Genotype x Environment Interaction and Stability Analysis for Grain Yield of Advanced Tef Genotypes using GGE Biplots

Nigus Belay   Gethaun Bekana
Ethiopian Institute of Agricultural Research, P. O. Box: 2003, Addis Ababa, Ethiopia

Abstract
Selection of the most high-yielding and stable cultivars across environments are difficult because of the complex nature of genotype × environment interactions (G x E interaction). The study was conducted with the objectives to determine the G x E interaction of tef genotypes and to identify tef genotypes with high stability for grain yield using GGE biplot analysis. In the present study, 12 advanced tef genotypes and one standard check were evaluated at Holetta and Ginchi in 2014 and 2015 and at Adadi in 2014. Combined analysis of variance revealed the existence of significant G x E interaction for tef grain yield. Genotype, environment and G x E interaction explained 6.70 %, 81.03 % and 12.27 % of the variation in grain yield, respectively. The GGE model showed that the first and second principal component axis accounted for 43.2 % and 29.5 % of variability, respectively. The pattern of G x E interaction was a crossover type as revealed by differential yield ranking of the genotypes across environments. Genotype G4 (RIL- 77C) was both high yielding and stable across the test environments and could be considered as desirable genotype recommended for release as variety. The test environment E3 (Adadi-2014) was identified to be the most discriminating and representative environment to evaluate tef genotypes.
Keywords: Tef, GGE biplot, stability, genotype x environment interaction.

1. Introduction

Tef, Eragrostis tef (Zucc.) Trotter is the major staple cereal of Ethiopia. It is cultivated annually on more than three million hectares of land by over six million smallholder farmers, accounting one-third of the total cereal acreage (Kebebew et al., 2015). The production of tef in Ethiopia has increased from 1.74 million metric tons to 5.28 million metric tons in between in 2000/01 to 2017/18 at an estimated rate of 7.97 % per annum per annum whereas tef productivity has increased by 5.06% per annum during the same period to reach the current yield level of 1.73 tons/ha (Solomon et al., 2019). Utilization of tef grain as food crop has been limited to Ethiopia for several centuries. However, currently tef has been gaining global popularity as health food because of tef is gluten-free, which is suitable for peoples suffering from gluten protein allergy known as celiac disease (Spaenij-Dekkingetal., 2005) and many other health-related benefits.

Ethiopia is both the origin and center of diversity for tef (Vavilov, 1951). Eragrostis tef (Zucc.) Trotter, is an allotetraploid ($2n = 4x = 40$) and belongs to the Poaceae or grass family (Kebebew et al., 2015). The genus Eragrostis constitutes about 350 species of which only tef is cultivated for human consumption (Watson and Dallwitz, 1992).

Selection of high yielding cultivars with wide adaptability is a major objective of in a crop improvement program. However, attaining this goal challenged by the complex nature of genotype x environment interactions (G x E interaction) (Gauch and Zobel, 1996; Annicchiarico and Perenzin, 1994). Newly-developed tef cultivars need to be evaluated in a range of multi-location and multi-year trials before being recommended to be grown in given locations in order to select best cultivar and to determine if the target region can be subdivided into different mega-environments (Yan et al., 2000). G x E interaction reveals the inconsistent performance of a genotype in different environments and at the same time the information obtained from G x E interaction offers opportunities, especially in the selection and adoption of genotypes showing stable performance across environments and those with specific adaptation (Annicchiarico, 2002). Analysis of variance (ANOVA) only tests the significance of the G × E interaction but it does not provide information into particular pattern of genotype or environment that gives rise to G × E interaction (Samonte et al., 2005). When G × E interaction is significant, one of the options available to the breeder is to use stability analyses to identify the most high-yielding and stable cultivar. Several biometrical methods have been developed to analyze and visualize the nature and magnitude of G x E for multi‑environment trial data. Genotype plus genotype by environment interaction (GGE) has been suggested as the appropriate model for analyzing multi-environment trials Yan et al. (2000). The GGE biplot is based on principal component analysis and graphically displays the two-way (genotype x environment) data matrix and allows visualization mega‑environment analysis ('which‑won‑where' pattern), genotype evaluation (mean vs. stability), and test environment evaluation (discriminating power vs. representativeness) (Yan et al., 2007; Samonte et al., 2005). The objectives of this study were to determine the G x E interaction of tef genotypes and to identify tef genotypes with high stability for grain yield based on GGE biplot.

