Theor Appl Genet (1995) 91:1064–1073

G.-Y. Zhong · C. O. Qualset

Quantitative genetic diversity and conservation strategies for an allogamous annual species, *Dasypyrum villosum* (L.) Candargy (Poaceae)

Received: 21 February 1995 / Accepted: 12 May 1995

Abstract *Dasypyrum villosum* (L.) Candargy is a weedy annual diploid (2n = 14, VV genomes) allogamous grass species (Poaceae, Triticeae). Genetic variation for 12 traits was studied in 43 natural populations (31 from Italy and 12 from Croatia and Montenegro of former Yugoslavia) grown in a common field environment in California. Although 7 of 12 traits followed the theoretical prediction that a larger proportion of genetic variation was distributed within populations than among populations, exceptions were found for spike length, plant height, and days to flag-leaf emergence, heading, and anthesis. Covariate analysis showed that developmentally closely related characters were more likely correlated at both population and family within population levels. Geographically closer populations shared more genetic similarity than distant populations as indicated by mean coefficients of variation and cluster analysis of the Euclidean distances among populations. As few as five populations, each population with five or more half-sib seeds taken randomly from 5 plants, is expected to capture more than 95% of the total genetic variation of this species in the region sampled, but sampling a much larger number of seeds per population (> 1000) for long-term storage would supply research and plant breeding needs for several decades. If seed regeneration is required, populations can be sampled from clusters having similar genetic variation, and grown in reproductive isolation or bulked seed samples from all populations of each cluster group can be grown in isolation. The former is recommended if population integrity is desired while the latter is sufficient to provide genetic resources for plant-breeding purposes.

Key words Poaceae · Triticeae · Population structure · Sampling strategy · Genetic resource · Conservation biology

Introduction

Dasypyrum villosum (L.) Candargy (syn. *Haynaldia villosa* Schur) is an allogamous annual grass species (2n = 2x = 14, VV genomes) in the Triticeae tribe of the Poaceae. It is widely distributed in South Europe, the Mediterranean region, and West Asia. This species grows vigorously on disturbed and moisture-stressed sites and has several attributes of potential value for the improvement of common wheat (*Triticum aestivum* L.) (De Pace et al. 1990; Qualset et al. 1993; Murray et al. 1994).

Interspecific hybridizations showed very little homology of the V genome of *D. villosum* and the A, B, and D genomes of wheat (Frederiksen 1991). Despite the genetic differentiation of *D. villosum* from wheat, this species has been successfully hybridized with both hexaploid and tetraploid wheats (Jan et al. 1986; Blanco et al. 1987) and six of the seven *D. villosum* chromosomes have been disomically added to the wheat genome complement (Sears 1953). Spontaneous translocations between wheat and *D. villosum* chromosomes have been isolated in progenies of wheat × *D. villosum* hybrids (P. D. Chen, personal communication). Thus, it appears that genes of *D. villosum* can be integrated into wheat chromosomes.

The survey and documentation of genetic variation in natural populations are important for the efficient utilization of *D. villosum* as a genetic resource for wheat improvement. This species has revealed remarkable variability in several qualitative traits in natural populations, including at least 14 alleles at the *Glu-V1* locus, coding for high-molecular-weight glutenin subunits of seed storage proteins (Zhong and Qualset 1993), morphological markers (Zhong and Qualset, unpublished data), isozymes, and RFLP of rDNA loci (De Pace and Qualset, in press). Della Gata et al. (1984) showed vari-

---

Communicated by A. L. Kahler

G.-Y. Zhong · C. O. Qualset (✉)
Genetic Resources Conservation Program and Department of Agronomy and Range Science, University of California, Davis, CA 95616, USA

ation of seed protein concentration in natural populations of *D. villosum.* Variation in plant height and flowering time was observed by De Pace and Qualset (1995) in *in situ* and greenhouse studies of several Italian *D. villosum* populations.

D. villosum was collected at 181 sites in its natural habitats in Italy and Croatia and Montenegro of former Yugoslavia (referenced as originating from Yugoslavia throughout this paper) (Qualset et al. 1984; P. E. McGuire, C. De Pace, and C. O. Qualset, unpublished data) for *ex situ* conservation and for research and for wheat improvement. In the present study genetic variability was examined for 12 quantitative characters in the above-mentioned collections in a common environment in California USA. Two questions were addressed: 1) How is quantitative genetic variation in this species distributed within and among local populations? 2) What are the implications of the genetic differentiation patterns, if any, to local adaptation and for strategies for *in situ* and *ex situ* conservation of this species?

Materials and methods

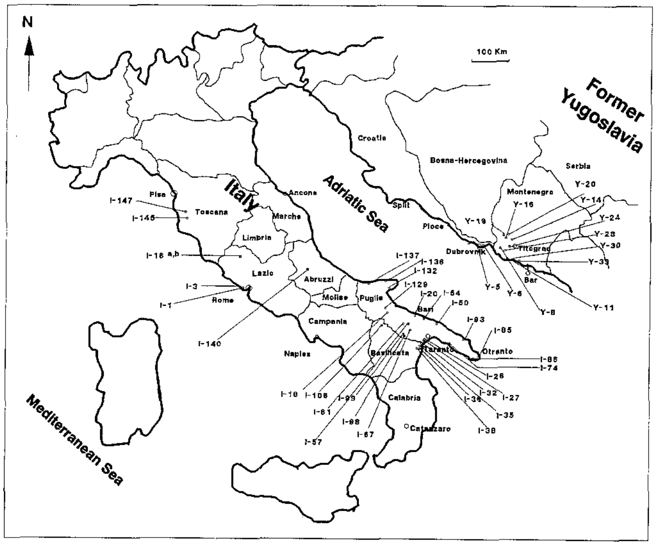
Materials

Forty-three populations, 31 from Italy and 12 from Yugoslavia, were chosen from the group of 181 that were sampled for this study (Fig. 1). About 1000 seeds were obtained at each collection site by taking three or four spikes from each of ten random plants. Seeds from all spikes of the same plant were composited and collectively identified as a half-sib family since *D. villosum* is a highly outcrossed species (De Pace and Qualset 1995). All the plants collected at one site are designated as a population.

