CBCS
http://agrobiol.sggw.waw.pl/cbcs

Communications in Biometry and Crop Science
Vol. 2, No. 1, 2007, pp. 41–48

International Journal of the Faculty of Agriculture and Biology,
Warsaw University of Life Sciences, Poland

NOTE

Genetic diversity in Pakistani populations of *Avena fatua* revealed by seed storage protein polymorphism

Bushra Mirza[1]*, Mohammad Shoaib[1], Munawar Ahmad[1], Yong-Bi Fu[2]

[1] Department of Biochemistry, Quaid-i-Azam University, Islamabad, 45320, Pakistan.
[2] Plant Gene Resources of Canada, Saskatoon Research Centre, Agriculture and Agri-Food Canada, 107 Science Place, Saskatoon, SK, Canada S7N 0X2.
* Corresponding author: Bushra Mirza, E-mail: bushra_dr@yahoo.com

CITATION: Mirza, B., Shoaib, M., Ahmad, M., Fu Y.B. (2007). Genetic diversity in Pakistani populations of *Avena fatua* revealed by seed storage protein polymorphism. *Commun. Biometry Crop Sci.* 2 (1), 41–48.

Received: 8 February 2007, Accepted: 31 May 2007, Published online: 11 June 2007

ABSTRACT

Avena fatua L. is a common weed of wheat fields and causes huge losses of grain yield in Pakistan. A preliminary study was conducted to assess the genetic diversity of 199 *A. fatua* plants collected from 13 districts in Pakistan using seed storage protein polymorphism. A total number of 34 clearly distinguishable protein bands were detected. The number of bands detected for each population ranged from 24 to 34, and the percentage of the polymorphic bands ranged from 37.9 to 97% with an average of 73.4%. These variations were not associated with latitude, longitude, elevation, or sample size. Analysis of molecular variance revealed 33.1% of the total protein variation resided among the 13 populations, 6% between white and black seed types and 10.3% between plants collected from irrigated and rain fed fields. The most diverse (and distinct) population was from the Swabi district and the least diverse population from the Lakkimarwat district. More variation was observed for irrigation-associated plants than rain-associated plants. Similarly, more variation was detected in plants with white seeds than black seeds. These findings are useful for developing effective strategies to control *A. fatua* for sustainable agricultural production systems in Pakistan. A more comprehensive assessment is currently being performed with microsatellite markers to cover more *A. fatua* populations under various weed management practices.

Key Words: *Avena fatua; genetic diversity; Pakistan; seed storage protein; weed, wild oat.*

INTRODUCTION

Avena fatua L. (wild oat) is one of the worst weeds in Pakistani wheat fields (Ahmad and Shaikh, 2003) and also is considered the 13th most important weed worldwide (Holm et al., 1977). Efforts have been made to control *A. fatua* largely by applying herbicides such as

Topic 15 WP (clodinafop propargyl) and Puma Super 75 EW (fenoxaprop-p-ethyl) (Marwat, 2002; Khan et al., 2003). However, little attention has been paid toward understanding the weed biology and genetic variation in *A. fatua* populations, particularly in Pakistan. Diversity studies of weeds can enhance the development and maintenance of effective weed management (Dekker, 1997; Marshall et al., 2003; Basu et al., 2004; Mengistu et al., 2005).

A. fatua is known to be an annual, predominantly self-fertilizing, hexaploid (2n=6x=42) grass with large morphological variations (Iman and Allard, 1965; Miller et al., 1982). An early allozyme survey of *A. fatua* populations revealed many polymorphic markers and substantial genetic variability (Jain and Marshall, 1967). A recent study on herbicide-resistant and herbicide–susceptible *A. fatua* populations showed up to 16% population differentiation and relatively large variation (up to 91%) residing within populations (Mengistu et al., 2005). However, no studies on genetic variation of Pakistani *A. fatua* populations have been found in the literature. Pakistani *A. fatua* plants grow in a variety of agro-ecological conditions and may harbour different levels of genetic variation, especially with respect to herbicide resistance.

