Multivariate Phenotypic Structures in the Batini barley landrace from Oman

A. Y. Al-Maskri [1*], M. Shahid [2] and A. A. Jaradat [3]

[1]*College of Agriculture and Marine Sciences, Sultan Qaboos University, P.O.Box 34, Al-Khod, Postal Code 123, Muscat, Oman.* [2]*Plant Genetic Resources Program, International Center for Biosaline Agriculture (ICBA), P.O. Box 14660, Dubai, United Arab Emirates.* [3]*Agricultural Research Service, 803 Iowa Avenue, Morris, MN, 56267 USA.*
e-mail: maskri99@squ.edu.om, aajaradat2002@yahoo.com

Abstract

We identified 14 phenological classes in three (grain, forage and dual-purpose) end-use types and formulated discriminant functions to help select elite germplasm for breeding purposes. Biological yield, pre- and post-anthesis thermal time, grain yield, plant height and tillers per plant, in decreasing order, were most influential in discriminating among phenological classes. Extensive divergence in the phenotypic covariance matrices among phenological classes suggests that directional selection, especially in the phenological traits, resulted in large, idiosyncratic changes in the principal components' structure, and that some of the changes are attributed to shifts in the mean phenotype. Knowledge of genetic co-variation of these traits will be useful for plant breeders by targeting traits that have a disproportionately large influence on differences in the mean covariance.

Key words: Phenotypic structure, phenology, landrace, Oman.

Introduction

Barley is an important crop in the Sultanate of Oman and its significance has increased with the renewed interest in local sources for forage and biomass production [6]. The Batini barley landrace persisted in patches and islands of farming systems in Oman and neighboring countries for millennia [11] and was found to harbor large genetic diversity for a number of agronomic traits, including tolerance to salinity [12].

Livestock numbers in Oman and neighboring countries continued to increase during the last 20-30 years, however, forage production failed to meet the growing needs [6]. Consequently, growing livestock number accelerated the depletion of natural forage sources and intensified competition for feed grain and crop residues. Yau [24] pointed out that barley landraces are the feed source of choice for sheep throughout the Middle East. Farmers may have favored barley phenotypes adapted to the prevailing climatic and edaphic conditions [4], and specifically selected for certain phonological and agronomic traits such as earliness, large number of tillers per plant and large forage yield [13]. Visual selection by farmers [5] for certain phenotypic traits (e.g., high biomass, tall plants, profuse tillering and high grain yield) may have resulted in the development of diverse genotypes with different combinations of phonological stages (i.e., days to anthesis, filling period and days to maturity [1]).

The level of variation for biomass, grain yield and phenological traits in this landrace is unknown but would aid in determining selection and breeding priorities. The objectives of this study were to: (1) examine whether farmers' and natural selection have altered the mean phenotype among and within subpopulations and (2) identify grain, forage or dual-purpose genotypes and formulate discriminant functions to help select elite germplasm from the Batini barley landrace for breeding purposes.

Materials and Methods

Details of a field experiment conducted during the 2001/2002 growing season at the Experiment Station of The International Center for Biosaline Agriculture (ICBA), Dubai, United Arab Emirates (25°13'N and 55°17'E), are available in Jaradat *et al.* [12]. For the purpose of this study, data on phenological development were recorded on the primary tillers of 2,040 accessions of the Batini barley landrace according to Tottman [21]. An accession was deemed to have attained a particular stage when approximately 50% of the plants in that accession reached that particular developmental stage. Single plants were harvested at full maturity, and data were collected on a minimum of five plants per accession of only 2,040 accessions for plant height, number of tillers, biological yield and grain yield. Each accession was visually classified by three independent raters, and the combined visual classification was used to classify the germplasm collection into grain, dual-purpose and forage types.

Statistical analyses: Pre- and post-anthesis developmental stages were converted to thermal time (°C d^{-1}) using a base temperature of 0.0°C for the period (in days) from sowing to anthesis and 9.0°C for the period from anthesis to maturity [8]. Quantitative data for the three phenological stages were used to identify phenotypic classes[8] of accessions with similar phenological stages. Fourteen phenological classes were delineated within the three end-use types using the likelihood ratio test [9]. This "optimal" number of phenological classes was selected when the log-likelihood function reached its highest increase and was verified through a two-stage clustering analysis procedure to minimize the Euclidean sum of squares using the K-means cluster module in Clustan Graphics [23].

