Published June 15, 2016

RESEARCH

# Value of Locations for Representing Mega-Environments and for Discriminating Yield of Watermelon in the U.S.

Mahendra Dia, Todd C. Wehner,* Richard Hassell, Daniel S. Price, George E. Boyhan, Stephen Olson, Stephen King, Angela R. Davis, Gregory E. Tolla, Jerome Bernier, and Benito Juarez

## ABSTRACT

In a crop breeding program, multiple-location trials can be used to define target regions and mega-environments that, in turn, will help the breeder develop stable cultivars. In addition, locations can be chosen that are efficient for distinguishing among cultivars (genotypes) and that are good representatives of the target regions. The objectives of this study were to study mega-environments and identify test locations that were both discriminating and representative of target regions. Watermelon [*Citrullus lanatus* (Thunb.) Matsum. & Nakai] fruit yield and yield components were evaluated in 3 yr and eight locations using replicated, multiple-harvest trials. Data were analyzed using genotype main effect and genotype × environment interaction (GGE) biplot model as well as other methods for stability analysis. Marketable yield and percentage early fruit had a nonrepeatable crossover pattern and thus, formed a single and complex mega-environment. Two key locations, (Kinston, NC, and Charleston, SC) were efficient representatives of two mega-environments for fruit count. Locations at Woodland, CA, and College Station, TX, can be used interchangeably for identifying genotypes with high percentage cull fruit. There was only one mega-environment for fruit size. Identification of mega-environments for watermelon in the southern United States has implications for future breeding and genotype evaluation in the United States including the use of specialized genotypes for high performance in specific locations.

M. Dia and T.C. Wehner, Dep. of Horticultural Science, North Carolina State Univ., Raleigh, NC 27695-7609; R. Hassell, Clemson Univ., Coastal Research and Education Center, 2700 Savannah Hwy, Charleston, SC 29414; D.S. Price, Georgia County Extension, SW District, 110 West 13th Ave, Suite C, Cordele, GA 31015; G.E. Boyhan, Univ. of Georgia, Dep. of Hort., 1111 Miller Plant Science Bldg., Athens, GA 30602; S. Olson, North Florida REC, Univ. of Florida, 155 Research Road, Quincy, FL 32351-5677; S. King, Texas A&M Univ., Dep. of Hort. Sci., 1500 Research Pkwy, Ste A120, College Station, TX 77845, A.R. Davis, USDA–ARS, 911 East Highway 3, Lane, OK 74555; G.E. Tolla and J. Bernier, Monsanto/Seminis Veg Seeds, 37437 State Hwy 16, Woodland, CA 95695, B. Juarez, 2387 McNary Way, Woodland, CA 95776. Received 16 Nov. 2015. Accepted 26 Mar. 2016. *Corresponding author (tcwehner@gmail.com).

**Abbreviations:** AEC, average environment coordinate; CA, Woodland, CA; CI, Clinton, NC; FL, Quincy, FL; $G \times E$, genotype × environment interaction; $G \times L$, genotype × location interaction; GA, Cordele, GA; GGE, genotype main effect and genotype × environment interaction; GGL, genotype main effects plus genotypic × location interaction effect; KN, Kinston, NC; $L \times Y$, location × year interaction; OK, Lane, OK ; SC, Charleston, SC; TSS, total sum of squares; TX, College Station, TX.

The United States is the fifth largest watermelon producer, with 2 Tg harvested from 21,450 ha in 2007 and valued at $476 million (FAO, 2010; USDA, 2010). Watermelons are grown in almost all US states; however, the major producers are in southern and western states having a long frost-free season including Florida, Texas, Oklahoma, and California (Wehner, 2008). Growers are interested in watermelon cultivars that perform well over locations and years.

In studies where genotypes (breeding lines and cultivars) are evaluated in a set of environments (location–year combinations),

Published in Crop Sci. 56:1726–1735 (2016).
doi: 10.2135/cropsci2015.11.0698

the data can be used to identify broadly adapted genotypes for a wide range of target environments. A target environment is a production environment used by growers. Since it is impossible to test in all target environments, plant breeders do indirect selection using their own multiple-environment trials or test environments. Genotype × environment interaction ($G \times E$) reduces the predictability of the performance of genotypes in target environments based on their performance in test environments. An important element in the final phase of crop breeding is the selection of suitable test locations, since it accounts for $G \times E$ and maximizes gain from selection (Yan et al., 2011). A test location is efficient when it represents the target environment and discriminates among genotypes. Discriminating locations can detect differences among genotypes with few replications. Representative locations make it likely that genotypes selected will perform well in target environments (Yan et al., 2011).

Plant breeders are interested in development of cultivars adapted to a wide range of environments. However, it is not possible to identify genotypes that are superior in yield and yield components in all environments especially in future years. Furthermore, the same genetic system may not control yield over this diverse set of environments (Ceccarelli and Grando, 1993; Ceccarelli, 1989; El-Soda et al., 2014; El-Soda et al., 2015; Simmonds, 1991). Therefore, breeders often develop genotypes for a particular environment to take advantage of specific adaptations (Annicchiarico et al., 2005; Samonte et al., 2005). Breeding for specific adaptation is more efficient if production areas can be divided into mega-environments, each representing a target environment for breeding.

Several definitions have been proposed for mega-environments. For example, CIMMYT defined it as "a broad, not necessarily contiguous area, occurring in more than one country and frequently transcontinental, defined by similar biotic and abiotic stresses, cropping system requirements, consumer preferences, and, for convenience, by volume of production" (Braun et al., 1996). Yan (2006) defined mega-environment as a group of geographical locations that share the same (sets of) genotypes consistently across years. Gauch and Zobel (1997) defined mega-environment as 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. Other researchers have defined mega-environment as a group of growing areas that are similar in terms of genotype response and that show a repeatable relative performance of crop genotypes across years (Yan and Rajcan, 2002; Yan and Tinker, 2005).

Mega-environments are often identified through the analysis of multiple-environment trial data for a set of genotypes. The purpose of mega-environment analysis is to understand the $G \times E$ patterns responsible for specific adaptation within a target region so it can be divided into mega-environments (Yan et al., 2011). Several methods have been used to analyze multiple-environment trial data and to group the environments. The GGE model, with a biplot display, is becoming popular for analyzing multiple-environment trial data, determining mega-environments, evaluating test locations, and evaluating genotype stability (Luo et al., 2015; Naiying et al., 2014; Yan, 2015).

The objectives of this research were to (i) identify mega-environments for the main US watermelon production areas and (ii) identify test locations having high discriminating ability and representativeness for yield and yield components.

## MATERIALS AND METHODS

### Multiple-Environment Trials

Forty diverse genotypes of watermelon were evaluated for 3 yr (2009, 2010, and 2011) in eight locations across the United States. The pedigree and phenotype of 40 genotypes were described in Dia et al. (2016b). Locations were Kinston, NC (KN), Clinton, NC (CI), Charleston, SC (SC), Cordele, GA (GA), Quincy, FL (FL), College Station, TX (TX), Lane, OK (OK), and Woodland, CA (CA). The locations were chosen to represent the key watermelon production regions in the US.

The experiment was a randomized complete block design with four replications, eight locations, and 3 yr. Field plots were 3.7 m long with six plants per plot (Neppl, 2001). At each location, the 40 watermelon genotypes were evaluated for traits including marketable yield (Mg ha$^{-1}$), marketable fruit count (thousand fruit ha$^{-1}$), percentage cull fruit ( = 100 × cull fruit weight/total fruit weight), percentage early fruit (100 × marketable yield for the first harvest/marketable yield for all harvests), and fruit size (kg fruit$^{-1}$).

Fruit were determined to be ripe based on days from planting to harvest, and if there was a brown tendril nearest the fruit, a light-colored ground spot, and a dull sound of the fruit when thumped (Maynard, 2001). Fruit were graded into marketable and cull and then counted and weighed for each plot. Yield was calculated as total and marketable fruit weight (Mg ha$^{-1}$) and number (thousands ha$^{-1}$) for first harvest and for all harvested summed. All curved, bottle-necked, and deformed fruit were considered culls. Each trial had one to four harvests depending on year and location. Data were not collected on all traits from Oklahoma in 2009, Georgia in 2010, or Florida in 2011. Additionally, percentage cull fruit was not available for South Carolina in 2009, 2010, and 2011; and Florida in 2009 and 2010. A single harvest was done at California in 2009 and Georgia in 2011, so data for percentage early fruit was not available.

### Data Analysis

Data were analyzed for all the traits evaluated in the study using the SASGxE program (Dia et al., 2015) in SAS v9.4 software (SAS Institute, 2014) and RGxE program (Dia et al., 2016a) in R v3.2.4 software (R Development Core Team, 2007). Another SAS program was published by Hussein et al. (2000), but it is not available. RGxE was used to compute location descriptive statistics. SASGxE was used to compute analysis of variance

(ANOVA), input files that were used in R statistical software (R Development Core Team, 2007) for genotype evaluation, mega-environment identification, and test location evaluation. Environment (year × location combinations), location, replications, and genotypes were considered random effects. The GGE biplot analysis was computed using the GGEBiplotGUI package (Frutos et al., 2014) and R software with the support in the helper application RStudio (RStudio, 2014) to visually assess the presence of mega-environments, redundancy, discriminating ability and representativeness of the test locations (Yan et al., 2000; Yan and Kang, 2003). The *F*-statistics of genotypes for each location were computed using the GLM procedure of SAS v9.4. To determine appropriate means squares for testing the effects in the model, the TEST option in the RANDOM statement was used (SAS Institute, 2016). Similarly, Pearson correlations between test locations with overall location mean were computed using SAS v9.4.

# RESULTS

## Relative Magnitude of Location, Genotype, and Genotype × Environment Interaction

Results of combined ANOVA revealed significant environment (*E*), genotype (*G*), and *G* × *E* effects for all evaluated traits (Table 1). Marketable yield was controlled to a large extent by *E* (48% of total sum of squares [TSS]) followed by *G* × *E* (18% of TSS), and *G* (8% of TSS) (Table 1). Of the 48% *E* variation of marketable yield, 75% was attributable to location (*L*), 21% to *L* × year interaction (*L* × *Y*), and 3% to year (*Y*) (Table 1). However, fruit count was explained largely by *E*, *G* and *G* × *E* (35, 20, and 21% of TSS, respectively) (Table 1). The *E* portion (35%) of yield of fruit count was attributable to 68% to *L*, 27% to *L* × *Y*, and 5% to *Y* (Table 1).

ANOVA for percentage cull fruit and percentage early fruit showed that *E*, *G*, and *G* × *E* accounted for 26 and 38, 10 and 7, and 23 and 22%, respectively, of the TSS (Table 1). On the contrary, for fruit size, ~51% of total variance was due to *G* and a relatively small effect of *E* (14%) and *G* × *E* (13%) was observed (Table 1).

The significant *L* × *Y* for all the traits evaluated in this study warranted separate ANOVA for each year (Supplemental Table S1) (Fan et al., 2007). The results of ANOVA for the yearly data gave an overall picture of the relative magnitude of the *L*, *G*, and genotype × location interaction (*G* × *L*) variance terms. Within each year, *L* was significant for all the traits evaluated in this study (mean square not presented). Except for fruit size, location was the most important source of yield and yield components variation, accounting for 55 to 72, 36 to 49, 9 to 53, and 45 to 56% of TSS for marketable yield, fruit count, percentage cull fruit, and percentage early fruit, respectively (Supplemental Table S1).

