COMPARING THE PERFORMANCE OF F_1 TESTERS VERSUS THEIR INBRED LINE PARENTS IN EVALUATING EXPERIMENTAL SORGHUM R AND B LINES IN TESTCROSSES

A Thesis

by

DANIEL JACOB PACKER

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

December 2007

Major Subject: Plant Breeding

COMPARING THE PERFORMANCE OF F_1 TESTERS VERSUS THEIR INBRED LINE PARENTS IN EVALUATING EXPERIMENTAL SORGHUM R AND B LINES IN TESTCROSSES

A Thesis

by

DANIEL JACOB PACKER

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	William Rooney
Committee Members,	Juerg Blumenthal
	Stephen R. King
Head of Department,	David Baltensperger

December 2007

Major Subject: Plant Breeding

ABSTRACT

Comparing the Performance of F1 Testers Versus Their Inbred Line Parents in Evaluating Experimental Sorghum R and B Lines in Testcrosses.

(December 2007)

Daniel Jacob Packer, B.S., Brigham Young University

Chair of Advisory Committee: Dr. William L. Rooney

An appropriate tester correctly identifies the relative performance of experimental lines while maximizing the differences between lines. Most sorghum breeding programs use elite inbred lines testers. Inbred line testers evaluate experimental lines against a specific genetic background, possibly increasing the probability of incorrectly discarding material. A potential solution would be to use F_1 testers that combine two genetic backgrounds. The purpose of this research was to compare F_1 testers versus inbred line testers for evaluating experimental sorghum lines in testcrosses

Line x tester analyses were performed to assess tester consistency in assigning ranks. With one exception, all of the line x tester analyses were non-significant, indicating that the testers provided similar evaluations of the experimental lines.

Correlations between the ranking of the experimental lines by their average performance and the rank assignments of each tester were measured to further asses tester accuracy. In all cases, the rank correlations were highly significant, implying that all of the testers accurately ranked experimental lines. In addition, all of the testers consistently

identified the majority of the top performing experimental lines despite some important rank shifts.

F-ratios for variance among the experimental lines (entry effect) were compared with the Schumann-Bradley statistical test to compare efficiencies. With one exception, the F_1 testers always produced the largest or second largest entry effect F-ratio. Where the F_1 testers produced the second largest F-ratio, it was not declared statistically different from the largest F-ratio by the Schumann-Bradley test, indicating that the testers had similar discriminatory efficiencies.

Testcross variances were measured to further compare discriminatory efficiencies. With one exception, the F_1 testers consistently produced the largest variances, evidence that the F_1 testers were effective in maximizing differences among the experimental lines.

The results indicate that F_1 testers represent valid testers for evaluating experimental sorghum lines against two genetic backgrounds in a single testcross.

DEDICATION

First and foremost, I dedicate this thesis to my Heavenly Father, who has aided and accompanied me down many paths I was never expecting. And to my wonderful family, who has provided me with unconditional love, support, and happiness. Whatever may come, they are my bastion and refuge.

"Mas buscad primeramente el reino de Dios y su justicia, y todas estas cosas os eran añadidas"

ACKNOWLEDGEMENTS

I greatly appreciate the much needed guidance and knowledge that Dr. Bill Rooney has provided me. And I look forward to continue learning from his wealth of experience. I would like to thank my committee members Dr. Juerg Blumenthal and Dr. Stephen R. King who have provided me with knowledge and support in my endeavors.

I am also greatly indebted to Dr. Rick Jellen and Dr. Jeffrey Maughan of Brigham Young University and Dr. John Walker of Ricks College for placing me on the path I now find myself.

And to all the present and past members of the Texas A&M Sorghum Breeding program I have had the pleasure of working with, Delroy Collins, Karen Prihoda, Dustin Borden, Stephen Labar, Terry Felderhoff, Bill Lyles, Les Kuhlman, Ryan Bading, Gabriel Krishnamoorthy, and Rafael Mateo, thank you for your tremendous help and friendship.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

TABLE Page

TABLE Page

TABLE Page

CHAPTER I

INTRODUCTION

Sorghum (*Sorghum bicolor* (L.) Moench) is cultivated in many of the semi-arid regions of the world for both grain and forage production. The crop was first domesticated approximately 4000 to 6000 years ago in northeastern Africa (Kimber 2000). Subsequently, cultivated sorghums spread to other regions of Africa, India, and China through trade and migration routes. With time and isolation, farmer selection for local adaptation led to the development of the five main races of sorghum, as well as individual varieties within those races. Sorghum was introduced into North America during the 17th century and subsequent development of pure-line cultivars initially came from farmer selections (Rooney 2004).

Worldwide, grain sorghum ranks fifth among grain crops for production with approximately 40 million Ha grown worldwide in 2005 (USDA, 2007b). Within the United States, approximately 6.7 million acres of sorghum harvested for grain are forecasted for 2007, with 2.5 million of those acres located in Texas (USDA, 2007a).

With the commercial introduction of hybrid sorghum in 1956, and its subsequent rapid adoption by producers, most sorghum breeding programs shifted their emphasis from developing pure-line cultivars to the development of inbred parental lines for hybrid production. During the process of inbred line development, the need arises to evaluate new inbred lines for their potential as future hybrid parents. This is accomplished using a testcross; new inbred lines are crossed to a tester line to produce hybrid progeny, which

---

This thesis follows the style of Crop Science.

are then evaluated for their performance. New inbred lines will either be advanced or discarded based on the performance of the progeny produced with the tester line.

To ensure the most accurate evaluation of new inbred lines, it is essential that an appropriate tester be selected. Tester options range from broad genetic base populations to elite inbred lines; most sorghum breeding programs will have their own criteria for tester selection, based on their needs and goals. An effective tester should correctly rank inbred lines for performance in hybrid combination, and it should maximize the variance between testcross progeny to allow for efficient discrimination of new inbred lines (Rawlings and Thompson 1962).

Most sorghum breeding programs utilize elite inbred line testers, so as to evaluate experimental lines in realistic hybrid combinations. This approach provides breeders with valuable information regarding the performance of new lines with specific genetic backgrounds. However, utilizing a single elite inbred tester only allows for the evaluation of new lines with alleles from a single genetic background or adaptation type, which may increase the probability of incorrectly discarding material. This limitation may be circumvented by testcrossing new lines to several different elite inbred testers. However, this requires greater expenditures in time, resources, and space that are limited in any breeding program. This is particularly true during the early phases of testcross evaluation, when the amount of new material for consideration is greatest.

The use of a single F_1 tester is a compromise that allows evaluation of new material with alleles from more than one genetic background or adaptation type in a single testcross. F_1 testers combine alleles from two backgrounds in a single tester. New

lines identified as superior by F_1 testers could then be advanced and tested with the inbred parents of the F_1 tester to identify specific combining abilities.