Principal component analyses (PCA) were performed on the co-variance matrices [18] for subpopulations, end-use types and phenological classes. Co-variance matrices were used to evaluate

the extent and nature of differences among subpopulations, among end-use types and among phenological classes [7, 16]. For each of these three groupings, a common principal component analysis (CPCA) was employed to test whether pairs of co-variance matrices have completely unrelated structures, whether the matrices share one or more principal components (eigenvectors), whether the matrices differ by a simple constant of proportionality or whether the matrices are identical [16].

Similarity matrices, based on phenological and agronomic traits, of seven subpopulations, three end-use types and 14 phenological classes were used to test whether each pair of subpopulations, end-use types or phenological classes are independent [15]. A dimension reduction and perceptual mapping statistical procedure [9] was employed to reduce the dimensionality of a matrix of three end-use types, two phenological (days to heading and days to maturity), two phenotypic (plant height and number of fertile tillers per plant) and two productivity (grain and biological yield per plant) traits. A two-dimensional plot was developed where associations among end-use types and the plant traits can be identified. In addition, correlation coefficient among each phenological trait and the rest for each end-use type were calculated.

Discriminant functions for 80% of the whole collection were calculated to mathematically predict and verify visual group membership of each of the end-use types and of the phenological classes. The utility of the discriminant function was validated using the remaining 20% of the data. The *a priori* probabilities were set to be proportional to the size of the end-use types and to the phenological classes in the whole collection [18]. Discriminant loadings and their univariate F-ratios were used to rank traits according to their discriminating power among end-use groups and among phenological classes. All statistical analyses were conducted using several modules in the statistical packages STATISTICA [20], unless otherwise specified.
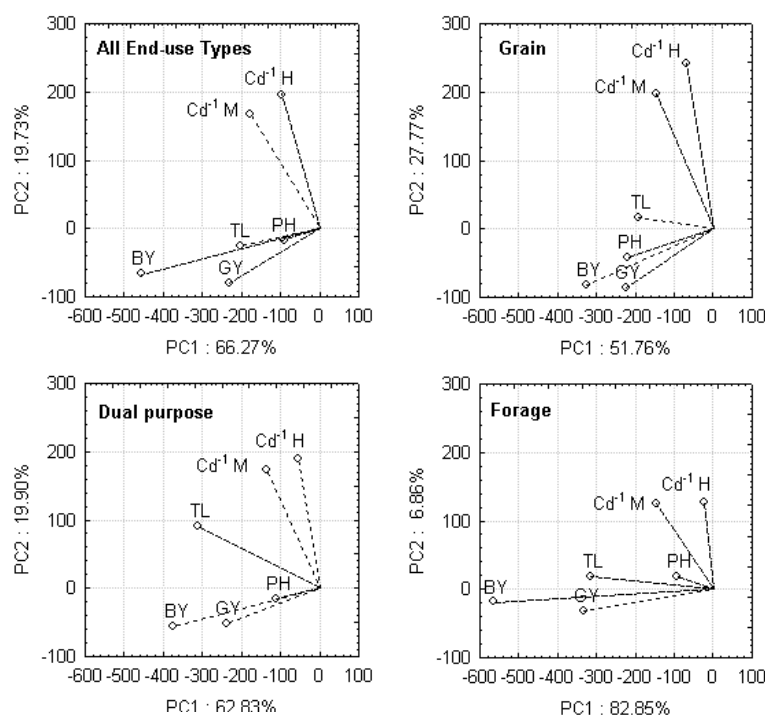
Figure 1. Combined and individual principal components analyses, based on the covariance matrix of three end-use types identified in the Batini barley landrace.

Results

Results of the likelihood ratio test [9], as confirmed by the K-means clustering procedure [23] separated the 2,040 accessions into 14 phenological classes with a combination of short, medium and long °C d^{-1} H (mean 2306 ± standard deviation 219), °C d^{-1} M (2649 ± 246) and °C d^{-1} Fp (343 ± 92). End-use types exhibited highly variable relationships among their phenological stages. The grain, dual purpose and forage end-use types on average allocated 89, 87 and 82% of their total °C d^{-1} to pre-anthesis growth stage, respectively, with a concomitant increase in biological yield and decrease in grain yield. Across all phenological classes, °C d^{-1}H and °C d^{-1}M were positively and significantly correlated (r = 0.91; *P*<0.01), whereas °C d^{-1} Fp was only negatively and significantly correlated (r = -0.74; *P*<0.05) with grain yield. Significant correlations among agronomic traits were limited to those between plant height and each of biological yield (r = 0.84; *P*<0.01) and number of tillers (r = 0.80; *P*<0.05) and between biological yield and number of tillers (r = 0.97; *P*<0.01).

