

GRAPHIC ANALYSIS OF TRAIT RELATIONS OF SPINACH (*SPINACIA OLERACEA* L.) LANDRACES USING THE BIPLLOT METHOD

Naser Sabaghnia¹, Farhad Behtash², Mohsen Janmohammadi¹

¹ Department of Agronomy and Plant Breeding, Faculty of Agriculture, University of Maragheh, Golshar, Madar Square, Amirkabir Highway, 55142-21111 Maragheh, Iran

² Department of Horticulture, Faculty of Agriculture, University of Maragheh, Golshar, Madar Square, Amirkabir Highway, 55142-21111 Maragheh, Iran

Abstract

SABAGHNIA NASER, BEHTASH FARHAD, JANMOHAMMADI MOHSEN. 2015. Graphic Analysis of Trait Relations of Spinach (*Spinacia oleracea* L.) Landraces Using the Biplot Method. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 63(4): 1187–1194.

Interest in growing spinach (*Spinacia oleracea* L.) in Iran is increasing due to its potential returns relative to other leafy vegetable crops. An experiment was conducted to examine the influencing characters on spinach by using 54 spinach landraces with application of the biplot methodology in visualizing research data. The study revealed that genotype by trait (GT) biplot can graphically display the interrelationships among traits and facilitate visual comparison of genotypes. The first two principal components (PC1 and PC2) of the site regression model accounted for 66% of the total variation in 54 landraces and accounted for 78% of the total variation in 16 clusters which are generated from cluster analysis of above landraces. The polygon view of GT biplot suggested seven sections for the 54 landraces and suggested four sections for the 16 clusters. The vertex landrace G38 had longest days to flowering and the vertex landrace G20 had most values for the other measured traits. The other vertex genotypes were not good in any traits. The vertex clusters were C1, C4, C8, and C14 which cluster C8 had the highest values of flowering period, leaf length and petiole diameter while, cluster C14 had the highest values for the other traits. The results can be used to design efficient breeding strategies in spinach.

Keywords: genotype by trait (GT) biplot, dry yield, polygon view, site regression

INTRODUCTION

Leafy vegetables such as spinach (*Spinacia oleracea* L.) are an important part in the human diets which are used as a salad, a cooked vegetable or as a component of many other cooked meat and vegetable dishes. The spinach is one of the dark green leafy vegetables which contain high beta carotene and folate, vitamin C, calcium, iron phosphorous, sodium and potassium (Dicoteau, 2000; Avsar, 2011). Spinach leaves are alternate, simple, from ovate to triangular-based, with larger leaves at the base of the plant and small leaves higher on the flowering stem (Vural *et al.*, 2000). Today, China, the United States and Indonesia are among the largest commercial producers of spinach (FAO, 2011). Iran's spinach production is

about 105 thousand tons per year and its average yield is 2096 kg ha⁻¹ while world's average yield is 2420 kg ha⁻¹ (FAO, 2011). Also, the spinach average yield of Iran is not low based on FAO statistics (2011), but it is low in comparison to the highest yield of world which is produced in Japan (12471 kg ha⁻¹), China (2768 kg ha⁻¹) and United States (2360 kg ha⁻¹).

The primary goal of a plant improvement program is to identify the most favorable genotypes for the desired characteristics which are evaluated on the basis of multiple traits. In fact, these breeding programs are conducted throughout the target region in which multiple traits are usually recorded (Yan and Rajcan, 2002; Dehghani *et al.*, 2008). Effective interpretation and utilization of such programs dataset is important at all stages of plant breeding, particularly when it is only

possible to select on yield components. Numerous statistical methods have been used in the explore for an understanding of the data patterns, although strategies may differ in overall appropriateness, different methods usually lead to the same or similar conclusions for a given dataset (Rubio *et al.*, 2004; Yan *et al.*, 2007).

