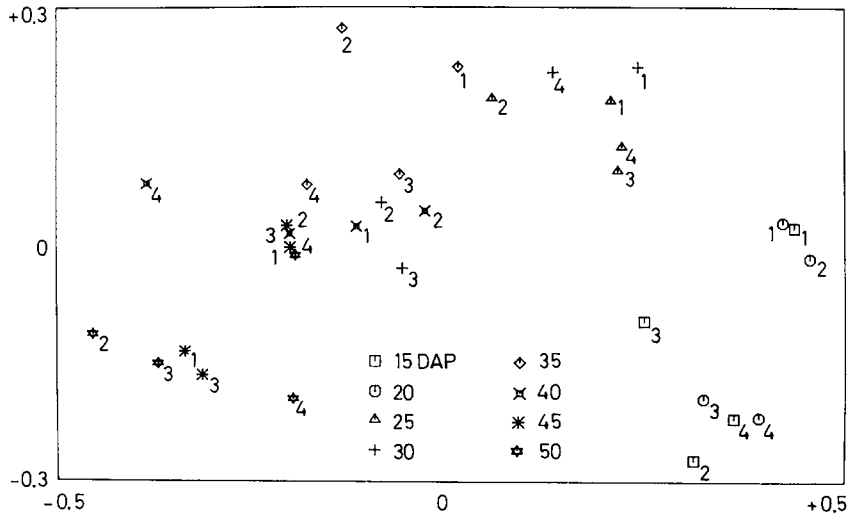

Fig. 3. Distribution of individuals within the plane described by the first two component axes from a PCoA of coded fluorographic data. 1. OH43, 2. OH43/M14, 3. M14/OH43, 4. M14

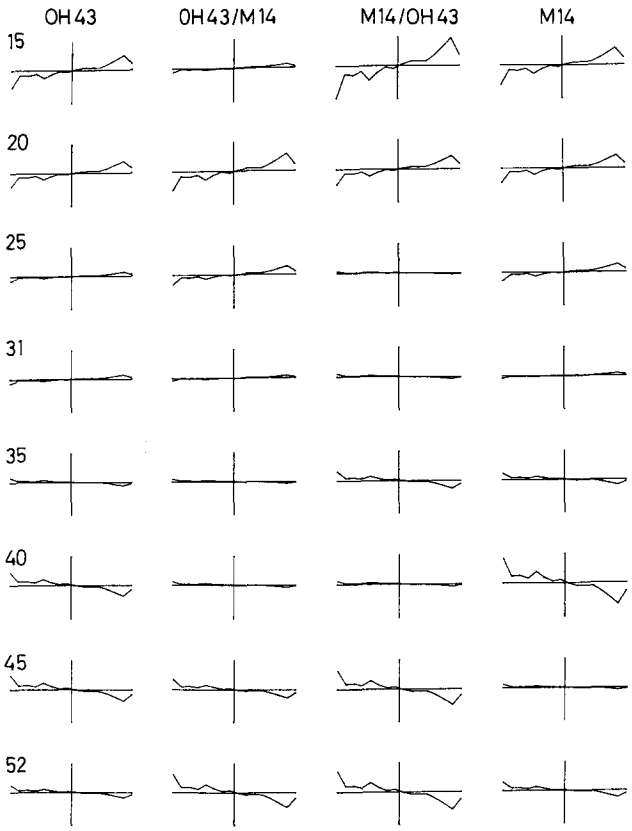

Fig. 4. Profiles obtained from a concentration analysis of coded fluorographic data. Values from the first lattice of deviations from random expectation are plotted on the vertical axis. Grid units are plotted on the horizontal axis in the order from left to right of V, XIII, XV, XIV, IX, XII, XVI, IV, II, VIII, VI, X, XI, VII, III and I