2. Materials and Methods

2.1 Experimental sites and materials

Thirteen tef genotypes consisting of 12 advanced tef genotypes (recombinant inbred lines) and one standard check (DZ-01-99) obtained from Debrezeit agricultural research center were evaluated at Holetta (altitude 2400 masl, 09°03"N, 38°30"E) and Ginchi (2200 masl, 09°30"N, 38°30"E) during 2014 and 2015 main production seasons and in Adadi (2383 masl, 08°31"N, 38°13"E) during the 2014 main cropping season, which are located in the central highlands of Ethiopia. The 12 advanced recombinant inbred lines (RILs) were generated through F2 derived single-seed descent (SSD) method from two different crosses; eight of them from the hybridization of DZ-Cr-387 and Gealmie while the remaining four from Gomade and Gealmie (Table 1). Each combination of year and location were considered as one environment for statistical analysis, totaling five environments, namely Holetta-2014 (E1), Ginchi-2014 (E2), Adadi-2014 (E3), Holetta-2015 (E4) and Ginchi-2014 (E5).

2.2 Experimental Design and Data Collected

The experimental design was a randomized complete block design with four replications of 2 m x 2 m (4m^2) plot size during the two main seasons of 2014 and 2015. The field experiment was managed as per the research recommendation of agronomic practices of the respective test locations. Grain yield (g) of each plot was measured on clean, sun dried seed and the measured grain yield value (g) has converted to kilogram per hectare for data analysis.

Table1. Description of tef genotypes used for the experiment

Genotypes	Genotype code
DZ-01-99 (standard check)	G1
DZ-Cr-387 x Gealmie (RIL- 96)	G2
DZ-Cr-387 x Gealmie (RIL- 185)	G3
DZ-Cr-387 x Gealmie (RIL- 77C)	G4
DZ-Cr-387 x Gealmie (RIL -124A)	G5
DZ-Cr-387 x Gealmie (RIL- 222B)	G6
DZ-Cr-387 x Gealmie (RIL- 239)	G7
DZ-Cr-387 x Gealmie (RIL- 129)	G8
Gomade x Gealmie (RIL -94)	G9
Gomade x Gealmie (RIL -6)	G10
Gomade x Gealmie (RIL- 100)	G11
Gomade x Gealmie (RIL -102)	G12
DZ-Cr-387 x Gealmie (RIL -25B)	G13

3. Statistical Analysis

3.1 Combined Analysis of Variance

The grain yield data were subjected to combined analysis of variance using PROC GLM in SAS using a RANDOM statement with the TEST option (SAS Institute, 2011) to determine the effects of genotype, environment and genotype × environment interaction. The combined ANOVA was done considering year-location combination as the environment. Genotype was considered as the fixed effect while environment was considered as a random effect.

3.2 GGE Biplot Analysis

GGE biplot was computed using the "GGEBiplotGUI" package of R statistical software in RStudio (Frutos et al., 2014; R Core Team, 2019) to analyze the multi-environment trial data, and evaluate the adaptability and stability of the cultivars and the effects of genotype, environment, and G × E interaction. A GGE biplot is a biplot that displays the genotypic main effect and G x E interaction of a multi-environment trial based on principal component analysis (PC1 and PC2) derived from subjecting a two way data (Genotype x environment data) to singular value decomposition (Yan et al., 2000). The GGE model used was:

$$Y_{ij} - \mu - \beta_j = \lambda_1 \xi_{i1} \eta_{j1} + \lambda_2 \xi_{i2} \eta_{j2} + \varepsilon_{ij}$$

where Y_{ij} was the measured mean of ith genotype in j^{th} environment; μ was the grand mean; β_j was the main effect of jth environment; $\mu + \beta_j$ was the average trait over all genotypes in j^{th} environment; λ_1 and λ_2 were the singular values for the first and second principal component (PC1 and PC2), respectively; ξ_{i1} and ξ_{i2} were eigenvectors of ith genotype for PC1 and PC2; η_{j1} and η_{j2} were eigenvectors of jth environment for PC1 and PC2; and ε_{ij} was the residual of the model associated i^{th} genotype in j^{th} environment.

