GENETIC VARIATION OF SOME TOBACCO (*NICOTIANA TABACUM* L.) GENOTYPES BY MORPHOLOGICAL TRAITS

Z. Porkabiri[1], N. Sabaghnia[1], R. Ranjbar[2], H.H. Maleki[1]

[1]*University of Maragheh, Faculty of Agriculture, Department of Plant Production and Genetics, Maragheh, Iran*
[2]*Urmia Tobacco Research Center, Urmia, Iran*

Simple lattice design with two replicates was used to evaluate 25 tobacco genotypes. The following eight traits were measured and evaluated: plant height (PH), leaf length (LL), leaf width (LW), leaf number per plant (LN), stem girth (SG), leaf area (LA), dry leaf yield (DLW), and fresh leaf yield (FLW). The genotype by trait biplot for tobacco dataset explained 73% of the total variation of the standardized data. The polygon view of genotype by trait presented for the eight different traits of tobacco genotypes showed six vertex cultivars (G4, G5, G11, G12, G18, and G21). Genotype G21 had the highest values for most of the measured traits. Based on a vector view biplot, strong positive associations among LL, DLW, FLW, and LA were detected while near zero correlations between LL and SG, as well as between LL and PH were observed. Generally, based on ideal genotype and ideal tester biplots, it was demonstrated that the selection for high leaf yield should be performed via leaf length, leaf width, and leaf area traits identified as effective selection criteria for high yielding cultivars.

genotype-by-trait interaction, principal components analysis, trait associations

doi: 10.2478/sab-2019-0001
Received for publication on February 23, 2018
Accepted for publication on June 2, 2018

INTRODUCTION

Tobacco (*Nicotiana tabacum* L.) is one of the most commercially valued crops and its major breeding goal is to develop cultivars with acceptable yield performance. It is one of the most important commercial types for tobacco production in the world and it is one of the most important nonfood crops that are widely cultivated worldwide. It belongs to family *Solanaceae* which has many species (Ren, Timko, 2001). Tobacco is natural amphidiploid (2n = 4x = 48) arisen by hybridization of N. *sylvestris* and N. *tomentosiformis* as wild progenitor species (Murad et al., 2002). Numerous types of tobacco are defined by different indices such as region of production, intended usage, method of curing, biochemical characteristics, and morphological traits. Tobacco is an intensively tilled crop because cultivation has been used to control weeds and improve yields. Its traditional production systems comprise up to seven cultivation operations for soil preparation for tobacco plantlet transplanting. Such an intensive tillage leaves the soil bare and, therefore, it can contribute to soil losses by wind and runoff erosion (Benham et al., 2007).

Genetic diversity of crop resources provides an opportunity for breeders to develop new varieties with desirable traits, which include farmer-preferred traits. Natural genetic diversity has been explored within crop species to meet consumption requirements via introducing desirable genes and eliminating undesirable ones slowly, altering in the process of underlying heredity principle for several decades. The study of genetic diversity in tobacco is needed for breeding programs and several properties such as morphological (Wenping et al., 2009; Zeba, Isbat, 2011), and chemical traits (El-Morsy et al., 2009; Darvishzadeh et al., 2011) have already been used. The study of the genetic diversity of tobacco genotypes is important not only for germplasm studies, but also for the choice of parents in tobacco breeding. The results of genetic

Table 1. Name and origin of the studied tobacco genotypes.

Code	Name	Type	Curing method	Origin
G1	North Carolina	Virginia	flue-cured	USA
G2	Montcalm Brum 258	Virginia	flue-cured	Switzerland
G3	Pee Dee	Virginia	flue-cured	Germany
G4	Madar	Oriental	sun-cured	Iran
G5	Hicks 26-110	Virginia	flue-cured	USA
G6	Vir REE	Virginia	flue-cured	USA
G7	NC TC 52	Virginia	flue-cured	unknown
G8	T.R 1	Oriental	sun-cured	Iran
G9	K 394	Virginia	flue-cured	USA
G10	Rosecan Nela	Virginia	flue-cured	Canada
G11	Coker 254 36-150	Virginia	flue-cured	USA
G12	All Purpase	Virginia	flue-cured	USA
G13	South Carolina	Virginia	flue-cured	USA
G14	Coker 176-97	Virginia	flue-cured	UK
G15	X 55	Oriental	sun-cured	Turkish
G16	Bell	Virginia	flue-cured	USA
G17	Samatra 9	Semi-Oriental	sun-cured	unknown
G18	Coker 319	Virginia	flue-cured	USA
G19	Coker 258	Virginia	flue-cured	USA
G20	Vir Aurea	Virginia	flue-cured	unknown
G21	Vir Yold	Virginia	flue-cured	Germany
G22	C 319	Virginia	flue-cured	USA
G23	Coker 411	Virginia	flue-cured	USA
G24	Coker 55	Virginia	flue-cured	unknown
G25	Coker 411 26-130	Virginia	flue-cured	USA

