Theor Appl Genet (1986) 72:296–301

Quantitative variation in the kernel proteins among 841 accessions of *Triticum dicoccoides* estimated by SDS-PAGE*

L. Mansur-Vergara [1], C. F. Konzak [1], Z. K. Gerechter-Amitai [2], A. Grama [2] and A. Blum [2]

[1] Department of Agronomy and Soils, Washington State University, Pullman, WA 99164-6420, USA
[2] The Volcani Center, Agricultural Research Organization, Bet-Dagan, Israel

Received August 19, 1985; Accepted December 16, 1985
Communicated by K. Tsunewaki

Summary. The relative proportion and amount of proteins in five defined molecular weight (MW) regions (A1 = above 71,000 = 71K, A2 = 71K-49K, A3 = 49K-31K, A4 = 31K-20K, A5 = 20K and less) were estimated by densitometric analyses of the amount of dye bound by kernel proteins (Fullington et al. 1980) of *Triticum dicoccoides* SDS-PAGE gels. These MW regions roughly correspond to the wheat protein solubility classes (Cole et al. 1981; Fullington et al. 1983). One purpose of the study was to select accessions whose seed proteins bind relatively high amounts of dye in the glutenin and albumin globulin regions. These accessions will be used for further in-depth studies as possible candidate donors of genes to improve the baking and nutritional quality of wheat. Marked differences in the quantitative relationships were found among the proteins in the five MW regions. Coefficients of variation (CV's) for the highest peak (i.e., most abundant protein) MW in different protein MW regions were similar for A1, A2 and A3, at 11.4, 11.7, and 11.1%, respectively, but only 4.1 for A4, and 10.6% for region A5. The CV for the highest peak MW overall was 29.8. Accession BP0649, for example, had over 44% of its protein in region A5, whereas BP0566 (lowest among the top 10%) had only 21.4% of its protein in that region. Over 37% of the proteins of accessions BP0649 and 0001 to 0005 was in region A5. At least 84 accessions with the highest amount of protein in region A5, and 13 accessions with more protein in region A1 than Chinese Spring may merit further evaluation as possible protein gene donors. High amounts of protein in A1 may be of importance in bread-baking quality, and in A4 and A5 for high lysine wheat. Accessions in both extremes were selected to test these hypotheses. All accessions are now or will be available in the USDA Wheat Collection.

Key words: Electrophoresis – Endosperm – Proteins – Wild emmer – Wheat

Introduction

Improving protein quality is a major challenge for wheat breeders, especially if one considers that the pool of potentially useful genes for quality traits has become dangerously narrow and little use has been made of exotic gene pools (Johnson et al. 1972). Therefore, breeders have become increasingly interested in the gene reservoir present in the wild-type relatives of wheat. High protein accessions of *Triticum dicoccoides* are being utilized by some breeders as sources of genes to increase protein content of wheat (Grama et al. 1982), although little is known about the relative distribution and kinds of kernel proteins present in the many accessions assembled in collections.

The 841 accessions used in this study ranged from 15–25% protein content with some exhibiting high protein content along with relatively large kernel size. In addition, considerable genetic variability exists among their soluble and storage proteins (Nevo et al. 1982; Mansur-Vergara 1983). This rich base of variability makes wild emmer a valuable gene source to improve our cultivars.

Fullington et al. (1980, 1983) and Cole et al. (1981) showed that dye binding provides a useful quantitative measure of the proteins present in the wheat kernel and that the molecular weight (MW) distribution of the proteins roughly corresponds to the Osborne solubility classes (albumin = water soluble; globulin = salt water soluble; gliadin = alcohol soluble; glutenin = residue left) (Osborne 1907).