**Table 1. ANOVA for marketable yield (Mg ha$^{-1}$), fruit count, percentage cull fruit, percentage early fruit, and fruit size of 40 watermelon genotypes (kg fruit$^{-1}$) tested in 3 yr and eight locations.**

| Source | df | Mean square | Percentage total sum of squares |
|---|---|---|---|
| Marketable yield | | | |
| Environment (*E*) | 20 | 106,825.19* | 47.83 |
| Location (*L*) | 7 | 231,712.42 * | 75.92 |
| Year (*Y*) | 2 | 3,240.24 | 3.00 |
| *L* × *Y* | 11 | 40,726.21* | 20.97 |
| Replication within *E* | 63 | 2,366.27* | 3.33 |
| Genotype (*G*) | 39 | 9,227.34* | 8.05 |
| *G* × *E* | 780 | 1,023.33* | 17.87 |
| *G* × *L* | 273 | 1,322.62* | 45.23 |
| *G* × *Y* | 78 | 870.00 | 8.50 |
| *G* × *L* × *Y* | 429 | 850.71* | 45.72 |
| Pooled Error | 2442 | 419.77 | 22.95 |
| Fruit count | | | |
| Environment (*E*) | 20 | 1,554.37* | 34.51 |
| Location (*L*) | 7 | 3,036.82* | 68.38 |
| Year (*Y*) | 2 | 107.95 | 5.01 |
| *L* × *Y* | 11 | 752.03* | 26.61 |
| Replication within *E* | 63 | 30.36* | 2.12 |
| Genotype (*G*) | 39 | 456.12* | 19.75 |
| *G* × *E* | 779 | 24.27* | 20.99 |
| *G* × *L* | 273 | 34.52* | 49.84 |
| *G* × *Y* | 78 | 21.31 | 8.79 |
| *G* × *L* × *Y* | 428 | 18.23* | 41.26 |
| Pooled Error | 2436 | 8.37 | 22.52 |
| Percentage cull fruit | | | |
| Environment (*E*) | 18 | 10,327.63* | 25.52 |
| Location (*L*) | 6 | 17,835.89 | 57.57 |
| Year (*Y*) | 2 | 10,925.12 | 11.75 |
| *L* × *Y* | 10 | 5,801.61* | 31.21 |
| Replication within *E* | 57 | 359.93* | 2.82 |
| Genotype (*G*) | 39 | 1,792.39* | 9.59 |
| *G* × *E* | 699 | 234.49* | 22.50 |
| *G* × *L* | 234 | 306.34* | 4.73 |
| *G* × *Y* | 78 | 293.43* | 13.96 |
| *G* × *L* × *Y* | 387 | 186.02* | 43.92 |
| Pooled Error | 2163 | 133.89 | 39.75 |
| Percentage early fruit | | | |
| Environment (*E*) | 19 | 40,143.06* | 37.67 |
| Location (*L*) | 7 | 63,207.46* | 58.01 |
| Year (*Y*) | 2 | 11,355.56* | 2.98 |
| *L* × *Y* | 10 | 4,879.89* | 6.40 |
| Replication within *E* | 57 | 1,301.27* | 3.66 |
| Genotype (*G*) | 39 | 3,498.84* | 6.74 |
| *G* × *E* | 661 | 648.49* | 21.17 |
| *G* × *L* | 273 | 743.35* | 47.34 |
| *G* × *Y* | 78 | 782.52* | 14.24 |
| *G* × *L* × *Y* | 310 | 542.10* | 39.20 |
| Pooled Error | 2019 | 305.96 | 30.51 |
| Fruit size | | | |
| Environment (*E*) | 20 | 226.17* | 14.45 |
| Location (*L*) | 7 | 345.91 | 53.53 |
| Year (*Y*) | 2 | 21.28 | 0.94 |
| *L* × *Y* | 10 | 180.26* | 43.84 |
| Replication within *E* | 63 | 9.70* | 1.95 |
| Genotype (*G*) | 39 | 412.78* | 51.46 |
| *G* × *E* | 776 | 5.42* | 13.45 |
| *G* × *L* | 273 | 6.18* | 40.10 |
| *G* × *Y* | 78 | 8.58* | 15.91 |
| *G* × *L* × *Y* | 386 | 4.13* | 41.74 |
| Pooled Error | 2374 | 2.21 | 16.77 |

* Significant at the 0.01 level of probability.

## Mega-Environment Investigation

The two-dimensional polygon (which-won-where) view of GGE biplot of multiple-environment trial data of 3 yr, 40 genotypes, and eight locations was based on environment-standardized data and environment-focused singular-value partition (Fig. 1). The straight line originating from the biplot origin and being perpendicular to the each side of the polygon divides the biplot into sectors. The sectors were labeled from 1 to 4 for all the traits evaluated in this study (Fig. 1). The winning genotype for a sector is usually the vertex genotype at the intersection of the two polygon sides. However, it is not necessary that the winning genotype is within its winning sector (Yan, 2002). If all environment markers fall into a single sector, this indicates that a single genotype has the highest yield in all environments. Conversely, if environment markers fall into different sectors, then different genotypes won in different sectors and thus, $G \times E$ or a crossover patterns exists. This crossover $G \times E$ indicates that target environments may be divided into different mega-environments.

Thus, there could be two mega-environments (represented by shaded area) for marketable yield (Fig. 1A), percentage cull fruit (Fig. 1C), and percentage early fruit (Fig. 1D) and one mega-environment for fruit count (Fig. 1B) and fruit size (Fig. 1E). For marketable yield, the biplot Sectors 1 and 2 represented the first and second mega-environments (Fig. 1A). The first mega-environment for marketable yield contained locations CA, GA, KN, OK, and TX. Mega-environment 2 contained location FL, CI, and SC (Fig. 1A). Likewise, for percentage cull, Mega-environment 1 contained location KN. Mega-environment 2 contained locations CA, CI, GA, OK, and TX (Fig. 1C). For percentage early fruit, Mega-environment 1 contained locations CA, GA, FL, KN, OK, and TX and Mega-environment 2 contained locations CI and SC (Fig. 1D).

Since a mega-environment is defined as a group of locations that consistently shares the best set of genotypes across years, it is recommended that a target region be divided into mega-environments if crossover patterns are repeatable across years (Yan and Tinker, 2005). Therefore, genotype main effect plus genotypic × location interaction effect (GGL) biplots for individual years were constructed (Supplemental Fig. S1–S5). Test locations were grouped based on the polygon (which-won-where) view of GGL biplots for individual year for all the traits, and results were summarized in Table 2.

## Test Environment Evaluation: Representative versus Discriminative Ability

After identifying mega-environments, the next step was to determine whether test locations were representative of the target environment as well as discriminating among genotypes. The discriminating power vs. representativeness view of GGE biplot distinguishes test environments that are able to discriminate among genotypes and represent the mega-environment (Fig. 2). An ideal test environment is represented by the circle on an arrowhead on the average environment coordinate (AEC) abscissa (near parallel to the horizontal axis) (Fig. 2). Therefore, the test locations that best represent mega-environment in most years and that are also good at discriminating genotypes should have a long vector and a small angle with the AEC (represented in Fig. 2A). The length and the angle of the vector for each environment with AEC are the measure of its discriminating power and representativeness, respectively. A biplot judges correlation using the cosine of the angle between two vectors (Yan and Holland, 2010).

Based on repeatable which-won-where or crossover pattern across years (GGL biplots), locations were classified into two and one mega-environment for fruit count and percentage cull fruit, respectively. The discriminating power vs. representativeness view of GGE biplots was constructed for each mega-environment. For fruit count, the locations FL, SC, and TX represented the first mega-environment and CA, CI, and KN represented the second mega-environment. Locations SC and KN were close to the circle on the AEC abscissa (Fig. 2A, 2B), as were the most discriminating of genotypes, and each represented their mega-environment well. Similarly, mega-environment for percentage cull fruit contained the locations CA and TX. Location TX had large vector length and relatively small angle with AEC (Fig. 2C). Likewise, yearly discriminating power vs. representativeness view of GGE biplots were constructed for fruit count and percentage cull fruit (Supplemental Fig. S6, S7). If a pattern repeats across years, then it can be concluded that some test locations are redundant and can be dropped.

## Genotype *F*-Ratio and Correlation of Test Location with Location Mean

The genotype *F*-statistic is the ratio of genotype variance to the appropriate error variance, which the GLM procedure determines based on the expected mean squares (SAS Institute, 2016). When the mean of all genotypes are equal, then the *F*-ratio will be close to 1. If analysis of variance is run by location, then a high genotype *F*-ratio indicates high discriminating ability for that location. For marketable yield, the locations CI, KN, and SC had the highest *F*-ratio, while location TX had the lowest (Table 3). The locations CI and KN had consistently high *F*-ratio for fruit count, percentage cull fruit, percentage early fruit, and fruit weight, so they were the most discriminating (Table 3). Location OK had low *F*-ratio for fruit count, percentage cull fruit, percentage early fruit, and fruit weight (Table 3).

Correlation of genotype performance between test location and overall location mean is presented in Table 3. High and significant correlation value reflected strong

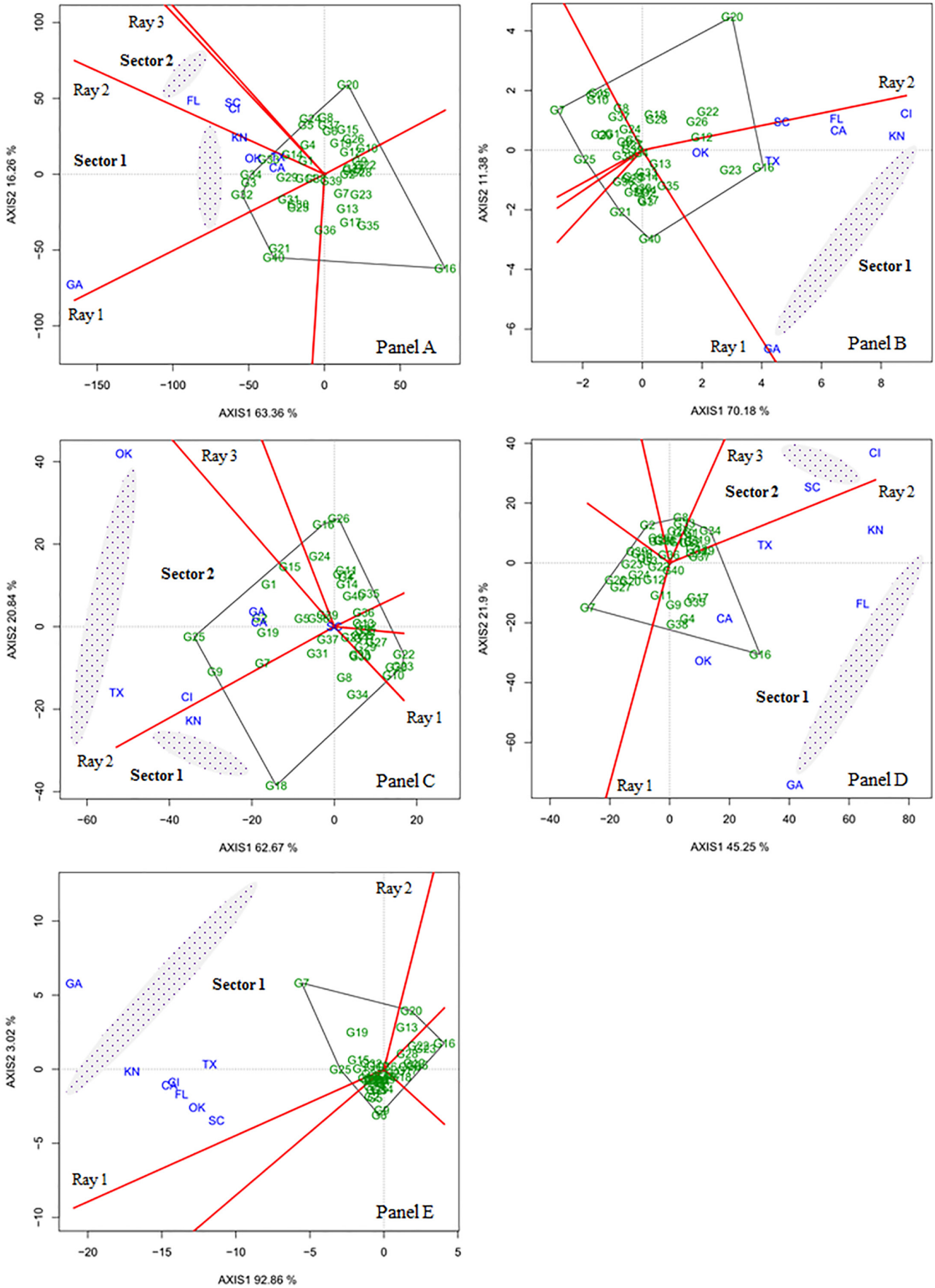



Fig. 1. The polygon (which-won-where) view of GGE biplot of 40 watermelon genotypes tested in 3 yr and eight locations for (A) marketable yield, (B) fruit count, (C) percentage cull fruit, (D) percentage early fruit, and (E) fruit size. The biplots were based on scaling = 0, centering = 2, and singular-value partitioning = 2.