Yan *et al.* (2000) developed a site regression methodology which contains genotype main effect plus genotype by environment (GGE) using biplot procedure as a graphical tool for analysis of multi-environment trials dataset. A biplot is a plot that simultaneously displays the effects of genotypes and the environment (Gabriel, 1971). The biplot according to GGE model is a biplot that displays the GGE of multi-environment trials dataset. It is constructed by plotting the first two principal components (PC1 and PC2) derived from singular value decomposition of the environment-centered data (Yan *et al.*, 2001). However, it can also be equally used for all types of two-way data that assume a two-way structure. The genotypes can be generalized as rows and the multiple traits as columns. Yan and Rajcan (2002) used a genotype by trait (GT) biplot, which is an application of the GGE biplot technique to study the genotype by trait data. The objective of this study was to evaluate spinach landraces base on the multiple traits and to study the interrelationships among spinach traits using GT biplot technique.

MATERIALS AND METHODS

Field Experiment

Fifty-four spinach landraces were collected as seed from different geographical regions of Iran, and then evaluated in the field according to a randomized complete block design (RCBD) replicated four times. The geographical properties of the 54 locations of the collected spinach landraces

are given in Tab. I. Field soil was loamy structure and had low organic matter which was poor for nitrogen and phosphorous. Fertilization was carried out by spreading 80 kg N ha⁻¹ (half of N at sowing stage and half of N at seedling emergence). Each field plot consists on six 3 m long rows with 25 cm between rows and plot size was 4.5 m². Sowing was done manually at the rate of 50 kg seed ha⁻¹. Control by hand was carried out twice when the weed density was high, in the pre-flowering and post-flowering stages. The harvested plot size was 2.5 m² (four 2.5 rows at the center of each plot). Several traits spinach traits were measured on 54 landraces and most of them were measured on 10 random (vying plant) points plot⁻¹. The measured quantitative traits were consist on: leaf length (LL), leaf width (LW), petiole length (PL), petiole diameter (PD), leaf area (LA), leaf numbers in flowering (LN), days to flowering (DF), female plants percent (FP), fresh yield (FY) and dry yield (DY).

Statistical Analysis

The Anderson and Darling normality test was used to normality test of dataset using MINITAB version 16 (2010) software. Cluster analysis is a tool for classifying objects into groups. The GT biplot analysis was performed on the sixteen cluster groups and nine quantitative traits. The genotype by trait (GT) biplot method (Yan and Rajcan, 2002) was used to show the spinach landrace by trait two-way data in a biplot. These statistical methods have been described in detail by Yan *et al.* (2000) and Yan and Rajcan (2002). Simple correlation coefficients between all possible combinations of traits were worked out according to Freund *et al.* (2010). To estimate the standard error of correlation coefficients, bootstrap analysis (Efron and Tibshirani, 1993) was performed by S-Plus (Springer-Verlag Inc., 2000) statistical package. All biplots presented in this study were generated using the software GGEbiplot package (Yan, 2001).

I: Names of the 54 collected spinach landraces

No.	Name	No.	Name	No.	Name	No.	Name
1	Arak	15	Rahnan 1	29	Gochan	43	Mashahad
2	Ardestan	16	Rahnan 2	30	Kashan	44	Malekan 1
3	Urmia	17	Zabol	31	Karaj 1	45	Malekan 2
4	Esfahan 1	18	Zanjan	32	Karaj 2	46	Minandab
5	Esfahan 2	19	Saveh	33	Karaj 3	47	Mianeh
6	Bojnord	20	Salmas	34	Kerman	48	Hamadan
7	Brojerd	21	Sanandaj	35	Kermanshah	49	Varamin 1
8	Beenab	22	Sirjan	36	Lahijan 1	50	Varamin 2
9	Birjand	23	Shiraz 1	37	Lahijan 2	51	Varamin 3
10	Tabriz	24	Shiraz 2	38	Langrood	52	Varamin 4
11	Chamkahriz	25	Shirvan	39	Mako	53	Varamin 5
12	Khoramabad	26	Salehabad	40	Mobarake	54	Yazd
13	Drood	27	Ajabsher	41	Maragheh 1		
14	Rahimabad	28	Qum	42	Maragheh 2		