* Scientific Paper No. 7092. College of Agriculture Research Center, Washington State University, Pullman, Project No. 1568. This work was supported in part by the U.S.-Israel Binational Agricultural Research and Development Fund (BARD)

We used their approach to estimate the relative distributions and phenotypic variation of kernel proteins in *T. dicoccoides* as a basis for selecting those high in glutenin and those high in albumin and globulin proteins. Accessions with high relative amounts of albumin and globulin proteins were identified for further study because these proteins are known to be rich in basic essential amino acids, such as lysine, for which wheat is notably deficient. The high MW glutenin proteins (region A1) have been implicated in wheat baking quality (Payne et al. 1981), therefore, accessions of interest were selected for studies aimed toward the transfer of these genes by conventional breeding methods for improving wheat protein processing and nutritional properties.

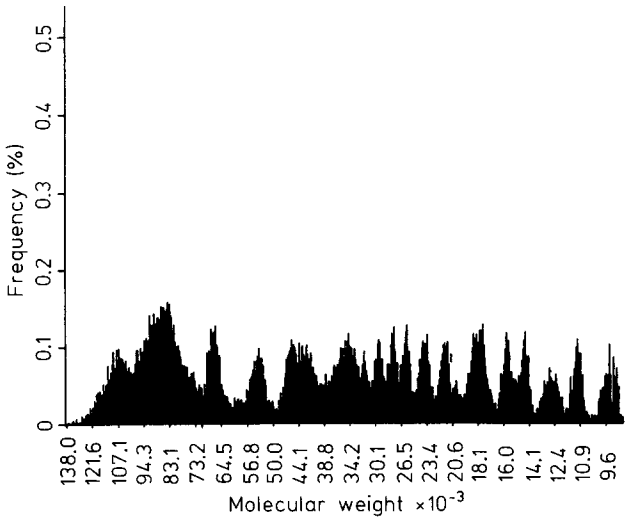

Fig. 1. Computer generated graph of frequency versus MW of all accessions

Materials and methods

Details regarding the materials and methods employed in preparing and scanning the SDS-PAGE gels were described in a previous paper (Mansur-Vergara et al. 1984). The computer bank of densitometer scan data accumulated for the 841 accessions of the *T. dicoccoides* analyzed was processed using specially constructed software (Mansur-Vergara et al. 1984) to provide the results for this paper. Average percentage total area and adjusted total area were calculated for proteins grouped in five molecular weight regions, which correspond to the solubility classes mentioned previously. The adjusted values relate to corrections based on the fact that the region 5 proteins (Tables 1 and 2) are overestimated due to their higher proportion of basic amino acid proteins, which have a high affinity for Coomassie Blue stain.

Results

Quantitative analysis of densitometer scans of SDS-PAGE gels

The total area under each densitometer scan was made equal to 100% and then subdivided into five regions (Figs. 1 and 2): A1, 71K MW and above; A2, 71K-49K MW; A3, 49K-31.15K MW; A4, 31.15K-20K MW; A5, 20K MW and under. Analysis of the normalized total area for each region makes it possible to draw some conclusions which do not depend on the relative amounts of protein loaded onto the gels, i.e., the analysis is made in relative rather than absolute terms as developed earlier by Cole et al. (1981) and Fullington et al. (1983). In this study the MW's defining areas A1 through A5 differ somewhat from those used by the above-cited authors, although a very similar SDS-PAGE system was used. The differences may be due to a different grouping of proteins among their hexaploid wheats and the wild tetraploids of this study.