Table 2. Grouping of test locations with winning genotypes based on genotype × location patterns for marketable yield (Mg ha$^{-1}$), fruit count (thousand ha$^{-1}$), percentage cull fruit, percentage early fruit, and fruit size (kg fruit$^{-1}$) of 40 watermelon genotypes tested in 3 yr and eight locations.

| | | Year | | | | | |
|---|---|---|---|---|---|---|---|
| | | 2009 | | 2010 | | 2011 | |
| Trait | Group | Location† | Genotype‡ | Location† | Genotype‡ | Location† | Genotype‡ |
| Marketable yield | 1 | GA | G8 | CI, OK, SC | G14 | CA, GA, OK, TX | G32 (G3) |
| | 2 | FL, KN, SC | G1 (G34) | KN, TX | G32 (G34) | KN, SC | G34 |
| | 3 | CI | G32 | CA, FL | G4 | CI | G20 |
| | 4 | CA, TX | G4 | | | | |
| Fruit count | 1 | FL, SC, TX | G20 (G22, G23) | FL, SC, TX | G20 (G23) | CA, CI, KN, OK, SC, TX | G16 (G23) |
| | 2 | CA, CI, GA, KN | G16 | CI, CA, KN | G16 | GA | G40 |
| | 3 | | | OK | G14 | | |
| Percentage cull fruit | 1 | CI, GA, KN | G18 | CA, CI, KN, TX | G25 (G18) | CI, GA, KN, TX | G9 (G5, G19) |
| | 2 | CA, TX | G25 | OK | G19 | OK | G16 (G26) |
| | 3 | | | | | CA | G25 |
| Percentage early fruit | 1 | CI, TX | G28 | OK, TX | G22 (G12, G13) | CA, OK | G16 |
| | 2 | CA, GA, FL, KN, SC | G16 | CA, CI, FL, KN, SC | G21 (G16, G34) | KN, SC, TX | G19 (G5) |
| | | | | | | CI | G33 |
| Fruit size | 1 | CA, CI, FL, GA, KN, SC, TX | G7 | CI, FL, GA, OK, SC, TX | G7 | CA, CI, GA, KN, OK, SC, TX | G7 |
| | 2 | | | CA, KN | G25 | | |

† KN, Kinston, NC; CI, Clinton; SC, Charleston, SC; GA, Cordele, GA; FL, Quincy, FL; TX, College Station, TX, OK, Lane, OK; CA, Woodland, CA. The locations were chosen to represent the key watermelon production regions in the United States.

‡ Cultivars: G1, AU-Jubilant; G2, Allsweet; G3, Big Crimson; G4, Black Diamond; G5, Calhoun Gray; G6, Calsweet; G7, Carolina Cross #183; G8, Charleston Gray; G9, Congo; G10, Crimson Sweet; G11, Desert King; G12, Early Arizona; G13, Early Canada; G14; Fiesta F1; G15, Georgia Rattlesnake; G16, Golden Midget; G17, Graybelle; G18, Hopi Red Flesh; G19, Jubilee; G20, King & Queen; G21, Legacy; G22, Mickylee; G23, Minilee; G24, Mountain Hoosier; G25, NC Giant; G26, Navajo Sweet; G27, Peacock WR-60; G28, Quetzali; G29, Regency F1; G30, Royal Flush F1; G31, Sangria F1; G32, Starbrite F1; G33, Stars-N-Stripes F1; G34, Stone Mountain; G35, Sugar Baby; G36, Sugarlee; G37, Sweet Princess; G38, Tendersweet OF; G39, Tom Watson; G40, Yellow Crimson. Genotypes in parenthesis are either on the vertex or close to the vertex where winning genotype is located.

representation of the average location (overall location mean). Except for OK, all locations had significant positive correlation with overall location mean for marketable yield, fruit count, percentage cull fruit, and percentage early fruit. The location OK had either weak (significant) or poor correlation (nonsignificant) and thus, had the lowest correlation value for marketable yield, fruit count, percentage cull fruit, and percentage early fruit. For fruit size, all locations had strong and significant correlation with average location. Similarly, correlations of genotype performance among test locations were computed, and results are presented in Supplemental Table S2. These findings further confirm that location OK had weak or no correlation with other locations for marketable yield, fruit count, percentage cull fruit, and percentage early fruit (Supplemental Table S2).

## DISCUSSION

Other than fruit size, all yield traits had large variance as a result environment, with large differences among environmental means causing most of the variation in genotypic performance. Furthermore, $G \times L$ across years for watermelon fruit size was small compared with those of the $L$ and $G$ (Supplemental Table S1). This suggests that breeding watermelon for location-adapted fruit size would not be much of an advantage as pointed out by Annicchiarico (2002). However, for marketable yield, fruit count, percentage cull fruit, and percentage early fruit, $G \times L$ must be exploited to identify mega-environments and watermelon genotypes that are high performing in specific locations or over many locations.

The visualization of which-won-where patterns of GGE biplot identifies the existence of different mega-environments in watermelon growing regions. However, a definitive conclusion on existence of mega-environment must be based on a repeatable which-won-where pattern rather than merely a repeatable environment-grouping pattern (Yan and Rajcan, 2002; Yan and Kang, 2003).

Yearly GGL biplot for marketable yield and percentage early fruit suggested that test locations had different winning genotypes that were not repeatable across years (Table 2). Thus, the $G \times E$ that causes the crossovers among winning genotypes cannot be exploited or converted into $G$ (Yan et al., 2007). Therefore, for marketable yield and percentage early fruit in the target environment consist of a single but complex mega-environment. Watermelon breeders must select widely adapted genotypes for the whole region based on both mean performance and stability analysis using multiple-environment trial data. In this study, 'Stars-N-Stripes' F1 (G33), 'Fiesta' F1 (G14),

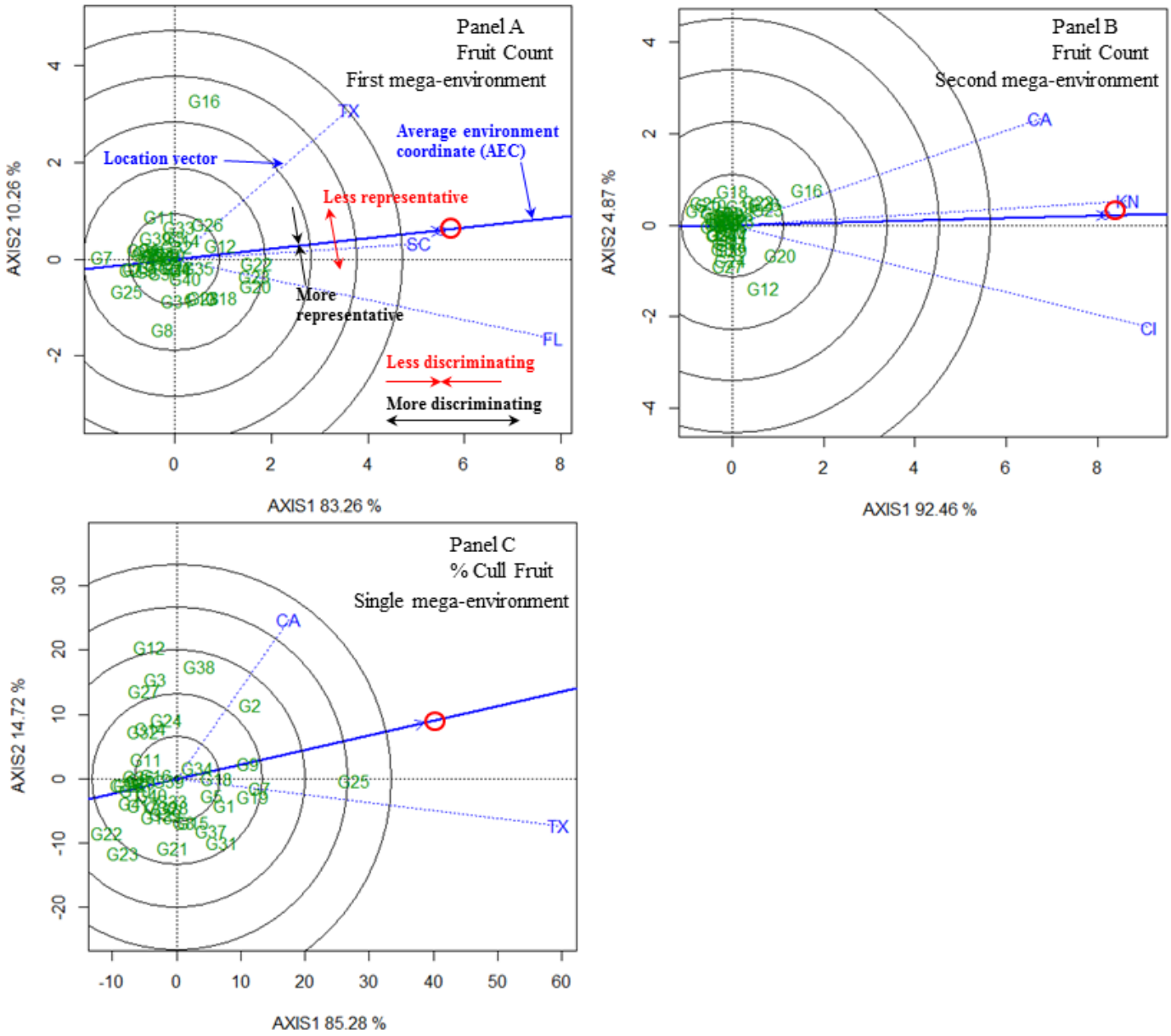



Fig. 2. The discriminating power vs. representativeness view of the GGE biplot of 40 watermelon genotypes tested in 3 yr and eight locations for two mega-environments of fruit count and single mega-environments of percentage cull fruit. (A) Fruit count for first mega-environment. (B) Fruit count for second mega-environment. (C) Percentage cull fruit in single mega-environment. The biplots were based on scaling = 0, centering = 2, and singular-value partitioning 2. The arrows (→← = short; ↔ = long) in (A) indicate the length of location vector that represent less or more discriminating ability, respectively, of location. The red (↕) and black arrows (↓) on average environment coordinate (AEC) indicate the angle of location vector on AEC, which represent less or more representative ability, respectively, of location.

