GGE Biplot Analysis of Genotype × Environment Interaction in Wheat-Agropyron Disomic Addition Lines

EZATOLLAH FARSHADFAR*1,2 AND MAHSA SADEGHI2
1Campus of Agriculture and Natural Resources, Razi University, Kermanshah, Iran.
2Department of Agronomy and Plant Breeding, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran.
*Corresponding Author: e_farshadfar@yahoo.com (Accepted: 29 May 2014)

ABSTRACT
Identification of the genetic architecture of phenotypic stability and management of adaptational genes is a prerequisite for improvement of adaptation. To locate the genes controlling adaptation, disomic addition lines of agropyron (agropyron elongatum, 2n=2x=14) into the genetic background of bread wheat (Triticum aestivum L., 2n=6x=42, cv. Chinese Spring=CS) were used in a randomized complete block design with three replications under three different conditions (irrigated, pre-anthesis water stress and post-anthesis water stress). The GGE [genotype main effect (G) and genotype by environment interaction (GE)] biplot graphical tool was applied to analyze multi-environment trials (MET) data. Combined analysis of variance showed that the environment (E) effect was a predominant source of variation and accounted for 71.93% of the total sum of squares (TSS), while G and GE interaction sources of variation accounted for 5.97% and 22.10% of the total variation, respectively indicating that the GE interaction was more complex. GGE biplot revealed that PC1 and PC2 accounted for 81.98 of the G + GE variation for the grain yield of the genotypes evaluated in 3 environments. According to which-won-where pattern of GGE biplot the vertex genotypes were G1, G3, G4, G5, G7 and G8. These genotypes were the best or the poorest genotypes in some or all of the test environments since they had the longest distance from the origin of the biplot. The G4, G5, well performed in environment E2 (pre-anthesis stress condition), while the other addition lines showed the lowest performance. Genotype (G2) was more stable as well as high yielding followed by G1 indicating that QTLs controlling yield and stability in agropyron are located on chromosomes 2E and 1E. Comparison of the genotypes with the ideal genotype discriminated addition line G1 (1E) and G2 (2E) as more favorable than all the other genotypes. Negative relationship (obtuse angle) was found between the irrigated (E1), pre-anthesis water stress (E2) and post-anthesis water stress (E3) environments indicating that these environments were different and independent in genotype rankings.

Keywords: Abiotic stress, adaptation, adaptational genes, multi-environment trials, QTL, stability.

Abbreviations:
AEC: average environment coordination; AMMI: additive main effect multiplicative interaction; DAL: disomic addition lines; E: environment; G: genotype main effect; GE: genotype by environment; GEI: genotype × environment interaction; MET: multi-environment trials; SVD: singular value decomposition; TSS: total sum of squares.

INTRODUCTION
Genotypes tested in different locations or years often have significant fluctuation in yield due to the response of genotypes to environmental factors such as soil fertility or the presence of disease pathogens (Kang et al., 2006). These fluctuations are often referred as genotype × environment interaction (GEI) and are common. Genotype × environment interactions have been studied in many crops, including common bean (Mekbib, 2003), corn (Zea mays L.) (Fan et al., 2007), cowpea (Padi, 2007), rice (Haryanto et al., 2008), soybean (Yan and Rajcan, 2002), tomato (Ortiz and Izquierdo, 1994), and wheat (Yuksel et al., 2006).

GEI results from a change in the relative rank of genotype performance or a change in the magnitude of differences between genotype performances from one environment to another. GEI affects breeding progress because it complicates the demonstration of superiority of any genotype across environments and the selection of superior genotypes (Magari and Kang, 1993; Ebdon and Gauch, 2002). Another undesirable effect of GEI includes low correlation between phenotypic and genotypic values, thereby
reducing progress from selection. This leads to bias in the estimation of heritability and in the prediction of genetic advance (Comstock and Moll, 1963). Therefore, the magnitude and nature of GEI determine the features of a selection and testing program.