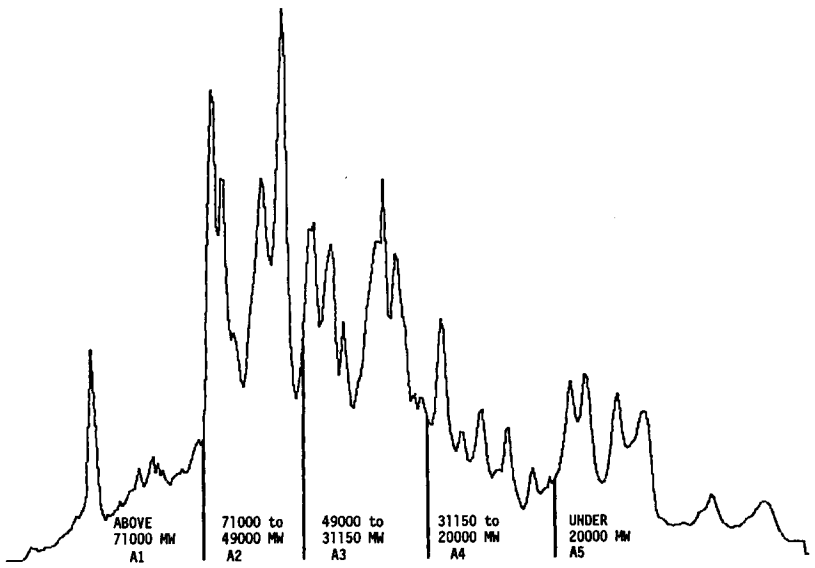

Fig. 2. Scan of BP0150 depicting regions A1, A2, A3, A4, A5

The cut-off points 71K, 49K, 31.15K and 20K were necessary to avoid splitting major peaks. This was done after printing a graph of frequency versus MW and then choosing those points in the area of interest that had a low frequency (Fig. 2). Consequently, the specific separation points may be more natural for the particular groups of wheats studied. Moreover, the large number of samples of *T. dicoccoides* involved in the present work, compared with the relatively few in other studies, provide a more firm basis for protein group divisions specific to the proteins present in accessions of this species.

Area analysis

The data indicate that the average percentage total area is greatest in A3, with 34.6% of the total, then A2 with 22.5%, A5 with 21.4%, A4 with 14.5% and A1 with 7.1% (Table 1). Furthermore, regions A1 and A5 had the greatest quantitative variability. Average percentage and adjusted values are presented based on evidence from Cole et al. (1981) and Fullington et al. (1983). These workers found that proteins in the lower MW (LMW) region (below 30,000 MW) bind two to three times more dye than the remaining proteins, depending upon the destaining procedure used. The staining method in this study was similar to that of Cole et al. (1981), except that the Coomassie Blue R-250 concentration in this study was higher (1% vs. 0.02%) and a methanol/acetic acid/water solution was used to dissolve the dye instead of a six percent trichloroacetic acid. Assuming that the low MW area (A5) binds twice as much dye, the distribution for the area under each region after adjusting (Table 1) would be as follows: A1 = 8.02%, A2 = 25.54%, A3 = 39.27%, A4 = 16.45% and A5 = 10.68% of the total area on the average. Regions A1, A2, and A3 include mostly glutenins and gliadins while A4 and A5 contain globulins and albumins (Cole et al. 1981). Those proteins in regions A1-A3 comprise 73% of the total area under in reasonable agreement with the 75–85% range attributed to gliadin plus glutenin in solubility studies (Simmonds 1981). This finding also agrees with the observations of Fullington et al. (1980, 1983) and Cole et al. (1981) that solubility fractions roughly correspond to SDS-PAGE MW groupings of protein bands. The correlation coefficient matrix (Table 2) indicated that when area A1 increases, A2 also increases, but A3, A4 and A5 decrease. In turn, the quantity of proteins in area A3 is negatively correlated with all other areas, especially A5; area A4 is also somewhat strongly correlated negatively with all other areas except A5. The area of region A5 was strongly and negatively correlated with all other areas except A4 with which it was strongly positively correlated. All correlation coefficients were significant because of the large sample size.