The objective of this research is to compare the performance of F_1 testers for evaluating experimental R and B sorghum lines versus using the inbred parents of the F_1 testers alone as the testers. The comparison of these two types of testers is based on the following criteria:

1. Analyses comparing the accuracy of the F_1 testers versus their respective inbred line testers in correctly identifying the relative performance of the experimental lines.
2. Analyses comparing the efficiency of the F_1 testers versus their respective inbred line testers in discriminating among the experimental lines.

The hypothesis of this research is that the F_1 testers will provide similar levels of accuracy and efficiency in evaluating the experimental lines as their inbred line counterparts and thus, represent a valid option for the evaluation of new material against two genetic backgrounds or adaptation types, in a single testcross.

CHAPTER II

LITERATURE REVIEW

Testcrossing

With the commercial introduction of hybrid maize in the 1930s, methods for evaluating the combing ability of inbred lines began to be developed. Jenkins and Brunson (1932) compared the correlation between the ranking of inbred lines by their average performance in multiple single crosses with the ranks produced by individual testcrosses. They found strong correlations and concluded that testcrosses could be used to identify inbred lines with good combing ability with an acceptable degree of accuracy. These results were further confirmed by the work of Johnson and Hayes (1936) and Cowan (1943).

To ensure the most accurate evaluation of new inbred lines, it is essential that an appropriate tester be selected. Tester options range from broad genetic base populations to elite inbred lines, and most breeding programs will have their own criteria for tester selection, based on their needs and goals. However, any tester selected must meet certain requirements to be considered effective. An effective tester should correctly rank inbred lines for performance in hybrid combination, and it should maximize the variance between testcross progeny to allow for efficient discrimination of new inbred lines (Rawlings and Thompson 1962). Allison and Curnow (1966) state that the use of any tester followed by selection will lead to an increase in mean yield but they will do so at different rates depending on gene action of the trait and allele frequencies in the tester.

Initial testcrosses are designed to identify genotypes with good general combing ability and wide adaptation. For estimating the GCA of a new inbred line, Matzinger

(1953) found that testers with increasing levels of genetic heterogeneity performed better. According to his results, as the heterogeneity of a tester increases, the component of variance due to a Line x Tester interaction decreases. Thus in theory, an F_1 tester provides a more accurate ranking of new lines for GCA than an inbred line tester. However, he also found that using inbred line testers produced larger testcross variances, indicating they provide a more efficient discrimination between lines. Zambezi et al. (1986) found that using unrelated inbred lines as testers in maize produced rankings for GCA that were very similar to the rankings provided by using the population from which the inbred testers were derived from as the tester. This provides evidence that unrelated inbred line testers and population testers are similar in their ability to correctly rank lines. Hallauer and Lopez-Perez (1979) compared the testcross variances produced by several different types of maize testers to identify those providing the largest variances. For related testers, the variance produced by inbred line testers was highly dependent on the performance of the tester for the trait being measured. As the performance of the inbred line testers improved, the testcross variances they produced diminished. However, when an unrelated elite inbred line tester was used, large variances were produced in early generation testing. While an unrelated elite tester will be fixed for many favorable alleles, the allele frequencies are sufficiently different to allow more of the genetic effect of new lines to be observed, thereby increasing the testcross variance.

Shebeski (1966) evaluated the performance of unrelated F_1 testers for estimating the GCA of five wheat varieties versus the average performance of the varieties in hybrid combinations with each other. The F_1 tester and the average hybrid performance of the varieties were identical in ranking the varieties for GCA. In addition, both methods were

consistent in their level of statistical significance for the differences between varieties. He proposed that unrelated F_1 testers could be used to accurately and efficiently evaluate wheat varieties for combining ability.

Gebrekidan and Rasmussen (1970) compared the combining ability estimates of barley inbred lines when evaluated with F_1 testers, inbred line testers, and *per se* evaluation. They were unable to identify clear differences between the three evaluation methods, and concluded that *per se* evaluation should be the preferred method because of convenience, rather than superiority in estimating barley inbred line combining abilities. In addition, barley is grown as a cultivar, and hybrid performance is of little importance.

Sorghum Testcrossing

Karper and Quinby (1937) produced F_1 hybrids from several pure-line sorghum varieties, and identified a high degree of heterosis for several traits in the hybrids. All the varieties exhibited heterosis in hybrid combination, but they differed substantially in the amount of heterosis. Grain yield was the trait that demonstrated the highest level of heterosis, and the hybrids consistently matured earlier than the parental pure-line cultivars. Similar reports of sorghum heterosis were also identified by Bartel (1949) and Stephens and Quinby (1952), who also stated that the efficient capture of sorghum heterosis awaited a mechanism for efficiently producing hybrid sorghum seed.

Stephens and Holland (1954) identified a cytoplasmic male sterility system that allowed for the efficient production of hybrid sorghum seed. This CMS system is based on the interaction of male-sterile cytoplasm from the sorghum “milo” race and nuclear male fertility restoration genes in the “kafir” race.

With the introduction of CMS in sorghum, hybrid sorghum was rapidly adopted by farmers and necessitated the development and evaluation of inbred lines for use as hybrid parents. Kambal and Webster (1965) crossed 10 male-sterile sorghum lines to 19 sorghum fertility restoring lines to compare estimates of components of variance for general and specific combining ability for five traits. They concluded that while both types of combing ability are important, general combing ability was more important in determining hybrid sorghum yield.

Ross (1969) compared the performance of related single-cross and three-way sorghum hybrids over a period of four years. Over the test period, the two types of hybrids had similar averages for yield, but the three-way hybrids were more variable for other traits such as height and flowering. The three-way hybrids had smaller variance components and mean squares, and the author proposed that using F_1 testers could be used to provide more reliable estimates of GCA.

Ross and Kofoid (1978) crossed 42 experimental R-lines to a male-sterile F_1 tester and to the parents of the F_1 tester alone to compare their performance as testers. The three types of hybrids differed significantly for all traits except yield. The three testers performed similarly in identifying superior R-lines, and a single "best" tester could not be identified. However, they did suggest that the F_1 tester may provide higher stability in the testcrosses and that they yield higher quantities of testcross seed.

In the initial phases of testcrossing, male-sterile versions of experimental B-lines are not available. To test male-fertile B-lines, Schertz and Johnson (1984) proposed crossing experimental B-lines to standardized A-lines to yield sterile F_1 hybrids, which could then be crossed to R-line testers. In this manner, experimental B-lines could be

evaluated with an R-line tester before producing a sterile A-line counterpart. Their results indicated that this method was consistent in identifying B-lines with superior combing ability, and would be a valid testcrossing system for evaluating B-lines.

