Theor Appl Genet (1987) 74: 827–836

Wheat storage proteins: diversity of HMW glutenin subunits in wild emmer from Israel

1. Geographical patterns and ecological predictability

E. Nevo [1] and P. I. Payne [2]

[1] Institute of Evolution, University of Haifa, Haifa 31999, Israel
[2] Plant Breeding Institute, Trumpington, Cambridge CB2 2LQ, UK

Received March 26, 1987; Accepted April 2, 1987
Communicated by H. F. Linskens

Summary. The diversity of HMW glutenin subunits in the tetraploid wild progenitor of wheat, *Triticum turgidum* var. *dicoccoides* was studied electrophoretically in 231 individuals representing 11 populations of wild emmer from Israel. The results show that (a) The two HMW glutenin loci, *Glu-A1* and *Glu-B1*, are rich in variation, having 11 and 15 alleles, respectively. (b) Genetic variation in HMW glutenin subunits is often severely restricted in individual populations, supporting an island population genetic model. (c) Significant correlations were found between glutenin diversity and the frequencies of specific glutenin alleles and physical (climate and soil) and biotic (vegetation) variables. Our results suggest that: (a) at least part of the glutenin polymorphisms in wild emmer can be accounted for by environmental factors and (b) the endosperm of wild emmer contains many allelic variants of glutenin storage proteins that are not present in bread wheat and could be utilized in breeding varieties with improved bread-making qualities.

Key words: *Triticum turgidum* var. *dicoccoides* – HMW glutenin polymorphism – Genetic model – Wheat quality – Environmental influence

Introduction

Endosperm proteins of wheat grain consist predominantly of the two prolamin storage proteins, gliadin and glutenin, which are present in roughly equal amounts in the endosperm. Glutenin consists of approximately 20% high-molecular-weight (HMW) glutenin subunits and 80% low molecular weight (LMW) subunits. The biochemistry and genetics of endosperm proteins in bread wheat have been studied intensively in recent years (Kasarda et al. 1976; Konzak 1977; Payne et al. 1980, 1984; Galili and Feldman 1985) due to their great importance in the processing of wheat flour into bread (Wall 1979). The genes that code for the prolamins occur at nine complex loci on six different chromosomes (Payne et al. 1984).

In bread wheat, *Glu-A1*, *Glu-B1* and *Glu-D1* contain the genes for HMW subunits of glutenin and are close to the centromere on the long arms of chromosomes 1A, 1B and 1D, respectively (Payne et al. 1980). Each of these loci displays allelic variation (Payne et al. 1981a; Galili and Feldman 1983, 1985) which is partly responsible for differences among varieties in protein quality for making bread (Payne et al. 1981b). Novel storage proteins are also being screened in landraces of primitive agriculture for possible incorporation into the genomes of commercial wheats for developing new varieties with improved bread-making quality (Payne et al. 1984). Promising sources of genes coding for novel proteins are also the diploid and tetraploid wild relatives of wheat (Law and Payne 1983).

The diverse, wild gene pools of cultivated plants contain many genes of economic importance that can be transferred to and improve crops (reviewed in Feldman 1983). The use of wild genetic resources in wheat improvement was discussed by Feldman (1979) and Feldman and Sears (1981). Modern plant breeding practices have reduced the range of genetic variability amongst cultivars of many crops, including wheat. As indicated by several authors (e.g. Feldman and Sears 1981; Nevo et al. 1982; Nevo 1983, 1987a) a major hope for future crop improvements lies in the utilization of the rich gene pool of the plant's wild relatives. Wild emmer wheat, *Triticum turgidum* var. *dicoccoides*, genome constitution AB, is the immediate progenitor of most tetraploid and hexaploid wheats and appears to

be a suitable donor for the transfer of desirable genes into cultivated gene pools.

The wild gene pool of emmer wheat, in Israel, was investigated in a multidisciplinary study at the Institute of Evolution, University of Haifa (Nevo 1983). The Near East Fertile Crescent in general, and Israel in particular, are the centers of origin and diversity of wild emmer. In this area, wild emmer developed, during its long evolutionary history, wide and adaptive genetic diversity. Wild emmer contains rich genetic diversity for multiple disease resistances, agronomic traits of economic significance, variation in protein quantity and composition, and resistance to diverse ecological stresses. This genetic diversity is geographically structured and is partly predictable by ecology and allozyme markers (Nevo 1987b). Consequentally, conservation and utilization programs should maximize sampling strategies by using the ecological-genetic factors and allozyme and DNA markers as effective predictors.

In view of the great importance of wild emmer as a potential genetic resource for wheat improvement, we have now tested it for glutenin subunit diversity across its major ecogeographical regions in Israel. The objective of this communication is to describe the genetic polymorphism of HMW glutenin subunits of wild emmer in Israel, and to correlate the revealed diversity with ecological factors, including physical (soil and climate) and biotic (plant communities) factors. We demonstrate here that glutenin diversity of wild emmer is rich and partly predictable by ecological factors.

Materials and methods

Ecological and genetic background

Wild tetraploid emmer wheat, *Triticum turgidum dicoccoides*, is the immediate progenitor of most cultivated wheats. Wild emmer is distributed over the Near East Fertile Crescent (Harlan and Zohary 1966). Its center of distribution is in the drainage area of the Upper Jordan Valley in Israel, where it grows as an annual in steppe-like herbaceous formations, and in the *Quercus ithaburensis* park forest belt. Its ecological background was described by Zohary (1969, 1970) and its ecological genetics by Nevo et al. (1982), which tested allozyme variation encoded by about 50 gene loci in 457 individuals representing 12 populations from Israel.

Sampling

Individual plants of *T. turgidum dicoccoides* were collected at random from 12 populations in April-June of the successive years 1975–1978 across the major ecogeographical regions of wild emmer in Israel. The localities and their ecogeographical background appear in Table 1, and in Fig. 4. We tested the seeds collected originally in the field for allozymic variation encoded by about 50 gene loci. Here we analyzed 231 plants belonging to 11 populations out of the 12 populations reported in Nevo et al. (1982), for glutenin subunit diversity (Table 1, Fig. 4). The plants analyzed were selected from the original 457 original seeds.