Table 1. Mean (%), standard deviation (STD DEV) and coefficient of variation (CV) of total area under regions A1, A2, A3, and A5

Region	% Mean area		CV	STD DEV
	Unadjusted	Adjusted		
A1	7.1	8.0	32	2.3
A2	22.5	25.5	14	3.1
A3	34.6	39.3	11	3.9
A4	14.5	16.5	15	2.2
A5	21.4	10.7	17	3.6

Table 2. Correlation coefficient matrix for area under regions A1, A2, A3, A4, and A5*

Region	A1	A2	A3	A4	A5
A1	1.00	0.33	−0.24	−0.40	−0.41
A2		1.00	−0.19	−0.53	−0.54
A3			1.00	−0.38	−0.54
A4				1.00	0.52
A5					1.00

* All significant because of the large number of samples studied ($P < 0.01$)

Table 3. Number or percent of times region contains largest peak

Region	Times	%
A1	5/842	0.59
A2	259/842	30.76
A3	574/842	68.17
A4	3/842	0.36
A5	1/842	0.12

Peak analysis

It was found that 68% of the time A3 contained the most intense peak. Region A2 had the most intense peak about 31% of the time. The largest peak was seldom present in the other areas (Table 3). Note that in Table 3 the denominator shows 842 accessions instead of 841. This is because the hexaploid standard, 'Chinese Spring', was included in the analysis and was designated as BP900.

The mean MW values, standard deviations, and coefficient of variation (CV) for the largest peak within each area and overall are shown in Table 4.

The results show the largest peak in A4 is at 29,670 MW and is remarkably conserved (CV = 4.06%) among the accessions analyzed. In contrast, the largest peak overall was highly variable, with an average MW of 44,250 and a CV = 29.8%. These results suggest that changes in the quantitative expression of genes coding for kernel proteins vary not only among species, as

Table 4. Mean MW, standard deviation, and coeffiicient of variation of largest peak within each region and overall

Region	Mean MW	STD DEV	CV
A1	85,480	9,720	11.4
A2	60,920	6,810	11.2
A3	36,750	4,080	11.1
A4	29,670	1,210	4.1
A5	15,290	1,620	10.6
Overall	44,250	13,190	29.8

Table 5. Comparison of "average" *T. dicoccoides* with 'Chinese Spring' for areas under regions A1 to A5

Region	% Mean area			
	Adjusted		Unadjusted	
	T. dicoccoides	'Chinese Spring'	*T. dicoccoides*	'Chinese Spring'
A1	8.0	14.0	7.1	12.0
A2	25.5	17.5	22.5	15.0
A3	39.3	36.2	34.6	31.0
A4	16.5	18.7	14.5	16.0
A5	10.7	12.5	21.4	25.0

Table 6. List of accessions with higher relative percent area in A1 than 'Chinese Spring'

% area	% protein	Accession no.
12.52	18.7	BP0095
12.57	19.4	BP0126
12.64	18.7	BP0370
12.84	17.5	BP0148
12.88	*	BP0519
12.99	19.7	BP0032
13.36	*	BP0516
13.91	*	BP0531
15.12	20.4	BP0501
15.13	16.4	BP0096
15.35	17.8	BP0123
15.47	18.7	BP0088
17.46	18.2	BP0125

* Not available

found by Cole et al. (1981), but also can be detected within a species if a large number of samples are analyzed. Although there is within-species variability, *T. dicoccoides* definitely shows a particular pattern. This pattern is evident when a large number of samples are observed. The data show that, 68% of the time, area A3 contains the largest peak with a mean MW of 36,750 and, in the majority of the cases, region A3 contained the most protein.

The SDS-PAGE pattern of the hexaploid variety Chinese Spring was, in many respects, similar to that of the average *T. dicoccoides* accessions shown in Fig. 1. However, the proportions of protein subunits in 'Chinese Spring' is remarkably different from that of most *T. dicoccoides* accessions, particularly in the high MW region (7% vs. 12% for 'Chinese Spring'). Interestingly, specific *T. dicoccoides* accessions appear to be similar to 'Chinese Spring' or even exceed it in the proportion of proteins present in the high MW region and low MW regions (Tables 6 and 7). Table 5 shows a comparison made between the "average" *T. dicoccoides* accession with 'Chinese Spring' showing the adjusted (for basic amino acids) and nonadjusted values for the area under each region.