The objective of this study was to assess the patterns of genetic variability within and among *A. fatua* populations using seed storage proteins. Seed storage protein polymorphism has been successfully detected in *A. fatua* (Peterson and Brinegar, 1986; Gregová et al., 1996; Dvoracek et al., 2003) and other plants (Gepts, 1990). It is our hope that this assessment yielded some baseline information for the development of effective weed control strategies for sustainable agricultural production systems in Pakistan.

MATERIALS AND METHODS

PLANT MATERIALS

The *A. fatua* germplasm used for this study was collected from five districts in Punjab province and eight districts in North West Frontier province, representing different agro-ecological conditions (Table 1). In each district, a field of approximately 2000–8000 m^2 was selected. About 10 single plants at least ten meters apart were randomly sampled, and their seeds were collected individually. For the field where there were two types of *A. fatua* seeds (black and white), sampling was made for both seed types, as seed color might reflect herbicide resistance. In this collection there were seven districts with both seed types and six districts with either black or white types. Thus, a total number of 20 populations were characterized in this study. Based on the habitat of a sampling field, the collected germplasm was also categorized into two subgroups designated as irrigated and rain-fed (Table 1).

PROTEIN EXTRACTION AND ELECTROPHORESIS

One seed from each plant sample was randomly selected for protein extraction. The selected seeds were individually ground to a fine powder using a mortar and pestle, and 0.01g of the powder was added into an Eppendorf tube for each seed with 400 uL of a protein extraction buffer. The extraction buffer contained 0.05 M Tris-HCl pH 8.0, 2.5% SDS, 10% Glycerol, 5 M Urea, 5% β–mercaptoethanol, and 0.2% Bromophenol (Ghafoor et al., 2002). The Eppendorf tubes were vortexed thoroughly using an Automatic Lab Mixer DH-10 to homogenize, and the homogenate samples were purified by centrifuging at 13000 rpm for 10 min at room temperature. The extracted crude proteins were recovered as clear supernatant, transferred into new 1.5 mL Eppendorf tubes and stored at 2°C for electrophoresis.

Electrophoresis was performed using a discontinuous sodium dodecylsulphate polyacrylamide gel electrophoresis (SDS-PAGE) system of Laemmli (1970) with 15% (w/v) separating gel and 4.5% stacking gel (Scheibe et al., 2001). A 15 µL of supernatant was loaded with the micropipette into the gel wells and run at a constant current of 90 Volts at room temperature (25°C) till the tracking dye migrated to the gel bottom. The gels were later stained in 0.04% Coomassie Brilliant Blue R-250 in methanol, acetic acid and distilled water

(45:10:45 v/v) for 8-10 hours. Next they were destained in the same solution without Coomassie Brilliant Blue R-250 with occasional shaking till the gels became clear. The resulting gels were photographed to visualize the protein band patterns.

Table 1. Description of sampling sites and variation patterns of seed storage protein in *A. fatua* populations with respect to sample site, habitat type, and seed color.