4. Results and discussion

4.1 Analysis of variance

The combined analysis of variance for grain yield of the 13 tef genotypes tested in five environments showed significant differences among genotypes (G), environments (E) and G x E interaction (Table 2). A large grain yield variation revealed by environments which explained 81.03 % of the total G + E + G x E variation, while the effects of the genotypes and G x E interaction contributed 6.70 % and 12.27 % of the total variation, respectively. These results in agreement with previous findings on tef (Habte et al., 2019), maize (Thokozile et al., 2014) and sorghum (Asfaw et al., 2011). Genotype mean grain yield (averaged across environments) ranged from 1874 kg ha^{-1} for G6 to 2429.5 kg ha^{-1} for G8 and Environment mean grain yield (averaged across genotypes) ranged from 1595 kg ha^{-1} at E4 to 3009.6 kg ha^{-1} at E3 (Table 3).

Table 2. Combined analysis of variance for grain yield (kgha^{-1}) of 12 tef genotypes grown at five environments.

Source	DF	SS	MS	F	Probability	% SS
Genotype (G)	12	6190036.15	515836.35	2.18	0.0281	6.7
Environment (E)	4	74870175.38	18717543.85	79.22	<.0001	81.03
Rep (E)	15	20612992.31	1374199.49	12.82	<.0001	
G X E	48	11341494.62	236281.14	2.2	0.0001	12.27
Error	180	19301007.7	107227.8			
TOTAL	259	132315706.2				
G+E+G x E	64	92401705.53				

G x E = G x E interaction; Rep (E) = Replication nested in each location; DF = Degree of freedom; SS = Sum squares; MS = Mean squares; F = F value; and % SS = Percent sum squares explained

The presences of significant G x E interaction effect indicate variable phenotypic performance of the tested genotypes across environments because of the impact of environment on trait expression. A large G x E interaction effect compared to genotype effect suggests the possible existence of diverse mega-environments with different winner genotypes (Yan and Kang, 2003). Mega-environment was defined as group of locations that consistently share the same best cultivars (Yan and Rijcan, 2002).

Table 3. Mean grain yield (kgha^{-1}) of tef genotypes tested at five environments

Genotypes	Environments					
	E1	E2	E3	E4	E5	Mean
G1	2577.5	2150	3082.5	1685	2570	2413
G2	2327.5	1702.5	3352.5	1725	2070	2235.5
G3	2972.5	1277.5	3055	1752.5	1615	2134.5
G4	2800	1765	3295	1492.5	2442.5	2359
G5	2192.5	1452.5	3180	1657.5	2627.5	2222
G6	2332.5	1425	2142.5	1337.5	2132.5	1874
G7	2132.5	1497.5	3047.5	1775	2460	2182.5
G8	2477.5	2162.5	3070	1900	2537.5	2429.5
G9	2757.5	1425	3010	1400	1912.5	2101
G10	2417.5	1390	3042.5	1467.5	2135	2090.5
G11	2095	1687.5	2795	1357.5	1962.5	1979.5
G12	2477.5	1447.5	2830	1500	2260	2103
G13	2440	1497.5	3222.5	1685	1990	2167
Mean	2461.54	1606.15	3009.62	1595	2208.85	2176.23

4.2 GGE biplot analysis

4.2.1 Polygon view of GGE biplot (which-won-where)

Figure 1 presents a polygon view of 13 tef genotypes tested at five environments. The first two principal components (PC1 and PC2) obtained by singular value decomposition of environment-centered data of grain yield explained 72.6 % of the total effect it had on the grain yield variation with PC1 and PC2 accounted for 43.2 % and 29.5 % of variability, respectively (Figure 1) using environment centered data. The polygon view GGE biplot indicates best genotype(s) in each environment and groups of environment (Yan and Hunt, 2002). The plot is formed by connecting the vertex genotypes (located farthest away from the biplot origin) while the rest are inside

the polygon with perpendicular lines radiating from the origin of the biplot divide the biplot into different sectors. The highest yielding genotype (winning genotype) for an environment or set of environments in a sector is the vertex genotype (Yan et al., 2010). From polygon view of GGE biplot, the polygon divided into four sectors. Three environments, E2, E4 and E5, fell in the first sector with vertex genotype G8 implying that this genotype was the winning genotype for these environments. Sector 2 comprised one environment (E3) with two genotypes (G2 and G4) where G4 was the highest yielder. The remaining environment (E1) was contained in sector 3 with G3 being high yielding genotype. Sector 4 in the polygon consisted of G6 as vertex genotype had no test environment indicating that the genotype was poor performer in all test environments. Thus, the G × E interaction was a crossover type where a change in performance ranking of the genotypes across environments observed.