'Stone Mountain' (G34), and 'Calhoun Gray' (G05) were found to be high yielding and stable genotypes (Dia et al., 2016b). Watermelon breeders can select high performing locations using multiple-environment trial data (Table 3). The locations CI, KN, and SC had high mean, high genotype *F*-ratio, and high correlation with average location and so could be used to improve the efficiency of testing (Table 3). Historically, high performing locations are used for trials rather than marginal locations that are associated with large error, less discrimination, and less repeatability over years (Braun et al., 1992).

The GGL biplot for fruit count revealed that the locations FL, SC, and TX tended to be grouped separately from the locations CA, CI, and KN. This pattern was repeatable in 2009 and 2010 but not exactly in 2011 (Table 2). Similarly, two groupings of locations (FL, SC, and TX and CA, CI, and KN) were found with values of genotype *F*-ratio and correlation of locations (Table 3). This location grouping suggested there were two mega-environments for fruit count. The objective of mega-environment analysis is to subdivide the target environment into subregions (mega-environment) so that repeatable $G \times E$

**Table 3. Location mean, standard deviation, genotype *F*-ratio, and correlation of location with overall location mean for marketable yield (Mg ha$^{-1}$), fruit count (thousand ha$^{-1}$), percentage cull fruit, percentage early fruit, and fruit size (kg fruit$^{-1}$) of 40 watermelon genotypes tested in 3 yr and eight locations.**

| Location† | Mean | Genotype discrimination (SD) | Genotype *F*-ratio‡ | Correlation (location mean) |
|---|---|---|---|---|
| Marketable yield | | | | |
| CA | 46.27 | 08.45 | 3.59 | 0.66*** |
| CI | 74.58 | 14.86 | 5.36 | 0.80*** |
| FL | 99.58 | 18.04 | 3.59 | 0.80*** |
| GA | 81.19 | 28.89 | 3.90 | 0.76*** |
| KN | 65.64 | 11.56 | 5.89 | 0.84*** |
| OK | 28.33 | 11.97 | 3.05 | 0.63*** |
| SC | 71.96 | 14.42 | 6.06 | 0.78*** |
| TX | 29.17 | 07.83 | 2.21 | 0.74** |
| Fruit count | | | | |
| CA | 07.31 | 2.86 | 13.29 | 0.90*** |
| CI | 10.77 | 3.73 | 17.40 | 0.93*** |
| FL | 13.45 | 3.20 | 7.03 | 0.80*** |
| GA | 09.63 | 3.15 | 3.43 | 0.56*** |
| KN | 09.56 | 3.51 | 20.26 | 0.95*** |
| OK | 04.64 | 1.89 | 2.59 | 0.44* |
| SC | 10.04 | 2.22 | 6.35 | 0.84*** |
| TX | 5.24 | 1.91 | 4.78 | 0.90*** |
| Percentage cull fruit | | | | |
| CA | 11.74 | 4.85 | 2.30 | 0.64*** |
| CI | 06.37 | 6.75 | 6.50 | 0.89*** |
| FL | –§ | – | – | – |
| GA | 4.83 | 4.31 | 2.16 | 0.68*** |
| KN | 5.19 | 7.58 | 7.37 | 0.81*** |
| OK | 21.09 | 10.62 | 1.66 | 0.67*** |
| SC | – | – | – | – |
| TX | 12.50 | 9.56 | 3.78 | 0.88*** |
| Percentage early fruit | | | | |
| CA | –¶ | – | – | – |
| CI | 45.80 | 12.59 | 5.49 | 0.82*** |
| FL | 46.20 | 13.08 | 5.07 | 0.76*** |
| GA | 13.27 | 14.60 | 4.64 | 0.35*** |
| KN | 24.65 | 12.20 | 6.92 | 0.90*** |
| OK | 24.71 | 12.06 | 1.67 | 0.21 |
| SC | 47.09 | 08.97 | 3.41 | 0.80*** |
| TX | 16.36 | 08.83 | 1.92 | 0.63*** |
| Fruit size | | | | |
| CA | 7.27 | 2.41 | 23.63 | 0.97*** |
| CI | 7.59 | 2.30 | 30.16 | 0.99*** |
| FL | 7.89 | 2.26 | 13.57 | 0.97*** |
| GA | 9.25 | 3.48 | 27.18 | 0.94*** |
| KN | 7.83 | 2.76 | 37.83 | 0.98*** |
| OK | 6.67 | 2.17 | 6.31 | 0.92*** |
| SC | 7.46 | 1.94 | 27.85 | 0.94*** |
| TX | 5.80 | 1.95 | 16.50 | 0.97*** |

* Significant at the 0.05 probability level.

** Significant at the 0.01 probability level.

*** Significant at the 0.001 probability level.

† KN, Kinston, NC; CI, Clinton; SC, Charleston, SC; GA, Cordele, GA; FL, Quincy, FL; TX, College Station, TX, OK, Lane, OK; CA, Woodland, CA. The locations were chosen to represent the key watermelon production regions in the United States.

‡ *F*-ratio was significant at 0.001 level of probability for all the traits and locations.

§ Missing location.

¶ Not enough genotypes were harvested early.

effects can be converted to $G$ by selecting winning genotypes for each mega-environment (Yan et al., 2007).

For genotype evaluation of fruit count, locations SC and KN were identified as ideal test locations in each mega-environment, respectively. Both locations had a large environment vector and small angle with AEC (Fig. 2A, 2B). However, the ideal test location pattern was not consistent across years (Supplemental Fig. S6). It could be due to large differences in experimental error over years as well as differences in correlation coefficients between genotype values in an environment with genotype mean across environments (Yan et al., 2007). Therefore, we assume locations SC and KN were representative of their mega-environment and watermelon breeders can use them as test locations for the mega-environment but should not be used as the main test locations for fruit count.

For percentage cull fruit, yearly GGL biplot indicated that locations CA and TX exhibited repeated crossover pattern across years (Supplemental Fig. S7). However, locations in the eastern United States (CI, GA, and KN) had clear crossover $G \times L$ with varied location grouping across years (Table 2; Supplemental Fig. S7). Thus, eastern watermelon-growing regions cannot be further divided into meaningful subareas but should be regarded as a single complex mega-environment with unpredictable crossover pattern. Locations CA and TX consistently displayed large environment vector and angle with AEC across years (Supplemental Fig. S7). According to Yan et al. (2007), test locations having a large angle with AEC are less representative of the mega-environment and thus, can be used for eliminating inferior genotypes. Thus, locations CA and TX provided the same information (redundancy exists) and can be used interchangeably in identifying genotypes that produce less culls. Identification of redundant test locations can reduce testing cost and improve the efficiency of breeding programs.

For fruit size, all eight locations formed one large mega-environment. The mega-environment was considered simple because there was no major crossover $G \times E$ found, with 'Carolina Cross #183' (G7) being the top in all the locations (Table 2; Supplemental Fig. S5). Watermelon breeders can select the best genotypes based on testing at a single location in a single year. These findings were in agreement with positive and significant correlation between location and location mean (Table 3).

## FUTURE STUDIES

Division of target regions into mega-environments is interesting to plant breeders so that genotype and test-location evaluation become more useful. For marketable yield, percentage early fruit, and percentage cull fruit, the suggested mega-environment based on location grouping did not correspond with the usual regions. The magnitude of the $G \times L$ relative to $G$ suggested the existence of different

mega-environments. However, in most cases, a single complex mega-environment was identified and the crossover pattern did not repeat. It could be that GGL biplot ignores $G \times Y$, $G \times L \times$ Y, or missing values. Furthermore, a definitive conclusion on identification of mega-environment for major watermelon producing regions can be accomplished by adding more dense testing locations and using GGL plus GGE biplot methodology. The GGL plus GGE biplot approach uses missing-value estimation and considers $G \times Y$ and $G \times L \times$ Y. Future study of the spatial and temporal variation of environmental factors across locations may provide insight into location stratification.

## CONCLUSIONS

Watermelon breeders identify superior genotypes based on the $G \times E$ pattern within a target environment. Therefore, it is important for breeders to understand the relationship between target environment and mega-environments as well as the difference between genotype testing and decision making (cultivar release and recommendation) (Yan., 2015). Mega-environment analysis of multiple-environment data of watermelon yield using GGE biplot classified the target environment into three categories. First, marketable yield and percentage early fruit had nonrepeatable crossover $G \times E$ and thus, formed a single complex mega-environment. High yielding and stable genotypes were recommended across locations, for example, Stars-N-Stripes F1 (G33), Fiesta F1 (G14), Stone Mountain (G34), and Calhoun Gray (G05) (Dia et al., 2016b). Second, fruit count and percentage cull fruit had crossover $G \times E$ that was repeatable across years. It provided an opportunity to exploit repeatable $G \times E$ by classifying target environment into mega-environments. The $G \times E$ becomes $G$ when scope of environment was narrowed. Locations KN and SC were the ideal test locations for two mega-environments for fruit count. Locations CA and TX can be used interchangeably for culling genotype that produce high percentage cull fruit. Third, fruit weight did not exhibit crossover interaction and hence, target environment consisted of a single mega-environment. Testing at any single location for a single year was sufficient to select a best genotype for fruit weight.

Identification of mega-environments in US watermelon production regions has several implications for plant breeding. First, high-yielding and stable genotypes should be grown in complex mega-environments to achieve maximum yield. Second, crossover $G \times L$ can be minimized through genotype evaluation and selection focusing on genotype main effect or general adaptation.

## Supplemental Information Available

Supplement information is included with the online version of this article.


## Acknowledgments

The authors would like to thank the following individuals for assistance and support with field and laboratory experiments: Ms. Tammy L. Ellington at North Carolina State University, Raleigh, NC; Mr. Mark Schaffer at Clemson University, Charleston, SC; Mr. Jackie Snell at University of Florida, Gainesville, FL; Ms. Desiree Koehn at Texas A&M University, College Station, TX; Ms. Cathy White and Ms. Dolores Safran at Monsanto Vegetables, Woodland, CA; and Ms. Amy Helms and Mr. Cody Sheffield at USDA–ARS, Lane, OK. We are grateful to the anonymous reviewer for their comments.


## References


Annicchiarico, P. 2002. Genotype × environment interactions: Challenges and opportunities for plant breeding and cultivar recommendations. FAO Plant Production and Protection Paper No. 174. FAO, Rome.

Annicchiarico, P., F. Bellah, and T. Chiari. 2005. Defining subregions and estimating benefits for a specific-adaptation strategy by breeding programs: A case study. Crop Sci. 45:1741–1749. doi:10.2135/cropsci2004.0524

Braun, H.J., W.H. Pfeiffer, and W.G. Pollmer. 1992. Environments for selecting widely adapted spring wheat. Crop Sci. 32:1420–1427. doi:10.2135/cropsci1992.0011183X003200060022x

Braun, H.J., S. Rajaram, and M. van Ginkel. 1996. CIMMYT's approach to breeding for wide adaptation. Euphytica 92:175–183. doi:10.1007/BF00022843

Ceccarelli, S. 1989. Wide adaptation: How wide? Euphytica 40:197–205.

Ceccarelli, S., and S. Grando. 1993. From conventional plant breeding to molecular biology. In: D.R. Buxton, R. Shibles, R.A. Forsberg, B.L. Blad, K.H. Asay, G.M. Paulsen, and R.F. Wilson, editors, International Crop Science I. CSSA, Madison, WI. p. 533–537.

Dia, M., T.C. Wehner, and C. Arellano. 2015. Analysis of genotype × environment interaction (G×E) using SAS programming. http://cuke.hort.ncsu.edu/cucurbit/wehner/software.html (accessed 8 Mar. 2015).