Lee et al. (1992) proposed using male-sterile versions of R-line testers to evaluate B-lines prior to sterilization. They used R-lines that restore male fertility in A1 cytoplasm, but that had been sterilized in A3 cytoplasm. They compared their ability as testers for evaluating B-lines to reciprocal testcrosses made with the A-line counterpart of the B-lines and the standard fertility restoring version of the R-lines. They found few differences between the A1 hybrids and their respective reciprocal A3 hybrids. Indicating that B-line testcrosses made with male-sterile A3 R-lines can accurately predict the relative performance of the respective reciprocal cross of A/B lines to A1 fertility-restoring R-lines.

Gilbert et al. (1996) compared the performance of R-line testers sterilized in four alternate cytoplasms to their standard A1 male-fertility restoring versions. They found that hybrids derived from alternate cytoplasm R-lines generally did not yield as well as their standard cytoplasm counterparts, but the relative performance of the hybrids to each other were similar. They concluded that R-line testers sterilized in alternate cytoplasms can be used to accurately assess the relative performance of experimental B-lines prior to their sterilization.

Most sorghum breeding programs begin testcrossing in the F_4 generation, using at least two testcross hybrids in a limited number of environments. Experimental R-lines are crossed to standardized A-line testers to evaluate their GCA and male-fertility restoration capabilities. Some sorghum breeding programs test experimental B-lines by

crossing them to A3 cytoplasm versions of commonly used R-lines to evaluate their GCA before initiating sterilization. Experimental lines that perform well in initial testcrosses are advanced and the number of testers and environments used are expanded. Also at this point, the sterilization process for B-lines selected for advancement is initiated, with further testcrossing occurring during sterilization to further eliminate poor performing lines (Rooney 2004).

Conclusions

A valid tester should correctly identify the relative performance of new material and efficiently discriminate among experimental lines. F_1 testers have been shown to be valid, but not necessarily superior testers, in several crops. This includes sorghum, where F_1 testers have correctly identified the relative performance of experimental B and R lines. In addition, the increased heterogeneity of F_1 testers may provide increased stability in the testcrosses and more reliable estimates of GCA. However, the increased heterogeneity of the F_1 testers may reduce the variance among the experimental lines and therefore, be less efficient in discriminating among those lines. Also, R-line testers that have been male sterilized in A3 cytoplasm have been shown to correctly predict the relative performance of experimental B-lines when their A-line counterpart is crossed to a standard R-line that restores fertility in A1 cytoplasm.

CHAPTER III

R-LINE TEST

Introduction

In developing new sorghum inbred lines, it is essential to evaluate those inbred lines for performance in hybrid combination. This is done using a testcross, in which experimental lines are crossed to a common line, producing hybrids which are evaluated for performance. New inbred lines will either be advanced or discarded, based on the performance of the hybrid produced with the tester.

To ensure the most accurate evaluation of experimental lines, it is essential that an appropriate tester be selected. Tester options range from broad genetic base populations to elite inbred lines; most sorghum breeding programs will have their own criteria for tester selection, based on their needs and goals. However, any tester selected must meet certain requirements. An effective tester should correctly rank inbred lines for performance in hybrid combination, and it should maximize the variance between testcross progeny for efficient discrimination (Rawlings and Thompson 1962).

Most sorghum breeding programs utilize elite inbred line testers to evaluate new lines. Elite inbred testers provide breeders with valuable information regarding the performance of new lines with specific genetic backgrounds. However, utilizing an elite inbred tester only allows for the evaluation of new lines with alleles from a single genetic background. An alternative would be to use two inbreds, but this increases the numbers and is usually not economically or logically feasible. The use of a single inbred line tester increases the risk of incorrectly discarding material that may combine well with germplasm adapted to a different region or genetic background than that of the tester.

Most sorghum hybrids are adapted to either temperate or tropical environments. Few, if any, are well adapted to both types of environments. Because the state of Texas produces sorghum in both temperate and tropical environments, it is possible to evaluate experimental lines in both environments in a single state.

A potential solution for evaluating experimental lines against multiple groups of germplasm, would be to use a single-cross (F_1) tester. An F_1 tester derived from the cross of an elite temperately adapted line to an elite subtropical-tropical line could reduce the probabilities of incorrectly discarding material. The purpose of the R-line test is to determine whether the F_1 hybrid of two commonly used female inbred lines, ATx623 and ATx2752, is a valid tester for evaluating sorghum R-lines versus its inbred line parents alone. ATx623/BTx2752 combines subtropical adaptation from ATx623 and temperate adaptation from ATx2752. And thus, ATx623/BTx2752 may represent a viable tester option for estimating the GCA of sorghum R-lines against two adaptation germplasm pools and genetic backgrounds, in a single testcross.

Evaluation of hybrid testcrosses made with ATx623, ATx2752, and ATx623/BTx2752 allows for the detection of differences between them in their performance as testers. By comparing parameters that quantify their accuracy in ranking the R-lines, conclusions can be drawn regarding the validity of these testers for obtaining correct GCA estimates of new R-lines. And parameters that measure tester efficiency can be used to establish differences between the testers in maximizing the variance between experimental lines, for the efficient discrimination between those lines.

The correlation between the ranking of experimental lines by a tester and the ranks produced by the average performance of the same material in multiple hybrid

combinations is an indicator of a tester's accuracy in ranking experimental lines (Castellanos *et al.* 1998). A hypothesis of this test is that all three testers will rank the experimental lines with sufficient accuracy to produce significant correlations between their rank assignments and the ranks produced by the average performance of all testcross hybrids. However, differences in the size of the tester rank correlations should be observed based on the environment in which the test is grown. Using an unadapted tester may mask and confound performance and ranking of the material. Therefore, the temperately adapted tester, ATx2752, should produce the strongest rank correlation of the three testers in temperate environments, while ATx623 will have the highest correlation in tropical environments. Because the F_1 tester ATx623/BTx2752 contains alleles from both adaptation germplasm pools, the size of its rank correlations should be between those of the two inbred line testers in both adaptation regions.

When considering the data combined across locations, the differences in rank correlations between the testers should be reduced. The combined data will contain information from both adaptation regions, thereby reducing the effect of the environment in the tester rank assignments. In this manner, the advantage of one tester in its adapted environment will be offset by the advantage of the opposite tester in its respective environment, thereby reducing differences in rank correlations across environments.

Line x Tester analyses provide an additional parameter for evaluating the accuracy of a group of testers. A Line x Tester analysis is the interaction between the experimental lines and the testers for the dependant variable in a statistical model. A significant line x tester interaction provides evidence that the ranking of experimental lines differs depending on the tester used. In such cases, testcross evaluations made with

one tester will not be comparable to those made with one of the other testers. In this test, the hypothesis is that the three testers will not differ sufficiently in their assessment of the experimental R-lines to produce significant line x tester interactions. Such results would provide evidence that the testers have similar accuracies in ranking the experimental lines, and that testcross evaluations derived from the different testers, including the F_1, are comparable. But because ATx623 and ATx2752 represent different adaptation germplasm pools, a significant line x tester interaction is plausible. If this occurs, the previously mentioned tester rank correlations can provide insight regarding the differences amongst the testers in assigning ranks.