diversity study provide estimates on the level of genetic variation among diverse materials that can be used in germplasm management, varietal protection, and tobacco improvement. Morphological characters have already been used to study the genetic diversity of tobacco germplasm (Lei et al. 1997; Zhang et al. 2005).

Several methods have been used to understand the data structure which may differ in overall appropriateness; different methods usually lead to similar conclusions for a given dataset. Yan et al. (2000) have developed a genotype main effect (G) plus genotype by environment (GE) biplot methodology for the graphical analysis of multi-environment trial data. Yan, Rajcan (2002) have used a genotype by trait (GT) biplot, which is an application of the GGE biplot to study the genotype by trait data. The present study was performed to study the interrelationship of tobacco yield components and their contribution to yield using the GT biplot technique.

MATERIAL AND METHODS

Field experiment

Twenty-five genotypes of tobacco including commercial cultivars or breeding lines with different origins were studied under field conditions (Table 1). Tobacco seeds were sown at a rate of approximately 5 g m^{-2} in bed and after sowing the seeds, a fine layer of well fermented and sieved sheep manure was spread over top of beds. Then tobacco seedlings were transplanted to plots when plant averaged about 12 cm in height. The experiment was performed in a simple lattice design (5 × 5) with two replicates while each plot was comprised of three 5 m lines, with a spacing of 0.65 × 0.20 m. The measured morphological traits were plant height (PH), leaf length (LL), leaf width (LW), leaf number per plant (LN), stem girth (SG), and leaf area (LA) which were recorded on 10 random plants in total competition per plot. Also, dry leaf yield (DLW) and fresh leaf yield (FLW) were evaluated using total plants of plots with the exception of border effects (Kara, Esendal, 1995).

Data analysis

Analysis of the genotype × trait (GT) biplot was used to show the two-way pattern of rye genotypes' traits in a biplot based on the following formula:

$$\frac{\alpha_{ij}-\beta_j}{\sigma_j} = \sum_{n=1}^{2}\lambda_n\xi_{in}\eta_{jn} + \varepsilon_j = \sum_{n=1}^{2}\xi^*_{in}\eta^*_{jn} + \varepsilon_j$$

where:
α_{ij} = mean value of genotype *i* for trait *j*
β_j = mean value of all genotypes in trait *j*
σ_{ij} = standard deviation of trait *j* among the genotype means
λ_n = singular value for principal component n (PCn)
ξ_{in} , η_{jn} = scores for genotype *i* and trait *j* on PCn, respectively
ε_{ij} = residual associated with genotype *i* in trait *j*

To achieve symmetric scaling between the genotype scores and the trait scores, the singular value λ_n has to be absorbed by the singular vector for genotypes ξ_{in} and that for traits η_{jn} , i.e., $\xi^*_{in} = \xi_{in}\lambda_n^{0.5}$ and $\eta^*_{jn} = \eta_{jn}\lambda_n^{0.5}$. Because of n = 2, only PC1 and PC2 are retained in the model and such a model tends to be the best for extracting pattern and rejecting noise from the data. Because the traits were measured in different units, the biplots were generated using the standardised values of the trait means using GGE biplot software (Yan, 2001).