Table 1. The ecogeographical background of 11 populations of wild emmer wheat, *Triticum turgidum dicoccoides*, in Israel. Climatic values were taken from the Atlas of Israel (1970) and from multiple-year records of the Meteorological Service of Israel. Variables are as follows. Geographical: Ln = Longitude, in decimals; Lt = Latitude, in decimals; Alt = Altitude, in m. Temperature: Tm = mean annual temperature; Ta = mean August temperature; Tj = mean January temperature; Td = mean seasonal temperature difference; Tdd = mean day-night temperature difference; Trd = mean number of tropical days; Sh = mean number of Sharav days, i.e., hot and dry days; Ev = mean annual evaporation. Water availability: Dw = mean number of dew nights in summer; Rn = mean annual rainfall, in mm; Rd = mean number of rainy days; Hu14 = mean humidity at 14:00; Huan = mean annual humidity; Th = Thornthwaite's moisture index. Edaphic: So = soil type: 1 = Rendzinna; 2 = Terra Rossa; 3 = Basalt. Biotic: Pl = Climax plant community: 1 = marginal Mediterranean desert batha; 2 = tragacantic batha (phyrgana); 3 = marginal Mediterranean batha with *Zizyphus loti*; 4 = marginal Mediterranean batha with perennial weeds; 5 = primary climax of *Quercus ithaburensis* (sagital); 6 = primary climax of *Q. ithaburensis* (batha); 7 = *Quercus ithaburensis* Park forest; 8 = batha of *Sarcopterium spinosum* (primary evergreen climax); *Pistacia palaestina*; Sz = estimate of population size: 1 = small, 2 = intermediate, 3 = large

Populations	Ecogeographical variables																			
	Ln	Lt	Alt	Tm	Ta	Tj	Td	Tdd	Rn	Rd	Hu 14	Hu an	Dw	Sh	Th	Trd	Ev	So	Sz	Pl
1 Mt. Dov	35.73	33.30	1,300	11	21	3	18	6	1,400	66	48	60	60	80	–	0	150	2	2	2
2 Qazrin	35.67	33.02	350	18	26	10	16	12	530	50	43	58	58	50	–	60	155	3	3	5
3 Yehudiyya	35.70	32.93	200	19	27	11	16	12	550	47	42	58	58	50	–	100	160	3	3	7
4 Rosh Pinna	35.52	32.95	700	18	25	9	16	10	697	50	48	58	50	75	– 10	35	150	2	3	4
5 Tabigha	35.53	32.90	0	24	32	15	17	10	436	45	45	57	58	60	– 30	120	160	3	3	3
6 Bat Shelomo	35.02	32.60	75	20	26	13	13	10	650	55	58	68	77	40	– 10	30	150	1	2	6
7 Mt. Gerizim	35.28	32.20	800	17	23	8	15	9	700	47	45	60	42	–	10	0	155	2	2	8
8 Kokhav Hashahar	35.34	31.95	600	20	28	12	16	12	400	40	45	59	30	80	– 30	25	165	2	2	8
9 Taiyiba	35.35	31.92	450	19	26	10	16	12	400	40	44	58	30	80	– 10	25	165	2	2	1
10 Sanhedriyya	35.22	31.80	800	17	24	9	15	9	548	44	51	62	44	102	– 10	0	155	2	1	8
11 Bet Meir	35.03	31.80	500	19	26	11	15	9	582	44	47	60	61	70	– 10	100	160	2	1	9

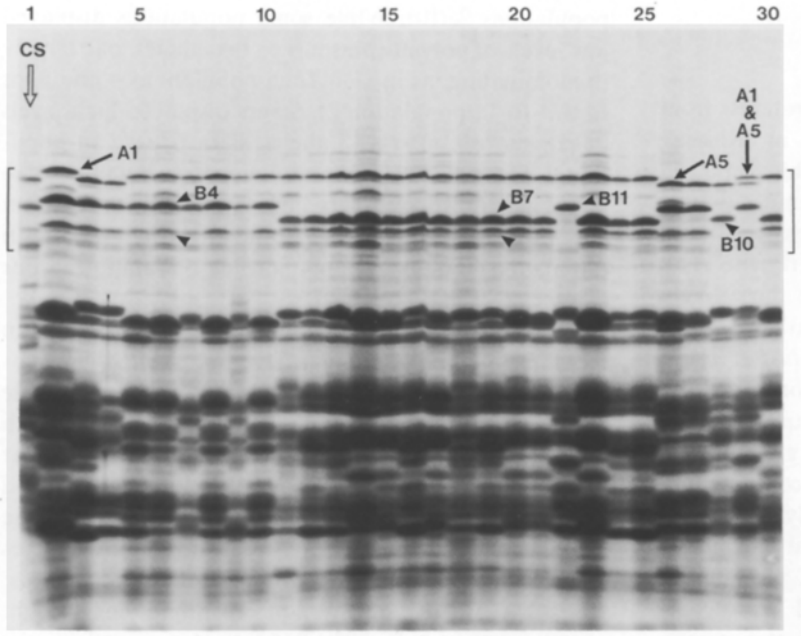


Fig. 1. SDS-PAGE of grain proteins found in a population of wild emmer from Kokhav Hashahar, near Jerusalem. The region of the gel containing the HMW glutenin subunits is marked by *brackets*. The proteins thought to be coded by the *Glu-A1* and *Glu-B1* loci are prefixed with letters A and B, respectively. The HMW glutenin subunits of bread-wheat cultivar Chinese Spring are separated in *slot 1*. They are, with increasing mobility, *2* (chromosome 1D), *7+8* (1B), and *12* (1D)