Multivariate phenotypic structures: The covariance matrix of all agronomic and phenological traits was used to identify principal components that account for most variation in all and each of the end-use types (Fig. 1). All agronomic traits were associated with PC1, whereas °Cd^{-1}H and °Cd^{-1}M were associated with PC2. Significant differences were found among end-use types for eigenvalues (MS = 48.52, F = 401.2, *P*<0.000) and for the first (MS = 1.68, F = 16.2, *P*<0.000) and second (MS = 0.62, F = 4.14, *P*<0.001) eigenvectors (PCs). Total variation explained by two PCs in the whole collection and in the grain, dual-purpose and forage types were 86.0, 79.5, 82.7 and 89.7%, respectively, however, loadings of the phenological traits on these two PCs were highly variable (Fig. 1). No such sharp differences were found in loadings of agronomic traits on PC1 except for plant height loading in the grain type.

Complex models (Table 1, below diagonal), ranging from CPC1 to UR, were found as a result of multivariate covariation among phenological classes. The majority of these models (70%) reflects large differences among and within end-use types. The Mantel test (Table 1, above diagonal) as illustrated by the pairwise matrix correlations among phenological classes indicates that 14 out of the 91 pairwise correlation coefficients were not significant and they reflect the level of shared principal components among the relevant phenological classes.

Discrimination among phenological classes: Percent correct classifications of the end-use types, based on classification functions, were 98.5, 99.8, and 95.5% for the grain, dual-purpose and forage types, respectively. However, misclassification of accessions within each phenological class ranged from 2.8 in MMS to 46.7% in MML, with most accessions being misclassified as part of the same end-use type, with two exceptions in the EEM and MES phenological classes (Table 2).

Canonical correlation and perceptual mapping: Canonical R between the phenological and agronomic traits was high for grain and forage types (0.93 and 0.92, respectively) and moderate (0.64) for the dual-purpose

Table 1. Pairwise matrix correlations (above diagonal) and common principal components models (below diagonal) among 14 phenological classes identified in three end-use types of the Batini barley landrace.

Type	Phelonological Class			Grain			Dual purpose					Forage		
		EES	EEM	MES	EMM	MMS	MMM	MLM	LMS	LLS	LLM	EEL	EML	MML	MLL
Grain	**EES**: early heading, early maturity, short filling period	0	0.536	0.976	0.944	0.843	0.469	0.742	0.968	0.726	0.753	0.949	0.964	0.926	0.925
	EEM: early heading, early maturity, medium grain filling period	UR	0	0.751	0.534	0.227	0.943	0.195	0.629	0.239	0.241	0.509	0.393	0.713	0.446
	MES: medium heading, early maturity, short filling period	CPC3	PR	0	0.921	0.831	0.390	0.698	0.924	0.650	0.698	0.869	0.972	0.876	0.895
Dual	**EMM**: early heading, medium maturity, medium filling period	CPC4	UR	CPC3	0	0.742	0.372	0.646	0.866	0.657	0.802	0.868	0.885	0.938	0.966
Purpose	**MMS**: medium heading, medium maturity, short filling period	CPC3	UR	PR	CPC3	0	0.237	0.904	0.828	0.862	0.725	0.824	0.876	0.574	0.273
	MMM: medium heading, medium maturity, medium filling period	UR	CPC3	UR	UR	UR	0	0.209	0.602	0.223	0.104	0.462	0.335	0.574	0.273
	MLM: medium heading, late maturity, medium filling period	PR	UR	CPC2	CPC3	CPC4	UR	0	0.772	0.981	0.757	0.814	0.756	0.649	0.690
	LMS: late heading, medium maturity, short filling period	CPC4	CPC1	PR	CPC3	CPC3	CPC2	CPC3	0	0.758	0.692	0.954	0.938	0.931	0.865
	LLS: late heading, late maturity, short filling period	PR	UR	CPC2	CPC3	CPC3	UR	PR	CPC3	0	0.809	0.827	0.711	0.672	0.706
	LLM: late heading, late maturity, medium filling period	CPC3	UR	CPC2	CPC3	CPC2	UR	CPC2	CPC2	CPC3	0	0.753	0.733	0.748	0.854
Forage	**EEL**: early heading, early maturity, long filling period	PR	UR	CPC3	CPC2	CPC2	UR	CPC4	CPC4	CPC3	CPC3	0	0.898	0.883	0.876
	EML: early heading, medium maturity, long filling period	CPC3	UR	CPC3	CPC4	CPC2	UR	PR	CPC4	CPC2	PR	CPC4	0	0.880	0.917
	MML: medium heading, medium maturity, long filling period	PR	CPC1	CPC3	PR	UR	UR	CPC1	CPC4	CPC2	CPC2	CPC3	PR	0	0.940
	MLL: medium heading, late maturity, long filling period	CPC4	UR	PR	PR	UR	UR	CPC2	PR	CPC3	CPC4	CPC3	CPC4	CPC4	0