Dia, M., T.C. Wehner, and C. Arellano. 2016a. RG×E: An R program for genotype × environment interaction analysis. http://cuke.hort.ncsu.edu/cucurbit/wehner/software.html (accessed 25 Feb. 2016).

Dia, M., T.C. Wehner, R. Hassell, D.S. Price, G.E. Boyhan, S. Olson, S. King, A.R. Davis and G.E. Tolla. 2016b. Genotype × environment interaction and stability analysis for watermelon fruit yield in the U.S. Crop Sci. doi:10.2135/cropsci2015.10.0625 (in press).

El-Soda, M., W. Kruijer, M. Malosetti, M. Koornneef, and M.G.M. Aarts. 2015. Quantitative trait loci and candidate genes underlying genotype by environment interaction in the response of *Arabidopsis thaliana* to drought. Plant Cell Environ. 38:585–599. doi:10.1111/pce.12418

El-Soda, M., M. Malosetti, B.J. Zwaan, M. Koornneef, and M.G.M. Aarts. 2014. Genotype × environment interaction QTL mapping in plants: Lessons from Arabidopsis. Trends Plant Sci. 1131:1360–1385.

Fan, X.M., M. Kang, H. Chen, Y. Zhang, J. Tan, and C. Xu. 2007. Yield stability of maize hybrids evaluated in multi-environment trials in Yunnan, China. Agron. J. 99:220–228. doi:10.2134/agronj2006.0144

FAO. 2010. Food and Agricultural commodities production. Food and Agriculture Organization of the United Nations: FAOSTAT. http://faostat.fao.org/site/339/default.aspx (accessed 30 Apr. 2010;).

Frutos, E., M.P. Galindo, and V. Leiva. 2014. An interactive biplot implementation in R for modeling genotype-by-environment interaction. Stochastic Environ. Res. Risk Assess. 28:1629–1641. doi:10.1007/s00477-013-0821-z

Gauch, H.G., and R.W. Zobel. 1997. Identifying mega-environment and targeting genotypes. Crop Sci. 37:311–326. doi:10.2135/cropsci1997.0011183X003700020002x

Hussein, M.A., A. Bjornstad, and A.H. Aastveit. 2000. SAS × ESTAB: A SAS program for computing genotype × environment stability statistics. Agron. J. 92:454–459. doi:10.2134/agronj2000.923454x

Luo, J., Y.B. Pan, Y. Que, H. Zhang, M.P. Grisham, and L. Xu. 2015. Biplot evaluation of test environments and identification of mega-environment for sugarcane cultivars in China. Sci. Rep. 5:15505. doi:10.1038/srep15505

Maynard, D.N. 2001. Watermelons. Characteristics, production, and marketing. ASHS Press, Alexandria, VA.

Naiying, X., F. Michel, Z. Guowei, L. Jian, and Z. Zhiguo. 2014. The Application of GGE biplot analysis for evaluating test locations and mega-environment investigation of cotton regional trials. J. Integr. Agric. 9:1921–1933.

Neppl, G.P. 2001. Efficient trailing methods for watermelon *Citrullus lanatus* (Thumb.) Matsum. & Nakai. M.S. thesis, North Carolina State Univ., Raleigh.

R Development Core Team. 2007. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/ (accessed 30 Sept. 2014).

RStudio. 2014. RStudio: Integrated development environment for R (Computer software v0.98.1074). http://www.rstudio.com/ (accessed 30 Sept. 2014).

Samonte, S.O., L.T. Wilson, A.M. McClung, and J.C. Medley. 2005. Targeting cultivars onto rice growing environments using AMMI and SREG GGE biplot analyses. Crop Sci. 45:2414–2424. doi:10.2135/cropsci2004.0627

SAS Institute. 2014. SAS system for Windows. v. 9.4m2. SAS Inst. Inc., Cary, NC.

SAS Institute. 2016. SAS/STAT 9.2 User's Guide, Second Edition: Random-effects analysis. https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_glm_sect037.htm (accessed 1 Mar. 2016).

Simmonds, N.W. 1991. Selection for local adaptation in a plant breeding program. Theor. Appl. Genet. 82:363–367. doi:10.1007/BF02190624

USDA. 2010. Statistics by subject. USDA National Agricultural Statistics Services. Washington, DC. http://www.nass.usda.gov/Statistics_by_Subject/?sector=CROPS (accessed 30 Apr. 2010).

Wehner, T.C. 2008. Watermelon. In: J. Prohens and F. Nuez, editors, Handbook of Plant Breeding; Vegetables I: Asteraceae, Brassicaceae, Chenopodiaceae, and Cucurbitaceae. Springer Science+Business LLC., New York. p. 381–418.

Yan, W. 2002. Singular-value partition for biplot analysis of multi-environment trial data. Agron. J. 94:990–996. doi:10.2134/agronj2002.0990

Yan, W. 2006. Exploring multi-environment trial data using biplots. http://www.ggebiplot.com/workshop.htm (accessed 11 Oct. 2011).

Yan, W. 2015. Mega-environment analysis and test location evaluation based on unbalanced multiyear data. Crop Sci. 55:113–122. doi:10.2135/cropsci2014.03.0203

Yan, W., and J.B. Holland. 2010. A heritability-adjusted GGE biplot for test environment evaluation. Euphytica 171:355–369. doi:10.1007/s10681-009-0030-5

Yan, W., L.A. Hunt, Q. Sheng, and Z. Szlavnics. 2000. Cultivar evaluation and mega-environment investigation based on GGE biplot. Crop Sci. 40:597–605. doi:10.2135/cropsci2000.403597x

Yan, W., and M.S. Kang. 2003. GGE biplot analysis: A graphical tool for breeders, geneticists, and agronomists. CRC Press, Boca Raton, FL.

Yan, W., M.S. Kang, B. Ma, S. Woods, and P.L. Cornelius. 2007. GGE biplot vs. AMMI analysis of genotype-by-environment data. Crop Sci. 47:643–653. doi:10.2135/cropsci2006.06.0374

Yan, W., D. Pageau, J. Frégeau-Reid, and J. Durand. 2011. Assessing the representativeness and repeatability of test locations for genotype evaluation. Crop Sci. 51:1603–1610. doi:10.2135/cropsci2011.01.0016

Yan, W., and I. Rajcan. 2002. Biplot analysis of test sites and trait relations of soybean in Ontario. Crop Sci. 42:11–20. doi:10.2135/cropsci2002.0011

Yan, W., and N. Tinker. 2005. An integrated biplot analysis system for displaying, interpreting, and exploring genotype × environment interaction. Crop Sci. 45:1004–1016. doi:10.2135/cropsci2004.0076

**Supplemental Material Description**

The supplemental tables provide yearly analysis of variance and correlation of test location performance for 40 genotypes and for all the traits evaluated in the study. Similarly, supplemental figures provide yearly GGL biplot for marketable yield, fruit count, % cull fruit, % early fruit, and fruit size; and yearly 'discriminating power vs. representativeness' view of the GGE biplot for fruit count and % cull fruits.

Supplemental Table Captions:

Supplemental Table 1. Genotype, location and genotype x location variance effect by year for marketable yield (Mg $ha^{-1}$), fruit count, % cull fruit, % early fruit, and fruit size of 40 watermelon genotypes tested in 3 years and 8 locations.

Supplemental Table 2. Correlation among test locations for marketable yield (Mg $ha^{-1}$), fruit count (thousand $ha^{-1}$), % cull fruit, % early fruit, and fruit size (kg $fruit^{-1}$) of 40 watermelon genotypes tested in 3 years and 8 locations.

Supplemental Figure Captions:

Supplemental Figure 1. The polygon ('which-won-where') view of GGL biplot of 40 watermelon genotypes tested in 2009 (Panel A), 2010 (Panel B) and 2011 (Panel C), and 8 locations for marketable yield. The biplots were based on 'Scaling = 0', 'Centering = 2' and 'SVP = 2'.

Supplemental Figure 2. The polygon ('which-won-where') view of GGL biplot of 40 watermelon genotypes tested in 2009 (Panel A), 2010 (Panel B) and 2011 (Panel C), and 8 locations for fruit count. The biplots were based on 'Scaling = 0', 'Centering = 2' and 'SVP = 2'.

Supplemental Figure 3. The polygon ('which-won-where') view of GGL biplot of 40 watermelon genotypes tested in 2009 (Panel A), 2010 (Panel B) and 2011 (Panel C), and 8 locations for % cull fruit. The biplots were based on 'Scaling = 0', 'Centering = 2' and 'SVP = 2'.

Supplemental Figure 4. The polygon ('which-won-where') view of GGL biplot of 40 watermelon genotypes tested in 2009 (Panel A), 2010 (Panel B) and 2011 (Panel C), and 8 locations for % early fruit. The biplots were based on 'Scaling = 0', 'Centering = 2' and 'SVP = 2'.

Supplemental Figure 5. The polygon ('which-won-where') view of GGL biplot of 40 watermelon genotypes tested in 2009 (Panel A), 2010 (Panel B) and 2011 (Panel C), and 8 locations for fruit size. The biplots were based on 'Scaling = 0', 'Centering = 2' and 'SVP = 2'.

Supplemental Figure 6. The 'discriminating power vs. representativeness' view of the GGE biplot of 40 watermelon genotypes tested in 3 years and 8 locations for two mega environments of fruit count. The biplots were based on 'Scaling = 0', 'Centering = 2' and 'SVP = 2'. The first

mega-environment for 2009, 2010 and 2011 is presented in Panel A, Panel B and Panel C, respectively. Similarly, the second mega-environment for 2009, 2010 and 2011 is presented in Panel D, Panel E and Panel F, respectively.

Supplemental Figure 7. The ‘discriminating power vs. representativeness’ view of the GGE biplot of 40 watermelon genotypes tested in 3 years and 8 locations for single mega environments of % cull fruit. The biplots were based on ‘Scaling = 0’, ‘Centering = 2’ and ‘SVP = 2’.

Table S1. Genotype, location and genotype x location variance effect by year for marketable yield (Mg $ha^{-1}$), fruit count, % cull fruit, % early fruit, and fruit size of 40 watermelon genotypes tested in 3 years and 8 locations.

| | Marketable yield | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Degrees of freedom | | | Sum of squares (in thousands) | | | % Sum of squares* | | |
| Source | 2009 | 2010 | 2011 | 2009 | 2010 | 2011 | 2009 | 2010 | 2011 |
| Location (L) | 6 | 6 | 6 | 389.21 | 775.52 | 908.18 | 55 | 72 | 63 |
| Genotype (G) | 39 | 39 | 39 | 148.32 | 118.17 | 165.78 | 21 | 11 | 12 |
| GxL | 234 | 234 | 234 | 173.81 | 186.97 | 365.12 | 24 | 17 | 25 |

| | Fruit count | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Degrees of freedom | | | Sum of squares (in hundreds) | | | % Sum of squares* | | |
| Source | 2009 | 2010 | 2011 | 2009 | 2010 | 2011 | 2009 | 2010 | 2011 |
| Location (L) | 6 | 6 | 6 | 68.00 | 92.64 | 136.38 | 36 | 49 | 48 |
| Genotype (G) | 39 | 39 | 39 | 79.74 | 46.77 | 68.14 | 42 | 24 | 24 |
| GxL | 234 | 233 | 234 | 40.41 | 50.91 | 80.87 | 21 | 27 | 28 |

| | % cull fruit | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Degrees of freedom | | | Sum of squares (in hundreds) | | | % Sum of squares* | | |
| Source | 2009 | 2010 | 2011 | 2009 | 2010 | 2011 | 2009 | 2010 | 2011 |
| Location (L) | 5 | 5 | 6 | 245.03 | 90.46 | 1126.04 | 36 | 9 | 53 |
| Genotype (G) | 39 | 39 | 39 | 152.37 | 405.97 | 365.78 | 22 | 41 | 17 |
| GxL | 195 | 194 | 232 | 292.71 | 488.45 | 632.06 | 42 | 50 | 30 |

| | % early fruit | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Degrees of freedom | | | Sum of squares (in thousand) | | | % Sum of squares* | | |
| Source | 2009 | 2010 | 2011 | 2009 | 2010 | 2011 | 2009 | 2010 | 2011 |
| Location (L) | 6 | 6 | 6 | 260.06 | 173.98 | 177.86 | 56 | 45 | 53 |
| Genotype (G) | 39 | 39 | 39 | 76.24 | 69.61 | 55.49 | 17 | 18 | 17 |
| GxL | 234 | 193 | 156 | 125.28 | 139.87 | 101.83 | 27 | 36 | 30 |

| | Fruit size | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Degrees of freedom | | | Sum of squares (in hundreds) | | | % Sum of squares* | | |
| Source | 2009 | 2010 | 2011 | 2009 | 2010 | 2011 | 2009 | 2010 | 2011 |
| Location (L) | 6 | 6 | 6 | 11.66 | 26.37 | 6.06 | 15 | 31 | 7 |
| Genotype (G) | 39 | 39 | 39 | 57.41 | 45.04 | 68.84 | 74 | 54 | 78 |
| GxL | 234 | 232 | 232 | 8.73 | 12.73 | 12.88 | 11 | 15 | 15 |

* Percent sum of square were calculated based on L, G and GxL (Replication within location effect was not included).