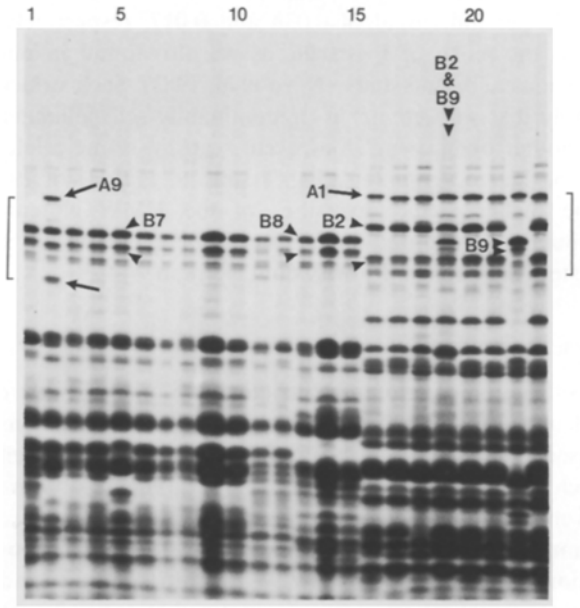


Fig. 2. SDS-PAGE of grain proteins in two wild emmer populations: *slots 1–15:* Qazrin; *slots 16–23:* Yehudiyya

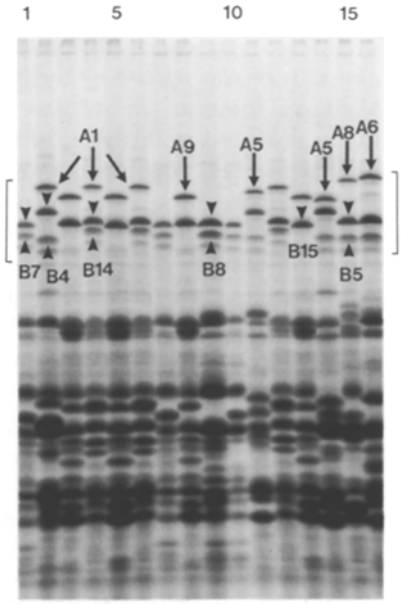


Fig. 3. SDS-PAGE of a range of wild emmers to show the extent of allelic variation in HMW glutenin subunits. Grains were taken from populations Qazrin (*slot 1*). Yehudiyya (*slots 2, 7, 8,* and *10*). Tabigha (*slot 12*). Bet-Meir (*slot 14*). Rosh Pinna (*slot 11*). Bat Shelomo (*slot 9*), Mt Hermon (*slot 15*), Sanhedriyya (*slots 3–6,* and *13*), and Taliyiba (*slot 16*)

Electrophoresis

Allelic variation was determined by analyzing half grains by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE), as described previously (Payne et al. 1980, 1981b). The method fractionates proteins mainly according to their molecular weight. Samples from different locations were analyzed many times in different orders on the gels to establish the relative mobilities of all the principle subunits.

Statistical analysis

Spearman rank correlations were computed between all variables. Stepwise multiple regression analysis (SPSSx 1986) was employed to determine whether environmental factors were associated with gene frequencies of HMW glutenin subunits coded as *Glu-A1* and *Glu-B1*. Multivariate analysis was conducted on 10 populations for *Glu-A1* and for *Glu-B1*.

Results

Pattern of variation

The banding patterns of endosperm proteins from individual grains of emmer, collected in Kokhav Hashahar, are shown in Fig. 1. Appreciable variation occurs, but only that due to HMW glutenin subunits will be discussed here. On the basis of extensive studies on this group of proteins in bread wheat (Payne et al. 1981a), the subunits were subdivided into those likely to be coded by genes on chromosome 1A at the *Glu-A1* locus and those on chromosome 1B at *Glu-B1*. These tentative conclusions have since been confirmed by Levy and Feldman (1987) by direct genetical analysis of wild emmer. The HMW subunits of glutenin are marked by brackets in Fig. 1, and they contain both major and minor components. Very recently it has been shown by Levy and Feldman (1987) that the minor components are derived from the major proteins by post-translational modification in accord with complementary work with bread wheat. For this reason only the major components have been studied here.

Amongst the 31 individuals shown in Fig. 1, there are two alleles of *Glu-A1* described, coding for protein groups A1 and A5, and four of *Glu-B1* coding for B4, B7, B10 and B11. Allelic variation at *Glu-A1* and at *Glu-B1* is therefore rather restricted in this population of emmer. Similarly, when other populations were studied, for instance at Yehudiyya and Qazrin (Fig. 2), the number of HMW subunits was also small but the particular alleles present were highly distinctive for a population. The range of variation in HMW subunits from different populations of wild emmer is shown in Fig. 3. As in bread wheat, there are usually two major subunits produced from *Glu-B1* (B2, B4, B7, B8, and B9), although occasionally only one was detected (B10, B11). The normal situation for *Glu-A1* is the presence of one subunit (eg. A1, A5) or none at all. Unlike bread wheat, two subunits (e.g. A9) are also produced sometimes (see also Levy and Feldman 1987).

Altogether from all the 11 populations studied here we found 11 alleles for *Glu-A1* and 15 for *Glu-B1*. The frequencies of alleles at *Glu-A1* and *Glu-B1*, for all the 11 populations of wild emmer tested, are presented in Table 2 and illustrated in Fig. 4.