Breeders should be made aware that the averages shown in Table 5 hide the fact that great variability exists in the wild emmer collection. Some wild emmer accessions greatly exceed the hexaploid wheat in both the high and low MW regions (Tables 6 and 7).

Generally, as the percent kernel protein increases, the proportion of albumins and globulins decrease and the glutenins and gliadins increase (Fullington et al. 1983). The *T. dicoccoides* accessions had a minimum of 15% protein and a maximum of 25% (av. 18%), whereas the 'Chinese Spring' sample had 12% protein. When regions A1, A2, and A3 were combined, glutenins and gliadins made up 73% of the total area in *T. dicoccoides* and only 68% in 'Chinese Spring' (based on adjusted values in Table 5). These data tend to support the validity of classifying solubility fractions in terms of SDS-PAGE patterns, especially if standard techniques were adopted by cereal chemists.

Accessions with a higher percentage of proteins in area A1 may be of special interest as gene donors because of the importance this class of proteins have on the baking properties of wheat. Similarly, accessions with an unusually high proportion of protein in region A5 may be of interest for breeding because they may serve as gene donors for developing higher lysine wheats. Accessions with a higher relative percent of proteins in A1 than 'Chinese Spring' are listed in Table 6; likewise, the top 10% of accessions with highest relative intensity in A5 are listed in Table 7.

Some caution must be exercised in the interpretation of data from this type of study, because it is possible that some of the quantitative variability observed may be due to the SDS-PAGE technique. However, the validity of the results is corroborated because the relationships held true within as well as between runs. In addition, the fact that each pattern was normalized to 100% allowed us to remove variability due to staining and amount of protein loaded. The analyses in this study seem an efficient way to focus attention on subsets of accessions which may prove of greater

Table 7. The top 10% of the accessions with highest proportion of area in A5

% area	Acc. no.	% area	Acc. no.	% area	Acc. no.	% area	Acc. no.
44.1	BP0649	28.5	BP0038	26.6	BP0442	26.0	BP1110
38.3	BP0004	28.4	BP0650	26.5	BP0014	26.0	BP0642
37.8	BP0002	28.3	BP0603	26.5	BP0406	26.0	BP0250
37.7	BP0005	28.2	BP0470	26.4	BP0720	26.0	BP0596
37.4	BP0001	28.0	BP0417	26.4	BP0467	25.9	BP0060
37.2	BP0003	27.8	BP0270	26.4	BP0061	25.9	BP0055
33.1	BP0188	27.8	BP0928	26.4	BP0723	25.9	BP0627
32.7	BP0407	27.7	BP0121	26.4	BP0289	25.9	BP0424
32.4	BP0040	27.6	BP0803	26.3	BP0237	25.9	BP0015
31.6	BP0075	27.6	BP0122	26.3	BP0636	25.9	BP0634
30.7	BP0721	27.6	BP0210	26.3	BP0073	25.7	BP0195
29.6	BP0008	27.4	BP0034	26.3	BP0299	25.6	BP0548
29.4	BP0039	27.3	BP0011	26.3	BP0178	25.6	BP0037
29.3	BP0213	27.2	BP0337	26.2	BP0620	25.6	BP0067
29.2	BP0521	27.1	BP0855	26.2	BP0222	25.5	BP0190
29.1	BP0006	27.0	BP0430	26.1	BP0686	25.5	BP0042
28.9	BP0643	27.0	BP0719	26.1	BP0009	25.5	BP0258
28.7	BP0415	26.9	BP0624	26.1	BP0611	25.4	BP0300
28.7	BP0318	26.8	BP0641	26.0	BP0813	25.4	BP0587
28.7	BP0428	26.8	BP0016	26.0	BP0193	25.4	BP0200
28.5	BP0429	26.7	BP0007	26.0	BP0722	21.4	BP0566

interest for further research, especially when considering that protein content and kernel weight data also can be included as additional criteria.