Often, plant breeders want to develop broadly-adapted genotypes for a wide range of environments. However, it is often not possible to identify genotypes that are superior in yield and yield components in all environments. Furthermore, the same genetic system may not control yield over a diverse set of environments (Ceccarelli and Grando, 1993). Therefore, breeders often develop genotypes for a particular environment to take advantage of specific adaptations (Annicchiarico et al., 2006). However, breeding for a specific adaptation is more efficient if production areas are divided into mega-environments, each representing a target environment for breeding. Mega-environment is a portion (not necessarily contiguous) of the growing region of a crop species having a fairly homogeneous environment that causes similar genotypes to perform best there (Gauch and Zobel, 1997).

Various statistics have been proposed to measure the stability of genotypes over environments. However, no single method can adequately explain cultivar performance across environments (Dehghani et al., 2006). Becker and Leon (1988) suggested two different concepts of stability: static (biological) and dynamic (agronomic). With the static concept, a stable genotype possesses an unchanged performance regardless of variation in the environmental conditions. Thus, genotypic variance among environments is zero. With the dynamic concept, response of a genotype to environments is predictable. Thus, a stable genotype has no deviation from response to environments. Both concepts of stability are useful, but their application depends on the trait considered. For qualitative traits such as resistance to diseases or stress, the static concept of stability is useful. For quantitative traits such as yield, the dynamic concept of stability is useful (Norden et al., 1986).

Statistical methods for measuring genotypic stability should partition the information from a genotype - environment data matrix into simpler components representing real responses vs. random variation (Gauch, 1992). These statistical methods can be classified into two groups: parametric (univariate and multivariate) and nonparametric. Univariate models ranged from parametric, such as environmental variance (Roemer, 1917), ecovalence (Wricke, 1962), stability variance (Shukla, 1972), regression slope (Finlay and Wilkinson, 1963), deviation from regression (Eberhart and Russell, 1966) and coefficient of determination (Pinthus, 1973). Non-parametric models include Nassar and Huehn (1987), Thennarasu (1995) and Kang's yield stability statistic (Kang, 1993). Multivariate models includes a wide range of methods such as principal component analysis (PCA) (Gower, 1967), cluster analysis (Mungomery et al., 1974), genotype main effect plus genotype by environment interaction (GE) biplot analysis (Yan, 2001), and additive main effects and multiplicative interaction models (AMMI) (Gauch, 1988).

Univariate, nonparametric stability statistics define environments and phenotypes relative to biotic and abiotic factors. Nonparametric stability statistics are based on rank order of genotypes and do not rely upon assumptions about distribution of observed values or of variance homogeneity. Univariate, parametric stability statistics involve relating observed genotypic responses to a sample of environmental conditions. With certain statistical assumptions, parametric stability methods exhibit beneficial properties, providing information about the normal distribution of error and of interaction effects (Huehn, 1990). For those reasons, parametric stability is more commonly used. Mut et al. (2009) reported that for many applications, including selection in breeding and testing programs, parametric stability statistics are useful but there is justification for the use of non-parametric measures for the assessment of the yield stability of crop genotypes.

The multivariate models, AMMI and GGE biplot, appeared to be able to extract a large part of the genotype - environment interaction and were efficient in analyzing interaction patterns (Zobel et al., 1988). Gauch (1992) reported that multivariate models captured a large portion of the genotype × environment interaction sum of squares clearly separating main and interaction effects, and the model often provided an agronomically meaningful interpretation of the data. Differences in genotype stability and adaptability to environment can be qualitatively assessed using the biplot graphical representation that scatters the genotypes according to their principal component values (Vita et al., 2010).