Underlined correlation coefficients are not significant at *P*=0.05.

Table 2. Discriminant analysis based on phenological and agronomic traits of 14 phenological classes identified in three end-use types of the Batini barley landrace.

Type/ Phenological class	Constant	$°Cd^{-1}h$	$°Cd^{-1}m$	Biological yield	Grain yield	Number of tillers	Plant height	Percent correct classification		Percent misclassified (80% of accessions) as:
								80% of collection	Validation (20%)	
Grain	-858.8	-0.032	0.235	1.059	-0.818	-83.7	0.844	98.5		1.5 Dual purpose
EES	-1454.0	-0.097	0.611	1.541	-1.43	-133.2	4.116	87.5	90.6	12.5 MES
EEM	-1491.8	-0.487	0.973	1.626	-1.60	-145.5	4.975	57.4	63.0	18.5 EMM, 24.1 MMS
MES	-1395.9	-0.491	0.987	1.499	-1.44	-131.9	4.439	58.5	60.2	18.6 MML, 22.9 MMS
Dual purpose	-1324.9	0.105	0.15	1.33	-0.99	-102.9	0.97	99.77		0.23 Grain
EMM	-1528.7	-0.13	0.707	1.511	-1.43	-131.5	4.458	77.0	81.7	23.0 MMM
MMS	-2120.9	0.10	0.502	1.911	-1.72	-162.8	4.781	97.2	98.3	2.8 LLS
MMM	-2042.0	-0.44	0.954	1.964	-1.87	-172.8	5.429	92.7	95.5	7.3 EMM
MLM	-2161.9	0.057	0.560	1.917	-1.72	-162.8	5.149	37.5	46.0	62.5 MMM
LMS	-2168.4	0.33	0.299	1.936	-1.75	-165.4	5.952	33.7	45.5	66.3 MMS
LLS	-2204.0	-0.04	0.670	1.893	-1.72	-161.8	5.413	91.2	95.6	8.8 LLS
LLM	-2275.6	-0.13	0.798	1.880	-1.75	-162.9	5.56	97.0	97.0	3.0 MMM
Forage	-1594.9	0.122	0.14	1.46	-1.09	-111.46	1.00	95.5		4.5 Dual purpose
EEL	-2515.4	0.260	0.380	2.110	-1.890	-178.08	5.390	90.5	95.3	9.5 EML
EML	-2642.9	0.060	0.600	2.130	-1.930	-180.40	5.390	95.9	97.3	4.1 MLL
MML	-2568.0	0.020	0.640	2.060	-1.860	-173.21	5.460	53.4	56.9	46.6 EML
MLL	-2281.0	0.030	0.590	1.870	-1.600	-151.90	3.540	72.2	78.9	27.8 EML
Multivariate test of significance										
Discriminant loading		0.75**†	0.65**	0.78**	0.51**	0.58**	0.58**			
Univariate F-ratio		137.5	127.2	143.6	147.3	109.2	112.4			

†Significant at the 0.01 level of probability.

type. The variance extracted by the first two canonical roots for each set of traits and end-use type ranged from high to extremely low. For example, variance extracted by the first canonical root from phenological traits was 0.52, 0.36 and 0.15 in the grain, dual-purpose and forage types, respectively, whereas the respective values for the agronomic traits were 0.59, 0.83 and 0.84.