Table S2. Correlation among test locations for marketable yield (Mg $ha^{-1}$), fruit count (thousand $ha^{-1}$), % cull fruit, % early fruit, and fruit size (kg $fruit^{-1}$) of 40 watermelon genotypes tested in 3 years and 8 locations.

| | Marketable yield | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Location | CA | CI | FL | GA | KN | OK | SC | TX |
| CA | 1.00 | | | | | | | |
| CI | 0.53*** | 1.00 | | | | | | |
| FL | 0.47** | 0.55*** | 1.00 | | | | | |
| GA | 0.49** | 0.42* | 0.52*** | 1.00 | | | | |
| KN | 0.40* | 0.66*** | 0.70*** | 0.57*** | 1.00 | | | |
| OK | 0.24 | 0.30 | 0.50** | 0.49** | 0.56** | 1.00 | | |
| SC | 0.31 | 0.63*** | 0.63*** | 0.41** | 0.63*** | 0.57*** | 1.00 | |
| TX | 0.72*** | 0.71*** | 0.47* | 0.47* | 0.57*** | 0.25 | 0.44* | 1.00 |

| | Fruit count | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Location | CA | CI | FL | GA | KN | OK | SC | TX |
| CA | 1.00 | | | | | | | |
| CI | 0.84*** | 1.00 | | | | | | |
| FL | 0.66*** | 0.66*** | 1.00 | | | | | |
| GA | 0.40* | 0.40* | 0.36* | 1.00 | | | | |
| KN | 0.89*** | 0.91*** | 0.68*** | 0.47** | 1.00 | | | |
| OK | 0.23 | 0.27 | 0.45** | 0.24 | 0.38* | 1.00 | | |
| SC | 0.66*** | 0.75*** | 0.80*** | 0.34* | 0.75*** | 0.52*** | 1.00 | |
| TX | 0.90*** | 0.84*** | 0.62*** | 0.53*** | 0.88*** | 0.22 | 0.64*** | 1.00 |

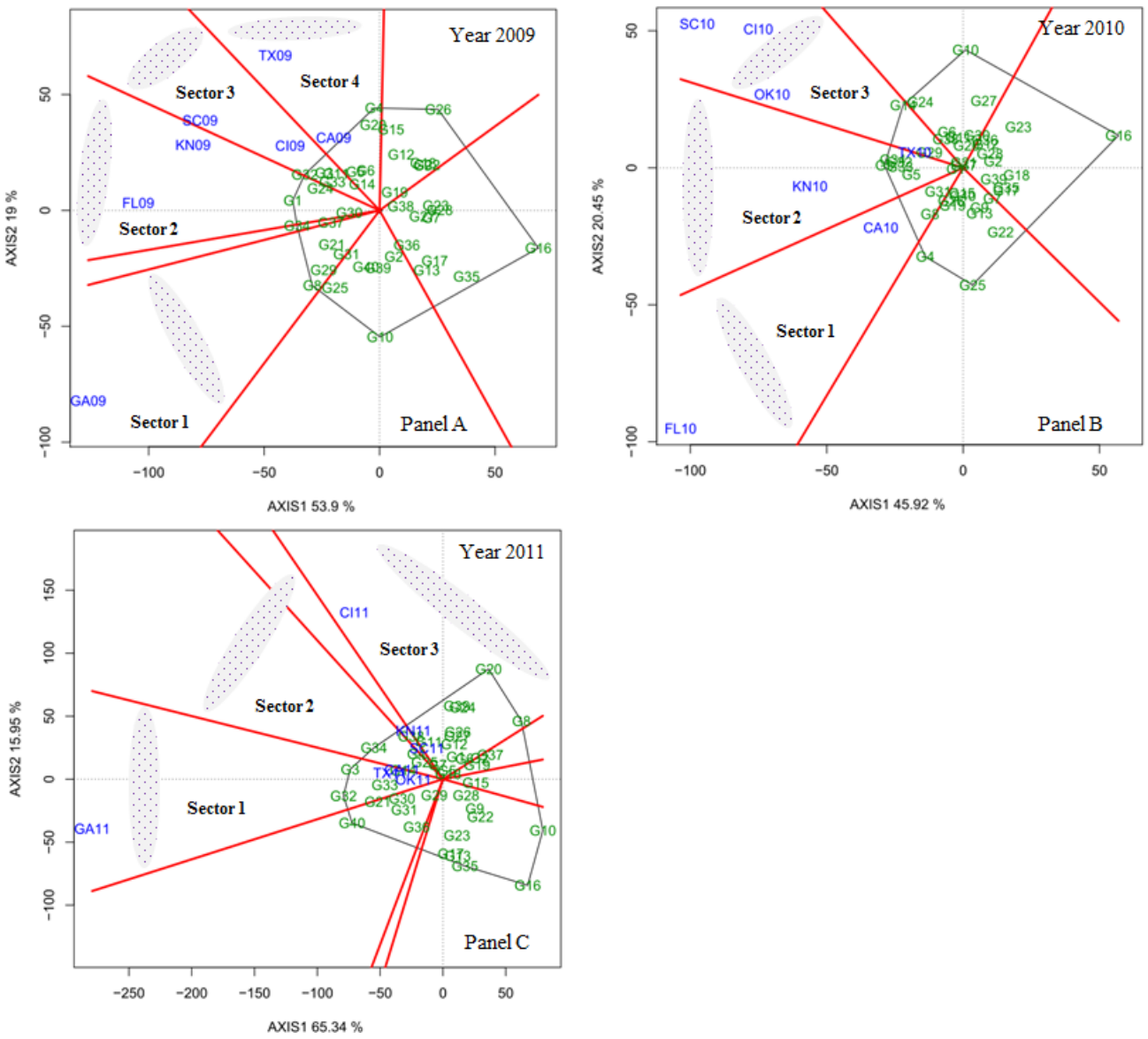
| | % cull fruit | | | | | |
|---|---|---|---|---|---|---|
| Location | CA | CI | GA | KN | OK | TX |
| CA | 1.00 | | | | | |
| CI | 0.44** | 1.00 | | | | |
| GA | 0.21 | 0.49** | 1.00 | | | |
| KN | 0.45** | 0.82*** | 0.48** | 1.00 | | |
| OK | 0.47** | 0.41* | 0.60*** | 0.29 | 1.00 | |
| TX | 0.47** | 0.80*** | 0.55** | 0.61*** | 0.48** | 1.00 |

| | % early fruit | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Location | CA | CI | FL | GA | KN | OK | SC | TX |
| CA | 1.00 | | | | | | | |
| CI | 0.16 | 1.00 | | | | | | |
| FL | 0.56*** | 0.57*** | 1.00 | | | | | |
| GA | 0.49** | -0.02 | 0.37* | 1.00 | | | | |
| KN | 0.30 | 0.53*** | 0.36* | 0.28 | 1.00 | | | |
| OK | 0.27 | 0.07 | 0.21 | 0.18 | 0.19 | 1.00 | | |
| SC | 0.09 | 0.64*** | 0.48** | -0.01 | 0.65*** | 0.04 | 1.00 | |
| TX | 0.13 | 0.29 | 0.28 | 0.15 | 0.48** | 0.05 | 0.34* | 1.00 |