The Schumann-Bradley test provides a statistical method to elucidate differences in the discriminatory efficiencies of the testers. The Schumann-Bradley test compares the F-ratios of similar experiments to determine whether the experiments have significantly different efficiencies. This test has previously been used to compare tester efficiencies (Sharma et al., 1967). Within the R-line test, significant differences in efficiencies should be observed between the testers when evaluated with the Schumann-Bradley test. In particular, ATx623 and ATx2752 will likely demonstrate significantly different efficiencies based on environmental adaptation. In these environments, an unadapted tester will mask some of the differences between the experimental lines, and reduce the testcross variance and efficiency that otherwise would be seen with an adapted tester. Because the F_1 tester combines alleles from both adaptation types, this reduction in efficiency should not be as pronounced, and the F_1 tester is not expected to differ significantly from the adapted inbred line tester in efficiency, as measured by the Schumann-Bradley test.

Another method for evaluating tester efficiencies is to compare the testcross variances produced by each tester. A larger testcross variance implies an increase in the variance between experimental lines and therefore, greater efficiency in the discrimination among those lines. Testcross variances are maximized when the genetic potential of each experimental line is allowed to be fully expressed with minimal interference from the tester. For the R-line test, ATx623 is expected to produce the largest testcross variances in subtropical-tropical environments and ATx2752 is expected to produce the largest testcross variances in temperate environments. With the combination of alleles from two adaptation regions, ATx623/BTx2752 is expected to produce a testcross variance between that of the two inbred line testers.

Although the F_1 tester is not expected to produce the largest testcross variances, it should provide a sufficiently large testcross variance to allow for the detection of significant differences between the experimental lines, if significant differences are to be had. If the tester with the largest testcross variance produces a significant entry effect in a statistical model, the F_1 tester should produce a smaller, but sufficiently large, testcross variance to also detect a significant entry effect. Confirmation of this would lend validity to the use of F_1 testers for the efficient discrimination of experimental lines.

The R-line test objectives are to compare the accuracies of ATx623, ATx2752, and ATx623/BTx2752 in evaluating the experimental lines using rank correlations and line x tester analyses, as well as comparing their discriminatory efficiencies with the Schumann-Bradley test and comparisons of testcross variances. The results of these analyses will be interpreted to assess the validity of ATx623/BTx2752 as a tester for sorghum R-lines.

Materials and Methods

Testcross Development

Thirty-three experimental sorghum R-lines were randomly selected from a set of F_5 breeding lines in the Texas A & M sorghum breeding program. Two commonly used female parents, ATx623 and ATx2752, along with their F_1 hybrid, ATx623/BTx2752, were used as the testers to evaluate the experimental R-lines. ATx623 represents tropically-subtropically adapted germplasm while ATx2752 represents temperately adapted germplasm. During the summer of 2005, each experimental R-line was hybridized to ATx2752, ATx623, and their F_1 hybrid, ATx623/BTx2752 to yield a total of 99 hybrid testcross entries. Each experimental R-line was represented by three entries, each derived from a cross to one of the three testers.

Experimental Design

The 99 entries were arranged using the "sets in rep" design with 11 sets, and 9 entries/set and three replications. The three entries representing each experimental lines were assigned into the same set. During the summer of 2006, the R-line test was grown in three locations: College Station, TX; Weslaco, TX, and Halfway, TX. In each environment, standard agronomic management practices were followed for fertilization. Supplemental irrigation was used at all three locations. In each environment, a plot was defined as two rows 6 meters in length and spaced 76 cm apart. Each row within a plot was planted with four grams of seed that had been pre-treated with fungicides.

Data Collection

Standard agronomic notes were taken for all the plots at each location. These included height and days to mid-anthesis. Plant height was measured in inches from the base of the plant to the tip of the panicle as an average for the plot. Days to mid-anthesis was recorded as the Julian date that 50% of the plot reached 50% anthesis. Data for days to mid-anthesis was not collected for Halfway, TX.

Two heads per plot were covered with bags before flowering, to confirm the male fertility of the entries. Prior to combine harvest, two randomly selected panicles were harvested from each single cross hybrid plot, while four panicles were harvested from the three-way hybrids (due to segregation in these hybrids). These samples were used to obtain average panicle lengths and 100 seed sizes. Panicle length was measured in centimeters from the from the bottom pedicel to the tip of the panicle as an average for the plot. One hundred seed size is measured as the weight of 100 seeds in grams as an average for the plot.

Grain yield was measured using a modified John Deere 3300 plot combine, with plot grain weight, grain moisture, and test weight collected by the HarvestMaster HM-1000 weigh system onboard the combine. Total plot yield (kg/ha) was obtained by adjusting the plot weight with the following formula:

$$[(((100 - \%\ \text{Moisture})/87) * \text{Plot Weight}) * 385] * 1.115 = \text{Total Plot Yield (kg/ha)}$$

The plot yield obtained with this formula was further adjusted by multiplying it by the stand rating for each plot to account for missing plants. The adjusted yield data for

each location was organized into datasets using Microsoft Excel, and subsequently imported into the SAS 9.1 statistical analysis software as text files for analysis.

Statistical Analysis

The initial analyses of the R-line test were performed on an individual location basis using the PROC GLM procedure available in SAS 9.1 using the following statistical model:

$$\text{Yield} = \mu + \text{Tester} + \text{Male} + \text{Rep} + \text{Set(Rep)} + \text{Tester*Male} + \text{Error}$$

The Male effect represents the proportion of the variance attributable to the differences among the experimental R-lines. All the effects were analyzed using the SAS 9.1 defaults, and were considered fixed effects. The size and significance value of the Tester*Male interaction term represents the Line x Tester interaction, and was used in evaluating the accuracy of the testers in each location.

Rank correlations were based on average performances in each environment. The ranks based on average performances were then compared to the rank assignments generated by each tester to derive rank correlations using the CORR function in Microsoft Excel spreadsheets. The ranks produced by each individual tester were also compared to each other using the Microsoft Excel CORR function to determine whether any of the testers, particularly the F_1 tester, assigned ranks more like one tester than to another.

Data for each location of the R-line test was analyzed by tester with the following statistical model:

$$\text{Yield} = \mu + \text{Male} + \text{Rep} + \text{Set(Rep)} + \text{Error}$$

In this model, an F-ratio for the effect due to the variance between the R-lines (Male effect) was obtained for each tester. These F-ratios were then compared against each other via the Schumann-Bradley test to determine whether they were statistically different from each other (Sharma et al. 1967).