Discussion

The relative proportion of protein present in defined MW regions (as determined by densitometry of the amount of dye bound by proteins) is a less reliable indicator of protein genetic variability than the presence or absence of a protein band determined by qualitative analyses. There is strong evidence that the presence or absence of bands in SDS-PAGE is in fact a genetic characteristic, but it is well known that the amount of nitrogen available in the soil and, hence, of protein in the kernel, will cause some variation in the quantitative distribution of the protein subunits (Fullington et al. 1983). Another factor known to affect the amount of dye bound to a protein is the number of basic amino acids it contains, i.e., lysine and arginine. Therefore, as either the protein content or the number of basic amino acids increase, the dye-binding capacity of the protein also increases.

The accessions used in this study varied in protein content from 15 to 25% so it would be expected that variation in the relative distribution of their proteins also would occur. It was observed that regions designated A1 and A2 (region above 49,000 MW) increased in relative percentage when regions A4 and A5 decreased. The proportions of protein in regions A1 and A2 were positively correlated, as were those in A4 and A5. However, protein in A3 was negatively correlated with all others. Fullington et al. (1983) also found that storage proteins (their regions A1, A2, A3, A4) increase more than soluble proteins (their region A5, ours A4 and A5) as kernel protein percentage increases. Our data are generally consistent with their finding and it appears to suggest that a common regulator mechanism exists for the quantitative expressions of glutenin and gliadin genes. This behavior may be explainable from evidence that the lower molecular weight proteins (albumins and globulins) may be differently regulated from the storage proteins, perhaps involving feedback inhibition mechanisms. Mitra and Bhatia (1973) for example, showed that the soluble protein component of endosperm N had a much lower rate and an earlier accumulation end point during endosperm development than that of the storage protein components. How the relative distribution of protein subunits is determined at the gene level is not yet known (Cole et al. 1981).

The observation that region A3 contained the largest peak with a MW of 36,750 (most concentrated band) in almost 70% of the accessions (Table 3) and that the relative distribution of protein was similar (high density bands 70K and above, dense bands between 70K and 33K and medium density bands in the region below 33K) (Fig. 2) suggests there is a "typical" *T. dicoccoides* regardless of the environment from which it came.

There also was variation for the largest peak (most dense band) within each MW region (Table 4). We were

surprised to find that the CV of the largest peak within each of the regions A1, A2, A3 and A5 was about the same (near 11%) but not for A4 (4.0%). We have no ready explanation for this behavior except to say that the largest peaks may be highly conserved protein bands. This is especially true for the largest peak of region A4 which was observed in at least 80% of the samples. Of interest to breeders is that an unusually high staining band of approximately 60,000 MW corresponding to the most intense peak of region A2 shows up in advanced hexaploid lines derived from crosses with an accession of *T. dicoccoides* and selected for high quality and high protein (Grama et al. 1984).

This study demonstrates that considerable phenotype variability exists for the kernel proteins of *T. dicoccoides*. Of the five regions defined, A1 and A5 had the greatest quantitative variation, CV=32% and 17%, respectively. Region A3 (49,000 to 31,150) had the greatest relative amount of protein (34.6%). The region with the smallest protein contribution (7.06%) was A1 (above 70K MW). We strongly suspect region A5 to have bound a disproportionate amount of dye because these proteins are mainly albumin and globulins, rich in basic amino acids such as lysine. After adjusting for their greater dye binding, the contribution from region A5 decreased from 21.4% to 10.7%. Accessions with the highest relative amounts of protein in A5 were identified as possible sources of genes to improve the nutritional quality of our modern varieties. These accessions are listed in Table 7. Another group of interest includes those accessions with a higher proportion of their proteins in A1 (high MW glutenin) compared to 'Chinese Spring'. This group of accessions could be of interest for improving the baking quality of bread wheat. Payne and Corfield (1979) showed a positive correlation between the MW of native glutenin and the proportion of HMW subunits (region A1) in glutenin. This is important to baking quality because previous work by Huebner and Wall (1976) had shown that the higher the ratio of HMW to LMW the better was the bread-making quality of the flour. Accessions with a higher relative area in A1 than 'Chinese Spring' also are listed (Table 7).