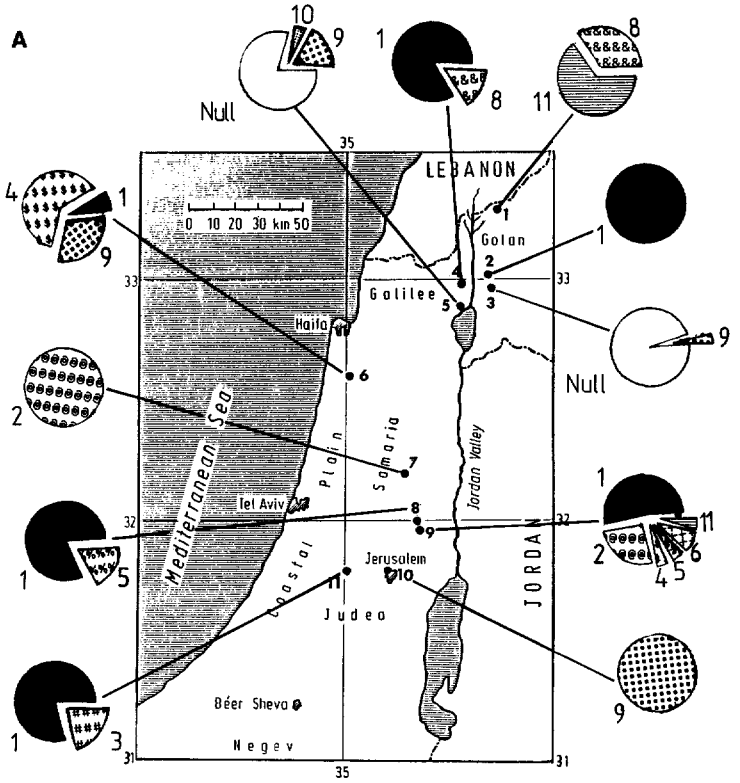
A wide range of polymorphic levels characterizes the HMW glutenin subunits in the 11 populations, ranging from monomorphism to high polymorphism. As indicated by the estimate of gene diversity (He or expected heterozygosity), which appears in the bottom of each of the 2 loci of each population, some populations are purely monomorphic in both systems (e.g. population 7: however since only two individuals were analyzed in this population this must be a sample-biased example), or in one only (e.g., *Glu-A1* in populations 2, 10). While some populations displayed low levels of polymorphism (i.e., two alleles, one largely predominating, as in *Glu-A1* in populations 4 and 8, or in *Glu-B1* in population 2), others displayed high levels of polymorphism (i.e., 3 alleles as in *Glu-A1* in population 6, or *Glu-B1* in population 5). Finally, some populations displayed very high levels of polymorphism (e.g. six alleles in *Glu-A1* and five in *Glu-B1* in population 9, the most polymorphic population in our study.

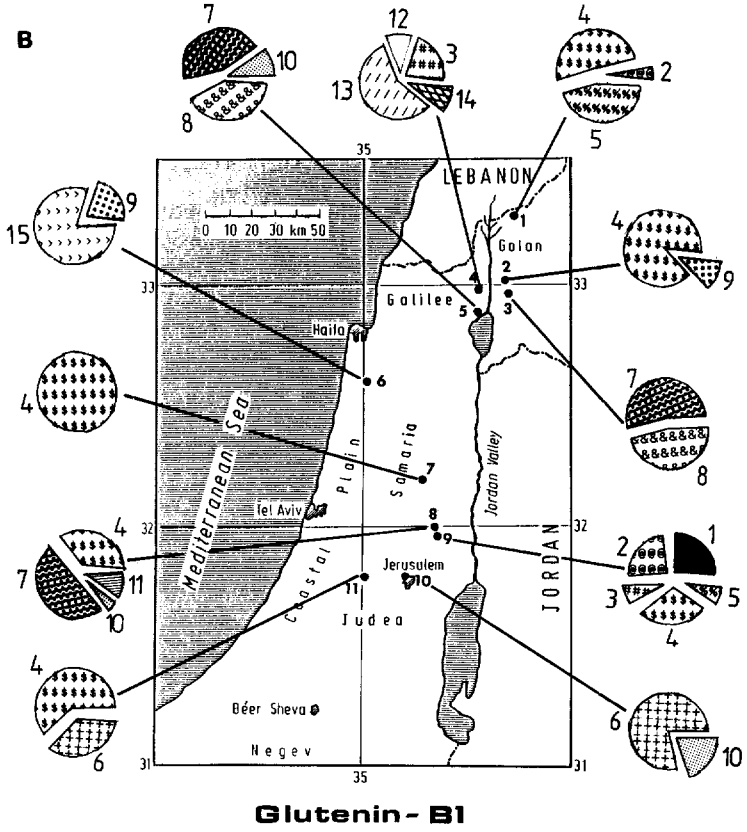
A summary of the genetic data of the two glutenin loci for the 11 populations is given in Table 3, in terms of three genetic indices. Clearly, there was appreciable variation which greatly varied geographically, for each of the indices: mean number of alleles per locus per population (A), polymorphism per population (P), and expected heterozygosity or genic diversity (He) (i.e., the average per locus probability that two random gametes from the population carry different alleles at a locus; Nei 1975). The mean of Wright's fixation index (Wright 1965) was $F = 1.0$ in most populations, except in populations 2 and 8 which had observed heterozygosity levels per individual of 0.038 and 0.017, respectively. Thus the mean of F is 0.96, as we also found in our previous allozyme study (Nevo et al. 1982). Such values, imply that wild emmer is predominantly self pollinated although outcrossing does occur, causing some alleles to be heterozygous (eg., Fig. 1 slot 29, Fig. 2 slot 19). The highest genic diversities of the HMW glutenin subunits were found in the steppic populations: 9 in eastern Samaria, and 1 on Mount Hermon.

Genetic distances

Coefficients of glutenin genetic similarity (I) were calculated for paired comparisons of all 11 populations based on the normalized identity of all loci between each pair of populations (Nei 1972). The results are given in Table 4. Clearly, the estimates in the table are commensurate only for comparing the 11 wheat populations studied here for HMW glutenin subunits, and not as a general estimate of genetic similarity, as was done on a multilocus allozymic basis for wild emmer (Nevo et al. 1982). The mean value of I here for glutenin subunits was 0.357 (range, 0.000–0.983). Populations with $I = 0$ (e.g., populations 1–3) displayed alternative alleles at both glutenin loci. Alternatively, populations with $I = 0.918$ (e.g., 2–11) shared mostly the same range of allelic variation. Only two pairs of populations displayed such a high glutenin similarity in our study (i.e., 3–5 and 2–11).