The Schumann-Bradley test compares the efficiencies of similar experiments by deriving a *w* statistic using the F-ratios of two experiments having the same degrees of freedom and number of replications (Bradley and Schumann 1957).

$$w = \text{Exp. 1 F-ratio/Exp. 2 F-ratio}$$

This *w* statistic is compared against tabulated significance values using transformed degrees of freedom and parameters derived from the two experiments being compared. The first step in conducting a Schumann-Bradley test is to calculate two parameters, **a** and **b**. The **a** parameter is equal to half the degrees of freedom of the effect being tested by the F-ratios of the experiments. In the case of the R-line test, this is equal to half the degrees of freedom for the Male effect. The **b** parameter is equal to half the error degrees of freedom shared by both experiments.

$$\mathbf{a} = \tfrac{1}{2}(\text{Male effect d.f.})$$

$$\mathbf{b} = \tfrac{1}{2}(\text{Error d.f.})$$

The **a** and **b** parameters are used to calculate a third parameter, λ_i, for each experiment using the following formula:

$$\lambda_i = \mathbf{a}(\text{F-ratio}_i - 1)$$

For the R-line test, a λ_i was calculated for each tester at each location. When comparing two experiments, the λ_is' for both experiments are added together to produce a total λ.

$$\lambda = \lambda_1 + \lambda_2$$

The total λ and the individual λ_is' form the basis for the hypothesis tests of the Schumann-Bradley test.

$$H_0: \lambda_1 = \lambda_2 = \lambda$$

$$H_A: \lambda_1 > \lambda_2$$

In more basic terms, the Schumann-Bradley test is a one-sided test of significance with the null hypothesis that both experiments have equal efficiencies and with the alternative hypothesis that Experiment 1 has a greater efficiency than Experiment 2. Experiment 1 represents the F-ratio used as the numerator in calculating the *w* statistic.

Upon the calculation of a total λ, it can then be used to calculate the a' parameter that is used for interpreting the tabulated w_0 values in the Schumann-Bradley table. The a' is calculated in the following manner:

$$a' = (\mathbf{a} + \lambda)^2/[\mathbf{a} + 2(\lambda)]$$

In conjunction with **b**, a' is used to identify the cutoff w_0 in the Schumann-Bradley table that is used for determining the significance of the observed *w*. If the observed *w* exceeds w_0, the null hypothesis is rejected and the alternative hypothesis that the experiment providing the numerator F-ratio in calculating *w* is more efficient than the experiment providing the denominator F-ratio is accepted. In the R-line test, the F-ratios for the Male effect produced by each tester were compared against each other per location using the Schumann-Bradley test to identify significant differences in their efficiencies.

Testcross variances were estimated using the same model. The testcross variances were obtained by subtracting the Error mean square from the Male mean square produced by each tester within a location and then dividing that number by the number of replications and sets. Standard errors for the testcross variances were derived by calculating a 95% confidence interval around the variance estimate (Bernardo, 2002) and dividing it by two. After the derivation of the testcross variances, they were compared against each other to make inferences regarding the discriminatory efficiencies of the testers within a location.

Upon completing the analyses on a per location basis, the yield data for the locations was combined to analyze the testers across locations. This was deemed

appropriate using the HOVTEST command within the PROC GLM procedure of SAS 9.1 with the combined yield data. The HOVTEST command performs a Levene's test for the homogeneity of variances, with the null hypothesis that the variances at all locations are equal. With the total combined yield data, a p-value of 0.05 was obtained, theoretically sufficient to reject the null hypothesis and indicates that the variances are not equal and that the data should not be combined. But to further investigate the appropriateness of combining the data, the combined yield data was segregated by tester to determine whether the variances for each tester were equal across locations. This was also done using the HOVTEST command within the SAS PROC GLM procedure, but on a per tester basis. The results in Table 1 indicate that the Levene's test failed to reject the null hypothesis that the variances for each tester are equal across locations. In addition, the data was normally distributed with no obvious outliers. With these results, analyses evaluating the testers with the data combined across the locations were performed.

Table 1. Probability values of the Levene's test for the homogeneity of location variances for each tester and the combined tester data in the R-line test.

Analysis	F-value	Pr < F
Combined Tester Data	3.07	0.05
ATx2752	1.96	0.14
ATx623	1.31	0.27
ATx623/BTx2752	1.17	0.31

The combined data was analyzed in a manner similar to that used on a per location basis. Using the PROC GLM procedure of SAS 9.1, the following model was used:

$$\text{Yield} = \mu + \text{Location} + \text{Tester} + \text{Male} + \text{Rep} + \text{Set(Rep)} + \text{Tester*Male} + \text{Location*Male} + \text{Error}$$

As before, all the effects were considered fixed per the SAS 9.1 defaults, and the Male effect represents the variance between the R-lines. The Line x Tester analysis and rank correlations were completed using the same methodology as previously described. As in the individual environment analysis, the combined data was analyzed by tester using the following model:

$$\text{Yield} = \mu + \text{Location} + \text{Male} + \text{Rep} + \text{Set(Rep)} + \text{Location*Male} + \text{Error}$$

The reduced model for the combined data produces F-ratios for testing the effect due to the variance between the R-lines (Male effect) for each tester. These F-ratios were compared against each other via the Schumann-Bradley test (as previously described) to determine whether they were statistically different from each other and elucidate differences in efficiencies among the testers. If the Schumann-Bradley test declared two F-ratios statistically different, the tester providing the F-ratio used as the numerator in calculating the *w* statistic was considered as being more efficient in discriminating amongst the R-lines across the locations.

The reduced combined model was used to estimate testcross variances across locations using a fixed effects model. The testcross variances for each tester were obtained by subtracting the Error mean square from the Male mean square produced by each tester across locations and then dividing that number by the number of replications, locations, and sets. Estimates of testcross variances were compared to make inferences regarding the discriminatory efficiencies of the testers across locations. Testers with larger testcross variances were considered superior in maximizing the variance between the lines and therefore, more efficient in discriminating among the lines.

Results and Discussion

Individual and Combined Analysis

Significant variation was detected for most traits in the combined analysis and in individual environments (Tables 2-5). While there is merit in pursuing the detailed analysis of agronomically important traits such as plant height and days to mid-anthesis, for the purposes of this study and subsequent breeding efforts, emphasis and further analysis will be primarily focused on grain yield. Yield data was relatively consistent as indicated by C.V., R-square values, and standard deviations (Tables 2-5).

Table 2. Yield, Height, and Days to Mid-Anthesis analysis of variance results for hybrid testcrosses of sorghum R-lines to ATx2752, ATx623, and ATx623/BTx2752 across locations in 2006.