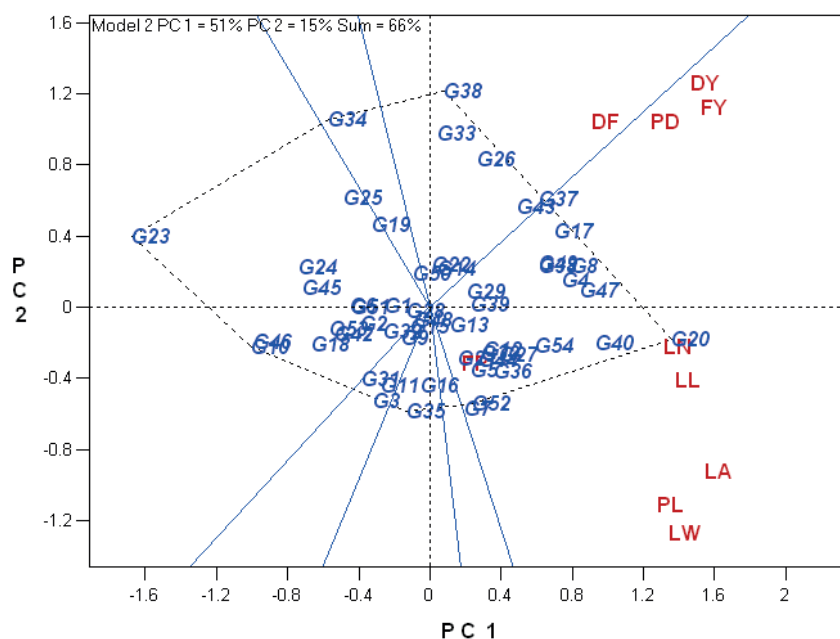
RESULTS AND DISCUSSION

The GT biplot for spinach dataset, based on proposed method of Yan and Rajcan (2002), explained 66% of the total variation of the standardized data. This relatively moderate percentage reflects the complexity of the relationships among the measured traits. The first two principal components (PC1 and PC2) explained 51% and 15%, respectively. In the GT biplot, a vector is drawn from the biplot origin to each marker of the genotype to facilitate visualization of the relationships between and among the traits as well as genotypes.

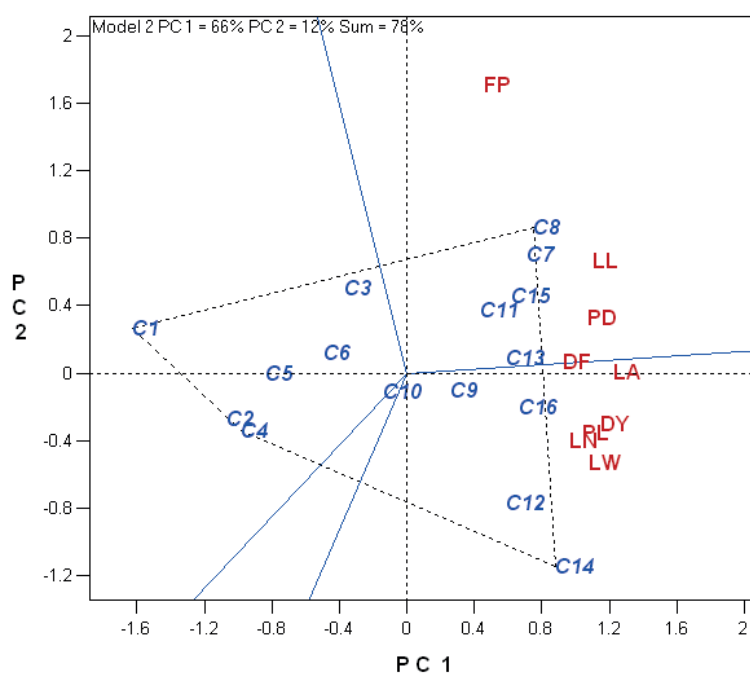
The biplot Fig. 1, as polygon view, presents data of 54 spinach landraces in ten different traits. The following can be seen from Fig. 1: The vertex landraces or genotypes in this investigation are G3, G16, G20, G23, G34, G35, and G38. These landraces are the best or the poorest landraces in some or all of the traits since they had the longest distance from the origin of biplot. Therefore, it seems that G20 had the highest values for all of the measured traits except days to flowering (DF). The G20 and the other genotypes of this sector had good amounts of leaf length (LL), leaf width (LW), petiole length (PL), petiole diameter (PD), leaf area (LA), leaf numbers in flowering (LN), female plants percent (FP), fresh yield (FY) and dry yield (DY). The vertex landrace G38 and its related genotypes which fall in its sector were good for days to flowering (DF). The other vertex genotypes (G3, G16, G23, G34 and G35) and related sectors were not good performance for the measured traits. The total variation explanation of biplot Fig. 1 was not high and so a cluster analysis was done and a dendrogram generated from landraces distance matrices could distinguished clearly sixteen clusters