4.2.2 Mean vs. stability and genotype comparison with ideal genotype views of GGE biplot
Ranking of 14 tef genotypes based on their mean yield and stability performance are presented in figure 2. The average environment coordinate (AEC) view based on genotype-focused singular value partitioning (SVP = 1) and mean value can be referred as the "mean vs. stability" view of GGE biplot (Yan et al., 2007). The single arrowed line shown on AEC abscissa points to higher mean yield across environments. Hence, genotype G8 had the highest mean grain yield followed by G1 and G4 while genotype G6 had the lowest. The stability of the genotypes are determined by their projection on to the AEC vertical axis with the most stable genotype was located on the AEC horizontal axis and had minimum projection on the AEC vertical axis. Thus, genotype G10 and G12 were the most stable followed by G4 and G2. While genotype G3 followed by G6 and G9 were the least stable for grain yield. Yan and Tinker (2006) reported that stability is important only when coupled with high trait mean. Hence, an ideal genotype would be one that has both high mean yield performance and high stability across environment. The "comparison with ideal genotype" view of GGE biplot has concentric circles with the ideal genotype in the inner circle (Figure 3). It permits to visualize the distance between each genotype and the ideal genotype; a genotype is more desirable than others if it is located closer to ideal genotype. Therefore, G4 was the most desirable genotype and could be considered as widely adaptable genotype.
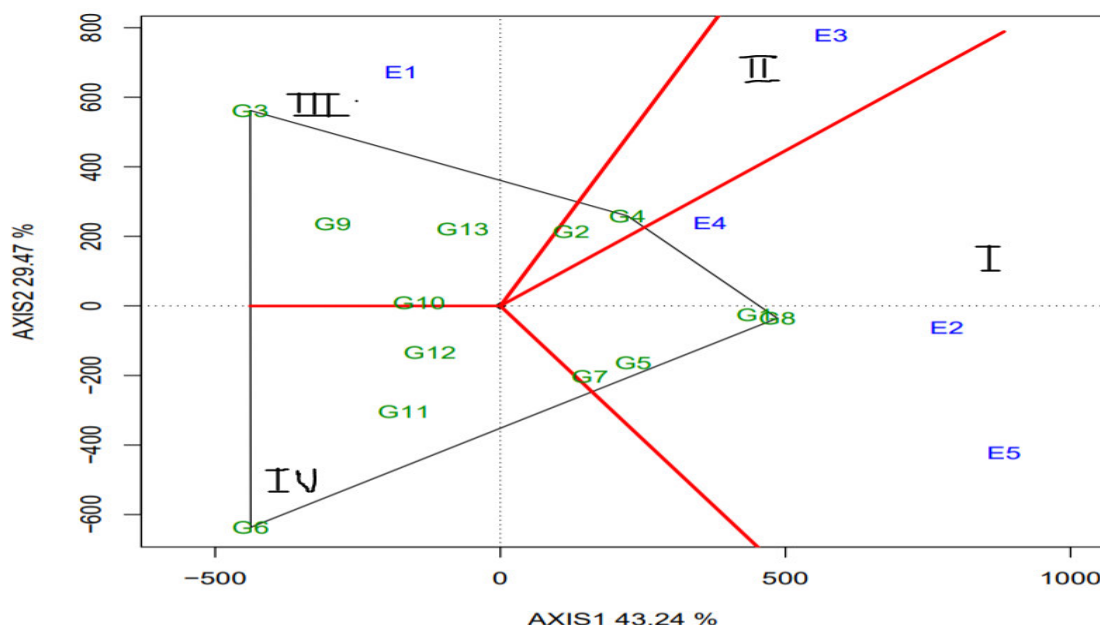

Fig. 1. The polygon (which–won–where) view of genotype main effects plus genotypic x environment interaction effect (GGE) biplot of 13 tef genotypes tested in five environments for grain yield

4.3.3 Discriminative vs. representativeness and ranking environments relative to an ideal environment
Evaluation of test environments is crucial to identify the most desirable genotypes for a mega environment in variety performance trial. Figure 4 shows the "discriminating ability vs representativeness" view of the GGE biplot. The distance between the markers of the environment to the biplot origin, is a measure of its discriminating ability (Frutos et al., 2013). Test environments with longer vectors are more discriminating of the genotypes whereas a test environment marker with a short vector provides little information about the genotypes differences (Yan et al., 2007). Hence, among the five environments evaluated, E3 followed by E5 were the most discriminating of the genotypes while E4 was the least discriminating of all test environments.