Field experiment

Seedlings from each population were grown in a greenhouse for about 3 weeks, and transplanted into the field in January with a modified randomized complete block design in two replicates at the University of California Agronomy Farm near Davis (38°32′ N, 121°46′ W, 17 m asl). The conditions at Davis are similar to the Mediterranean climate from where the samples were taken. Each replicate contained three to five plants from each family spaced 0.4 m apart in a single row 2.4 m long, with 0.3 m between rows. Four groups of 12 quantitative characters were evaluated for individual plants: (1) sequential development (days past April 30 to flag-leaf emergence, heading, and anthesis); (2) flag-leaf blade length and width (cm) on two leaves per plant; (3) spikes [lemma awn length, spike length and width (cm), awn/spike length ratio, spikelet number per spike, and spike internode (rachis) length (mm)]; and (4) plant height (cm) at maturity not including the awn. The mean of two spikes of average size taken from the same plant was used for the spike characters.

Fig. 1 Sites, identified by numbers, for the collections of *D. villosum* in Italy and Yugoslavia



Variance and covariance analysis

Analyses of variance and covariance were conducted using a random nested general linear model partitioned according to sources of variation between countries, among populations within countries, and among families within populations. The total genetic variation for a trait was estimated as the summation of the variance components between countries (σ_c^2), among populations within countries (σ_p^2), among families within populations (σ_f^2), and among plants within families (σ_w^2). These components were estimated as percentages of the total genetic variance as R_c, R_p, R_f, and R_w for the respective source of variation.

Because the plants-within-family variance component (σ_w^2) could not be separated from the experimental error (σ_e^2), σ_w^2 was estimated as $[(1-F)/2F]\sigma_f^2$ as an approximation of the additive genetic variance within families, where F is the average inbreeding coefficient among individuals within families. Based on the estimated mean multilocus outcrossing rate ($C = 0.75$) (De Pace and Qualset 1995), the mean inbreeding coefficient $F = (1 - C)/(1 + C)$ (Falconer 1981) among individuals within families in natural populations of *D. villosum* was estimated as 0.143, assuming that the populations were at inbreeding equilibrium, and that parents were equally inbred and uncorrelated.

Product-moment correlations were calculated for family means across all populations and on population means. Because of incomplete data from populations Y-8, Y-14, Y-28, and Y-30 for some characters, these populations were excluded in bivariate and multivariate analyses. Populations I-16a and I-16b from Italy were collected at the same site, and showed very similar means for all characters; thus these two collections were treated as a single population for analysis.

Multivariate analysis

The overall variability of a population, or groups of populations, with regard to multiple quantitative characters was assessed by a modification of Zhang's (1985) generalized coefficient of phenotypic variation (GCPV):

$$|\mathrm{GCPV}_j| = \sqrt{\sum_i \left(\frac{cv_{ij}}{\sum_j cv_{ij}} \right)^2}$$

where cv_{ij} is the coefficient of variation for ith character of the jth population or jth group of populations.

Euclidean distances (D_{ij}) with regard to multiple characters were calculated for each pair of populations using a formula similar to that suggested by Goodman (1972). A hierarchical cluster analysis of the Euclidean distances between populations was done according to the average linkage method (Sokal and Michener 1958).

Sampling strategy among and within populations

A probability formula, developed by Oka (1969, 1975), was extended in the present study to estimate the number of populations, families, and individuals necessary for conservation of a desired fraction G of the total genetic variation represented by i populations each with j families and k individuals within each family

$$G = 1 - \{(1-p) + [p(1-f) + pf(1-w)^k]^j\}^i,$$

where 1 = total genetic variation over all characters in a given region, including the genetic variation among populations, among families within populations, and among plants within families, and p = proportion of total genetic variation represented by a population and estimated as

$$1 - \frac{1}{n}\sum_{1}^{i} \frac{\sigma_{p_i}^2}{\sigma_{p_i}^2 + \sigma_{f_i}^2 + \sigma_{w_i}^2}$$

with $\sigma_{p_i}^2$, $\sigma_{f_i}^2$, and $\sigma_{w_i}^2$ as the estimates of variance for the character i due to variation from populations, families within populations, and plants within families, respectively, and n as the total number of characters investigated, and f = proportion of genetic variation of a population represented by a family and estimated as

$$1 - \frac{1}{n}\sum_{1}^{i} \frac{\sigma_{f_i}^2}{\sigma_{f_i}^2 + \sigma_{w_i}^2},$$

and w = proportion of genetic variation of a family represented by a half-sib individual, arbitrarily set at 0.05 because the proportion of genetic variation within a half-sib family explained by a half-sib individual must be very low, and i = number of populations sampled in given region, j = number of plants (half-sib families) sampled per population, and k = number of seeds (half-sib individuals) sampled per plant.