Site/frequency Code/habitat[a]	GIS information[b]			Sample size[c]		Band count[d]		Band frequency
	La	Lg	El	W	B	T	P	mean (range)
Abbotabad (AD)*	34.10	73.15	1255	10		26	50.0	0.49 (0.10-0.90)
Bunnu (BU)+	33.00	70.18	283	11	10	34	91.2	0.58 (0.05-0.95)
Charsadda (CA)+	34.07	71.45	276	10	10	34	88.2	0.53 (0.10-0.95)
Faisalabad (FD)+	31.30	73.05	171	10	9	34	85.3	0.63 (0.11-0.95)
Gujranwala (GA)+	32.10	74.12	223		10	24	50.0	0.66 (0.20-0.80)
Hafizabad (HD)+	32.05	73.40	193	10	10	33	84.8	0.55 (0.05-0.95)
Islamabad (ISD)*	33.40	73.10	77		10	30	70.0	0.41 (0.10-0.90)
Karak (KK)*	32.48	70.40	600	10		31	74.2	0.46 (0.10-0.80)
Lakkimarwat (LT)+	32.36	70.55	258	10		29	37.9	0.66 (0.10-0.90)
Mardan (MN)+	34.20	72.00	298	10	10	34	85.3	0.66 (0.20-0.95)
Peshawar (PR)+	34.02	71.37	288	9	10	34	94.1	0.51 (0.16-0.95)
Sialkot (ST)+	32.48	74.58	234	10		28	46.4	0.55 (0.20-0.90)
Swabi (SB)+	34.11	72.46	316	10	10	33	97.0	0.47 (0.15-0.95)
All white seeded				110		34	100.0	0.61 (0.34-0.98)
All black seeded					89	34	100.0	0.59 (0.36-0.94)
All irrigated				90	79	34	100.0	0.61 (0.48-0.96)
All rain-fed				20	10	34	91.2	0.50 (0.13-0.93)
Total				110	89	34	100.0	0.60 (0.47-0.96)

[a] Habitat type shown as irrigated with + and rain-fed with *.
[b] La – latitude (°N), Lg – longitude (°E), and El – elevation (meter).
[c] W – white seeded type and B – black seeded type.
[d] T – total number of bands observed and P – the percentage of the polymorphic bands.

DATA ANALYSIS

The number of monomorphic and polymorphic protein bands were counted for each sample and manually scored as 1 (present) or 0 (absent). The polymorphic bands were analysed for the level of polymorphism by counting the number of polymorphic bands and generating summary statistics on the band frequencies. To visualize the pattern of variation, the number of polymorphic bands was plotted with respect to their frequencies of occurrence in each district, habitat and seed types.

To assess the variation among samples of various groups (sample site, habitat and seed color), an analysis of molecular variance (AMOVA; Excoffier et al., 1992) was performed using Arlequin version 3.1 (Excoffier et al., 2005) on all 199 samples. This analysis not only allows the partition of the total protein variation into within- and among-group variation components but also provides a measure of inter-group genetic distances as the proportion of the total protein variation residing between any two groups (called Phi statistic; Excoffier et al., 1992). Models involving three types of structuring (sample site, habitat type, and seed type) were applied. Significances of resulting variance components and inter-group genetic distances were tested with 10,100 random permutations.

To assess the genetic associations of the *A. fatua* plants representing 13 districts, the inter-district distance matrices of the Phi statistic were analyzed using NTSYS-pc 2.01 (Rohlf, 1997) and clustered with the algorithm of unweighted pair-group methods using arithmetic averages (UPGMA). An individual pairwise similarity matrix for 199 samples was generated using a simple matching coefficient (Sokal and Michener, 1958) and converted to the Euclidean distance matrix for a principal coordinate analysis using the NTSYS-pc program. The first three resulting principal coordinate scores were plotted to assess the genetic associations of individual *A. fatua* plants.

RESULTS AND DISCUSSION

SEED STORAGE PROTEIN POLYMORPHISM

This study revealed a total of 34 polymorphic bands of seed storage protein in 199 samples of *A. fatua* germplasm collected in 13 districts of Pakistan. The molecular weights for these bands ranged from 14.2 to 66 KDa. The band frequencies in these samples ranged from 0.47 to 0.96 and averaged 0.60. This level of seed storage protein variation observed for *A. fatua* appears to be relatively lower than those reported in wheat (Alvarez et al., 2006) and other species (Gepts, 1990; Ghafoor et al., 2002).