(results are not shown). The number of clusters was verified by multivariate analysis of variance test through Wilks' Lambda statistics. Also, regarding high association among fresh and dry yield and for preventing of bias in biplot analysis, the fresh yield was omitted from further analysis. The GT biplot for spinach dataset, based on proposed method of Yan and Rajcan (2002), explained 78% of the total variation of the standardized data. This relatively moderate percentage reflects the complexity of the relationships among the measured traits. The first two principal components (PC1 and PC2) explained 66% and 12%, respectively. Therefore, reducing entry numbers from 54 landraces to 16 clusters and eliminating one of the two completely associated traits (FY) could increase the description of total variation using biplot model.

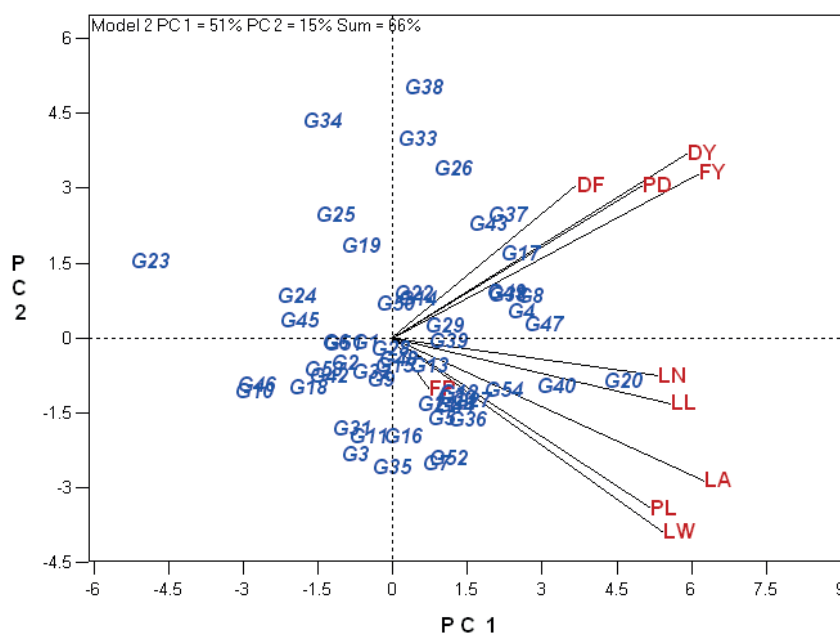
With knowing that higher dry yield is desirable for spinach varieties, the purpose of this exercise was to determine the most favorable genotype (cluster) for both breeding and cultivation proposes. The following can be seen from Fig. 2: The vertex clusters in this investigation are C1, C4, C8, and C14. These clusters are the best or the poorest landraces in some or all of the traits since they had the longest distance from the origin of biplot. Therefore, it seems that C8 had the highest values of flowering period (FP), leaf length (LL) and petiole diameter (PD) while, C14 had the highest values of days to flowering (DF), leaf area (LA), petiole length (PL), leaf number at flowering (LN), leaf width (LW) and dry yield (DY). The cluster C14 contains C9, C12 and C16 clusters while the cluster C8 contains C7, C8, C11, C13 and C15 clusters. The other two vertex clusters (C1 and C4) and their related clusters have



1: Polygon view genotype by trait biplot showing which genotype had the highest values for which traits for 54 spinach genotypes. For traits abbreviations, refer to the text



2: Polygon view genotype by trait biplot showing which genotype had the highest values for which traits for 16 spinach clusters. For traits abbreviations, refer to the text



3: Vector view genotype by trait biplot, showing the interrelationship among measured traits for 54 spinach genotypes. For traits abbreviations, refer to the text

not any traits and so all of these clusters were not high for the measured traits (Fig. 2).