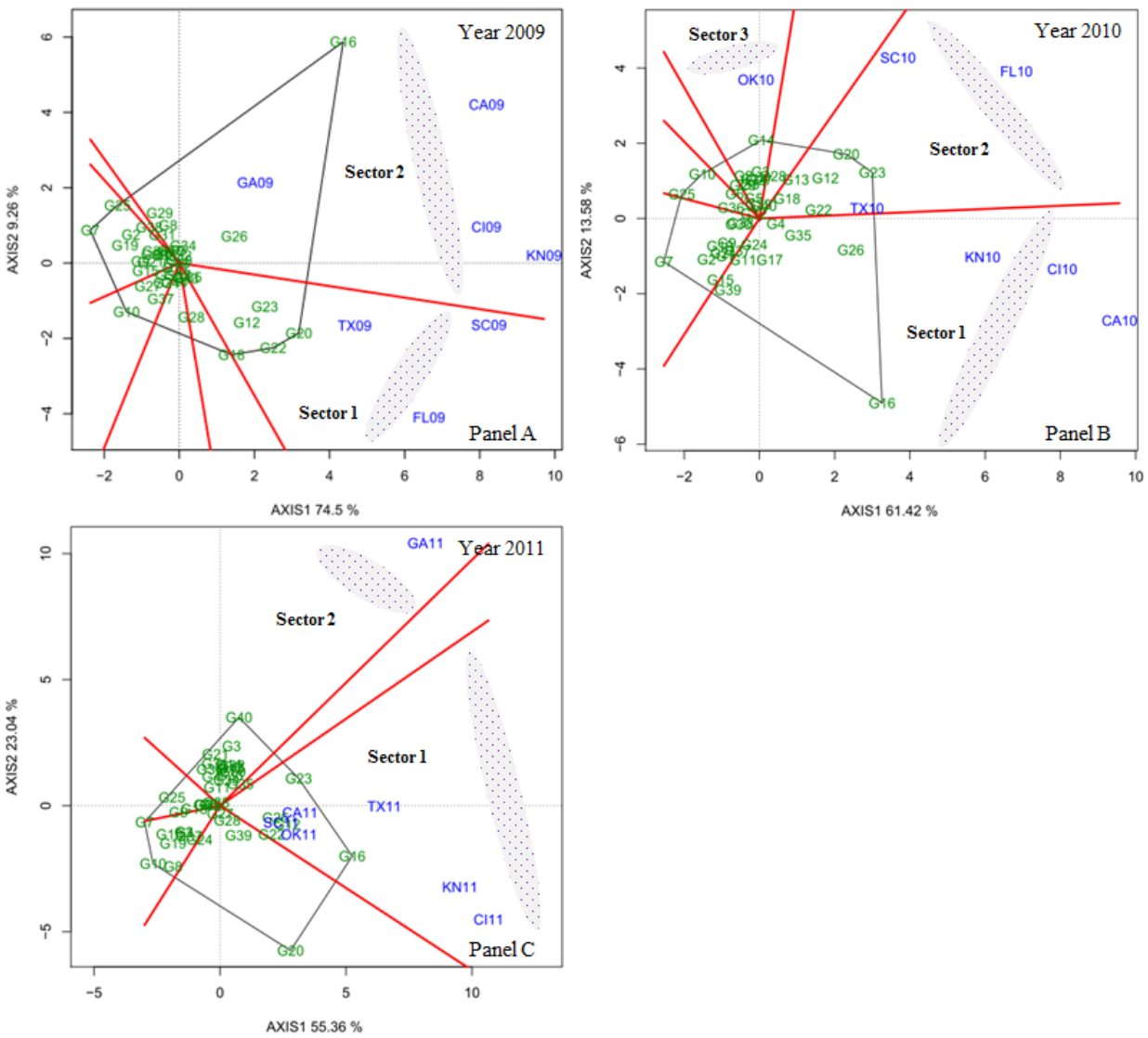
| Location | Fruit size |
|---|---|

| | CA | CI | FL | GA | KN | OK | SC | TX |
|---|---|---|---|---|---|---|---|---|
| CA | 1.00 | | | | | | | |
| CI | 0.95*** | 1.00 | | | | | | |
| FL | 0.94*** | 0.96*** | 1.00 | | | | | |
| GA | 0.90*** | 0.93*** | 0.90*** | 1.00 | | | | |
| KN | 0.96*** | 0.98*** | 0.95*** | 0.93*** | 1.00 | | | |
| OK | 0.87*** | 0.92*** | 0.89*** | 0.86*** | 0.89*** | 1.00 | | |
| SC | 0.90*** | 0.94*** | 0.93*** | 0.83*** | 0.91*** | 0.91*** | 1.00 | |
| TX | 0.93*** | 0.94*** | 0.92*** | 0.93*** | 0.94*** | 0.88*** | 0.89*** | 1.00 |

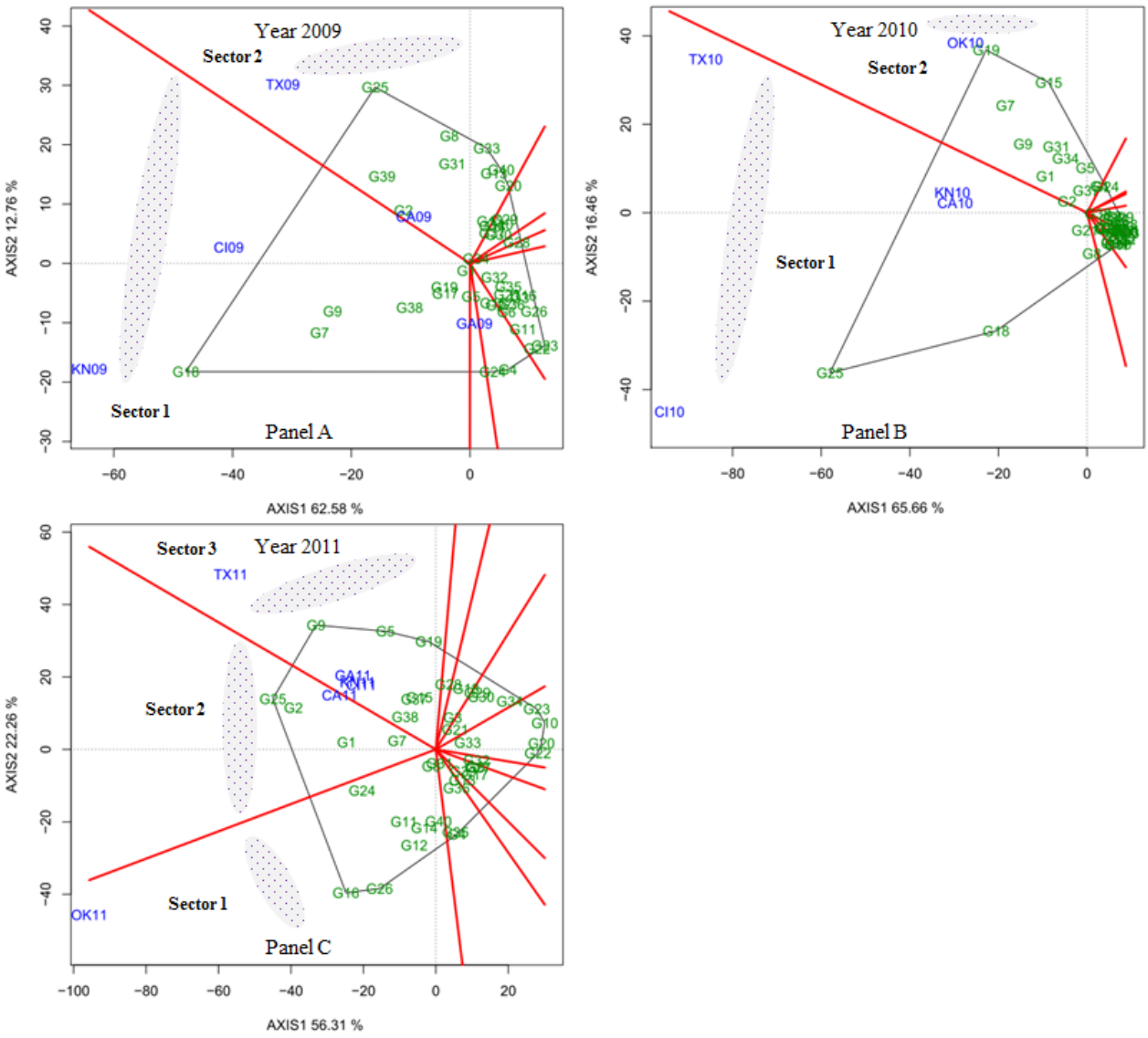
*, **, *** Significant at the 0.05, 0.01, and 0.001 level of probability, respectively



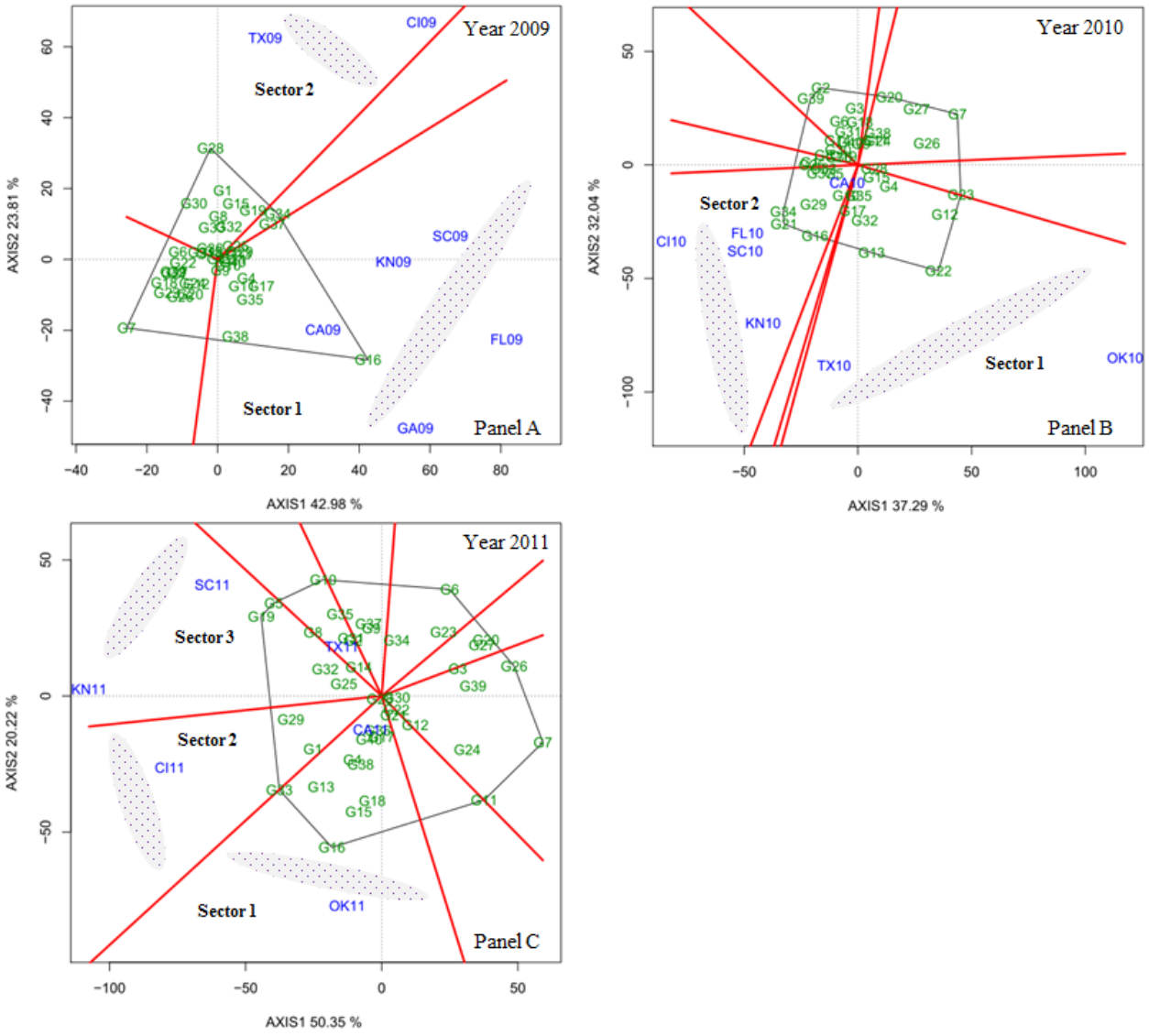

Supplemental Figure 1. The polygon ('which-won-where') view of GGL biplot of 40 watermelon genotypes tested in 2009 (Panel A), 2010 (Panel B) and 2011 (Panel C), and 8 locations for marketable yield. The biplots were based on 'Scaling = 0', 'Centering = 2' and 'SVP = 2'.