	Yield		Height		Days/Mid-Anth.	
Source of Variance	df	Mean Square	df	Mean Square	df	Mean Square
Location	2	470574212.1**	2	38501.86**	1	1324.9**
Tester	2	2430105.8	2	8997.04**	2	66.6**
Male	23	3358619.9**	23	1137.37**	22	33.6**
Rep	2	3512002.4	2	52.1	2	97**
Set(Rep)	21	2327376.8*	21	379.48**	20	3.6*
Tester*Male	64	1525494.3	64	77.4	64	4.4**
Location*Male	64	2888174.6**	64	173.19**	32	4.5**
Error	681	1400883.0	679	70505.3	416	1.9
	C.V. = 22.1 R^2 = 0.61		C.V. = 8.4 R^2 = 0.67		C.V. = 2.1 R^2 = 0.81	

*,** Significant at $p < .05$ and .01, respectively

† Days to Mid-Anthesis data was not collected for Halfway, TX

Table 3. Yield, Height, and Days to Mid-Anthesis analysis of variance results for hybrid testcrosses of sorghum R-lines to ATx2752, ATx623, and ATx623/BTx2752 in College Station, TX in 2006.

	Yield		Height		Days/Mid-Anth.	
Source of Variance	df	Mean Square	df	Mean Square	df	Mean Square
Tester	2	1747925.3	2	3550.4**	2	4.9**
Male	22	6013094.9**	22	488.6**	22	16.1**
Rep	2	29763739.9**	2	3322**	2	58.5**
Set(Rep)	20	2370809.9**	20	187.2**	20	1.8**
Tester*Male	64	1268511.6	64	81.5	64	2.2**
Error	176	1072416.3	176	75.6	176	0.79
	C.V. = 15.4 R^2 = 0.67		C.V. = 7.4 R^2 = 0.73		C.V. = 1.3 R^2 = 0.85	

*,** Significant at $p < .05$ and .01, respectively

Table 4. Yield, Height, and Days to Mid-Anthesis analysis of variance results for hybrid testcrosses of sorghum R-lines to ATx2752, ATx623, and ATx623/BTx2752 in Weslaco, TX in 2006.

	Yield		Height		Days/Mid-Anth.	
Source of Variance	df	Mean Square	df	Mean Square	df	Mean Square
Tester	2	8151.6	2	2186.3**	2	97.4**
Male	22	1334849.2	22	343.7**	22	18.6**
Rep	2	13657386.2**	2	19.9	2	50.2**
Set(Rep)	20	1033953.7	20	58.8**	20	5.1**
Tester*Male	61	1263390.3	61	36.3**	61	4.3**
Error	159	1290951.8	158	19.2		2.3
	C.V. = 21.7 R^2 = 0.45		C.V. = 3.3 R^2 = 0.86		C.V. = 2.2 R^2 = 0.81	

*,** Significant at $p < .05$ and .01, respectively

Table 5. Yield and Height analysis of variance results for hybrid testcrosses of sorghum R-lines to ATx2752, ATx623, and ATx623/BTx2752 in Halfway, TX in 2006.

	Yield		Height	
Source of Variance	df	Mean Square	df	Mean Square
Tester	2	3391601.1	2	3101.5**
Male	23	1744248.1	23	397.6**
Rep	2	15062728.4**	2	3317.7**
Set(Rep)	20	2817283.9**	20	422.4**
Tester*Male	64	1458500.0	64	127.9
Error	176	1238837.0	175	140.6
	C.V. = 26.9 R^2 = 0.57		C.V. = 10.7 R^2 = 0.63	

*,** Significant at $p < .05$ and .01, respectively

Table 6. Yield, Height, and Days to Mid-Anthesis analysis of variance results by tester (ATx623, ATx2752, and ATx623/BTx2752) for hybrid testcrosses of R-lines in College Station, TX in 2006.

		Dependent Variables		
ATx623	df	Yield	Height	Days/Anth.
Male	32	313246.3*	254.6**	6.6**
Rep	2	10458042.1**	897.9**	12.6**
Set(Rep)	20	1303628.6	79.8	1.2
Error	44	1462224.0	84.5	1.2
C.V.		17.7	7.5	1.6
ATx2752	df	Yield	Height	Days/Anth.
Male	32	1967710.6	131.1	8.0**
Rep	2	9041825.1**	1208.4**	24.0**
Set(Rep)	20	1031057.8	61.9	0.82
Error	44	1193826.6	80.8	0.65
C.V.		16.2	8.0	1.23
ATx623/BTx2752	df	Yield	Height	Days/Anth.
Male	32	2957529.2**	278.9**	4.4**
Rep	2	10537894.0**	1367.8**	23.3**
Set(Rep)	20	1431232.2	136.9	1.2*
Error	44	987018.6	89.4	0.56
C.V.		15.1	7.8	1.1

*,** Significant at $p < .05$ and .01, respectively

Table 7. Yield and Height analysis of variance results by tester (ATx623, ATx2752, and ATx623/BTx2752) for hybrid testcrosses of R-lines in Halfway, TX in 2006.

		Dependent Variables	
ATx623	df	Yield	Height
Male	32	946057.4	270.7
Rep	2	5694332.5*	968.7*
Set(Rep)	20	1977907.5	198.3
Error	44	1375081.7	196.0
C.V.		28.8	12.2
ATx2752	df	Yield	Height
Male	32	2142104.1*	172.1
Rep	2	4832399.4*	985.8**
Set(Rep)	20	3387267.3**	175.0
Error	44	1034129.1	124.4
C.V.		23.4	10.7
ATx623/BTx2752	df	Yield	Height
Male	32	1133302.1*	208.8
Rep	2	5794933.6**	1423.9**
Set(Rep)	20	1542810.1**	238.7
Error	44	625485.3	155.4
C.V.		19.8	11.0

*,** Significant at $p < .05$ and .01, respectively

Table 8. Yield, Height, and Days to Mid-Anthesis analysis of variance results by tester (ATx623, ATx2752, and ATx623/BTx2752) for hybrid testcrosses of R-lines in Weslaco, TX in 2006.

		Dependent Variables		
ATx623	df	Yield	Height	Days/Anth.
Male	32	1969954.3	160.7**	13.6**
Rep	2	10238610.6**	22.3	17.1**
Set(Rep)	20	979665.2	25.6*	4.1*
Error	44	1392895.8	13.3	1.8
C.V.		22.5	2.6	1.9
ATx2752	df	Yield	Height	Days/Anth.
Male	32	1283005.1	174.0**	6.1*
Rep	2	227597.3	38.7	27.2**
Set(Rep)	20	1113470.5	42.3*	4.3
Error	44	1057078.9	22.2	2.8
C.V.		19.8	3.6	2.5
ATx623/BTx2752	df	Yield	Height	Days/Anth.
Male	32	908098.0	110.1**	10.0**
Rep	2	6271623.2*	49.7	13.1**
Set(Rep)	20	1126661.1	611.1	2.3
Error	44	1522292.8	20.6	1.8
C.V.		23.5	3.3	1.9

*,** Significant at $p < .05$ and .01, respectively

In the BY tester analysis, significant variation was detected for most traits in the combined analysis and in most environments (Tables 6-8). While there is merit in pursuing the detailed analysis of agronomically important traits such as plant height and days to mid-anthesis, for the purposes of this study and subsequent breeding efforts, emphasis and further analysis will be primarily focused on grain yield. Yield data was relatively consistent as indicated by their C.V. values (Tables 6-8).