POPULATION VARIATION

The total number of seed storage protein bands observed for each population varied from 26 to 34 and averaged 31. The percentages of polymorphic bands over the total bands detected ranged from 37.9 (Lakkimarwat population) to 97% (Swabi population) and averaged 73.4%. The mean band frequency for each population ranged from 0.41 (Islamabad) to 0.66 (Gujranwala, Lakkimarwat and Mardan) and averaged 0.55. The most diverse populations measured by the average pairwise difference within a population were Swabi (15.1), followed by Preshawar (14.0) and Charsadda (13.7). The least diverse populations were Lakkimarwat (3.9), followed by Abbotabad (5.1) and Gujranwala (5.2). Thus, relatively large protein variation was observed for these populations. However, the regression analyses revealed that these population variations in terms of polymorphic bands (%) and mean band frequency were not significantly associated with latitude, longitude, elevation or sample size with one exception. The percentage of polymorphic bands observed for each population was significantly associated with the number of samples assayed. Thus, a larger sample size is needed to make the assessment more informative.

Partitioning of the total seed storage protein variation on the basis of sample site (districts) into within-population and among-population components by AMOVA showed that 33.1% of the total variation resided among 13 populations and 66.9% within populations. This difference was statistically significant ($P < 0.0001$) based on the permutation test. Such a level of within-population variation is relatively lower than those reported based on quantitative traits (Iman and Allard, 1965) and molecular markers (Mengistu et al., 2005). This discrepancy could be partly due to the use of the small sample sizes for these populations.

When measured by the average pairwise difference obtained from AMOVA, the largest between-population difference was observed between the populations (0.72) at Lakkimarwat and Abbotabad followed by the populations (0.69) at Gujranwala and Sialkot. The least between-population difference was observed between the populations (0.08) at Charsadda and Preshawar followed by the populations (0.13) at Charsadda and Hafizabad. These differences can also be visualized in the inferred genetic relationships of the 13 populations given in a dendrogram (Figure 1). The most distant populations away from others were Abbotabad, followed by Sialkot, Gujranwala, Lakkimarwat, and Islamabad (Figure 1). However, the principal coordinate analysis appears to show that Swabi and Charsadda were the most distinct populations (Figure 2A), as samples from these two sites were largely

located over the opposite edges of the plot (Figure 2B). Note that these first two principal coordinate components totally explained a relatively large protein variation (47.3%).