According to Yan and Tinker (2006), the representativeness of testing environment is visualized by the angle between environment vector and abscissa of average environment axis. The smaller the angle, the more representative of the test environment would be (Yan et al., 2007). Thus, E4 followed by E3 were identified to be more representative environments. The ideal test environment is one that is most discriminating for genotypes and

is representative of the target environments (Yan and Kang, 2003). The comparison with the ideal environment view of GGE biplot (Figure 5) has concentric circles with the ideal environment in the inner circle. An environment is more desirable and discriminating when located closer to the ideal environment (Naroui et al., 2013). Therefore, E3 was more representative and discriminating environment.
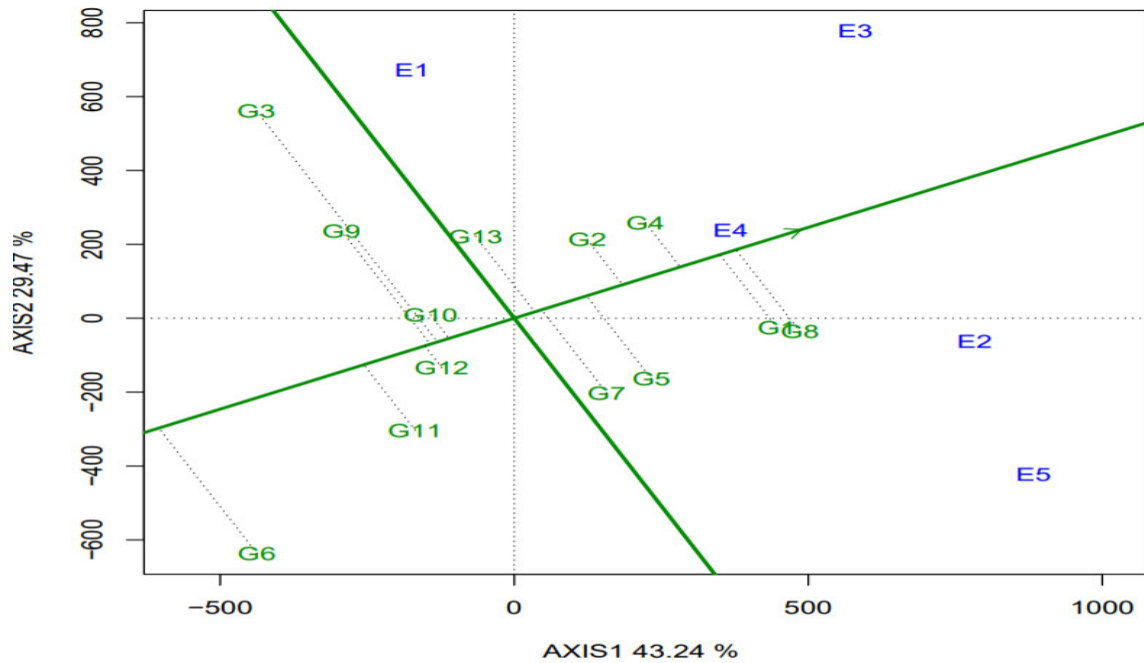