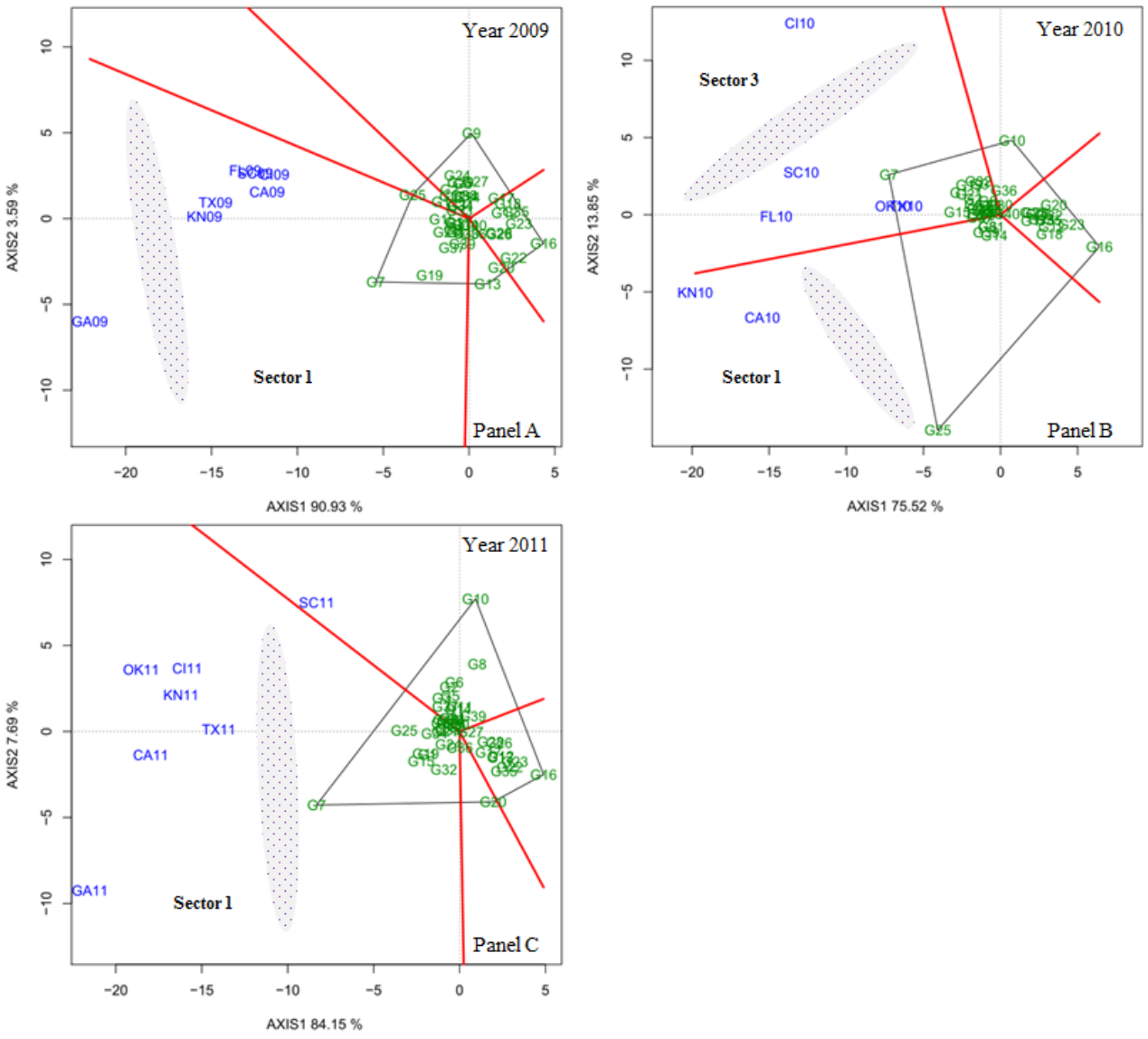

Supplemental Figure 2. The polygon ('which-won-where') view of GGL biplot of 40 watermelon genotypes tested in 2009 (Panel A), 2010 (Panel B) and 2011 (Panel C), and 8 locations for fruit count. The biplots were based on 'Scaling = 0', 'Centering = 2' and 'SVP = 2'.



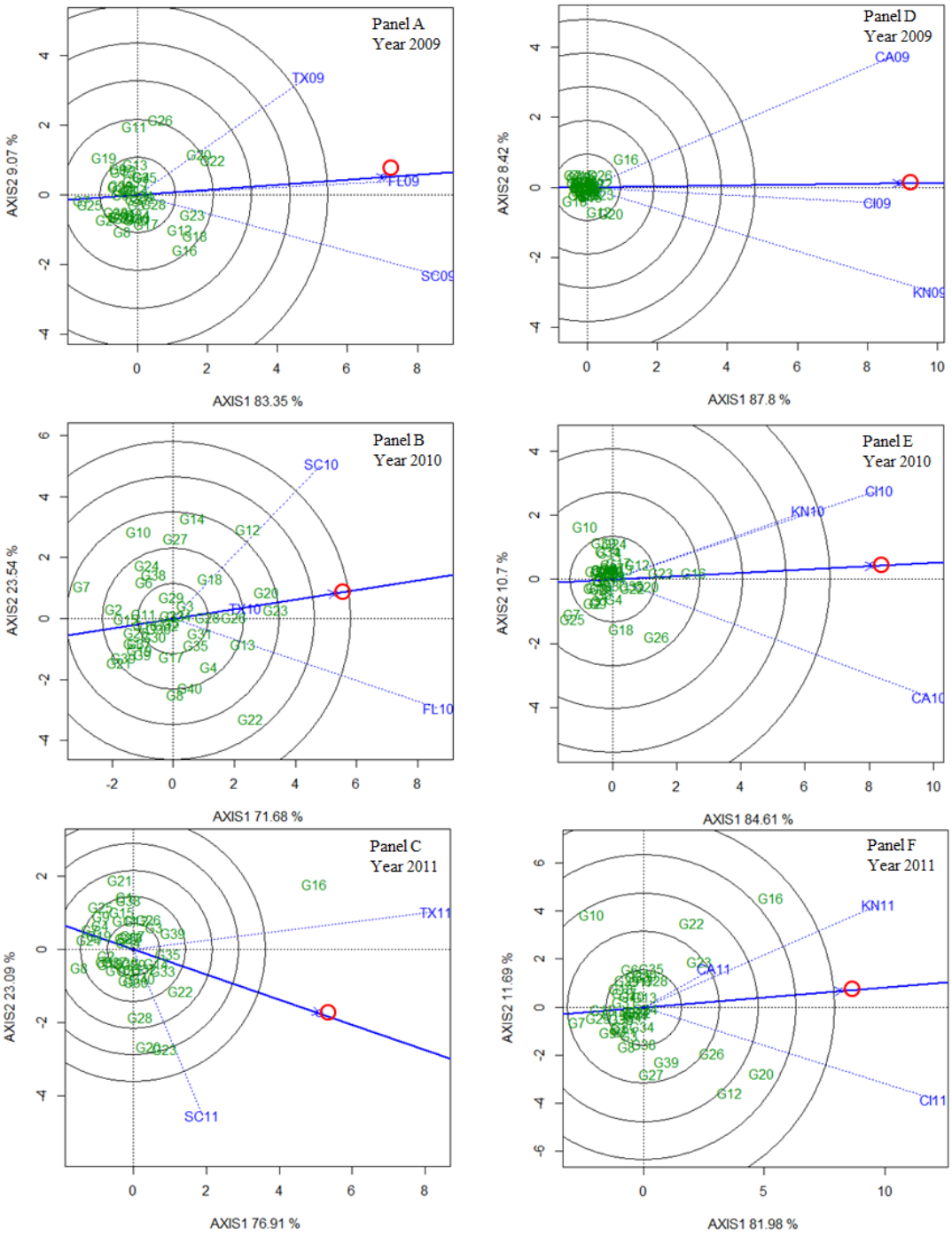

Supplemental Figure 3. The polygon ('which-won-where') view of GGL biplot of 40 watermelon genotypes tested in 2009 (Panel A), 2010 (Panel B) and 2011 (Panel C), and 8 locations for % cull fruit. The biplots were based on 'Scaling = 0', 'Centering = 2' and 'SVP = 2'.



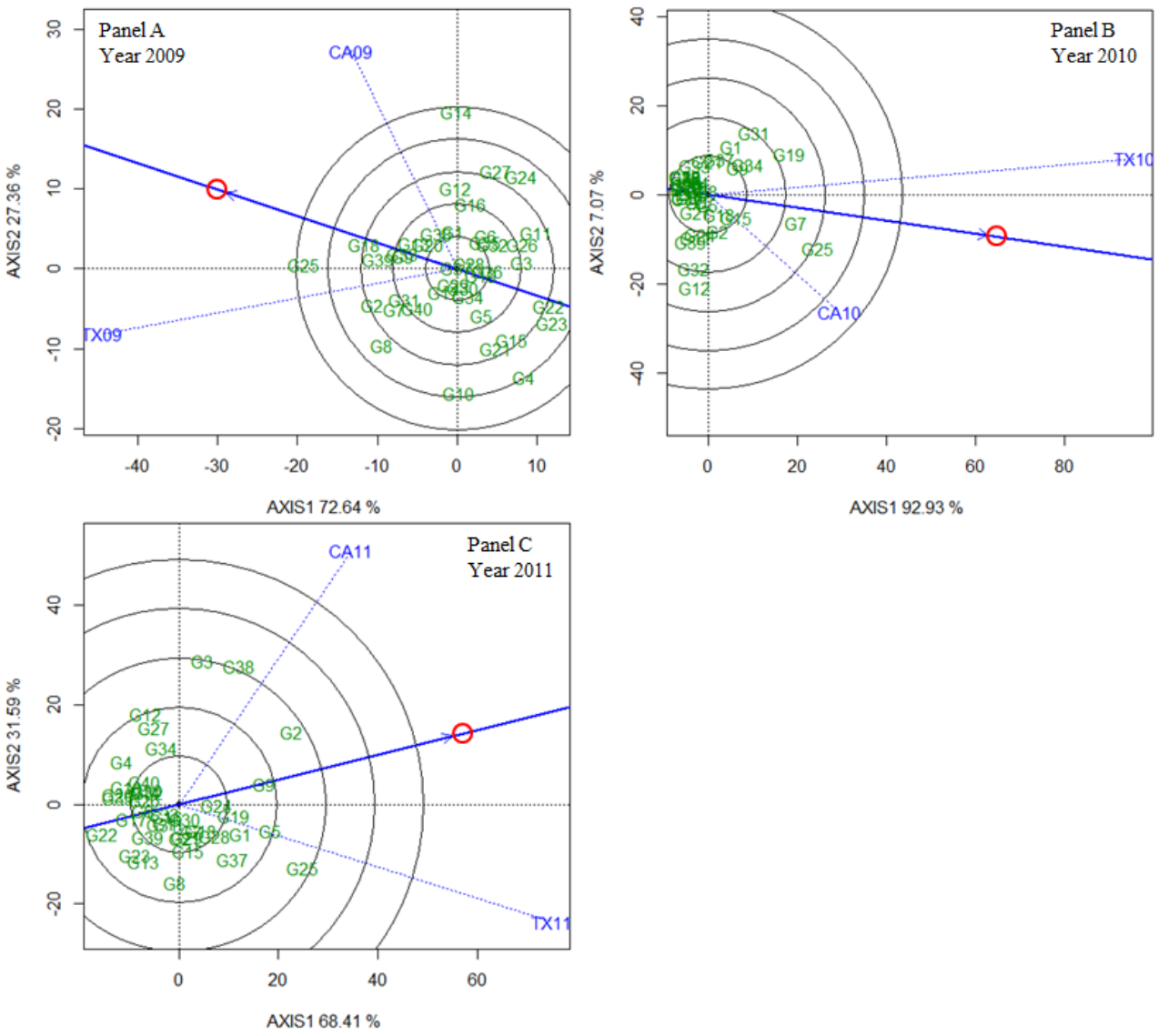

Supplemental Figure 4. The polygon ('which-won-where') view of GGL biplot of 40 watermelon genotypes tested in 2009 (Panel A), 2010 (Panel B) and 2011 (Panel C), and 8 locations for % early fruit. The biplots were based on 'Scaling = 0', 'Centering = 2' and 'SVP = 2'.




Supplemental Figure 5. The polygon ('which-won-where') view of GGL biplot of 40 watermelon genotypes tested in 2009 (Panel A), 2010 (Panel B) and 2011 (Panel C), and 8 locations for fruit size. The biplots were based on 'Scaling = 0', 'Centering = 2' and 'SVP = 2'.




Supplemental Figure 6. The 'discriminating power vs. representativeness' view of the GGE biplot of 40 watermelon genotypes tested in 3 years and 8 locations for two mega environments of fruit count. The biplots were based on 'Scaling = 0', 'Centering = 2' and 'SVP = 2'. The first mega-environment for 2009, 2010 and 2011 is presented in Panel A, Panel B and Panel C, respectively. Similarly, the second mega-environment for 2009, 2010 and 2011 is presented in Panel D, Panel E and Panel F, respectively.




Supplemental Figure 7. The 'discriminating power vs. representativeness' view of the GGE biplot of 40 watermelon genotypes tested in 3 years and 8 locations for single mega environments of % cull fruit. The biplots were based on 'Scaling = 0', 'Centering = 2' and 'SVP = 2'.