Recently, Yan et al. (2000) proposed a GGE biplot that allows visual examination of the GE interaction pattern of multi-environment trial (MET) data. The GGE Biplot emphasizes two concepts. First, although the measured yield is the combined effect of genotype (G), environment (E), and genotype by environment interaction (GE), only G and GE are relevant to, and must be considered simultaneously, in cultivar evaluation (hence the term 'GGE'). Second, the biplot technique developed by Gabriel (1971) was
employed to approximate and display the GGE of a MET, hence the term GGE biplot. This GGE biplot was constructed by the first two principal components (PC1 and PC2, also referred to as primary and secondary effects, respectively) derived from subjecting environment-centered yield data, i.e., the yield variation due to GGE, to singular value decomposition (SVD) (Yan, 1999; Yan et al., 2000). This GGE Biplot was shown to effectively identify the GE interaction pattern of the data. It clearly shows which cultivar won in which environments, and thus facilitates mega-environment identification.

In addition, the GGE biplot also has a usage in selecting superior cultivars and test environments for a given environment. Provided that the genotypic PC1 scores have a near-perfect correlation with the genotype main effects, ideal cultivars should have a large PC1 score (high yielding ability) and a small (absolute) PC2 score (high stability). Similarly, ideal test environments should have a large PC1 score (more discriminating of the genotypes in terms of the genotypic main effect) and small (absolute) PC2 score (more representative of the overall environment) (Yan, 1999; Yan et al., 2000).

Thus, the main objective of this study was to locate the genes controlling stability and yield performance in agropyron using wheat-agropyron disomic addition lines grown under different growing conditions by applying the GGE biplot approach.

MATERIALS AND METHODS
Plant Genetic Materials:
In order to identify QTLs controlling yield stability in Agropyron an experiment was conducted in three environments. The experiment was laid out with eight disomic addition lines (DALs) of *Agropyron elongatum* (2n=2x=14) into the genetic background of ‘Chinese Spring’ (CS) wheat (2n=6x=42) in a randomized complete blocks design with three replications. The DALs were named as: 1E to 7E indicating addition of chromosomes 1E to 7E from *Agropyron elongatum* into the genome of CS, respectively.

Experimental Condition:
The genotypes were cultivated in the field of Campus of Agriculture and Natural Resources, Razi University, Kermanshah, Iran (47° 20´ N latitude, 34° 20´ E longitude and 1351.6 m altitude). Climate of the region is classified as semi-arid with mean annual rainfall of 378 mm. Minimum and maximum temperature at the research station were -27°C and 44°C, respectively. Each genotype was planted in 2 m rows and at 15 × 25 cm inter-plant and inter-row distances, respectively. Each plot consisted of 100 seeds (each row 50 seeds). The environments were considered as random factors, while genotypes as fixed factors.

Treatments:
Irrigation was manipulated to create three different environments: (i) a fully irrigated control treatment, (ii) a mid-season water stress treatment where the crop was under progressive stress form approximately floral initiation (pre-anthesis) to flowering (post-anthesis) and rewatered thereafter until maturity and (iii) terminal stress, where irrigation was terminated at grain filling, and continuing until maturity.

Statistical Analysis:
The grain yield data were recorded for each genotype at each environment. Combined analysis of variance for grain yield data was performed to determine the effects of environment (E), genotype (G), and GEI. The mean values of genotypes at each experiment were used to analyze yield stability.

The GGE biplot methodology (Yan et al., 2000) was used to graphically analysis the GE interaction data attempting to identify the chromosomes of agropyron which carrying the genes controlling high yield and stability performance under different growing conditions.

To generate a GGE biplot (Yan et al., 2000), the genotype-environment two-way table of yield was first environment- standardized and then the environment-standardized table was decomposed into principal components (PC) via singular value decomposition (SVD). The first two PCs (PC1 and PC2) were used to generate a GGE biplot, whereas the rest were regarded as residuals (Yan and Tinker, 2006). All analyses were performed using the GGE-biplot software (Yan and Hunt, 2001).