Results

Variation among and within countries and populations

Significant variation was detected between countries (9 of 12 characters), among populations within countries (all characters), and among families within populations (11 of 12 characters) on the basis of the combined analysis (Table 1). Variance-component analysis showed that more than 50% of the total genetic variation was distributed between countries (R_c) for lemma awn length, spike length, and days-to-flag-leaf emergence, heading, and anthesis, and 30–40% for spike width, spikelet number per spike, spike internode length, and plant height (Table 1). Among populations within countries, variance components (R_p) showed that plant height and days-to-flag-leaf emergence, heading, and anthesis had a larger proportion of variation than within populations ($R_f + R_w$) (Table 1). About equal amounts of variation were found among and within populations for spike length and awn/spike length ratio. Non-additive genetic components were not included in the estimate of the within-population genetic variance, thus the actual amount of within-population genetic variation for spike length and awn/spike length ratio could be larger than the among-populations variance.

Large ranges for population means and intrapopulation coefficients of variation were found (Table 2), indicating that there was indeed differentiation among local *D. villosum* populations and that some populations were more variable than others with regard to specific characters. The mean intrapopulation coefficients of variation for plant height, days-to-flag-leaf emergence, heading, and anthesis were much smaller (10% or less) than that (13–25%) for other characters. Significant among-family variation was detected in more than 50% of the populations for plant height, about 30–40% for spikelet number per spike, width of flag-leaf blade, and days-to-flag-leaf emergence, heading, and anthesis, and less than 30% for the rest of the characters.

Correlations among characters

Some of the observed characters were developmentally more related than others. Four groups of characters were recognized: (1) sequential development characters

Table 1 Estimates of variance components (σ^2) and percentages (R) of total genetic variation attributable to countries (c), populations (p), families (f), and plants within families (w) for 12 characters of *D. villosum* grown in a common environment[a]

Character[b]	σ_c^2	σ_p^2	σ_f^2	σ_w^2	R_c	R_p	R_f	R_w	$R_f + R_w$
LAL	0.706**	0.084**	0.033**	0.099	76.6	9.1	3.6	10.7	14.3
SL	1.135**	0.290**	0.064**	0.192	67.5	17.3	3.8	11.4	15.2
ASLR	0.000	0.004**	0.001	0.003	0.0	50.0	12.5	37.5	50.0
SW	0.039**	0.031**	0.012**	0.036	33.1	26.2	10.2	30.5	40.7
SNPS	2.216**	3.584**	1.572**	4.711	18.3	29.7	13.0	39.0	52.0
SIL	0.078**	0.036**	0.019**	0.057	41.0	19.0	10.0	30.0	40.0
PH	132.080**	90.100**	19.172**	57.459	44.2	30.1	6.4	19.3	25.7
DEFL	75.690**	43.670**	4.340**	13.007	55.4	31.9	3.2	9.5	12.7
DH	78.140**	42.340**	3.850**	11.538	57.5	31.2	2.8	8.5	11.3
DA	65.310**	45.860**	3.350**	10.040	52.4	36.8	2.7	8.1	10.8
FLBL	0.000	0.800**	0.450**	1.349	0.0	30.8	17.3	51.9	69.2
FLBW	0.000	0.002**	0.003**	0.009	0.0	14.3	21.4	64.3	85.7
Mean	–	–	–	–	37.2	27.2	8.9	26.6	35.5

[a] ** $P < 0.01$, based on an analysis of variance with the degrees of freedom as: countries (1), populations (36 for first 6 characters, 41 for rest), families (293 for first 6 characters, 310 for rest), plants (1120 for first 6 characters, 1940 to 2017 for rest)

[b] LAL = lemma awn length, SL = spike length, ASLR = awn/spike length ratio, SW = spike width, SNPS = spikelet number per spike, SIL = spike internode length, PH = plant height, DEFL, DH, and DA = days-to-flag-leaf emergence, -heading, and -anthesis, FLBL and FLBW = length and width of flag-leaf blades

Table 2 Overall mean, ranges of population means and coefficients of phenotypic variation (CV), and percentage of populations in which significant among-family variation was detected

Character[a]	Mean			CV(%)			Percentage of populations with significant among-family variation
	Min.	Max.	Mean	Min.	Max.	Mean	
LAL, cm	3.3	5.7	4.7	7.6	28.5	15.5	5
SL, cm	3.8	6.8	5.5	9.1	23.5	15.8	16
ASLR	0.8	1.2	0.9	10.8	29.6	21.8	13
SW, cm	1.9	2.9	2.5	11.8	26.5	18.3	13
SNPS,	11.8	26.5	24.6	20.3	31.5	13.2	40
SIL, mm	1.5	2.7	2.3	12.8	28.2	18.2	24
PH, cm	59.3	117.9	93.6	6.6	16.2	8.2	54
DEFL, days	31.0	70.7	52.3	4.8	19.5	10.7	35
DH, days	37.1	79.2	61.2	3.5	15.1	9.7	37
DA, days	47.7	88.3	70.8	3.2	11.8	9.4	42
FLBL, cm	7.0	12.0	9.5	19.8	35.1	23.5	26
FLBW, cm	0.3	0.6	0.4	13.9	37.6	24.7	33

[a] See Table 1 for character abbreviations

Table 3 Phenotypic correlations for population (P) and family (F) means for 12 quantitative characters of *D. villosum*[a]