The correlation coefficient of glutenin diversity with geographic distance between populations was very low and nonsignificant. This indicates that geographically close populations might be very different in their





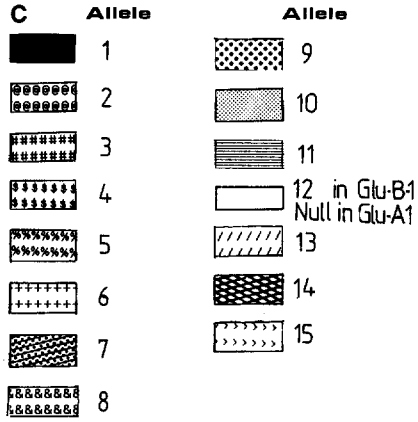


Fig. 4A–C. Pie diagrams displaying the percentage of the 11 alleles of *Glu-A1* and 15 alleles of *Glu-B1* in 11 populations of wild emmer wheat, *Triticum turgidum dicoccoides,* and their geographical location in Israel. Population numbers as in Table 1. **A** *Glu-A1;* **B** *Glu-B1.* **C** Key to diagrams; alleles are numbered as in Table 2

Table 2. Allele frequencies and gene diversity (He)[a] at 2 glutenin loci in 11 populations of *Triticum turgidum dicoccoides* in Israel

Locality		1[b]	2	3	4	5	6	7	8	9	10	11	Mean
n		20	26	23	19	24	26	2	30	23	19	19	231
Locus allele													
Glu-A1	1	0	1.000	0	0.842	0	0.077	0	0.817	0.522	0	0.789	0.413
	2	0	0	0	0	0	0	1.000	0	0.261	0	0	0.035
	3	0	0	0	0	0	0	0	0	0	0	0.211	0.017
	4	0	0	0	0	0	0.615	0	0	0.043	0	0	0.074
	5	0	0	0	0	0	0	0	0.183	0.043	0	0	0.028
	6	0	0	0	0	0	0	0	0	0.087	0	0	0.009
	8	0.350	0	0	0.158	0	0	0	0	0	0	0	0.043
	9	0	0	0.043	0	0.167	0.308	0	0	0	1.000	0	0.139
	10	0	0	0	0	0.042	0	0	0	0	0	0	0.004
	11	0.650	0	0	0	0	0	0	0	0.043	0	0	0.061
	null	0	0	0.957	0	0.792	0	0	0	0	0	0	0.177
He		0.455	0	0.083	0.266	0.344	0.521	0	0.299	0.647	0	0.322	0.268
Glu-B1	1	0	0	0	0	0	0	0	0	0.261	0	0	0.026
	2	0.050	0	0	0	0	0	0	0	0.261	0	0	0.030
	3	0	0	0	0.211	0	0	0	0	0.087	0	0	0.026
	4	0.500	0.885	0	0	0	0	1.000	0.367	0.304	0	0.623	0.281
	5	0.450	0	0	0	0	0	0	0	0.087	0	0	0.048
	6	0	0	0	0	0	0	0	0	0	0.789	0.368	0.095
	7	0	0	0.522	0	0.458	0	0	0.500	0	0	0	0.165
	8	0	0	0.478	0	0.417	0	0	0	0	0	0	0.091
	9	0	0.115	0	0	0	0.192	0	0	0	0	0	0.035
	10	0	0	0	0	0.125	0	0	0.033	0	0.211	0	0.035
	11	0	0	0	0	0	0	0	0.100	0	0	0	0.013
	12	0	0	0	0.105	0	0	0	0	0	0	0	0.009
	13	0	0	0	0.579	0	0	0	0	0	0	0	0.048
	14	0	0	0	0.105	0	0	0	0	0	0	0	0.009
	15	0	0	0	0	0	0.808	0	0	0	0	0	0.091
He		0.545	0.204	0.499	0.598	0.601	0.311	0	0.604	0.756	0.332	0.465	0.447

[a] The index of gene diversity (He) for each locus is recorded below each set of frequencies
[b] Population numbers refer to those listed in Table 1

Table 3. Glutenin diversity based on 2 loci in 11 populations of *Triticum turgidum dicoccoides* in Israel

Locality	Sample size (*n*)	Mean no. of alleles per locus (A)	Mean proportion of loci			Genic diversity (He)
			Polymorphic per population (P-5%)	Heterozygous per individual (H)		
				Mean	SE	
1. Mt. Dov	20	2.5	1.0	0.0	0.0	0.500
2. Qazrin	26	1.5	0.5	0.038	0.038	0.102
3. Yehudiyya	23	2.0	0.5	0.0	0.0	0.291
4. Rosh-Pinna	19	3.0	1.0	0.0	0.0	0.432
5. Tabigha	24	3.0	1.0	0.0	0.0	0.472
6. Bat-Shelomo	26	2.5	1.0	0.0	0.0	0.416
7. Mt. Gerizim	2	1.0	0.0	0.0	0.0	0.0
8. Kokhav Hashahar	30	3.0	1.0	0.017	0.017	0.452
9. Taiyiba	23	5.5	1.0	0.0	0.0	0.701
10. Sanhedriyya	19	1.5	0.5	0.0	0.0	0.166
11. Bet-Meir	19	2.0	1.0	0.0	0.0	0.399
Mean	231	2.5	0.773	0.006	0.002	0.357
Range		1.0–5.5	0–1	0–0.038		0–0.701

Table 4. Coefficients of genetic similarity (I), based on two glutenin loci, between 11 populations of *Triticum turgidum dicoccoides* in Israel. Mean 0.357; range 0.000–0.983