Fig. 2. The mean vs. stability view of genotype main effects plus genotypic x environment interaction effect (GGE) biplot of 13 tef genotypes tested in 5 environments for grain yield
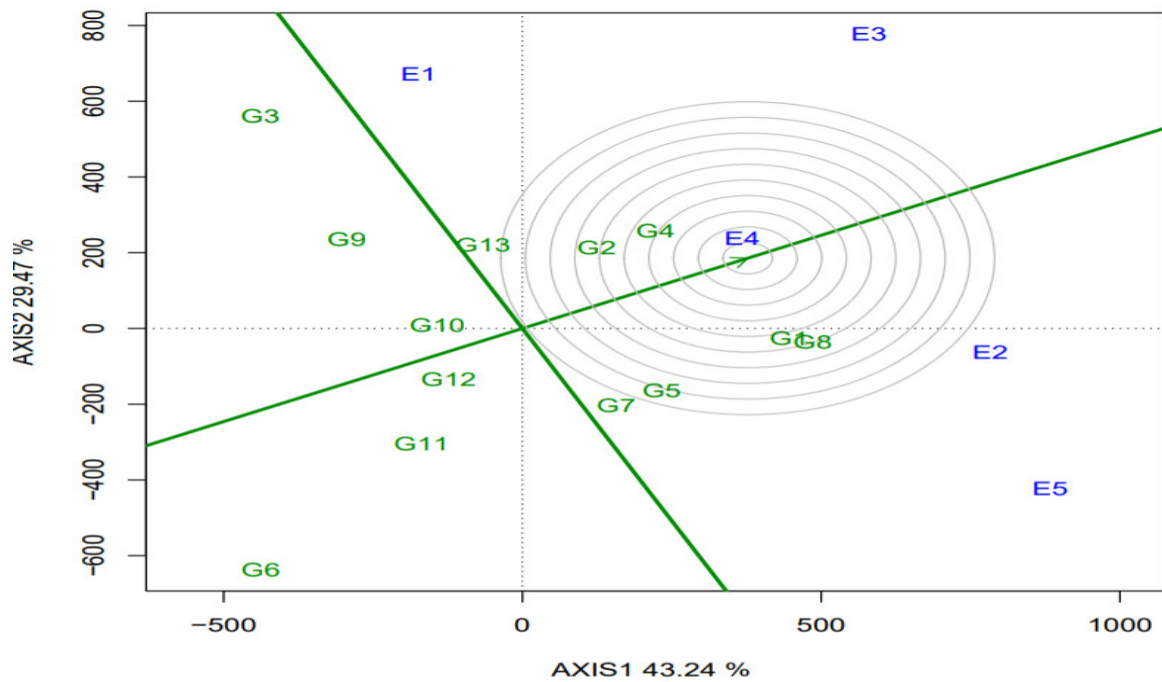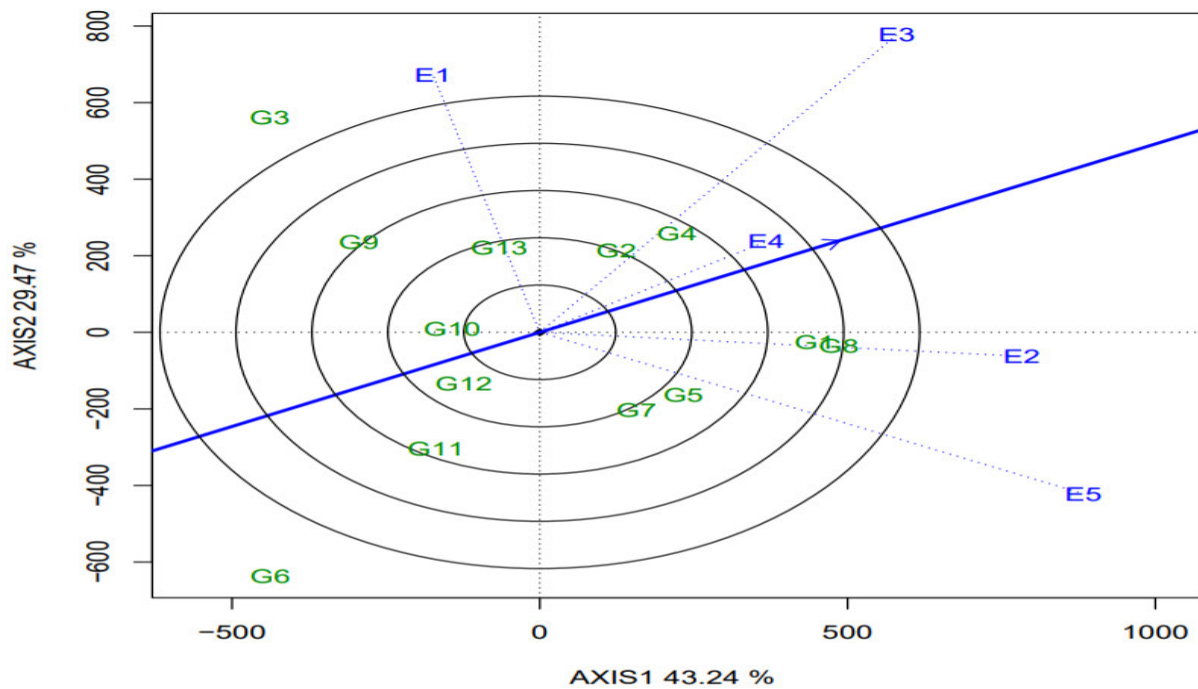