Visual observations of the gels and densitometer scans indicated a distinct pattern common to all of the *T. dicoccoides* accessions examined, as is evident from the summary scan shown in Fig. 2. A light region of mostly faint bands appeared down to about 70,000 MW, then the darkest region of thick bands appeared from 70,000 to about 33,300 MW. Finally, a highly homogeneous region of medium intensity was found below 33,300 MW. This observation, coupled with the finding that in almost 70% of the accessions the most intense peak was in region A3, indicated that there is a typical pattern for this species separate from environmental influences.

References

Cole EW, Fullington JG, Kasarda D (1981) Grain protein variability among species of *Triticum* and *Aegilops:* Quantitative SDS-PAGE studies. Theor Appl Genet 60:17–30

Fullington JG, Cole EW, Kasarda DD (1980) Quantitative SDS-PAGE of total proteins from different wheat varieties. J Sci Food Agric 31:43–53

Fullington JG, Cole EW, Kasarda DD (1983) Quantitative SDS-PAGE of total proteins extracted from different wheat varieties: effects of protein content. Cereal Chem 50:65–70

Grama A, Gerechter-Amitai ZK, Blum A, Rubenthaler GL (1984) Breeding bread wheat cultivars for high protein content by transfer of protein genes from *Triticum dicoccoides.* In: Cereal grain protein improvement. Int Atomic Energy Agency, Vienna, pp 145–153

Grama A, Ebercon A, Sinemena B, Blum A, Gerechter-Amitai ZK (1982) *Triticum dicoccoides* as a source of high protein in bread wheat. Wheat Newslett 28:74

Huebner FR, Wall JS (1976) Fractionation and quantitative differences of glutenin from what varieties varying in baking quality. Cereal Chem 53:258–269

Johnson VA, Allan RE, Gallun RL, Loegering WQ (1972) Wheat. In: Genetic vulnerability of major crops, chap 9. National Academy of Sciences USA, pp 119–154

Mansur-Vergara L (1983) Computer analysis of the kernel protein genetic variability in *Triticum dicoccoides* collections. Master's Thesis, Washington State University, Pullman

Mansur-Vergara L, Konzak CF, Gerechter-Amitai ZK, Grama A, Blum A (1985) A computer-assisted examination of the storage protein genetic variation in 841 accessions of *Triticum dicoccoides.* Theor Appl Genet 69:79–86

Mitra R, Bhatia CR (1973) Studies on protein biosynthesis in developing wheat kernels. In: Kaufmann B, Lewis M (eds) Nuclear techniques for protein improvement. STI/PUB/320, Int Atomic Energy Agency, Vienna, pp 379–389

Nevo E, Golenberg E, Beiles A (1982) Genetic diversity and environmental associations of wild wheat, *Triticum dicoccoides,* in Israel. Theor Appl Genet 52:241–254

Osborne TB (1907) The proteins of the wheat kernel. Carniege Inst, Washington, Publ 84, Washington, DC

Payne PI, Corfield KG (1979) Subunit composition of wheat glutenin proteins, isolated by gel filtration in a dissociating medium. Planta 145:83–89

Payne PI, Corfield KG, Holt LM, Blackman JA (1981) Correlations between the inheritance of certain high-molecular-weight subunits of glutenin and bread making quality in progenies of six crosses of bread wheat. J Sci Food Agric 32:51–60

Simmonds DH (1981) Wheat proteins: their chemistry and nutritional potential. In: Evans LT, Peacock WJ (eds) Wheat science: today and tomorrow. Cambridge University Press, Cambridge London, pp 149–166