Locality	2	3	4	5	6	7	8	9	10	11
1 Mt. Dov	0.330	0.0	0.052	0.0	0.0	0.354	0.175	0.301	0.0	0.288
2 Qazrin	*****	0.0	0.590	0.0	0.068	0.467	0.813	0.764	0.0	0.918
3 Yehudiyya	**********		0.0	0.983	0.010	0.0	0.209	0.0	0.028	0.0
4 Rosh-Pinna	**********			0.0	0.056	0.0	0.616	0.556	0.0	0.569
5 Tabigha	**********				0.046	0.0	0.217	0.0	0.145	0.0
6 Bat-Shelomo	**********					0.0	0.056	0.080	0.220	0.051
7 Mt. Gerizim	**********						0.248	0.517	0.0	0.407
8 Kokhav Hashahar	**********							0.674	0.005	0.763
9 Taiyiba	**********								0.0	0.713
10 Sanhedriyya	**********									0.205
11 Bet-Meir	**********									

Table 5. Differentiation of glutenin diversity within and between 11 populations of *Triticum turgidum dicoccoides,* in Israel. Ht = total gene diversity; Hs = average gene diversity within populations; Dst = average gene diversity between populations; Dm = average of interpopulational diversity only; Gst = gene diversity between population, relative to Ht; Rst = interpopulational diversity, relative to Hs

Locus	Alleles	Sample	Ht	Hs	Dst	Gst	Dm	Rst
Glu-A1	11	231	0.765	0.295	0.471	0.615	0.518	1.758
Glu-B1	15	231	0.859	0.486	0.373	0.434	0.410	0.844
Mean			0.812	0.390	0.422	0.519	0.464	1.189

glutenin structure (e.g., *Glu-B1* in populations 2–3, 10 km apart, have no alleles in common: similarly, both *Glu-A1* and *Glu-B1* in populations 9–10, 20 km apart, have no alleles in common, i.e., I = 0.00. By contrast, geographically distant populations might be very similar in their glutenin composition (e.g., populations 2–11, 145 km apart, I = 0.918). In other words, the structure of HMW glutenin subunits genes in wild emmer displayed genetic population differentiation, which was even more distinct than that of allozymic differentiation (Nevo et al. 1982). Combining the present study with allozyme diversity will clearly lead to very high genetic distances between populations.

Glutenin gene differentiation within and between populations

Gene diversity (Ht) of a subdivided population can be analysed into its components, i.e. the gene diversities within and between populations. The results provide measures of the average (Dst), absolute (Dm), relative (Gst) and inter-versus intrapopulational (Rst) degree of gene differentiation among subpopulations, where Hs is the mean gene diversity in a population, Ht is the total diversity and Ht = Hs + Dst (Nei 1975). Dm can be used for comparing the degrees of gene differentiation in different organisms. The analysis of glutenin gene diversity of the 11 populations of wild emmer is given in Table 5. On average, 48% of the total glutenin diversity in both systems exist within populations and 52% between populations (deduced from the Gst in Table 5). These results indicate, once again, substantial gene differentiation between populations, supporting an island genetic population model for wild emmer.

Environmental correlates of glutenin polymorphism

Correlation among environmental variables. The physical and biotic environmental variables (geographical, climatic, edaphic, and vegetational) appear in Table 1. Clearly, some of the variables are highly correlated, e.g., latitude-longitude, Spearman rank correlation, $r_s = 0.80$; mean annual temperature and mean January temperature, $r_s = 0.97$; number of rainy days and evaporation, $r_s = -0.88$. Intermediate level correlations occur between annual humidity and mean August temperature, $r_s = -0.51$. Low correlations occur between number of dew nights in summer and mean annual temperature, $r_s = 0.04$.

Correlations among glutenin diversity and environmental variables. Genic diversity (He) in HMW glutenin subunits increases towards the steppic, more arid regions (He-plant formation, $r_s = -0.66$, $P < 0.05$). Likewise, gene diversity of *Glu-B1* is positively correlated with evaporation and negatively correlated with number of dew nights in summer ($r_s = 0.56$ and -0.60, respectively, $P < 0.10$). The frequency of allele *8* of *Glu-A1* displayed high correlation with rainfall, Rn, and mean tempera-

Table 6. Coefficients of multiple regressions (R^2) of glutenin diversity and allele frequencies as the dependent variables and geographic, climatic, ecological, edaphic, and plant formation variables in 10 populations of *Triticum turgidum dicoccoides* as independent variables (for abbreviations see Table 1)

	n	Stepwise model		
Glu-A1 1	10	Tdd 0.172 ns	Tdd Alt 0.327 ns	
Glu-A1 9	10	Hu14 0.254 ns	Hu14 Sh 0.450 ns	Hu14 Sh So 0.760*
He of *Glu-A1*	10	So 0.306 @	So Pl 0.660*	So Pl Ev 0.828**
Glu-B1 7	10	Ta 0.467*	Ta Tm 0.657*	Ta Tm Pl 0.747*
Glu-B1 10	10	Sh 0.267 ns	Sh Alt 0.512 @	Sh Alt Rd 0.753*
He of *Glu-B1*	10	Dw 0.332 @	Dw Pl 0.508 @	
He of *Glu-A1 &B1*	10	Pl 0.323 @	Pl Lt 0.531 @	

* $P<0.05$; ** $P<0.01$; @ = $P<0.10$; ns = $P>0.10$

ture in January, Tj, ($r_s=0.70$ and -0.67, respectively; $P<0.05$ for both). Thus glutenin diversity is correlated with vegetation differentiation and the climatic regime. By contrast there was no correlation between population size (see estimates in Table 1) and glutenin He ($r_s=0.013$).

Multiple regression analysis

A test of the best predictors of glutenin diversity (He) and allele diversity of specific few alleles (those that occurred in at least 3 populations) in 10 populations was conducted by stepwise multiple regression analysis, MR (SPSSx 1986). The MR employed the genetic variables as dependent variables and geographic, climatic, edaphic, and vegetational variables as independent variables. Soil and plant formation were ranked according to texture and aridity, respectively (see values in Table 1). The results are given in Table 6.