Character[b]		SL	ASLR	SW	SNPS	SIL	PH	DEFL	DH	DA	FLBL	FLBW
LAL	P	0.85**	−0.21	0.63**	0.54**	0.63**	0.64**	−0.44**	−0.41**	−0.43**	0.33**	0.44**
	F	0.57**	−0.26**	0.42**	0.42**	0.30**	0.52**	−0.37**	−0.35**	−0.37**	0.19**	0.28**
SL	P		−0.61**	0.78**	0.66**	0.74**	0.54**	−0.28	−0.24	−0.25	0.41**	0.39**
	F		−0.62**	0.61**	0.50**	0.70**	0.46**	−0.17**	−0.13	−0.15**	0.28**	0.25**
ASLR	P			−0.54**	−0.44**	−0.46**	−0.06	−0.13	−0.17	−0.19	−0.28	−0.10
	F			−0.33**	−0.22**	−0.52**	−0.03	−0.16**	−0.20**	−0.20**	−0.13**	−0.03
SW	P				0.44**	0.65**	0.27	−0.18	−0.14	−0.11	0.38*	0.30*
	F				0.35**	0.41**	0.21**	−0.09	−0.07	−0.06	0.21**	0.24**
SNPS	P					−0.01	0.24	0.14	0.17	0.16	0.32*	0.25
	F					−0.24**	0.21**	0.15**	0.15**	0.14**	0.25**	0.28**
SIL	P						0.50**	−0.51**	−0.48**	−0.47**	0.26	0.31*
	F						0.34**	−0.32**	−0.27**	−0.28**	0.11*	0.06
PH	P							−0.53**	−0.52**	−0.57**	0.28	0.58**
	F							−0.48**	−0.48**	−0.52**	0.27**	0.42**
DEFL	P								1.00**	0.99**	0.24	−0.09
	F								0.98**	0.97**	0.10	−0.03
DH	P									0.99**	0.24	−0.07
	F									0.97**	0.10	−0.03
DA	P										0.22	−0.11
	F										0.10	−0.04
FLBL	P											0.61**
	F											0.57**

[a] Thirty-eight populations and 331 families were respectively used in the correlation analysis for populations and families

[b] See Table 1 for character abbreviations

*,** Significant at $P = 0.01$ and $P = 0.05$, respectively

related to reproductive biology (days-to-flag-leaf emergence, heading, and anthesis), (2) one aspect of leaf canopy structure (length and width of flag-leaf blades), (3) spike morphology, and (4) mature plant height. The characters within these groups, as expected, were more correlated than the characters from different developmental groups (Table 3). The largest correlation coefficients (0.97–1.00) were found between the traits related to flowering time. Among the intergroup correlations, lemma awn length, spike internode length, and plant height showed consistently significant correlations with most other characters from other groups at both the family and population levels (Table 3). Correlation coefficients among other intergroup characters were generally low and insignificant, or inconsistent, within and among populations. In general, the correlations over populations were larger than that over families within populations.

Multivariate variation

The generalized phenotypic coefficients of variation (GPCV) ranged from 5.9 for population I-147 to 11.7 for population I-16 with a mean value 8.5 ± 1.3 (Table 4). Yugoslavian populations showed larger GPCVs than most Italian populations.

Euclidean distances were estimated from the first five standardized principal component scores, which together explained more than 90% of the total variation (data not shown) and were used in cluster analysis (Fig. 2). Since the correlations between days-to-flag-leaf emergence, heading, and anthesis dates were very high, only heading date was used to represent that group of developmental characters. Six clusters or groups of populations were identified (Table 4). The various contributions of different characters to the clusters are represented in Fig. 3 as the percentage of deviations of the cluster means from the overall means.

Cluster 1 identified the one unique population, I-16, from Italy; this population was characterized by a very early heading date. The plants headed about 70 days after transplanting into the field and more than a month earlier than some late populations. This result was consistent with the observation in a greenhouse study (data not shown) and earlier investigations on this population, now known as ‘Bomarzo’, (De Pace and Qualset 1995). In addition, this population had the largest ratio of awn to spike length.

Cluster 2 included six Yugoslavian populations and showed the largest negative deviations of the cluster mean from the overall mean for lemma awn length, spike length, spike width, spike internode length, and plant height. The largest positive deviation was found for heading date.

Cluster 3 consists of the four Italian populations I-18, I-129, I-145, and I-147 and was characterized by several traits whose deviations differ in signs or magnitudes from all other cluster groups. All the characters, with the exception of awn/spike length ratio, showed positive deviations from the overall means.

Cluster 4 included populations Y-6, Y-11 and I-93 and is distinctive for its smaller number of spikelets, long spike internodes and longer than average flag leaves.

Cluster 5 included only two populations which were taller and had earlier heading dates, shorter but wider flag-leaf blades, smaller spikes and longer awns than the average means.

Cluster 6 included the remaining 22 populations, all from Italy, which showed no obvious deviations (less than 5%) from the overall means.

Table 4 The generalized coefficients of phenotypic variation (GCPV) for clusters and populations of *D. villosum*

Cluster		Population	
No.	GCPV	No.	GCPV
1	50.4		
		I-16	11.7
		Mean	**11.7**
2	63.6		
		Y-5	9.8
		Y-16	11.1
		Y-19	11.1
		Y-20	9.1
		Y-24	10.0
		Y-33	8.3
		Mean	**9.9**
3	42.2		
		I-18	7.6
		I-129	7.4
		I-145	6.7
		I-147	5.9
		Mean	**6.9**
4	60.9		
		Y-6	9.4
		Y-11	8.9
		I-93	8.8
		Mean	**9.9**
5	54.3		
		I-1	10.5
		I-3	10.2
		Mean	**10.3**
6	49.1		
		I-20	9.1
		I-26	8.3
		I-27	7.1
		I-32	8.1
		I-35	8.3
		I-36	8.5
		I-38	9.1
		I-50	7.3
		I-54	9.0
		I-57	7.9
		I-61	7.4
		I-67	8.3
		I-74	8.6
		I-85	6.5
		I-86	8.4
		I-98	8.2
		I-99	7.4
		I-106	7.2
		I-132	7.1
		I-136	7.3
		I-137	9.0
		I-140	7.9
		Mean	**8.0**