RESULTS AND DISCUSSION
Combined Analysis of Variance and Mean Comparisons:
The results of combined ANOVA for grain yield indicated that the differences among all sources of variation were highly significant (*P < 0.01*) (Table 1). The environment (E) effect was a predominant source of variation and accounted for 71.93% of the total sum of squares (TSS), while G and GE interaction sources of variation accounted for 5.97% and 22.10% of the total variation, respectively. The GE effect was about four times greater than the G effect, which suggesting the possible existence of different megaenvironments with different top-yielding genotypes (Yan and Kang 2003).

Mean comparisons exhibited that maximum yield was attributed to genotypes G1, G4 and G2 respectively (Table 2) indicating that QTLs controlling grain yield in agropyron are distributed on chromosomes 1E, 4E and 2E. Sadeghi and
Farshadfar (2014) reported the importance of chromosome 2E in controlling grain yield in agropyron.

### Table 1. Combined analysis of variance for wheat-agropyron disomic addition lines across 3 environments.

<table>
<thead>
<tr>
<th>S.O.V</th>
<th>Df</th>
<th>Sum of square</th>
<th>SS% (ð)</th>
<th>Mean of square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>71</td>
<td>620.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatments</td>
<td>23</td>
<td>568.3</td>
<td>91.57</td>
<td>24.71**</td>
</tr>
<tr>
<td>Genotypes</td>
<td>7</td>
<td>33.9</td>
<td>5.97</td>
<td>4.84**</td>
</tr>
<tr>
<td>Environments</td>
<td>2</td>
<td>408.8</td>
<td>71.93</td>
<td>204.40**</td>
</tr>
<tr>
<td>Interactions</td>
<td>14</td>
<td>125.6</td>
<td>22.10</td>
<td>8.97**</td>
</tr>
</tbody>
</table>

**: significant at 1% probability level

### Table 2. Genotypes number, mean yield and their ranks

<table>
<thead>
<tr>
<th>Genotypes No.</th>
<th>Mean yield</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1(1E)</td>
<td>4.98</td>
<td>1</td>
</tr>
<tr>
<td>G2(2E)</td>
<td>4.85</td>
<td>3</td>
</tr>
<tr>
<td>G3(3E)</td>
<td>3.87</td>
<td>6</td>
</tr>
<tr>
<td>G4(4E)</td>
<td>4.87</td>
<td>2</td>
</tr>
<tr>
<td>G5(5E)</td>
<td>4.10</td>
<td>4</td>
</tr>
<tr>
<td>G6(6E)</td>
<td>3.01</td>
<td>8</td>
</tr>
<tr>
<td>G7(7E)</td>
<td>3.32</td>
<td>7</td>
</tr>
<tr>
<td>G8(8E)</td>
<td>4.03</td>
<td>5</td>
</tr>
</tbody>
</table>

**GGE Biplot Analysis:**

1. **Which Won Where Pattern**

   GGE biplot was constructed by plotting the first two principal component PC1 and PC2 (referred to as the primary and secondary effects) derived from subjecting environment centered yield data to singular value decomposition (Yan et al., 2000). The PC1 and PC2 were accounted for 81.98 of the G + GE variation for the grain yield of the genotypes evaluated 3 environments. The polygon view of the GGE biplot was constructed to show which genotypes performed best in which environment (Fig. 1).

![Fig. 1. Polygon views of the GGE biplot based on symmetrical scaling for the which-won-where pattern of genotypes and environments.](image)

The vertices of the polygon were the genotype markers located farthest away from the biplot origin in various directions, such that all genotype markers were contained within the resulting polygon. Based on this, six genotypes identified as the markers farthest away from the biplot origin. The vertex genotype in each sector represented the highest yielding genotype in the environment that fell within that particular sector (Yan et al., 2000).