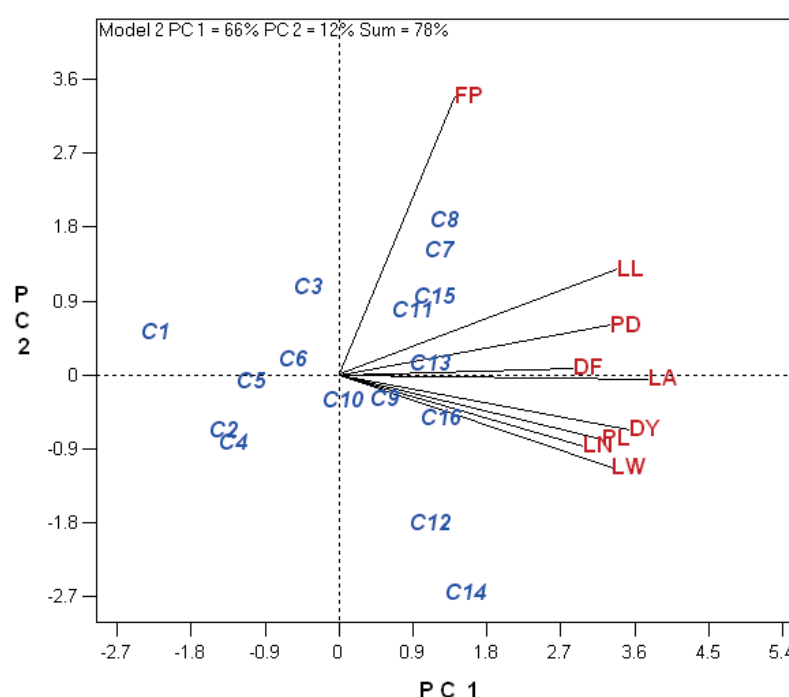
In the Fig. 3, the correlation coefficient between any two traits is approximated by the cosine of the angle between the vectors. The correlation coefficients among the traits indicate that the biplot currently shows relationship among the traits that had relatively large loading on both PC1 and PC2. Therefore, the most prominent relations by this figure (54 landraces and ten traits) are: a strong

positive association among DF, DY, PD and FY; and among LN and LL; and between PL and LW; and LA had positive correlation with LN and LL as well as PL and LW as indicated by the small obtuse angles between their vectors ($r = \cos 0 = +1$). There was a near zero correlation between DF, DY, PD and FY with PL and LW (Fig. 3) as indicated by the near perpendicular vectors ($r = \cos 90 = 0$). Some of mentioned results can be verified using correlation coefficients of Tab. II. Some discrepancies between

II: Simple correlation coefficients among ten spinach traits

	LL	LW	PL	PD	LA	LN	DF	FP	FY
LW	0.53**								
PL	0.50**	0.77**							
PD	0.64**	0.18	0.25						
LA	0.78**	0.87**	0.68**	0.46**					
LN	0.46**	0.53**	0.55**	0.41**	0.61**				
DF	0.10 ^{ns}	0.31*	0.19 ^{ns}	0.28*	0.27*	0.33*			
FP	0.20 ^{ns}	0.09 ^{ns}	0.02 ^{ns}	0.06 ^{ns}	0.12 ^{ns}	0.01 ^{ns}	0.16 ^{ns}		
FY	0.53**	0.42**	0.42**	0.69**	0.54**	0.53**	0.57**	-0.01 ^{ns}	
DY	0.49**	0.36*	0.40**	0.70**	0.48**	0.50**	0.57**	-0.04 ^{ns}	0.98**

Traits are: leaf length (LL), leaf width (LW), petiole length (PL), petiole diameter (PD), leaf area (LA), leaf numbers in flowering (LN), days to flowering (DF), female plants percent (FP), fresh yield (FY) and dry yield (DY)