Cluster analysis demonstrated a tendency for populations from similar geographic origins to cluster together. For example, most populations from Yugoslavia formed their own group (cluster 2), while Italian populations were grouped into other clusters. Populations I-1 and I-3 were collected from the same area (less than 10 km apart) in central Italy and these two populations

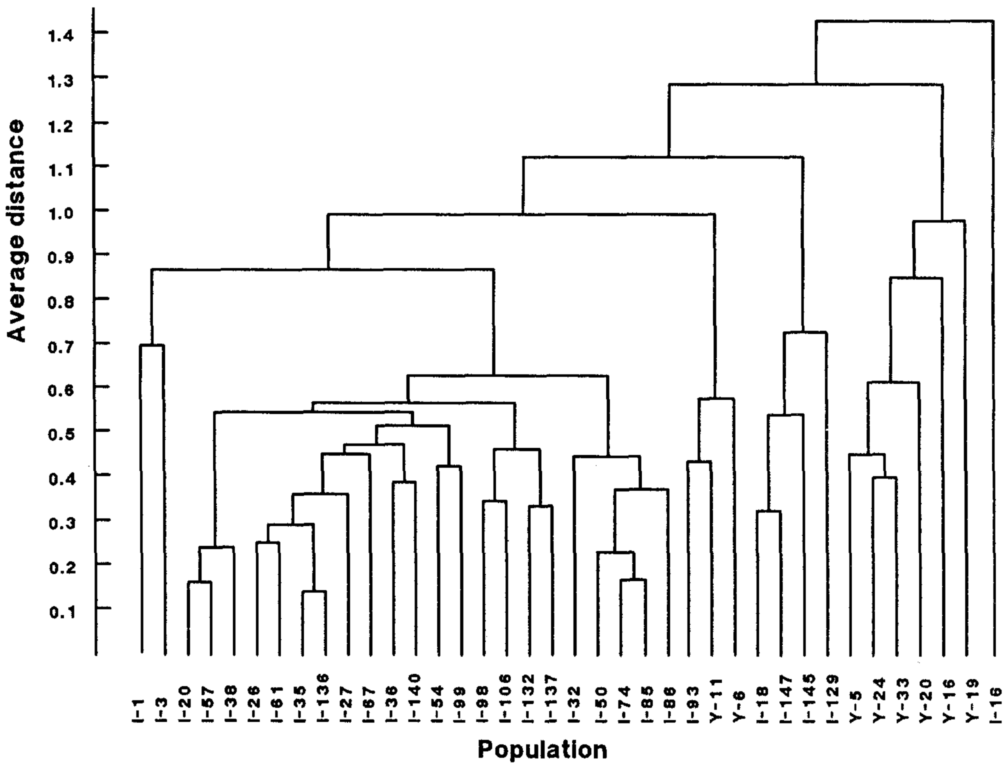


Fig. 2 Dendrogram of genetic distances among 38 *D. villosum* populations using the average linkage method

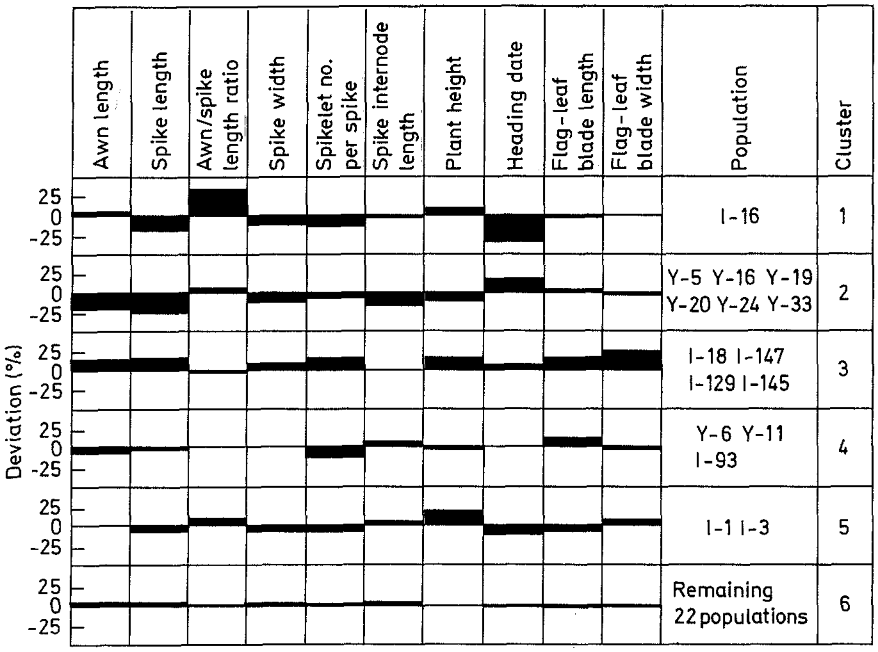


Fig. 3 Deviation (%) of cluster means of ten characters from the overall mean. Because of similar means and variation patterns among days-to-flag-leaf emergence, heading, and anthesis, only heading data was presented in this figure

formed a distinct group (cluster 5). Differentiation between groups of populations was usually dependent on several characters, although one or two of them were more discriminative than others. For example, populations from Yugoslavia (cluster 2) showed not only the shortest spike internode lengths but also were shortest in plant height, smallest in spike size, and latest in heading date. Furthermore, populations with similar GPCVs were grouped together (Table 4). This was well-demonstrated in cluster 2 in which all six populations had very similar large GPCVs, and in cluster 3 where the GPCVs were generally smaller.