RESULTS

The GT biplot for tobacco dataset explained 73% (59% and 14% for PC1 and PC2, respectively) of the total variation and this relatively moderate percentage reflects the complexity of the relationships among the traits. The polygon view of the GT biplot helps identify genotypes with the highest values for one or more traits and provides the best way for visualising the interaction patterns between genotypes and traits and to effectively interpret a biplot. For this purpose, the genotypes are connected with straight lines and are formed a polygon which the other traits contained within it. Fig. 1 presents data of 25 tobacco genotypes in eight traits and the following information can be understood: the vertex genotypes in this investigation are G4, G5, G11, G12, G18, and G21; these genotypes are the best or the poorest genotypes in some or all of the traits since they had the longest distance from the origin of biplot. Therefore, genotype G21 had the highest values for LL, LW, LA, DLW, and FLW, while genotype G4 had the highest values for PH. Genotype G12 had the highest values for SG, while genotype G18 had the highest values for LN. The vertex genotype and the other genotypes which fell in related sector had good amounts of the above mentioned traits. The other vertex cultivars (G5 and G11) and related sectors had not suitable performance for all of the measured traits (Fig. 1).

Provided that the GT biplot model described relatively a sufficient amount of the total variation, the correlation coefficient between any two traits is approximated by the cosine of the angle between their

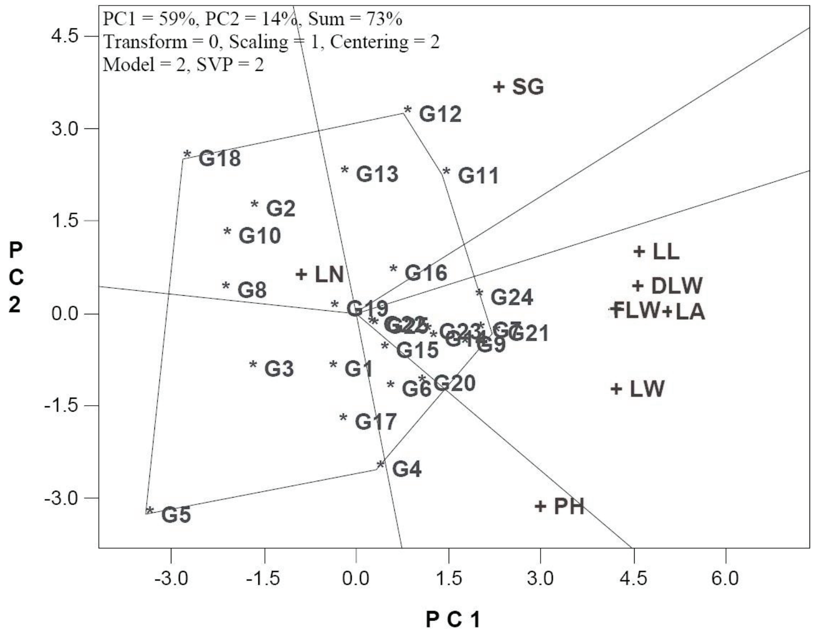


Fig. 1. Polygon view genotype by trait biplot, showing which genotype had the highest values for which traits for 25 different tobacco genotypes. Traits: PH = plant height, LL = leaf length, LW = leaf width, LN = leaf number per plant, SG = stem girth, LA = leaf area, DLW = dry leaf yield, FLW = fresh leaf yield

Table 2. Pearson's correlation coefficients between yield and morphologic traits of tobacco

PH	PH	LL	LW	LN	SG	FLW	DLW	LA
LL	1.00	0.35^{ns}	0.62^{**}	-0.10^{ns}	-0.01^{ns}	0.48^{*}	0.45^{*}	0.51^{**}
LW		1.00	0.68^{**}	-0.13^{ns}	0.47^{*}	0.82^{**}	0.83^{**}	0.93^{**}
LN			1.00	-0.10^{ns}	0.24^{ns}	0.66^{**}	0.61^{**}	0.89^{**}
SG				1.00	-0.03^{ns}	-0.13^{ns}	-0.25^{ns}	-0.13^{ns}
FLW					1.00	0.31^{ns}	0.42^{*}	0.41^{*}
DLW						1.00	0.92^{**}	0.83^{**}
LA							1.00	0.80^{**}
PH								1.00

PH = plant height, LL = leaf length, LW = leaf width, LN = leaf number per plant, SG = stem girth, LA = leaf area, DLW = dry leaf yield, FLW = fresh leaf yield, ns = not significant at $P > 0.05$, *significant at $P < 0.05$, **significant at $P < 0.01$

vectors (Yan, Kang, 2003). The most prominent relations by vector-view biplot (Fig. 2) are: a strong positive association among LL, DLW, FLW, and LA as indicated by the small obtuse angles between their vectors (r = cos 0° = +1). There was a near zero correlation between LL and SG, as well as between LL and PH (Fig. 2) as indicated by the near perpendicular vectors (r = cos 90° = 0). There was a negative correlation between LW and LN as indicated by the near angle of approximately 180° (r = cos 180° = –1). Some above discrepancies of the biplot predictions and original data were expected because the biplot accounted for < 100% of the total variation (Table 2).