Fig.3.The genotypes comparison with ideal genotype view of genotype main effects plus genotypic x environment interaction effect (GGE) biplot of of 13 tef genotypes tested in 5 environments for grain yield

Figure 4. The "discriminating power vs. representativeness" view of the GGE biplot based on 13 tef genotypes tested at five test environments.
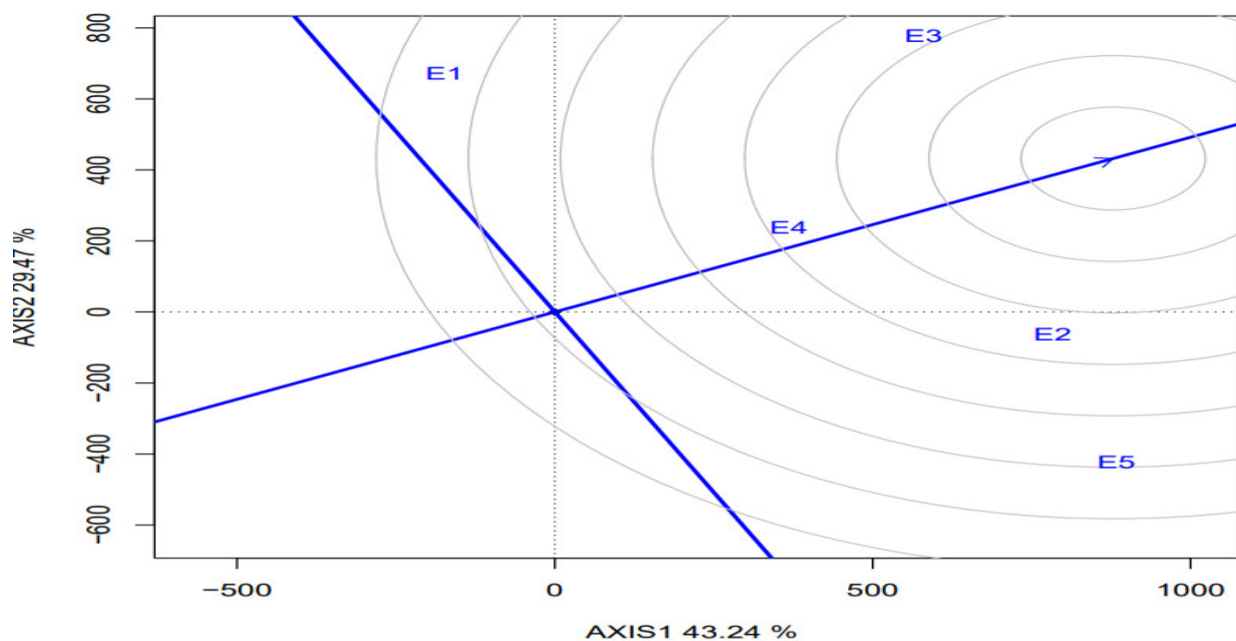

Figure 5. The comparison with the ideal environment view of the GGE biplot based on 13 tef genotypes tested at five test environments.

5. Conclusions

The present study showed that tef grain yield was highly impacted by environment followed by G x E interaction and by the differences among genotypic effects. Presence of G × E interaction for grain yield indicates influence of environment on the expression of the trait. GGE biplot model was effective for analyzing and visualizing pattern of G x E and identifying the most high-yielding and stable cultivar as well as discriminating ability and representativeness of the test environments. The G × E interaction was a crossover type where a change in performance ranking of the genotypes across environments observed. The test environment Adadi-2014 (E3) was more representative and discriminating environment. The GGE biplot showed that genotype G4 (DZ-Cr-387 x Gealmie (RIL- 77C)) was high yielding and stable across the test environments. Therefore, this genotype could be considered as widely adaptable genotype and can be recommended for release as variety.