Table 5 Expected percentage of total genetic variation obtained by sampling *i* populations, *j* plants, and *k* seeds per plant

No. plants (*j*)	No. seeds (*k*)	No. populations (*i*)			
		1	5	10	20
1	1	2.0	9.5	18.0	32.8
1	5	8.9	37.2	60.6	84.5
1	10	15.8	57.6	82.0	96.8
1	20	25.2	76.6	94.5	99.7
5	1	49.1	96.6	99.9	100
5	5	50.8	97.1	99.9	100
5	10	51.7	97.4	99.9	100
5	20	52.3	97.5	99.9	100
10	1	52.3	97.5	99.9	100
10	5	52.4	97.6	99.9	100
10	10	52.4	97.6	99.9	100
10	20	52.4	97.6	99.9	100
20	1	52.4	97.6	99.9	100
20	5	52.4	97.6	99.9	100
20	10	52.4	97.6	99.9	100
20	20	52.4	97.6	99.9	100

Sampling *D. villosum* among and within populations

The percentages of total genetic variation ($\sigma_p^2 + \sigma_f^2 + \sigma_w^2$) which could be captured by sampling with different combinations of the number of populations, plants, and individuals (seeds) for several levels of sampling effort were estimated (Table 5). Because the averaged variation was about equally distributed within and among populations ($t_p = 0.476$), both intra- and inter-population sampling are critical. It was estimated that about 50% of the total variation would be represented by sampling five plants, each with five or more half-sib seeds, from a single population (Table 5). The selected sampling strategies in Table 5 show that about 95% of the total variation can be captured by sampling five plants, each with as few as one seed, from five populations. While intrapopulational sampling is important, among-population sampling is more efficient to maximize the total variability than increasing the number of plants sampled within populations if more than one plant in each population is sampled. In fact, the sampling reward of more families is very little after five families have been sampled in a population. Similarly, sampling more plants will be more important than sampling more half-sib seeds from each plant.

Discussion

Patterns of genetic variation

Genetic variation in natural populations of plants can be detected by observing heritable polymorphisms in discrete characters, such as morphological markers, proteins and DNA sequences, and by investigating the heritable variation in quantitative characters. Variation at various discrete characters, including isozymes, the high-molecular-weight glutenin of storage proteins and rDNA, has been documented in natural populations of *D. villosum* (Zhong and Qualset 1993; De Pace and Qualset 1995) and the distribution patterns of variation for these characters were found to be typical for most allogamous species, i.e., variation is mainly distributed within populations rather than among populations (Brown 1979; Hamrick and Godt 1989). Similar variation patterns were also observed in the present study for various characters related to spikes and flag-leaf blades. However, exceptions were found for plant height, and days-to-flag-leaf emergence, heading, and anthesis, for which a larger proportion of variation was distributed among populations relative to that within populations. These results were not likely to be due to the possible downward biased estimates of the within-population genetic variation since the among-population genetic variance for those characters was greater than $4\sigma_f^2$, the theoretical maximum estimates of the within-population additive genetic variation. One possible alternative explanation is that those characters are highly critical for adaptation to local conditions and, through natural selection, resulted in the observed evolutionary divergences. The significance of plant height and days-to-heading and anthesis in natural adaptation has been described in both natural and cultivated grass species. For example, clinal variation for plant height and days-to-heading was found and attributed to local adaptation in natural populations of *Bromus mollis* (Knowles 1943) and *Hordeum spontaneum* (Snow and Brody 1984) and in cultivated *Secale cereale* L. (Hoskinson and Qualset 1967). A dramatic selective advantage for plant height was shown in experimental populations of wheat and rice (reviewed in Qualset 1978). These characters are often controlled by a small number of genes each with major effect and, therefore, may respond to natural selection more rapidly than characters which are controlled by a large number of minor genes. As a result, intrapopulational variation for these characters was reduced and a larger proportion of the variation was then expressed as genetic differentiation among localized populations.

Non-pleiotropic characters do not evolve independently in the course of evolution and some inter-character correlations are stronger than others in natural populations (Via 1984; Loeschke 1987). In the present study, characters, such as those involving time-to-flowering within a developmental sequence, showed similar among-family and among-population correlational relationships. This type of correlation would provide long-term conservation of character combinations during population differentiation of the species. Correlations were also found between developmental character groups. For example, spike length and plant height were positively correlated with flag-leaf size among families and populations. Tight gene linkage is one of the possible mechanisms that can maintain tight within-and among-population correlations which trace to mutational events. Another possible explanation is that these characters had strong selective values, and the correlations of these characters with others might have resulted in genic co-adaptation to the environments. Adaptive associations of characters that were believed to have formed through natural selection for adaptation have been reported in cultivated Ethiopian barley, a self-pollinated species, in which resistance to the barley yellow dwarf virus was non-randomly associated with late-maturing types (Qualset 1975). Multiple character associations could also be expected in *D. villosum*, although multilocus structural organization resulting from adaptive changes in outcrossing populations is usually not as distinct as that found in selfing populations (Allard 1988). The fact that geographically closer populations of *D. villosum* showed similar overall co-variation of traits and tended to cluster together suggests that co-adaptative associations have evolved in natural populations of *D. villosum.*

Although populations were generally differentiated with respect to several characters over a large area of the distribution of the species, an exception was found in the present study for the Bomarzo (I-16) population which was unique in several respects. This population was characterized by uniformity in very early heading date and large awn/spike length ratio. It also had larger seeds and, in a greenhouse study (data not shown), suffered little inbreeding depression in comparison with other populations. Earlier, we showed (Zhong and Qualset 1993) that Bomarzo had eight alleles at the glutenin seed storage protein locus *Glu-V1* compared to a mean of seven in all other populations (14 different alleles were observed) and none of its alleles were unique to the species. Bomarzo was, however, unique in that it did not have 2 of 8 widely distributed *Glu-V1* alleles. De Pace and Qualset (1995) also observed the uniqueness of this population and found that Bomarzo had alleles at isozyme loci that were absent in other populations; furthermore, it had high interfamily and low intrafamily variation, a departure from Hardy-Weinberg equilibrium for *Got* and *Est* genotypes, and geographic isolation. These and other observations caused De Pace and Qualset (1995) to suggest that Bomarzo was the product of a small number of seeds (founder effect) according to the expectations of Wright (1942).