Ideal genotypes should have large PC1 scores (high traits' means) and small (absolute) PC2 scores (low variability), therefore genotypes with above-average means were selected, whereas the rest were discarded. Genotype G24 was the most favourable genotype regarding all of the measured traits due to its low distance from horizontal axis (Fig. 3). Ranking of the best genotypes based on the ideal genotype was G7 > G21 > G9 > G14 > G23 and ranking of the most unfavourable genotypes based on the ideal genotype was G5 and G18 (Fig. 3).

In tobacco, improvement for achieving high dry leaf yield (DLW) as an important desirable character is the purpose of many breeding programs. In Fig. 4, DLW was compared with other measured traits and the ranking of measured traits based on DLW was FLW > LL > LA > LW. In other words, the most important traits for producing high yielding tobacco cultivars are leaf length, leaf width, and leaf area while the least important trait for tobacco yield were plant height, leaf number per plant, and stem girth.

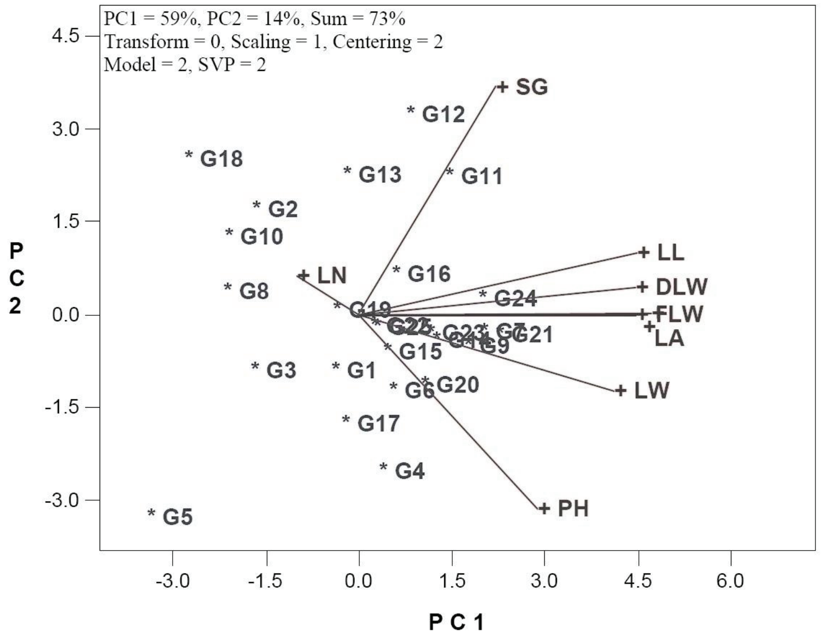


Fig. 2. Vector view genotype by trait biplot, showing the interrelationship among all measured traits for 25 different tobacco genotypes. Traits: PH = plant height, LL = leaf length, LW = leaf width, LN = leaf number per plant, SG = stem girth, LA = leaf area, DLW = dry leaf yield, FLW = fresh leaf yield

Ranking of tobacco genotypes based on DLW in Fig. 5 indicated that genotypes G7, G21, and G24 were the most favourable genotypes for producing high DLW and having high genetic potential for improving DLW in tobacco. Ranking of the other best genotypes based on the high DLW was G9 > G14 > G11 > G23 > G20 > G16> G15> G6 and ranking of the most unfavourable genotypes based on the DLW potential was G5 > G18 > G10 > G8 > G3 > G2 (Fig. 5).