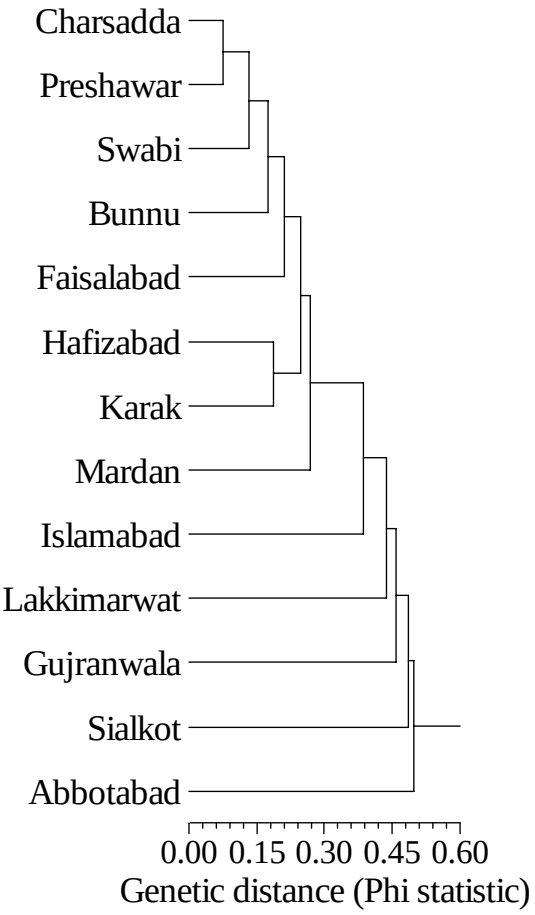


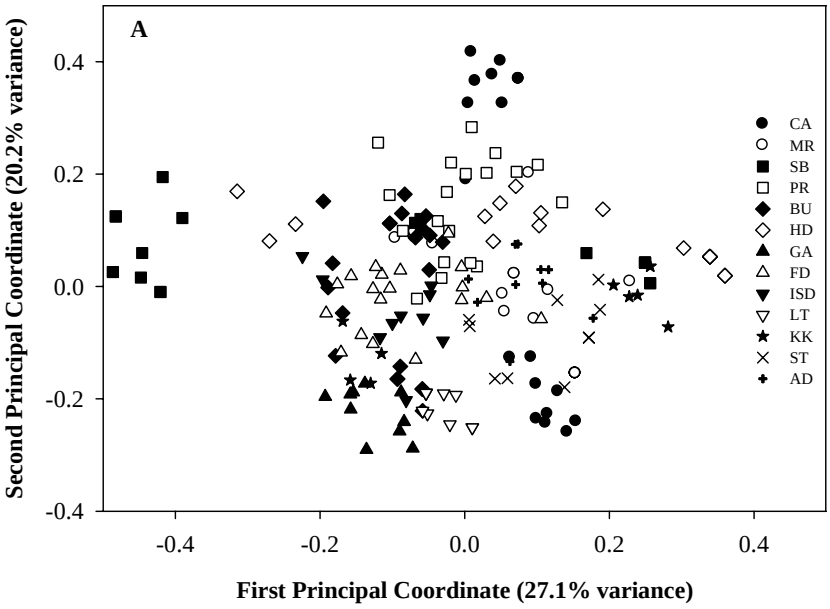
Figure 1. Dendrogram from UPGMA clustering of 199 *A. fatua* plants representing 13 districts in Pakistan.

VARIATION PATTERNS OVER SEED COLOR AND HABITAT TYPES

Comparisons of white and black seed germplasm revealed a significant 6% of the total protein variation explained by these two seed types (Table 2). Such a difference was not reflected in the total number of bands observed and the percentage of polymorphic bands, but rather in the band frequency distribution (Table 1). Measured by the average pairwise difference, the within-group variation for white seeds was slightly larger (15.2) than for black seeds (15.0). This difference was also obvious in the biplot of the two principal coordinate components, on which white seeds appear to spread more than black seeds (Figure 2B). As mentioned above, a marked difference existed between samples of two seed colors from Swabi and Charsadda. Further herbicide resistance studies on *A. fatua* plants with different seed colors may yield useful information for the weed management in Pakistani wheat fields.

Comparisons of *A. fatua* germplasm between irrigated and rain-fed habitats revealed more polymorphic bands observed (100%) and higher mean band frequency (0.61) for seeds collected in Irrigated field (Table 1). Irrigation status explained 10.3% of the total protein variation (Table 2). Such a large difference can also be seen in the biplot of the two principal coordinate components (Figure 2B), on which samples from irrigated fields spread much more widely than those from rain-fed fields. The within-group variation measured by

average pairwise difference for irrigated field was larger (15.4) than that for rain-fed fields (13.6). The reason for such population differentiation with respect to habitat is unknown, but the smaller rain-fed populations observed in the three districts may partly help to explain the finding of lower protein variation for plants grown in rain-fed habitat. Also, variable seed maturation conditions at the different districts may contribute to the difference in expressions of these storage proteins.