Temperature, water-availability, soil, and plant variables in combination, accounted for a significant proportion of the variation in glutenin diversity. For example a 3-variable combination consisting of soil type (So), plant formation (Pl) and evaporation (Ev) significantly explained 83% of the glutenin A1 variance expressed by He in 10 populations. Likewise, the variance in frequency of allele *9* in *Glu-A1* was explained by 3-variable combination of mean humidity at 14:00 (Hu14), number of sharav days (Sh) and soil type (So), $R^2=0.760$, $P<0.05$. The variance in frequency of allele *7* in *Glu-B1* was explained by a 3-variable combination of mean August temperature (Ta), mean annual temperature, (Tm) and plant formation (Pl), $R^2=0.75$. Thus, a substantial amount of the variance in HMW glutenin subunits can be accounted for ecologically.

Discussion

The HMW glutenin subunits are storage endosperm proteins that are important both theoretically as well as practically. Theoretically, they comprise two multiallelic gene families. As the level of polymorphism is very high, they can substantially contribute to the analysis of evolutionary forces causing population genetic structure and differentiation. Practically, the subunits studied are components of glutenin, which plays a fundamental part in the processing of such foods as chapatis and noodles.

Genetic differentiation of HMW glutenin subunits

Three features characterize glutenin differentiation in natural populations of wild emmer in Israel: (1) extensive variability, (2) extreme population differentiation, and (3) correlation with ecological factors. First, allele diversity is remarkable in the glutenin subunits, similar to that of hordein storage proteins in wild barley, *Hordeum spontaneum* in Israel (Nevo et al. 1983). Althogether, 15 hordein-1 phenotypes and 16 hordein-2 phenotypes have been described in wild barley in Israel. In line with this extremely high hordein diversity, we found 10 *Glu-A1* and 15 *Glu-B1* alleles in the present study. Similarly, our unpublished studies also reveal extensive allelic variation in the gliadin proteins of emmer. The high levels of allele diversities of hordein and glutenin storage proteins contrast sharply with the relatively lower levels of allozyme allele diversity per locus in each population of wild barley, and wild emmer in Israel. The mean number of alleles per locus per population, A, for allozymes in wild barley (Nevo et al. 1979) and wild emmer (Nevo et al. 1982), were 1.48 and 1.33 for 28 and 50 loci, respectively. By contrast, A = 2,500 for the 2 glutenin loci in wild emmer. Our estimates of allelic variation in the HMW glutenin subunits of emmer is also clearly an underestimation. The proteins have only been fractionated in one dimension and different subunits of very similar molecular weight will not be resolved. A good example of this are the subunits labelled A1 in Fig. 2. A careful examination reveals two alternative proteins of very similar mobility. As only exceptional separations showed this polymorphism, they were not subclassified. However, different proteins of more similar molecular weights would never be resolved by this technique.

The distribution of alleles for HMW glutenin subunits in wild emmer is nonrandom. Population differentiation is extreme, as clearly indicated by the analysis of genetic distance and glutenin differentiation within and between populations (Tables 4 and 5). Populations vary dramatically in the range of their HMW glutenin subunit alleles, regardless of geographic distance separating them, or population size. On the other hand, glutenin differentiation is associated with the environment. Climatic, edaphic, and vegetation factors in combination can account for most of the variance in overall genic diversity and in specific alleles. Most alleles are distributed locally or sporadically, suggesting an extreme population genetic island model as was previously shown by allozyme differentiation in wild emmer (Nevo et al. 1982; Nevo 1983).

The geographical patterns of HMW glutenin alleles and their ecological correlates suggest that they may in part be due to natural selection, and cannot be explained by migration and genetic drift as major explanatory models. The fact that nearby populations are very different in glutenin subunit content, and, by contrast, that distant populations are very similar, indicates that migration plays little or no role in population genetic differentiation of these proteins. Neither can genetic drift be easily invoked to explain allele distribution. Wild emmer distribution in Israel is characterized by discontinuities and by "archipelago" population structure. If HMW glutenin subunits were neutral, their spatial pattern would be expected to display marked interpopulation differentiation, independent of climate, soil, or vegetation. Yet glutenin differentiation is significantly correlated with these environmental factors. The sporadic and localized glutenin allele distribution of wild emmer are in accordance, however, with the ecological background. The island population genetic structure strongly suggest distinct local adaptations through diversifying selection, rather than drift.

The high diversity of HMW glutenin subunits in marginal steppic populations, is similar to that of allozyme (Nevo et al. 1982), protein (Nevo et al. 1986) and ribosomal DNA diversity (Flavell et al. 1986). We will devote another paper in this series to these remarkable interrelationships. It does appear that the same reasoning concerning the evolutionary forces operating in differentiating the above mentioned genes applies to the glutenins studied here, i.e., natural selection appears to be a major differentiating agent.

Utilization of glutenin diversity in bread making quality

Glutenins, like other wheat grain proteins, play a key role in food processing such as manufacture of bread, biscuits, breakfast cereals, and pasta products (Payne and Rhodes 1982). A major effort of plant breeders is to develop elite bread wheats with improved bread making quality. Significant associations were found for certain HMW subunits of glutenin and bread making quality (Payne et al. 1981b). These results have since been confirmed (Moonen et al. 1982) and expanded (Payne et al. 1984). Dough strength is the limiting factor in the bread making process in western European wheats. This character is controlled by the elasticity of glutenin, which in turn is determined by its specific composition of LMW and HMW subunits.

One approach to improving the quality of wheat cultivars for bread making is by utilizing genes from (1) landraces of primitive agriculture (Payne et al. 1984) and (2) the wild relatives of bread wheat (Law and Payne 1983). Wild emmer could provide a host of genetic variants of the endosperm storage proteins for improving bread-making quality, as it does for other agronomic traits (Nevo 1983; Feldman 1979, 1983). The range of allelic variation in the two glutenin loci of wild emmer analyzed here is remarkable, and all could be readily trasferred to bread wheat for testing by plant breeders.