DISCUSSION

According to different vertex genotypes of polygon view of biplot and their related genotypes, which are located in different sections, we found they are good candidates for examination heterosis (Yan et al. 2007) for hybrid production in tobacco using these genotypes or pure lines. Also, genotype G21 and its sector's genotypes such as G7, G9, G14, G15, G23, and G24

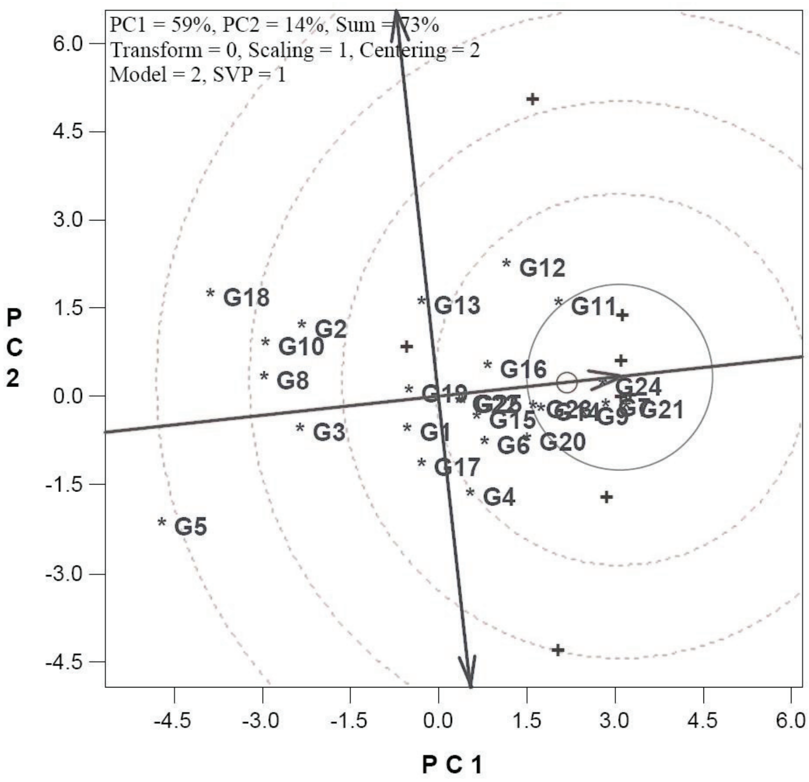


Fig. 3. Ideal entry view of genotype by trait biplot, showing the relationships of tobacco genotypes with ideal entry. Traits: PH = plant height, LL = leaf length, LW = leaf width, LN = leaf number per plant, SG = stem girth, LA = leaf area, DLW = dry leaf yield, FLW = fresh leaf yield

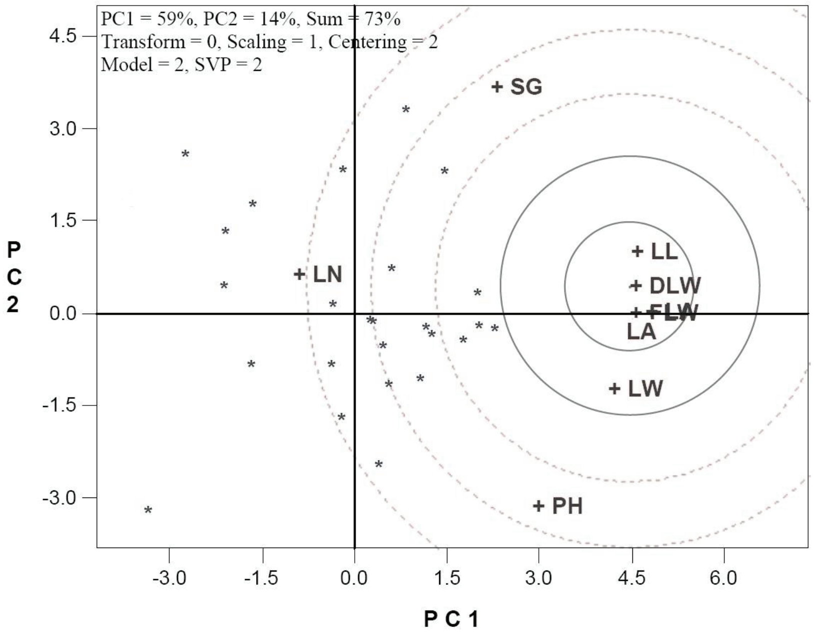


Fig. 4. Tester view of genotype by trait biplot, showing the relationships of tobacco genotypes with target tester as dry leaf yield. Traits: PH = plant height, LL = leaf length, LW = leaf width, LN = leaf number per plant, SG = stem girth, LA = leaf area, DLW = dry leaf yield, FLW = fresh leaf yield

could be used for improving fresh and dry leaf yield traits in tobacco breeding program.