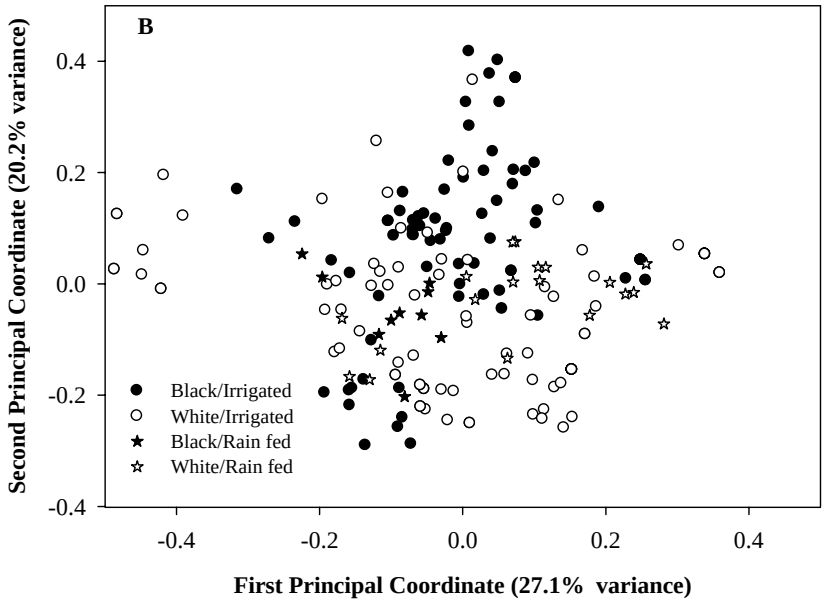


Figure 2. The genetic associations of 199 individual *A. fatua* samples as revealed by the principal coordinate analysis of 34 seed protein polymorphic bands with respect to seed color, habitat type and sampling site. Two graphs are the same, but labelled separately for (A) sample sites and (B) seed color and habitat types.

Table 2. Analysis of molecular variance (AMOVA) within and among the samples of *A. fatua* populations with respect to sample site, habitat type, and seed color.

Model and source of variation	df	SS	Variance component[a]	% variation
Sample site				
Among sites	12	547.23	2.65	33.1
Within sites	186	996.05	5.36	66.9
Seed color				
Among color types	1	55.26	0.48	6.0
Within color types	197	1488.01	7.55	94.0
Habitat type				
Among types	1	51.86	0.87	10.3
Within types	197	1491.41	7.57	89.7

[a] all the variance components were statistically significant at $P < 0.0001$ level, as calculated from random permutations.

IMPLICATIONS FOR WEED MANAGEMENT AND RESEARCH

This preliminary study represents our first attempt to address *A. fatua* control in Pakistani wheat fields. This study found 33.1% of the total protein variation resided among the 13 populations, 6% between white and black seed types and 10.3% between plants collected from irrigated and rain fed fields. The most diverse (and distinct) population was from the Swabi district and the least diverse population from the Lakkimarwat district. More variation was observed for irrigation-associated plants than rain-associated plants. Similarly, more variation was detected in plants with white seeds than black seeds. These findings help explain, at least partly, why current weed management practices were not always effective to control *A. fatua* populations. Such genetically diverse *A. fatua* populations may display various levels of herbicide resistance. The herbicide resistances of these samples are currently being assessed. These findings imply that various strategies may need to be considered for controlling *A. fatua* growing under different habitats. However, our present study is limited not only in the number of polymorphic bands detected, but also in the sampling scale of *A. fatua* distribution. A more comprehensive assessment is currently being performed with microsatellite markers to cover more *A. fatua* populations under various weed management practices.

REFERENCES

Ahmad, R., Shaikh, A.S. (2003). Common weeds of wheat and their control. *Pakistan J. Water Res.* 7, 73–74.

Alvarez, J.B., Moral, A., Martin, L.M. (2006). Polymorphism and genetic diversity for the seed storage proteins in Spanish cultivated einkorn wheat (*Triticum monococcum* L. ssp. *monococcum*). *Genet. Resour. Crop Ev.* 53, 1061–1067.

Basu, C., Halfhill, M.D., Muller T.C., Stewart, C.N. Jr. (2004). Weed genomics: new tools to understand weed biology. *Trends Plant Sci.* 9, 391–398.