Line x Tester Analysis

In all environments, a Tester*Male interaction was not detected (Tables 2-5). The non-significant Tester*Male interactions indicate that the testers were consistent in ranking the experimental lines. While there were some rank shifts, rank trends were similar for all testers at each location (Tables 9-12). They were consistently reliable in identifying poor performing lines, which is of particular importance.

Consistent with the results at the individual locations, the Tester*Male interaction for the data combined across locations was also non-significant (Table 5). These results provide evidence that in evaluating the R-lines for performance across locations, the testers assigned ranks to the experimental lines in a consistent manner (Table 8).

Table 9. Ranks, based on the grain yield of 33 experimental R-lines in testcross combination with ATx2752, ATx623, and ATx623/BTx2752 when evaluated in College Station, TX in 2006.

MALE	ATx2752	ATx623	ATx623/BTx2752	Average
R04184	3	1	2	1
R04104	5	4	1	2
RTx437	4	6	5	3
R04164	7	3	7	4
R04146	1	14	9	5
R04183	2	8	17	6
R04232	19	2	11	7
R04196	25	5	4	8
R04185	12	15	3	9
R04181	6	19	6	10
R04214	22	7	12	11
R04135	9	9	22	12
R04163	13	18	15	13
R04160	30	10	10	14
R04132	15	16	20	15
R04190	26	11	16	16
R04180	11	28	8	17
R04050	29	12	13	18
R04231	20	17	19	19
R04153	16	23	18	20
R04156	18	13	25	21
R04143	10	25	23	22
R04175	17	20	26	23
R04179	21	24	24	24
R04233	14	22	30	25
R04131	23	27	21	26
RTx436	24	21	28	27
R04081	32	29	14	28
R04165	28	26	27	29
R04234	27	30	31	30
RTx2783	8	32	32	31
R04083	31	33	29	32
R04047	33	31	33	33

Table 10. Ranks, based on the grain yield of 33 experimental R-lines in testcross combination with ATx2752, ATx623, and ATx623/BTx2752 when evaluated in Weslaco, TX in 2006.

MALE	ATx2752	ATx623	ATx623/BTx2752	Average
R04146	1	1	20	1
R04153	11	3	10	2
R04190	9	6	14	3
R04104	2	23	12	4
R04232	19	2	5	5
R04165	.	10	11	6
R04050	6	18	7	7
R04156	21	5	9	8
RTx436	3	30	6	9
R04131	10	13	13	10
R04233	4	11	24	11
R04181	15	4	23	12
R04231	7	14	17	13
R04135	12	15	8	14
R04184	23	24	1	15
R04083	5	20	16	16
R04164	.	9	28	17
RTx437	14	.	21	18
R04143	27	19	3	19
R04196	29	7	19	20
R04160	24	28	2	21
RTx2783	13	16	33	22
R04047	31	12	4	23
R04234	22	17	26	24
R04179	8	31	25	25
R04180	16	25	29	26
R04132	30	8	22	27
R04163	28	21	18	28
R04175	18	26	30	29
R04081	20	27	27	30
R04183	26	22	31	31
R04185	17	29	32	32
R04214	25	32	15	33

Table 11. Ranks, based on the grain yield of 33 experimental R-lines in testcross combination with ATx2752, ATx623, and ATx623/BTx2752 when evaluated in Halfway, TX in 2006.

MALE	ATx2752	ATx623	ATx623/BTx2752	Average
R04196	2	25	3	1
R04156	3	7	17	2
R04190	11	13	1	3
R04184	9	15	4	4
R04153	7	10	10	5
RTx436	19	3	13	6
R04232	5	4	27	7
R04160	20	1	22	8
R04185	8	2	26	9
R04143	4	21	14	10
R04135	10	11	16	11
R04146	1	31	15	12
RTx437	6	23	11	13
R04231	17	12	9	14
R04165	23	14	5	15
R04181	16	9	21	16
R04131	29	8	6	17
RTx2783	24	17	7	18
R04234	28	5	18	19
R04163	22	26	2	20
R04180	13	24	12	21
R04175	18	22	19	22
R04179	14	32	8	23
R04164	21	19	24	24
R04132	15	28	23	25
R04083	12	27	28	26
R04233	31	6	25	27
R04183	27	16	29	28
R04081	25	20	30	29
R04214	30	29	20	30
R04104	26	33	31	31
R04047	32	18	32	32
R04050	33	30	33	33

Table 12. Ranks, based on the grain yield of 33 experimental R-lines in testcross combination with ATx2752, ATx623, and ATx623/BTx2752 when evaluated across locations in 2006.

MALE	Average	ATx2752	ATx623	ATx623/BTx2752
R04184	1	4	3	1
R04146	2	1	13	8
R04232	3	14	1	13
RTx437	4	2	10	6
R04196	5	18	5	2
R04190	6	17	6	3
R04153	7	7	9	9
R04156	8	10	4	20
R04164	9	9	2	21
R04135	10	5	11	18
R04181	11	13	8	14
R04104	12	3	18	17
R04231	13	15	16	11
R04160	14	27	7	5
RTx436	15	6	19	19
R04185	16	8	15	24
R04143	17	12	24	10
R04131	18	24	21	7
R04163	19	25	25	4
R04180	20	11	29	16
R04165	21	28	20	15
R04183	22	20	14	29
R04233	23	23	12	28
R04132	24	26	17	23
R04179	25	16	32	22
R04175	26	22	27	26
R04214	27	29	31	12
R04234	28	31	22	31
RTx2783	29	19	30	32
R04081	30	30	28	25
R04083	31	21	33	27
R04050	32	32	23	30
R04047	33	33	26	33

Correlations

Rank correlations in each location and combined across locations were significant (Table 13). At the College Station location, ATx623 produced the largest rank correlation to the average performance ranks, and based on this method, it would be considered the most accurate tester for evaluating R-lines in College Station (Table 13). However, the F_1 tester produced a rank correlation almost identical to that of ATx623. While ATx2752 produced a sizeable rank correlation, its correlation was less than half that of the other two testers. The large rank correlations produced by both ATx623 and the F_1 tester indicate that they both provide accurate rankings of the experimental lines for GCA. In addition, the high degree of similarity between the size of the ATx623 and F_1 tester rank correlations to the average performance ranks demonstrate that both testers could be used for ranking R-lines in College Station to produce very similar results. Further evidence is provided by the large correlation, ($r = 0.70$), between the ranks produced by both testers (Table 13).