Discussion

Two-dimensional electrophoresis of radio-labelled polypeptides and multivariate analyses were used to characterize maize embryo development in two unrelated inbreds and their reciprocal hybrids. Variation among the patterns of polypeptides synthesized by embryos of different ages was resolved into an ordered distribution by PCoA of the numerically coded data. One of the properties of PCoA is a reduction in the dimensionality over which the majority of total variation in the data is represented. In the present study, the plane described by the first two component axes encompassed 59% of the total variation. Although all 16 of the original variables contribute to some extent to the variation represented in the first plane, the proportion attributable to each is not the same. In addition to facilitating the detection of trends or patterns, PCoA enables the identification of those variables which exhibit the greatest relative contribution to the variation residing along each axis and, hence, those which are most important in discriminating among individuals. The distribution observed here is most strongly influenced by variables 3, 5 and 7. Variable 5 shows an increase in value over the period of study, while decreasing values are observed for variables 3 and 7. This polarity is reflected in the separation of young (15–20 DAP) from mature (45–52 DAP) embryos along the first component axis (Fig. 3). Since the value assigned to a variable bears a direct relationship to the number and intensity of spots enclosed within the corresponding grid unit of the 2D fluorogram, variation may be interpreted in terms of its underlying cause: changes in the synthesis of polypeptides within the grid unit. Thus between 15 and 52 DAP, novel and/or enhanced polypeptide synthesis occur in the grid unit corresponding to variable 5 while repression and/or reduction in polypeptide synthesis occur in grid units corresponding to variables 3 and 7.

Concentration analysis of the data revealed a trend similar to that revealed by PCoA. A large proportion of the total variation, 36%, was represented in the first lattice of deviations from random expectation. Apart from confirming the results obtained from PCoA, concentration analysis provides a clear structural description of the 2D array of polypeptides for each variable, facilitating comparison among individuals. The gradual nature in which shifts in the patterns of protein snythesis occur during development was readily apparent in the succession of different profiles, as was the polarity between young and mature embryos.

The view of development which emerges from these studies is one of distinct, unidirectional physiological states as maize embryos mature. Each state is characterized by the prominent synthesis of a different subset of polypeptides and these are linked by a gradual transition phase. That quantitative variation in a number of polypeptides has a significant role in the process of development seems likely from the progressive changes in variables 3, 5 and 7 and is supported by visual inspection of the fluorograms. In a study of development of wheat seedling and plant leaf tissue proteins, Zivy et al. (1984) found that the number of polypeptides exhibiting intensity differences in fact exceeded that exhibiting presence/absence differences. It would appear from a number of studies that qualitative variation in gene products (mRNAs and polypeptides) and up/down modulation of gene product levels are general features of plant embryo development (reviewed in Dure 1985).

Little genotypic variation was revealed among the lines included in this study. The genotypic specificity observed for two polypeptides upon visual inspection of fluorograms did not constitute sufficient variability to be resolved in our analyses. Other studies have detected more extensive variation between the 2D patterns obtained from different lines (Zivy et al. 1983, 1984; Damerval et al. 1987). Two factors contribute to the lack of

genotypic variation uncovered in our study. The first of these is technical and concerns the method used to detect variation. The reduction in the number of variables (to 16 grid units) and the use of distinct scoring categories, each covering a range of variation, while easing the burden on the investigator, necessarily results in an underestimation of the variation. Furthermore, the multivariate techniques we employed were designed to reveal prominent trends or associations in the data, which in this study pertained to development. The second factor concerns the influence of tissue and stage of development on variation. Gene products exhibiting tissue or developmental specificity may differ in their extent of polymorphism. For example, in studies involving leaf tissues of wheat, considerable developmental and genotypic variation was detected for both nuclear and cytoplasmically encoded polypeptides (Zivy et al. 1984). However, most of the cytoplasmic variation was attributed to chloroplast-encoded products. A major source of polymorphism in seeds is associated with storage proteins (Kreis et al. 1985). Although polymorphism has been observed for the zeins in maize (Soave and Salamini 1984), zein polymorphism would not be expected to contribute to genotypic variation among embryos.