The present research has clearly shown that the site regression model can analyse patterns and relationships of genotypes and traits successfully as well as provide a valuable prediction. Also, most of the studied genotypes were good candidates for improving most of the measured traits and therefore there has been good genetic variability in our plant materials.

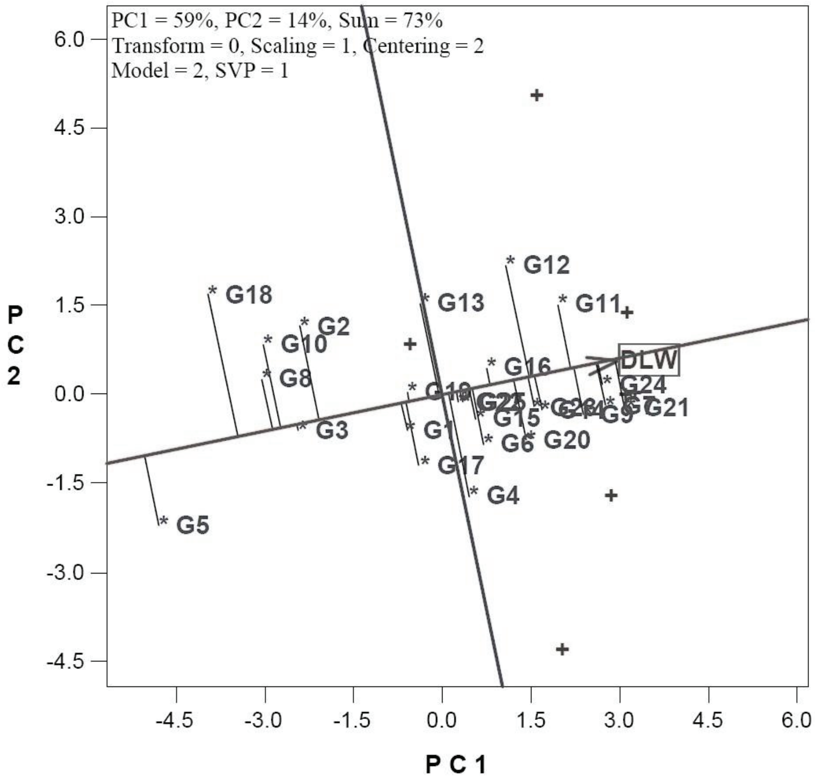
It is noticeable that the leaf properties (length, width, and area) are important for improving dry leaf yield and so it seems that defining breeding strategies for genetic improvement of dry leaf yield must be performed based on them. The relative contributions of different traits of economical crop 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 – soybean (Yan, Rajcan, 2002), white lupin (Rubio et al., 2004), and rapeseed (Sabaghnia et al., 2010).

The data obtained in this study could be useful for tobacco breeders in efforts to increase leaf yield. The correlation coefficients between the tobacco leaf yield and morphological traits showed good variation, and the results suggest that the leaf length, leaf width, and leaf area are the primary selection criteria for higher leaf yield in tobacco. It is clear that the GT biplot method is an excellent tool for visual genotype by trait data analysis because it is a powerful tool and can be used to graphically address research questions (Yan, 2014), and compared with conventional methods, the GT biplot approach has some advantages: (1) graphical presentation of the data greatly enhancing the ability to understand the patterns of the data, (2) it is more interpretative and facilitates pair-wise genotype or trait comparisons and effectively reveals the interrelationships among the tobacco traits, (3) it facilitates identification of possible which-won-where patterns or which-lost-where patterns, (4) it can be used in independent culling based on multiple traits and in comparing selection strategies (Yan, Rajcan, 2002; Yan et al. 2007).

CONCLUSION

In conclusion, based on the GT biplot, leaf length, leaf width, and leaf area were identified as traits suitable for selection for dry leaf yield improvement in tobacco. Thus, selecting for these traits is expected to improve dry leaf yield and this suggests that selection index that incorporates these traits will result in the development of not only high yielding cultivars, but those showing other desirable agronomic traits. From our observations, it appears possible to improve exotic

Fig. 5. Ranking of genotypes based on dry leaf yield (DLW) according to genotype by trait biplot method. Traits: PH = plant height, LL = leaf length, LW = leaf width, LN = leaf number per plant, SG = stem girth, LA = leaf area, DLW = dry leaf yield, FLW = fresh leaf yield



tobacco genotypes or lines by selecting for genotypes with higher values of the above mentioned traits. Also, almost all of the studied genotypes, especially G6, G9, G11, G14, G15, G16, G20, and G23, are good candidates for improving most of the measured traits due to the existence of good genetic variability in our plant materials.