Implications of the variation patterns to the collection and conservation of *D. villosum*

One of the goals of the present study was to provide guidance for future collections of this and similar grass species in the Triticeae. Sampling strategies for various plant species have been discussed by Allard (1970), Marshall and Brown (1975), Qualset (1975), and others as reviewed by Marshall (1989). The strategies recommended may vary with the species, the type of characters, and the primary interests, but it should be founded upon the knowledge of the patterns of genetic variability of populations in their natural habitats. The variation patterns revealed in the present study of *D. villosum* were based on 38 populations sampled over a wide geographical range, and can be regarded as representative of the variation patterns of *D. villosum* in the region where the collections were made.

Genetic variability associated with broad geographic regions is the most obvious variation pattern as observed in a number of species (Allard 1970). Geographic variation of characters was found between countries in the present study, presumably as a result of isolation by the Adriatic Sea. Cluster analysis also showed that geographically closer populations shared more genetic similarities with regard to multiple characters. This agrees with the general rule that sampling geographically distant populations may capture a large proportion of the extant variation in a species.

Maximizing overall genetic variability is a general requirement for genetic resource conservation programs. The present study showed that the overall variabilities, as measured by the generalized coefficients of variation, differed between countries and among populations. Populations from Yugoslavia, on average, showed larger overall variability, and therefore deserve more attention in the allocation of sampling efforts. Another critical question is how to allocate the sampling resources within and among populations in a given region so that the maximum amount of variability can be captured. Marshall and Brown (1975) suggested that, with no prior knowledge of the distribution of alleles in the target species, the optimum strategy to capture at least one copy of each common allele (frequency > 0.05) is to collect a bulk of 50 to 100 individuals from as many sites as possible. In the sampling of the present study as few as five populations each with five or more half-sib individuals obtained randomly from five plants (families) meets this goal. It should be kept in mind, however, that this strategy assumes that each population contains a constant proportion of the total variation of the species, and that the variation among populations is uniformly distributed. In reality, these assumptions may be violated. Some populations may have localized alleles which deserve special attention in collections. The Bomarzo population (I-16) is such an example in this study. In spite of these assumptions, this analysis suggests that a rather small number of geographically dispersed populations will capture the majority of genetic variation commonly found in this species. In a practical sense, *D. villosum* is not a rare species and populations are sufficiently large so that many seeds can be collected without risk to the natural populations. Further, it is easier to collect more seeds at each site than to sample the full geographic range. During the collection of seeds, bulk samples were taken at some sites to provide sufficient seeds for agronomic studies (De Pace et al. 1990). Our collection of *D. villosum* in the sampling areas of Italy and Yugoslavia, numbering nearly 200 sites, appears to be fully sufficient. We suggest that future expeditions in unsampled domains of *D. villosum* should sample at about the same density because it is relatively inexpensive and the possibility of finding unique populations, such as Bomarzo, should not be overlooked.

Another aspect of the conservation of an allogamous species such as *D. villosum* is that it is very difficult to retain the integrity of populations during off-site seed regeneration. For this reason, and also because it is generally not practical to re-sample the original sites, it is desirable to collect a large sample of seeds at each site. These seeds can then be placed in long-term storage conditions and viability retained for several decades. Thus, the original samples will be available for many uses over time. The sampling strategy used for the present study provided 900 to 1000 seeds at each collection site (ten plants, three spikes per plant), but larger samples could have easily been obtained.

If regeneration is required, it may be impractical to grow samples from each population. The cluster analysis may be used to great advantage to develop a regeneration strategy. For example, in this study, six groups of populations were identified in the cluster analysis. Thus, one strategy would be to generate populations with the highest diversity from each of these groups to conserve the maximum amount of variation. In our case, populations I-16, Y-16, I-18, Y-6, I-1, and I-20 from clusters 1, 2, 3, 4, 5, and 6, respectively, could be chosen. This strategy gives equal weight to each cluster. However, in practice, it may be more desirable to regenerate several populations from each cluster with reproductive isolation for each population according to the overall level of variability of each group. A second strategy is to prepare a bulk seed sample from all populations in each cluster and grow each "cluster population" under conditions of reproductive isolation in the field. Both strategies can avoid inbreeding and retain genetic integrity. The first strategy may be desirable for certain research purposes while the second is very appropriate for plant-breeding applications.

Acknowledgements G.-Y. Z. was supported by a Rockefeller Foundation Fellowship for this research. Populations were sampled with assistance of NSF grant No. INT8400743. The support and encouragement of Patrick E. McGuire and Ciro De Pace throughout this work was greatly appreciated.