Simple correlation coefficients among °C d⁻¹M and each of grain yield, biological yield and tillers per plant were significant for all phenological classes. Plant height, on the other hand, was not correlated with °C d⁻¹M. Also, the correlations among °C d⁻¹H and grain yield for all end-use types were not significant. We observed a steady increase in the magnitude of correlation coefficient between agronomic traits, except tillers/plant, and °C d⁻¹M, with the grain, dual-purpose and forage types having low, intermediate
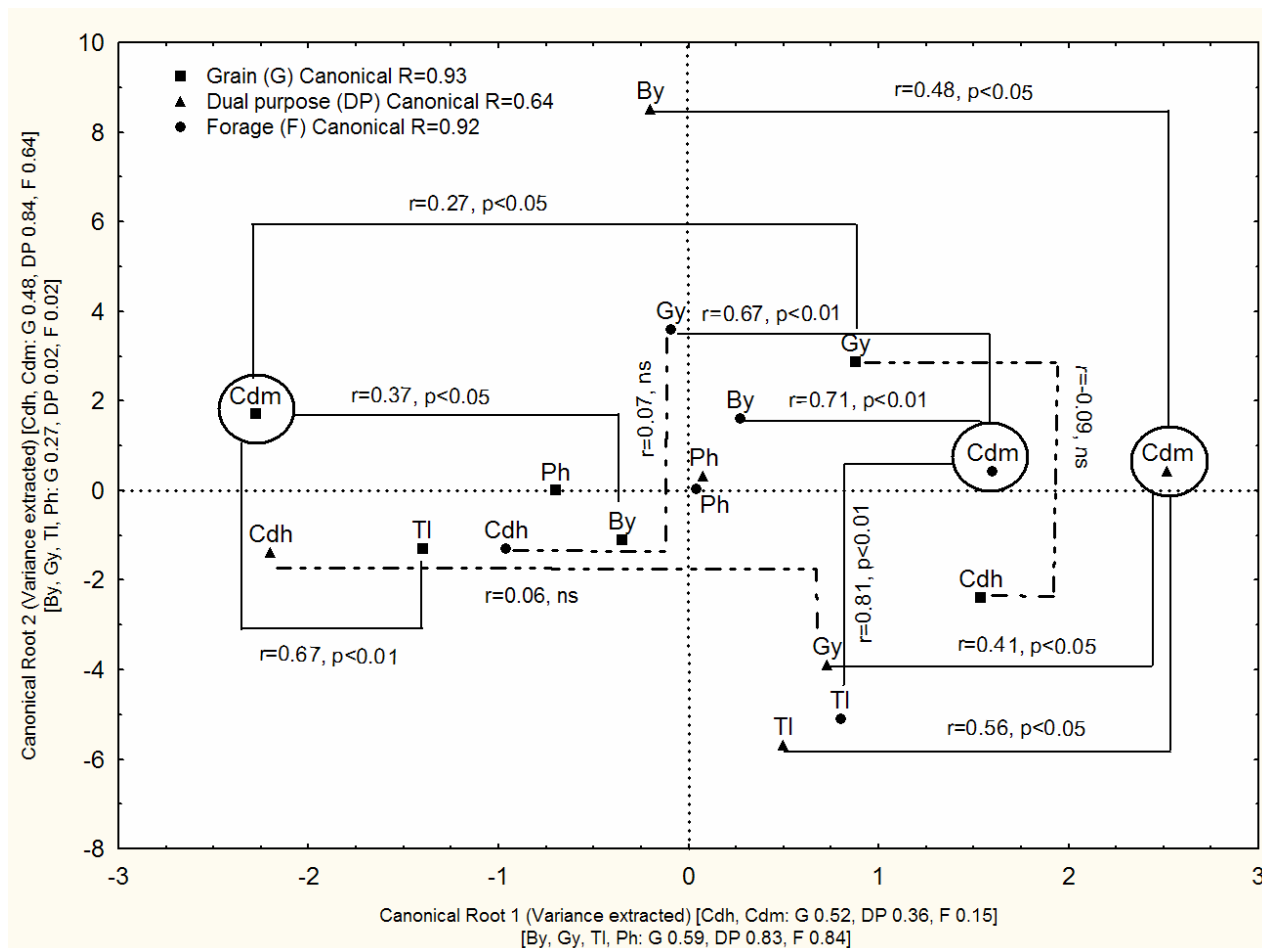
Figure 2. Biplot of six phonological, phenotypic and productivity traits in each of three end-use types of the Batini barley landrace from Oman.

and high values, respectively.