ACKNOWLEDGEMENT

We wish to thank kindly Professor Wei-Kai Yan (Eastern Cereal Oilseed Research Center of Agriculture and Agri-Food Canada) for making available a time-limited version of GGEbiplot as ‘testBiplotxlsx’.

REFERENCES

Benham BL, Vaughan DH, Laird MK, Ross BB, Peek DR (2007): Surface quality impacts of conservation tillage practices on burley tobacco production systems in southwest Virginia. Water Air Soil Pollution, 179, 159–166.

Darvishzadeh R, Alavi SR, Sarafi A (2011): Genetic variability for chlorine concentration in oriental tobacco genotypes. Archives of Agronomy and Soil Science, 57, 167–177.

El-Morsy SI, Dorra MDM, El-Hady EAA, Hiaba AA, Mohamed A (2009): Comparative studies on diploid and tetraploid levels of *Nicotiana alata*. Academic Journal of Plant Sciences, 2, 182–188.

Kara SM, Esendal E (1995): Heterosis and combining ability analysis of some quantitative characters in Turkish tobacco. Tobacco Research, 21, 16–22.

Lei YH, Xu ML, Huang XY (1997): Tobacco collection in the Yunnan Province. Science and Technology publishing company of Yunnan, Yunnan, China.

Murad L, Lim KY, Christopodulou V, Matyasek R, Lichtenstein CP, Kovarik A, Leitch AR (2002): The origin of tobacco’s T genome is traced to a particular lineage within *Nicotiana tomentosiformis* (Solanaceae). American Journal of Botany, 89, 921–928.

Ren N, Timko MP (2001): AFLP analysis of genetic polymorphism and evolutionary relationships among cultivated and wild *Nicotiana* species. Genome, 44, 559–571.

Rubio J, Cubero JI, Martin LM, Suso MJ, Flores F (2004): Biplot analysis of trait relations of white lupin in Spain. Euphytica, 135, 217–224.

Sabaghnia N, Dehghani H, Alizadeh B, Mohghaddam M (2010): Genetic analysis of oil yield, seed yield, and yield components in rapeseed using additive main effects and multiplicative interaction biplots. Agronomy Journal, 102, 1361–1368.

Wenping LI, Zhu L, Zhao S (2009): Correlation and path coefficient analysis and Euclidean distance clustering for several characters in tobacco germplasm resource. Chinese Tobacco Science, 30, 59–63.

Yan W (2001): GGE biplot – a Windows application for graphical analysis multienvironment trial data and other types of two-way data. Agronomy Journal, 93, 1111–1118.

Yan W (2014): Crop variety trials: data management and analysis. John Wiley and Sons, Inc.

Yan W, Kang MS (2003): GGE biplot analysis: a graphical tool for geneticists, breeders, and agronomists. CRC Press, Boca Raton.

Yan W, Rajcan I (2002): Biplot evaluation of test sites and trait relations of soybean in Ontario. Crop Science, 42 11–20.

Yan W, Hunt LA, Sheng Q, Szlavnics Z (2000): Cultivar evaluation and mega-environment investigation based on the GGE biplot. Crop Science, 40, 597–605.

Yan W, Kang MS, Ma B, Woods S, Cornelius PL (2007): GGE biplot vs. AMMI analysis of genotype-by-environment data. Crop Science, 47, 643–655.

Zeba N, Isbat M (2011) Multivariate analysis for yield and yield contributing traits in F0 and F1 generations in tobacco (*Nicotiana tabacum*). Journal of Experimental Bioscience, 2, 101–106.

Zhang H.Y., Liu X.Z., He C.S., Zheng C.M. (2005): Random amplified DNA polymorphism of *Nicotiana tabacum* L. cultivars. Biologia Plantarum, 49, 605–607.

Corresponding Author:

Naser Sabaghnia, Associate Professor, University of Maraghh, Faculty of Agriculture, Department of Plant Production and Genetics,
P.O. box 55181-83111, Maragheh, Iran, phone: +98 413 280 6781, e-mail: sabaghnia@maragheh.ac.ir ; sabaghnia@yahoo.com