Acknowledgements. We wish to thank Avigdor Beiles for statistical analysis and for commenting on the manuscript. We also would like to thank Dr. Levy for supplying us with the Levy and Feldman (1987) paper prior to print. This research was supported by grants from the Wolfson Foundation, the Israel Discount Bank Chair of Evolutionary Biology and the "Ancell-Teicher Research Foundation for Genetics and Molecular Evolution" established by Florence and Theodore Baumritter of New York.

References

Atlas of Israel (1970) Ministry of Labour, Jerusalem and Elsevier Publ, Amsterdam

Feldman M (1979) Genetic resources of wild wheats and their use in breeding. Monogr Genet Agrar 4:9–26

Feldman M (1983) Gene transfer from wild species into cultivated plants. Eucarpia workshop: prebreeding in relation to gene banks, April 1983. Beograd Zemun 12–15, pp 1–17

Feldman M, Sears ER (1981) The wild gene resources of wheat. Sci Am 244:102–112

Flavell RB, O'Dell M, Sharp P, Nevo E, Beiles A (1986) Variation in the intergenic spacer of ribosomal DNA of wild wheat, *Triticum dicoccoides* in Israel. Mol Biol Evol 3: 547–558

Galili G, Feldman M (1983) Genetic control of endosperm proteins in wheat: 11. Variation of high molecular weight glutenin and gliadin subunits of *Triticum aestivum* L. Theor Appl Genet 66:77–86

Galili G, Feldman M (1985) Structural homology of endosperm high molecular weight glutenin subunits of common wheat (*Triticum aestivum* L.) Theor Appl Genet 70:634–642

Harlan JR, Zohary D (1966) Distribution of wild wheat and barley. Science 153:1074–1080

Kasarda DD, Bernardin JE, Nimmo CC (1976) Wheat proteins. In: Pomeranz Y (ed) Advances in cereal science and technology. Am Assoc Cereal Chem, St Paul Minn, pp 158–236

Konzak CF (1977) Genetic control of the content, amino acid compositon and processing properties of proteins in wheat. In: Caspari EW (ed) Advances in genetics, vol XIX. Academic Press, New York London, pp 407–582

Law CN, Payne PI (1983) Genetical aspects of breeding for improved grain protein content and type in wheat. J Cereal Sci 1:79–93

Levy AA, Feldman M (1987) Polymorphism and genetic control of high molecular weight glutenin subunits in wild tetraploid wheat *Triticum turgidum* var. *dicoccoides*. Genetics (submitted)

Moonen JHE, Scheepstra A, Graveland A (1982) Use of the SDS-sedimentation test and SDS-polyacrylamide gel electrophoresis for screening breeder's samples of wheat for bread-making quality. Euphytica 31:277–690

Nei M (1972) Genetic distance between populations. Am Nat 106: 283–292

Nei M (1975) Molecular population genetics and evolution. North Holland, Amsterdam

Nevo E (1983) Genetic resources of wild emmer wheat: structure, evolution and application in breeding. Proc 6th Int Wheat Genet Symp, 3–28 December 1983. Kyoto, Japan, pp 421–431

Nevo E (1987a) Genetic resources of wild cereals and crop improvement: Israel, a natural laboratory. Isr J Bot 35: 255–278

Nevo E (1987b) Plant genetic resources: prediction by isozyme markers and ecology. In: Rattazi MC, Scandalios JG, Whitt GS (eds) Isozymes: current topics in biological research, vol 16. Agriculture, physiology and medicine. pp 247–267

Nevo E, Zohary D, Brown AHD, Haber M (1979) Genetic diversity and environmental associations of wild barley *Hordeum spotaneum* in Israel. Evolution 33:815–833

Nevo E, Golenberg EM, Beiles A, Brown AHD, Zohary D (1982) Genetic diversity and environmental associations of wild wheat *Triticum dicoccoides* in Israel. Theor Appl Genet 62:241–254

Nevo E, Beiles A, Storch N, Doll H, Andersen B (1983) Microgeographic edaphic differentiation in hordein polymorphisms of wild barley. Theor Appl Genet 64:123–132

Nevo E, Grama A, Beiles A, Golenberg EM (1986) Protein resources in wild wheat, *Triticum dicoccoides* in Israel: predictive method by ecology and allozyme markers. Genetica 68:215–228

Payne PI, Rhodes AP (1982) Cereal storage proteins: structure and role in agriculture and food technology. Encyc Plant Physiol 14:346–369

Payne PI, Law CN, Mudd EE (1980) Control by homoeologous group 1 chromosomes of the high molecular weight subunits of glutenin, a major protein of wheat endosperm. Theor Appl Genet 58:113–120

Payne PI, Holt M, Law CN (1981a) Structural and genetical studies on high-molecular-weight subunits of wheat glutenin. Part 1. Allelic variation in subunits amongst varieties of wheat (*Triticum aestivum*). Theor Appl Genet 60: 229–236

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

Payne PI, Holt M, Jackson EA, Law CN (1984) Wheat storage proteins: their genetics and their potential for manipulation by plant breeding. Philos Trans R Soc London, Ser B 304:359–371

SPSSx (1986) User's guide, 2nd edn. McGraw Hill, New York

Wall JS (1979) The role of wheat proteins in determining baking quality. In: Laidman DL, Wyn Jones RG (eds) Recent advances in the biochemistry of cereals. Academic Press, New York London, pp 275–311

Wright S (1965) The interpretation of population structure by F-statistics with special regard to systems of matings. Evolution 19:395–420

Zohary D (1969) The progenitors of wheat and barley in relation to domestication and agricultural dispersal in the old world. In: Ucko PJ, Dimbley GW (eds) The domestication and exploitation of plants and animals. Duckworth, London, pp 47–66

Zohary D (1970) Wild wheats. In: Frankel OH, Bennet E (eds) Genetic resources in plants: their exploration and conservation. Blackwell Oxford Edinburgh, pp 239–247