References

Allard RW (1970) Population structure and sampling methods. In: Frankel OH, Bennett E (eds) Genetic resources in plants. IBP Handbook No 11, Blackwell Scientific Publications, Oxford Edinburgh, pp 97–107

Allard RW (1988) Genetic changes associated with the evolution of adaptedness in cultivated plants and their wild progenitors. J Hered 79:225–238

Blanco A, Simeone R, Resta P (1987) The addition of *D. villosum* (L.) Candargy chromosomes to durum wheat (*Triticum durum* Desf). Theor Appl Genet 74:328–333

Brown AHD (1979) Enzyme polymorphism in plant populations. Theor Pop Biol 15:1–42

De Pace C, Qualset CO (1995) Mating system and genetic differentiation in *Dasypyrum villosum* (L.) Candargy in Italy. Plant Syst Evol (in press)

De Pace C, Paolini R, Scarascia Mugnozza GT, Qualset CO, Delre V (1990) Evaluation and utilization of *Dasypyrum villosum* as a genetic resource for wheat improvement. In: Srivastava JP, Damania AB (eds) Wheat genetic resources: meeting diverse needs. John Wiley and Sons, UK, pp 279–290

Della Gata C, Tanzarella OA, Resta A, Blanco A (1984) Protein content in a population of *Haynaldia villosa* and electrophoretic pattern of the amphiploid *T. durum* × *H. villosa.* In: Porceddu E (ed) Breeding methodologies in durum wheat and Triticale. University of Tuscia, Viterbo, Italy, pp 39–43

Falconer DS (1981) Introduction to quantitative genetics, 2nd edn. Longman, New York

Frederiksen S (1991) Taxonomic studies in *Dasypyrum* (Poaceae). Nord J Bot 11:135–142

Goodman MM (1972) Distance analysis in biology. Syst Zool 21:174–186

Hamrick JL, Godt MJ (1989) Allozyme diversity in plant species. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (ed) Plant population genetics, breeding, and genetic resources. Sinauer Associates Inc Publishers, Sunderland, Massachusetts, pp 43–63

Hoskinson PE, Qualset CO (1967) Geographic variation in Balbo rye. University of Tennessee Agricultural Experimental Station: Tennessee Farm and Home Science Progress Report 62:8–9

Jan CC, De Pace C, McGuire PE, Qualset CO (1986) Hybrids and amphiploids of *Triticum aestivum* L. and *T. turgidum* L., with *Dasypyrum villosum* (L.) Candargy. Z Pflanzenzuchtg 96:97–106

Knowles PF (1943) Improving an annual brome grass, *Bromus mollis* L. for range purposes. J Am Soc Agron 35:584–594

Loeschcke V (1987) Genetic constraints on adaptive evolution. Springer-Verlag, Berlin Heidelberg

Marshall DR (1989) Crop genetic resources: current and emerging issues. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (eds) Plant population genetics, breeding, and genetic resources. Sinauer Associates, Sunderland, Massachusetts, pp 367–388

Marshall DR, Brown AHD (1975) Optimum sampling strategies in genetic conservation. In: Frankel OH, Hawkes JG (eds) Crop genetic resources for today and tomorrow. Cambridge University Press, London, pp 53–80

Murray TD, De La Pena RC, Yildirim A, Jones SS (1994) A new source of resistance to *Pseudocercosporella herpotrichoides*, cause of eyespot disease of wheat, located on chromosome 4V of *Dasypyrum villosum.* Plant Breed 113:281–286

Oka HI (1969) A note on the design of germplasm preservation work in grain crops. SABRAO Newslett 1:127–134

Oka HI (1975) Consideration of the population size necessary for conservation of crop germplasms. In: Matsuo T (ed) Japanese Committee for the International Biological Program, Synthesis 5, University of Tokyo Press, Tokyo, pp 57–63

Qualset CO (1975) Sampling germplasm in a center of diversity. An example of disease resistance in Ethiopian barley. In: Frankel OH, Hawkes JG (eds) Crop genetic resources for today and tomorrow. Cambridge University Press, London, pp 81–96

Qualset CO (1978) Mendelian genetics of quantitative characters with reference to adaptation and breeding in wheat. In: Ramanujam S (ed) Proc 5th Int Wheat Genet Symp, New Delhi, India, pp 577–590

Qualset CO, McGuire PE, Stupar M, Pavicevic L (1984) Collection of *Dasypyrum villosum* (syn. *Haynaldia villosa*), *Aegilops* spp., and *Hordeum bulbosum* in Yugoslavia. University of California, Davis Agricultural Experiment Station: Agronomy Progress Report 157

Qualset CO, Zhong G-Y, De Pace C, McGuire PE (1993) Population biology and evaluation of genetic resources of *Dasypyrum villosum.* In: Damania AB (ed) Biodiversity and wheat improvement. John Wiley and Sons, UK, pp 227–232

Sears ER (1953) Addition of the genome of *Haynaldia villosa* to *Triticum aestivum.* Am J Bot 40:168–174

Snow L, Brody T (1984) Genetic variation in *Hordeum spontaneum* in Israel: eco-geographical races, detected by trait measurements. Pl Syst Evol 145:15–28

Sokal RR, Michener CD (1958) A statistical method for evaluating systematic relationships. Univ Kansas Sci Bull 38: 1409–1438

Via S (1984) The quantitative genetics of polyphagy in an insect herbivore. II. Genetic correlations in larval performance within and among host plants. Evolution 38:896–905

Wright S (1942) Statistical genetics and evolution. Bull Am Maths Soc 48:223–246

Zhang Q (1985) Interrelationships of allozyme, morphological markers, disease resistance and quantitative characters in the genetic systems of barley *Hordeum vulgare* L. PhD Dissertation, University of California, Davis

Zhong G-Y, Qualset CO (1993) Allelic diversity of high-molecular-weight glutenin protein subunits in natural populations of *Dasypyrum villosum* (L.) Candargy. Theor Appl Genet 86:851–858