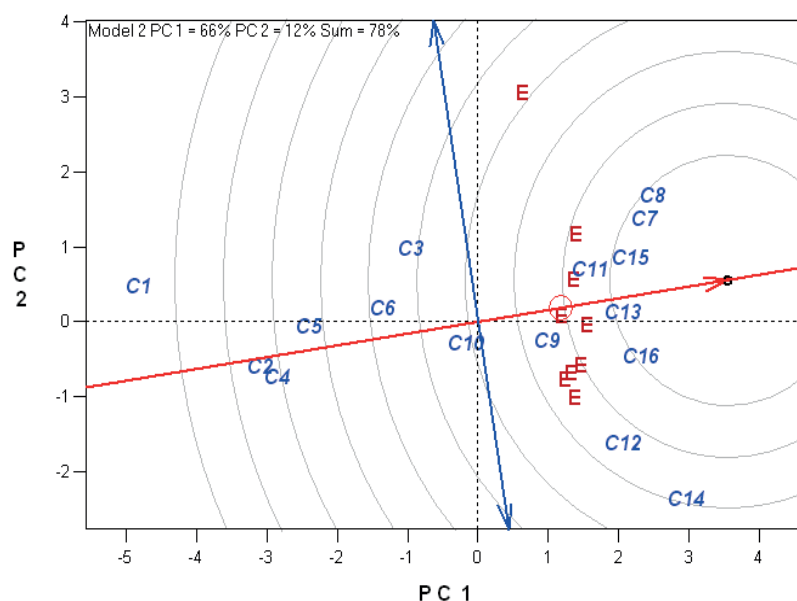


4: Vector view genotype by trait biplot, showing the interrelationship among measured traits for 16 spinach clusters. For traits abbreviations, refer to the text

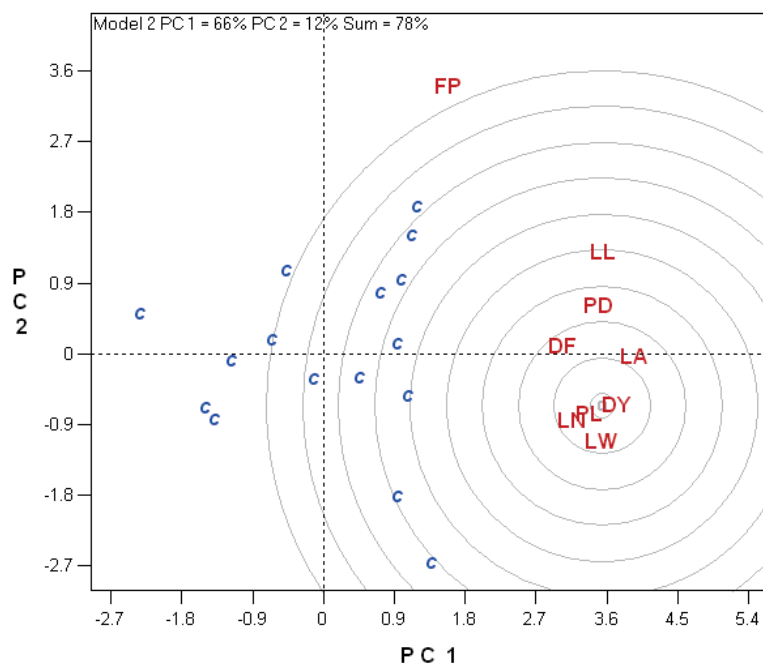
the biplot predictions and original data were expected because the biplot accounted for < 100% of the total variation (about 66%).

Similar to Fig. 3, the vector view of biplot was used to estimate the correlation coefficient among traits in sixteen clusters (Fig. 4). The most prominent relationships of Fig. 4 (16 clusters and nine traits) are: a strong positive association among DY, LW, PL and LN; and among DF and LA; and between LL and PD; and PD had positive correlation with DF and LA as indicated by the small obtuse angles between their vectors ($r = \cos 0 = +1$). There was a near zero correlation between DY, LW, PL and LN with FP (Fig. 4) as indicated by the near perpendicular vectors ($r = \cos 90 = 0$). Most of mentioned results can be verified using correlation coefficients of Tab. II. Some discrepancies between the biplot