Our results gave no indication of differential rates of development for inbreds and hybrids. This is surprising since morphological observations have demonstrated that, in many cases, maize hybrids mature more rapidly than their inbred parents (Sass 1976). However, it has been shown that the rate of development may vary with the phase of development (Tollenaar et al. 1984) and, therefore, our results must be interpreted with respect not only to the particular genotypes used, but also to the developmental 'window' examined. A wider selection of inbreds and hybrids would permit a more comprehensive evaluation of the extent of variability within different phases of embryo development.

Several recent studies have employed multivariate analysis to investigate protein variability using either isozyme frequencies in different populations (Smith 1984; Bretting et al. 1987; Smith and Smith 1987) or computed cumulative variation among electrophoretically or chromatographically separated samples (Damerval et al. 1987; Smith and Smith 1987). Our approach differs in that multivariate analyses are used to directly identify variation within a 2D gel. In this respect, it is more similar to the approach taken by Anderson et al. (1984) and Aliaga-Morell (1987), except that the variables we employ consist of grid units within a 2D fluorogram as opposed to individual spots or bands, eliminating the need for sophisticated hardware and software. Our method is applicable to any study involving 2D separation of macro-molecules, requiring only that a grid structure be constructed which is reproducible for all individuals included in the analysis. As indicated (Fewster and Walden 1987), care must be taken in optimizing the scoring procedure and number of grid units (variables) employed for each data set, in order to maximize the resolution among individuals. The opportunity to assess some variability is lost when changes occur simultaneously within a grid unit which cancel each other out or are too small to significantly affect the assigned score (as was the case with the genotypic differences detected through visual inspection of our fluorograms). However, the advantage of being able to detect trends amidst complex patterns of variation, thus aiding the interpretation of results, combined with the speed of analysis and clarity of presentation, should make this technique amenable to a wide variety of studies.

Acknowledgements. The authors gratefully acknowledge Operating and Strategic Grant support from the Natural Sciences and Engineering Research Council of Canada. We are grateful to D. Greyson and P. Fewster for many helpful discussions.

References

Aliaga-Morell JR, Culianez-Macia FA, Clemente-Marin G, Primo-Yufera E (1987) Differentiation of rice varieties by electrophoresis of embryo protein. Theor Appl Genet 74:224–232

Anderson NL, Hofmann J-P, Gemmell A, Taylor J (1984) Global approaches to quantitative analysis of gene expression patterns observed by use of two-dimensional gel electrophoresis. Clin Chem 30:2031–2036

Bahrman N, Thiellement H (1987) Parental genome expression in synthetic wheats (*Triticum turgidum* sp. × *T. tauschii* sp.) revealed by two-dimensional electrophoresis of seedling proteins. Theor Appl Genet 74:218–223

Baszczynski CL, Walden DB, Atkinson BG (1983) Regulation of gene expression in corn (*Zea mays* L.) by heat shock. II. In vitro analysis of RNAs from heat-shocked seedlings. Can J Biochem Cell Biol 61:395–403

Bonner WM, Laskey RA (1974) A film detection method for tritium-labeled proteins and nucleic acids in polyacrylamide gels. Eur J Biochem 43:83–88

Bretting PK, Goodman MM, Stuber CW (1987) Karyological and isozyme variation in West Indian and allied American mainland races of maize. Am J Bot 74:1601–1613

Damerval C, Herbert Y, de Vienne D (1987) Is the polymorphism of protein amounts related to phenotypic variability? A comparison of two-dimensional electrophoresis data with morphological traits in maize. Theor Appl Genet 74:194–202

Dunbar BD, Bundman DS, Dunbar BS (1985) Identification of cultivar-specific proteins of winter wheat (*T. aestivum* L.) by high resolution two-dimensional polyacrylamide gel electrophoresis and color-based silver stain. Electrophoresis 6:39–43

Dure L (1985) Embryogenesis and gene expression during seed formation. In: Miflin BJ (ed) Oxford surveys of plant molecular and cell biology, vol 2. Oxford University Press, New York, pp 67–69