The display of $°C\ d^{-1}\ M$ and $°C\ d^{-1}\ H$ for each end-use type in the biplot (Fig. 2) suggests that these two phenological attributes are separated by the second canonical root for all end-use types, whereas $°C\ d^{-1}\ M$ and $°C\ d^{-1}\ H$ for the grain type were separated from those of the dual-purpose and forage types by the first canonical root. On the other hand, grain yield and biological yield of the forage type were the least separated, followed by those of the grain type and the dual purpose type.

Discussion

Total biomass was little affected by earliness in the grain type, as the negative correlation between $°C\ d^{-1}\ H$ and $°C\ d^{-1}\ Fp$ (r= -0.25, $P<0.05$) suggests that gains from longer $°C\ d^{-1}\ H$ are lost by a shorter $°C\ d^{-1}\ Fp$. However, the respective r-value in the forage-type was positive and highly significant (r = 0.63, $P<0.001$). Extreme negative and significant correlation (r = -0.95, $P<0.001$) between $°C\ d^{-1}\ H$ and $°C\ d^{-1}\ Fp$ in barley resulted in low biological yield and was due not only to shorter $°C\ d^{-1}\ Fp$, but also to longer $°C\ d^{-1}\ H$, associated with taller but fewer tillers per plant.

Agronomic diversity in the Batini barley landrace, similar to other landraces [1, 11] can be impacted by a combination of natural and anthropogenic factors. The high level of diversification in this and other landraces [1] within subsistence farming systems can be explained on the basis of environmental, socioeconomic and cultural factors, or as a result of farmers' selection being in favor of heterozygosity [4, 5].

Patterns of covariation of the set of phenological and agronomic traits, in support of earlier findings[2], proved to be more relevant for phenotypic evaluation than patterns of univariate variation. The former offers an effective method of stratifying and sampling variation in germplasm collections, and identifying germplasm for breeding programs [1, 4, 5]. The discriminating power of two principal components formed on the basis of six phenologic and agronomic traits explained >80% of total variance in this landrace as compared to 51% in a Spanish barley collection [19].

CPC models (Table 1) were specific to certain pairwise comparisons, ranging from no common components (i.e., UR) to CPC4 and proportionality. These results can be attributed to two major patterns in the covariance matrices [22]. The first: different phenological classes or end-use types showed high or low co-variances for different traits, especially $°C\ d^{-1}\ H$, and $°C\ d^{-1}\ M$ (e.g., Fig. 2 for end-use types). The second: other traits, especially plant height and biological yield, in addition to $°C\ d^{-1}\ H$, and $°C\ d^{-1}\ M$, had a disproportionately large influence on differences in the mean covariance (Fig. 2) and as depicted by their loadings on the first two principal components. These results support the suggestion that farmers' or natural selection has caused large, idiosyncratic changes in the principal component structure [22], and that some of the changes can be attributed to shifts in the mean phenotype [17].

The extensive divergence in the phenotypic covariance matrices among subpopulation and among phenological classes, as supported by the significant differences among eigenvalues and

among the first two eigenvectors, suggest that simultaneous changes involved the principal components structure and eigenvalues [10]. Widen *et al.* [22] demonstrated extensive population divergence in the genetic covariance matrix of *Brassica cretica*, and indicated that the first changes involved PC structure rather than the eigenvalues. However, Alonso and Herrera [2] showed a conservation of PC structure, despite large changes in eigenvalues in *Prunus mahaleb*. Biologically, these differences might correspond to changes in the regulation of developmental pathways involved in trait construction [10].

Discriminant analysis correctly predicted membership of end-use types (>95.5%) and most phenological classes (up to 97.2%) using a small number of variables measured on numerous individual plants, with an overall ratio of 1:340. This ratio is much smaller than the one (1:20) used by Ruiz *et al.* [19] and resulted in 98% discrimination among barley cultivars, or the one (1:46.5) used by Baker and Johnson [3] in their morphometric analysis of three species in the Cactaceae which resulted in correct classification of >93% of accessions. Discriminant analysis was useful in identifying the most influential traits and their discriminating power (Table 2). This is indicated [18] by the discriminant loading and the univariate F-ratios, the joint ranking of which suggests that biological yield, thermal time to anthesis, thermal time to maturity, grain yield, plant height and number of tillers per plant, in decreasing order, have the most discriminating power among phenological classes in this barley landrace.

References

[1]Alemayehu, F. and Parlevliet, J.E. 1997. Variation between and within Ethiopian barley landraces. Euphytica **94**:183-189.