Dekker, J. (1997). Weed diversity and weed management. *Weed Sci.* 45, 357–363.

Dvoracek, V., Curn, V., Moudry, J. (2003). Suitability of oat seed storage protein markers for the identification of cultivars in grain and mixed flour samples. *Plant Soil Environ.* 11, 486–491.

Excoffier, L., Laval, G., Schneider, S. (2005). Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol. Bioinformatics Online* 1, 47–50.

Excoffier, L., Smouse, P.E., Quattro, J.M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131, 479–491.

Gepts, P. (1990). Genetic diversity of seed storage proteins in plants. In: Brown, A.H.D, Clegg, M.T., Kahler, A.L. Weir, B.S. (Eds.): *Plant Population Genetics, Breeding and Genetic Resources*. Sinauer Associates Inc., Sunderland, Masschusetts, USA, 64–82.

Ghafoor, A., Ahmad, Z., Qureshi, A.S., Bashir, M. (2002). Genetic relationship in *Vigna mungo* (L.) Hepper and *V. radiate* (L.) R. Wilczek based on morphological traits and SDSPAGE. *Euphytica* 123, 367–378.

Gregová, E., Longauer, I., Žák, I., Kraic, J. (1996). Electrophoretic distinguishing of cultivated oats (*Avena sativa* L.) by seed storage proteins. *Rostl. Výroba* 42, 169–172.

Holm, L.G., Plunknett, D.L., Pancho, J.V., Herberger, J.P. (1977). *The world's worst weeds: distribution and biology*. Hawaii University Press, Honolulu, Hawaii, USA.

Iman, A.G., Allard, R.W. (1965). Population studies in predominantly self-pollinated species. VI. Genetic variability between and within natural populations of wild oats from differing habitats in California. *Genetics* 51, 49–62.

Jain, S.K., Marshall, D.R. (1967). Population studies in predominantly self-pollinating species. X. Variation in natural populations of *Avena fatua* and *A. barbata*. *Am. Nat.* 101, 19–33.

Khan, N., Gassan, G., Khan, M.A., Khan, I. (2003). Efficacy of different herbicides for controlling weeds in wheat crop at different times of application–I. *Asian J. Plant Sci.* 2, 305–309.

Laemmli, U.K. (1970). Cleavage of structural proteins assembly of the head of bacteriophage T4. *Nature* 227, 680–685.

Marshall, J.P., Brown, V.K., Boatman, N.D., Lutman, J.W., Squire, G.R., Ward, L.K. (2003). The role of weeds in supporting biological diversity within crop fields. *Weed Res.* 43, 77–89.

Marwat, M.I. (2002). Effect of weed management practices on different traits of wheat. *Online J. Biol. Sci.* 2, 719–721.

Mengistu, L.W., Messersmith, C.G., Christoffers, M.J. (2005). Genetic diversity of herbicide-resistant and susceptible *Avena fatua* populations in North Dakota and Minnesota. *Weed Res.* 45, 413–423.

Miller, S.D., Nalewaja, J.D., Mulder, C.E.G. (1982). Morphological and physiological variation in wild oat. *Agron. J.* 74, 771–775.

Peterson, D.M., Brinegar, A.C. (1986). Oat storage proteins, In: Webster, F.H. (Ed.): *Oats: Chemistry and Technology*. American Association of Cereal Chemists, St Paul, MN, USA 153–203.

Rohlf, F.J. (1997). *NTSYS-pc 2.1. Numerical Taxonomy and Multivariate Analysis System*. Setauket, NY: Exeter Software.

Scheibe, B., Weiss, W., Bauer, C., Gorg, A. (2001) Electrophoretic and immunochemical determination of bioindustrial enzymes in apple juice. *Eur. Food Res. Technol.* 212, 691–695.

Sokal, R.R., Michener, C.D. (1958). A statistical method for evaluating systematic relationships. *Kansas Univ. Sci. Bull.* 38, 1409–1438.