Dure L, Greenway SC, Galau GA (1981) Developmental biochemistry of cottonseed embryogenesis and germination: changing messenger ribonucleic acid populations as shown by in vitro and in vivo protein synthesis. Biochemistry 20:4162–4168

Feoli E, Orlóci L (1979) Analysis of concentration and detection of underlying factors in structured tables. Vegetatio 1:49–54

Feoli E, Orlóci L (1985) Species dispersion profiles of anthropogenic grasslands in the Italian Eastern Pre-Alps. Vegetatio 60:113–118

Fewster PH, Walden DB (1987) Multivariate analysis of data from two-dimensional electrophoretic separation of macromolecules. Comput Biol Med 17:29–35

Gower JC (1966) Some distance properties of latent root and vector methods in multivariate analysis. Biometrika 53:315–328

Hughes WG, Baszczynski CL, Ketola-Pirie C (1981) Improved conditions for the extraction of maize polypeptides. Maize Genet Coop Newslett 55:116–117

Kreis M, Shewry PR, Forde BG, Forde J, Miflin BJ (1985) Structure and evolution of seed storage proteins and their genes with particular reference to those of wheat, barley and rye. In: Miflin BJ (ed) Oxford surveys of plant molecular and cell biology, vol 2. Oxford University Press, New York, pp 253–317

Kriz AL (1982) In vivo labeling of embryo and endosperm proteins in intact kernels with 35-S methionine. Maize Genet Coop Newslett 56:14–15

Laskey RA, Mills AD (1975) Quantitative film detection of ^{3}H and ^{14}C in polyacrylamide gels by fluorography. Eur J Biochem 56:335–341

Mans RJ, Novelli DG (1960) A convenient, rapid and sensitive method for measuring the incorporation of radioactive amino acids into protein. Biochem Biophys Res Commun 3:540–548

Misra S, Bewley JD (1985) Reprogramming of protein synthesis from a developmental to a germinative mode induced by desiccation of the axes of *Phaseolus vulgaris*. Plant Physiol 78:876–882

O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. J Biol Chem 250:4007–4021

Orlóci L (1981) Probing time series vegetation data for evidence of succession. Vegetatio 46:31–35

Sanchez-Martinez D, Puigdomenech P, Pages M (1986) Regulation of gene expression in developing *Zea mays* embryos. Plant Physiol 82:543–549

Sass JE (1976) Morphology. In: Sprague GF (ed) Corn and corn improvement. American Society of Agronomy, Madison/WI, pp 89–110

Sawney VK, Chen K, Sussex IM (1985) Soluble proteins of the mature floral organs of tomato (*Lycopersicon esculentum* Mill). J Plant Physiol 121:265–271

Soave C, Salamini F (1984) The role of structural and regulatory genes in the development of maize endosperm. Dev Genet 5:1–25

Smith JSC (1984) Genetic variability within US hybrid maize: Multivariate analysis of isozyme data. Crop Sci 24:1041–1046

Smith JSC, Smith OS (1987) Associations among inbred lines of maize using electrophoretic, chromatographic and pedigree data. Theor Appl Genet 73:654–664

Tellenaar M, Muldoon JF, Daynard TB (1984) Differences in rates of leaf appearance among maize hybrids and phases of development. Can J Plant Sci 64:759–763

Zivy M, Thiellement H, Vienne D de, Hofmann J-P (1983) Study on nuclear and cytoplasmic genome expression in wheat by two-dimensional gel electrophoresis. 1. First results on 18 alloplasmic lines. Theor Appl Genet 66:1–7

Zivy M, Thiellement H, Vienne D de, Hofmann J-P (1984) Study on nuclear and cytoplasmic gene expression in wheat by two-dimensional gel electrophoresis. 2. Genetic differences between two lines and two groups of cytoplasms at five developmental stages or organs. Theor Appl Genet 68:335–345