[2]Alonso, C. and Herrera, C.M. 2001. Patterns made of patterns: Variation and covariation of leaf nutrient concentrations within and between populations of *Prunus mahaleb*. New Phytol. **150**:629-640.

[3]Baker, M.A. and Johnson, R.A. 2000. Morphometric analysis of *Escobaria sneedii* var. *sneedii, E. sneedii* var. *leei*, and *E. guadalupensis* (Cactaceae). Systematic Botany **25**:577-587.

[4]Ceccarelli, S., Grando, S. and van Leur, J.A.G. 1987. Genetic diversity of barley landraces from Syria and Jordan. Euphytica **98**:269-280.

[5]Ceccarelle, S., Grando, S., Tutwiler, R. Baha, J. Martini, A.M., Salahieh, A., Goodchild, A. and Michael, M. 2000. A methodological study on participatory barley breeding. I. Selection phase. Euphytica **111**:91-104.

[6]El-Karbotly, A., Mahgoub, O., Al-Subhi, A. and Al-Halhali, A. 2003. Indigenous grass species with potential for maintaining rangeland and livestock feedingin Oman. Agric. Ecosyst. Environ. **95**:623-627.

[7]Flury, B, 1988. Common Principal Components and Related Multivariate Models. John Wiley and Sons, New York.

[8]Giunta, F., Motzo, R. and Deidda, M.1999. Grain yield analysis of a triticale (x *Triticosecale* Wittmack) collection grown in a Mediterranean environment. Field Crops Res. **63**:199-210.

[9]Hair, J.F. Jr., Anderson, R.E., Tathman, R.L. and Black, W.C. 1998. Multivariate Data Analysis. 5th ed. Printice Hall, NJ.

[10]Houle, D., Mezey, J. and Galpern, P. 2002. Interpretation of the results of common principal components analyses. Evolution **56**:433-440.

[11]Jaradat, A.A., Shahid, M. and Al-Maskri, A.Y. 2004. Genetic diversity in the 'Batini' barley landrace from Oman I. Spike and seed qualitative and quantitative traits. Crop Sci. **44**:304-315.

[12]Jaradat, A.A., Shahid, M. and Al-Maskri, Y.A. 2005. Biomass production potential in the Batini barley landrace from Oman. JFAE **3**(2):249-253.

[13]Kebebew, F., Tsehaye and McNeilly, T. 2001. Morphological and farmer cognitive diversity of barley at Bale and North Shewa of Ethiopia.

Genet. Resour. Crop Evol. **48**:1-10.

[14]Lakew, B., Semeane, Y., Alemayehu, F., Gebre, H., Grando, S., van Leur, J.A.G. and Ceccarelli, S. 1997. Exploiting the diversity of barley landraces in Ethiopia. Genetic Resources and Crop Evolution **44**:109-116.

[15]Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. **27**:209-220.

[16]Phillips, P., Whitlock, M.C. and Fowler, K. 2001. Inbreeding changes the shape of the genetic covariance matrix in *Drosophila melanogaster*. Genetics **158**:1137-1145.

[17]Rieseberg, L.H., Widmer, A., Arntz, A.M. and Burke, J.M. 2002. Directional selection is the primary cause of phenotypic diversification. PNAS **99**:12242-12245.

[18]Rohlf, F.J. 2000. NTSYSpc Numerical taxonomy and multivariate analysis system. Version 2.1. User guide. Exeter software.

[19]Ruiz, M., Varela, F. and Carillo, J.M. 1997. Analysis of the discriminating power of agro/morphological and biochemical descriptors in a sample of Spanish collection of barley (*Hordeum vulgare* L.).Genetic Resour. Crop Evol. **44**:247-255.

[20]StatSoft, Inc., 2005. STATISTICA (data analysis software systems), version 7, www.statsoft.com.

[21]Tottman, D.R. 1987. The decimal code for the growth stages of cereals, with illustrations. Ann. Appl. Biol. **110**:441-454.

[22]Widen, B., Andersson, S., Rao, G.-Y.and Widen, M. 2002. Population divergence of genetic (co)variance matrices in a subdivided plant species, *Brassica cretica*. J. Evol. Biol. **15**:961-970.

[23]Wishart, D. 2003. ClustanGraphics Primer: A guide to Cluster Analysis. 2nd ed., Edinburgh, U.K.

[24]Yau, S-K. 2003. Yield of early planted barley after clipping and grazing in a semiarid area. Agron. J. **95**:821-827.