predictions and original data were expected because the biplot accounted for < 100% of the total variation (about 78%). Ideal genotypes are those that should have large PC1 scores (high traits' means) and small (absolute) PC2 scores (low variability). Genotypes or clusters with above-average means (i.e., from G7 to G9 and from G11 to G16) was selected, whereas the rest (i.e., from G1 to G6 and G10) were discarded. Genotypes G7, G8, G13, G15 and G16 were the most favorable genotypes regarding all of the measured traits. Among these genotypes, G13 was the least variable genotype due to its low distance from horizontal axis while G8 was the most variable genotype due to its relatively high distance from horizontal axis. Conversely, G1 was the least favorable genotype or cluster among all studied clusters for its low mean yield for measured traits and high variability. The requirement for the use of



5: Ideal entry view of genotype by trait biplot, showing the relationships of 16 spinach clusters with ideal entry. For traits abbreviations, refer to the text



6: Ideal tester view of genotype by trait biplot, showing the relationships of traits of 16 spinach clusters. For traits abbreviations, refer to the text

site regression based GT biplots in the identification of most superior genotypes is to facilitate the identification of such genotypes (Crossa *et al.*, 2002). The study has clearly shown that the SREG model can analyze patterns and relationships of genotypes and traits successfully as well as provide a valuable prediction. Sabaghnia *et al.* (2010), however, mentioned that multivariate methods and their graphical tools are too sophisticated to provide a simple measure of genotypes favorability.

In spinach, improvement for achieving high leaf yield as an important desirable character is the purpose of many breeding programs. In the Fig. 6, spinach dry yield (DY) was compared with other measured traits. Ranking of other traits based on the DY was $PL > LN > LW > LA > DF > PD > LL > FP$. In other words, the most important traits for producing high yielding spinach varieties are PL, LN, LW and LA while the least important trait on spinach dry yield was FP. The relative contributions of different traits

of economical crop yield to the identification of desirable genotype found in this study by the traits comparing biplot procedure of the GT biplot are similar to those found in other crop studies (i.e. soybean (Yan and Rajcan, 2002), white lupin (Rubio *et al.*, 2004), rapeseed (Dehghani *et al.*, 2008), and canola (Sabaghnia *et al.*, 2010).

The GT biplot technique accounted only 66% of total variation for ten traits of 54 landraces while this technique accounted 78% of total variation for nine traits of 16 clusters (54 landraces clustered significantly in 16 groups). The results of the GT biplot analysis for the spinach in each dataset (54 landraces or 16 clusters) showed variable ranges of interactions, but regarding both biplots, at least one of the traits which related to leaf properties (leaf length, leaf width, leaf area and leaf numbers at flowering) or petiole properties (petiole length and petiole diameter) were identified as most influencing traits on dry yield. Therefore, for future spinach breeding programs it is essential to define selection indices based on above traits. Each one of the Fig. 1 and Fig. 3 suggest that there exist two

possible sections for spinach traits while Fig. 1 suggest seven possible sections for 54 spinach landraces and Fig. 3 suggest four possible sections for 16 spinach clusters. However, this section patterns needs verification through other spinach trials for this target region.

It also identifies superior entries with favorable traits effects. Sub-grouping of testers and entries by the GT biplot method will be useful for the breeding of lines for each target entry because it will help researchers exploit the interactions among entries and subsets of testers (Dehghani *et al.*, 2008). For several reasons, we prefer to the use of GT biplot: (i) by this method, the main effect of one of the genotype is not omitted, as in the case of GGE analysis; (ii), biplots are easy to use and to interpret; and (iii) other information such as the best mating partner for each genotype, groups of similar genotypes and the best crosses that are superior to their genotypes can be graphically visualized to rapidly identify suitable genotypes and crosses for further investigation.

CONCLUSION

The GT biplot analysis of 54 genotype showed variable ranges of interactions, and proposed one of the leaf-related traits (leaf length, leaf width, leaf area and leaf numbers at flowering) or petiole-related properties (petiole length and petiole diameter) as the most influencing traits on dry yield and so they could be used for future spinach breeding programs. Also the GT biplot had some benefits including; easy to use and to interpret; obtain other useful information such as the best mating partner for each genotype, and show visualization pattern to identify suitable genotypes.