6. References

Annicchiarico P, Perenzin M (1994). Adaptation patterns and definition of macro environment for selection and recommendation of common wheat in Italy. Plant Breed. 113:197–205.

Annicciarico, P. (2002). Genotype x Environment Interactions: Challenges and Opportunities for Plant Breeding and Cultivar Development. FAO Plant Production and Protection Paper 174. FAO, Rome.

Asfaw A, Tesfaye T, Erenso D, Taye T, Feyera M, Wasihun L, Alemu T, Haileselassie K, Andualem W, Chemeda D (2011). Genotype-by-environment interaction and yield stability analysis in finger millet (Elucine coracana L. Gaertn) in Ethiopia. American Journal of Plant Sciences.2: 408-415.

Frutos E, Galindo MP, Leiva V (2013). An interactive biplot implementation in R for modeling genotype-by-environment interaction. Stochastic Environ. Res. Risk Assess. 28: 1629 – 1641.

Gauch HG, Zobel RW (1996). AMMI Analysis of yield trials. In: Genotype-by-Environment Interaction, pp. 85-122, (Kang MS and Gauch HG, eds). Boca Raton: New York, USA, CRC.

Habte J, Kebebew A, Kassahun T, Zerihun T (2019). Genotype x environment interaction and stability analysis in grain yield of tef (Eragrostis tef) evaluated in Ethiopia. JEAI. 35 (5): 1-13.

Kebebew A, Gina C, Dejene G, Rizqah K, Solomon C, Sonia P, Regula B, Abiel R, Suhail R and Zerihun T (2015). Genetic diversity in tef [Eragrostis tef (Zucc.) Trotter].Front.Plant Sci. 6:177.

Naroui Rad MR, Abdul Kadir M, Rafii MY, Jaafar HZE, Naghavi MR, Ahmadi F (2013). Genotype x environment interaction by AMMI and GGE biplot analysis in three consecutive generations of wheat (Triticum aestivum) under normal and drought stress conditions. Australian Journal of crop science 7(7):956 - 961.

R core team (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Samonte SOPB, Wilson LT, McClung AM, Medley JC (2005).Targeting cultivars onto rice growing environments using AMMI and SREG GGE biplot analysis. Crop Sci. 45:2414 – 2424.

SAS Institute (2011). SAS® 9.3 for Windows. SAS Inst., Cary, NC.

Solomon C, Setotaw F, Tebkew D, Tsion F, Yazachew G, Worku K, Kidist T, Zerihun T, Kebebew A (2019). Significance and prospect of an orphan crop tef. Planta 250: 753- 767.

Spaenij-Dekking L, Kooy-Winkelaar Y and Frits K (2005). The Ethiopian cereal tef in in celiac disease. The New England Journal of Medicine 353: 1748-1749.

Thokozile N, Liezel H, Cosmos M, Peter S, Charles M, Maryke L(2014). Genotype x environment interaction of maize grain yield using AMMI biplots. Crop Sci. 54:1992 – 1999

Vavilov, I (1951). The origin, variation, immunity and breeding of cultivated plants. Translated from the Russian. Ed. K.S. by K. Chester (New York: Ronald Press Co).

Watson L, Dallwitz MJ (1992). The grass genus of the world. CAB International, Wallingford.

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

Yan W, Hunt LA (2002). Biplot analysis of diallel data. Crop Sci. 42:21–30.

Yan W, Rajcan IR (2002). Biplot analysis of test sites and trait relations of soybean in Ontario. Can. J. Plant Sci. 42: 11–20.

Yan W, Kang MS (2003). GGE Biplot Analysis: A graphical tool for breeders, geneticists, and agronomists. RC Press, Boca Raton, FL.

Yan W, Tinker NA (2006). Biplot analysis of multi-environment trial data: Principles and applications. Canadian Journal of Plant Science. 86:623 - 645.

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

Yan W, Fregeau-Reid J, Pageau D, Martin R, Mitchell-Fetch J, Etienne M, Sparry E (2010). Identifying essential test locations for oat breeding in Eastern Canada. Crop Sci. 50: 505 – 515.