According to Fig. 1, the vertex genotypes were G1, G3, G4, G5, G7 and G8. These genotypes were the best or the poorest genotypes in some or all of the test environments since they had the lowest performance. The other vertex genotypes (G2 and G6) without any environment in their sectors were not the highest yielding genotypes at any environment; thus, they were the poorest genotypes at all or some environments (Yan, 2001). The vertex genotype in each sector is the best genotype at environments whose markers fall into the respective sector (Yan et al., 2000). Environments within the same sector share the same winning genotype, and environments in different sectors have different winning genotypes. Thus, the polygon view of a GGE biplot indicates the presence or absence of crossover GE interactions involving the most responsive genotypes, and is suggestive of the existence or absence of different mega-environments among the tested environments (Yan and Rajcan, 2002).

2. **Mean vs. Stability View of the GGE Biplot**

   Performance and stability of genotypes were visualized graphically through the GGE biplot (Fig. 2). This can be evaluated by average environment coordination (AEC) method (Yan, 2001; 2002). In Fig. 2 the line with single arrow head is the AEC (average environment coordinate) abscissa. AEC abscissa passes through the biplot origin and marker for average environment and points towards higher mean values. The average environment has average PC1 and PC2 scores across environments (Yan, 2001). The perpendicular lines to the AEC passing through the biplot origin are referred to as AEC ordinate. The greater the absolute length of the projection of a genotype indicates more instability. Furthermore, the average yield of genotypes is approximated by the projections of their markers to the AEC abscissa (Yan and Kang, 2003).

According to Fig. 2, genotypes with above-average means were from G2, G8, G1, while genotypes below-average means were from G4, G5 and G7. However, the length of the average environment vector was sufficient to select genotypes based on yield mean performances.
Genotypes with above-average means (G1, G2 and G8) could be selected, whereas the rest were discarded. A longer projection to the AEC ordinate, regardless of the direction, represents a greater tendency of the GE interaction of a genotype, which means it is more variable and less stable across environments or vice versa. For instance, genotype G2 was more stable as well as high yielding followed by G1. Conversely, G8 was instable, but high yielding. The G5 was stable with low yield. It can be concluded that QTLs controlling yield and stability in agropyron are located on chromosome 2E (G2) and 1E (G1).

The cosine of the angle between their vectors. Acute angles indicates a positive correlation, obtuse angles a negative correlation and right angles no correlation (Yan and Kang, 2003). A short vector may indicate that the test environment is not related to other environments. According to Fig. 4, negative relationship (obtuse angle) was found between the irrigate (E1), preanthesis water stress (E2) and post anthesis water stress (E3) environments indicating that these environments were different and independent in genotype rankings.

3. Comparison of the Genotypes with the Ideal Genotype

An ideal genotype have the highest mean performance and be absolutely stable (i.e., perform the best in all environments). Such an ideal genotype is defined by having the greatest vector length of the high-yielding genotypes and with zero GE, as represented by the small circle with an arrow pointing to it (Yan, 2001). Although such an ideal genotype may not exist in reality, it can be used as a reference for genotype evaluation. A genotype is more desirable if it is located closer to the ideal genotype. Thus, using the ideal genotype as the center, concentric circles were drawn to help visualize the distance between each genotype and the ideal genotype (Fig. 3). In Fig. 3 the genotypes are ranked relative to the ideal genotype. A genotype is more favorable if it is closer to the ideal genotype. Accordingly, addition line of G1 (1E) and G2 (E2) were more favorable than all the other genotypes. The other genotypes were unfavorable because they were far away from the ideal genotype.

4. Relationships Among Test Environments

In GGE biplot, the correlation coefficient between any two environments is approximated by the distance between two environments measures their dissimilarity in discriminating the genotypes. Thus, the presence of close associations among test environments suggests that the same information about the genotypes could be obtained from fewer test environments, and hence the potential to reduce testing cost. If two test environments are
closely correlated consistently across years, one of them can be dropped without loss of much information about the genotypes (Farshadfar et al., 2011).

REFERENCES


Thennarasu, K. 1995. On certain non-parametric procedures for studying genotype-environment