Acknowledgement

We wish to thank Professor Dr. W. Yan (Eastern Cereal Oilseed Research Center of Agriculture and Agri-Food Canada) for making available a time-limited version of GGE biplot software as "Test Biplot.xlsx". Also, this research received funding from the Research and Technology Department of University of Maragheh, which is acknowledged kindly.

REFERENCES

- ASADI, H. A., HASANDOKHT, M. R. 2007. An evaluation of genetic diversity of Iranian spinach landraces. *Iran J. Hort. Sci.*, 38: 257–265.
- AVSAR, B. 2011. *Genetic diversity of Turkish spinach cultivars (Spinacia oleracea L.)*. A Master dissertation. Izmir, Turkey: Izmir Institute of Technology.
- CROSSA, J., CORNELIUS, P. L., YAN, W. 2002. Biplot linear-bilinear model for studying cross over genotype $\times$ environment interaction. *Crop Sci.*, 42: 619–633.
- DEHGHANI, D., OMIDI, H., SABAGHNIA, N. 2008. Graphic analysis of trait relations of canola (*Brassica napus* L.) using biplot method. *Agron. J.*, 100: 760–764.
- DICOTEAU, D. R. 2000. *Vegetable Crops*. New Jersey. USA: Prentice Hall.
- EFRON, B., TIBSHIRANI, R. J. 1993. *An Introduction to Bootstrap*. London, UK: Chapman and Hall.
- EFTEKHARI, S. A. A., HASANDOKHT, M. R., FATAHI-MOGHADAM, M. R., KASHI, A. A. K. 2010. Genetic diversity of some Iranian Spinach (*Spinacia oleracea* L.) landraces using morphological traits. *Iran J. Hort. Sci.*, 41: 83–93.
- FAO. 2011. FAOSTAT data of Food and Agriculture Organization of the United Nations. Rome, Italy: Food and Agriculture Organization. Available at: <http://faostat.fao.org/>. [Accessed: 2015. 7. 1].
- FREUND, R. J., MOHR, D., WILSON, W. J. 2010. *Statistical Methods*. 3rd edition. Amsterdam: The Netherlands. Academic Press.
- GABRIEL, K. R. 1971. The biplot graphic display of matrices with application to principal component analysis. *Biometrika*, 58: 453–467.
- MINITAB. 2010. MINITAB 16. MINITAB User's guide. Harrisburg, Pennsylvania, USA: Minitab Institute.
- RUBIO, J., CUBERO, J. I., MARTÍN, L. M., SUSO, M. J., FLORES, F. 2004. Biplot analysis of trait relations of white lupin in Spain. *Euphytica*, 135: 217–224.
- SABAGHNIA, N., DEHGHANI, H., ALIZADEH, B., MOGHADDAM, M. 2010. Genetic analysis

- of oil yield, seed yield, and yield components in rapeseed using additive main effects and multiplicative interaction biplots. *Agron. J.*, 102: 1361–1368.
- MATHSOFT. 2000. MathSoft, S-Plus 2000. Guide to Statistics. Seattle, USA: MathSoft Institute.
- VURAL, H., ESIYOK, D., DUMAN, I. 2000. *Vegetable Production*. Izmir, Turkey: Aegean University Press.
- YAN, W. 2001. GGEbiplot—A Windows application for graphical analysis multienvironment trial data and other types of two-way data. *Agron. J.*, 93: 1111–1118.
- YAN, W., RAJCAN, I. 2002. Biplot evaluation of test sites and trait relations of soybean in Ontario. *Crop Sci.*, 42: 11–20.
- YAN, W., HUNT, L. A., SHENG, Q., SZLAVNICS, Z. 2000. Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Sci.*, 40: 597–605.
- YAN, W., KANG M. S., MA, B., WOODS, S., CORNELIUS, P. L. 2007. GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop Sci.*, 47: 643–655.
- YAN, W., CORNELIUS P. L., CROSSA J., HUNT, L. A. 2001. Two types of GGE Biplots for analyzing multi-environment trial data. *Crop Sci.*, 41: 656–663.

Contact information

Naser Sabaghnia: sabaghnia@maragheh.ac.ir, sabaghnia@yahoo.com