In Weslaco, all three testers produced similar rank correlations to the average performance ranks (Table 13). Because ATx2752 represents temperately adapted germplasm, it was not expected to produce a rank correlation of 0.63, virtually identical to that of ATx623 (0.62) in a subtropical-tropical location such as Weslaco. The F_1 tester produced a slightly lower rank correlation (0.52) than both of the inbred line testers. However, all three testers were within a similar range, providing evidence that they all performed similarly in ranking the R-lines for GCA in Weslaco.

In the Halfway location, the rank correlations fit expectations (Table 13). ATx2752 produced the highest rank correlation (0.75) to the average performance

ranking of the R-lines. ATx623 produced a smaller rank correlation (0.51), and the F_1 tester produced a rank correlation lying between the values of the inbred line testers (0.60). While the F_1 tester did not produce the best correlation, it was sufficiently accurate to evaluate R-lines in the Halfway location.

In the combined analysis, all three testers were very similar to each other in their rank correlations with the average performance rankings of the R-lines across locations (Table 13). ATx2752 produced a rank correlation of 0.77, ATx623 produced a rank correlation of 0.83, and the F_1 tester had a rank correlation of 0.7

Table 13. Correlations between the R-line rank assignments of each tester with their average performance in 2006 across locations, in College Station, TX; Halfway, TX; and Weslaco, TX.

Locations Combined	Average	ATx2752	ATx623	ATx623/BTx2752
Average	1.00	0.76**	0.83**	0.74**
ATx2752		1.00	0.46**	0.38*
ATx623			1.00	0.45**
ATx623/BTx2752				1.00
College Station	Average	ATx2752	ATx623	ATx623/BTx2752
Average	1.00	0.65**	0.89**	0.86**
ATx2752		1.00	0.4*	0.42*
ATx623			1.00	0.7**
ATx623/BTx2752				1.00
Halfway	Average	ATx2752	ATx623	ATx623/BTx2752
Average	1.00	0.75**	0.51**	0.60**
ATx2752		1.00	0.03	0.36*
ATx623			1.00	0.03
ATx623/BTx2752				1.00
Weslaco	Average	ATx2752	ATx623	ATx623/BTx2752
Average	1.00	0.62**	0.62**	0.51**
ATx2752		1.00	0.05	-0.04
ATx623			1.00	0.12
ATx623/BTx2752				1.00

*,** Significant at $p < .05$ and .01, respectively

Because combining data from both adaptation regions negates a large portion of the differences between testers due to adaptation, the testers were expected to produce similar rank correlations. Although the F_1 tester produced the smallest rank, the difference between it and the remaining testers is sufficiently small so as to consider it approximately equal in performance and accuracy to the inbred line testers for evaluating the R-lines across locations.

All testers were consistent in the identification of the top performing lines. Of the top seventeen R-lines (averaged across all testers) in College Station, ATx2752 identified twelve, ATx623 identified ten, and ATx623/BTx2752 identified twelve (Table 14). In Weslaco, ATx2752, identified twelve of the lines, ATx623 identified twelve of the lines, and ATx623/BTx2752 identified thirteen of the lines (Table 15). In Halfway, all three of the testers identified thirteen of the lines (Table 16). Across all locations, ATx2752, ATx623, and ATx623/BTx2752 identified 15, 14, and 12 of the top 17 lines (Table 17). These results indicate that despite rank shifts (Tables 9-12) and differences in rank correlations, all three of the testers identified the majority of the top performing R-lines in all the locations and across locations. In addition, there were no large differences between the three testers, implying that they performed with similar degrees of accuracy.

However, some important rank shifts between the two inbred line testers were seen. For example, R04146 was the top performing line with ATx2752 in College Station, but was ranked fourteenth by ATx623 (Table 9). ATx623/BTx2752 ranked the same line ninth. Use of the F1 tester in this case would reduce the probability of discarding R04146 as compared to using ATx623 alone as the tester. Other similar examples can be found in tables 9-12.

Table 14. Inclusion of the seventeen top performing R-lines (based on average performance) for College Station, TX in 2006 in the top seventeen selections of each tester.

	Testers		
Male	ATx2752	ATx623	ATx623/BTx2752
R04184	x	x	x
R04104	x	x	x
RTx437	x	x	x
R04164	x	x	x
R04146	x	x	x
R04183	x	x	x
R04232		x	x
R04196		x	x
R04185	x	x	x
R04181	x		x
R04214		x	x
R04135	x	x	
R04163	x		x
R04160		x	x
R04132	x	x	
R04190		x	x
R04180	x		x
TOTAL	12	14	15

Table 15. Inclusion of the seventeen top performing R-lines (based on average performance) for Weslaco, TX in 2006 in the top seventeen selections of each tester.

	Testers		
Male	ATx2752	ATx623	ATx623/BTx2752
R04146	x	x	
R04153	x	x	x
R04190	x	x	x
R04104	x		x
R04232		x	x
R04165	missing	x	x
R04050	x		x
R04156		x	x
RTx436	x		x
R04131	x	x	x
R04233	x	x	
R04181	x	x	
R04231	x	x	x
R04135	x	x	x
R04184			x
R04083	x		x
R04164	missing	x	
TOTAL	12	12	13

Table 16. Inclusion of the seventeen top performing R-lines (based on average performance) for Halfway, TX in 2006 in the top seventeen selections of each tester.

	Testers		
Male	ATx2752	ATx623	ATx623/BTx2752
R04196	x		x
R04156	x	x	x
R04190	x	x	x
R04184	x	x	x
R04153	x	x	x
RTx436		x	x
R04232	x	x	
R04160		x	
R04185	x	x	
R04143	x		x
R04135	x	x	x
R04146	x		x
RTx437	x		x
R04231	x	x	x
R04165		x	x
R04181	x	x	
R04131		x	x
TOTAL	13	13	13

Table 17. Inclusion of the seventeen top performing R-lines (based on average performance) across locations in 2006 in the top seventeen selections of each tester.

	Testers		
Male	ATx2752	ATx623	ATx623/BTx2752
R04184	x	x	x
R04146	x	x	x
R04232	x	x	x
RTx437	x	x	x
R04196		x	x
R04190	x	x	x
R04153	x	x	x
R04156	x	x	
R04164	x	x	
R04135	x	x	
R04181	x	x	x
R04104	x		x
R04231	x	x	x
R04160		x	x
RTx436	x		
R04185	x	x	
R04143	x		x
TOTAL	15	14	12

Table 18. Results of the R-line Schumann-Bradley test for statistically testing differences in the discriminatory efficiencies of the testers by comparing the Entry (Male) effect F-values they produced in analyses of variance.

	Entry (Male) Effect F-value		
	ATx2752	ATx623	ATx623/BTx2752
Across Locations	1.43**a**	1.08**a**	1.29**a**
College Station	1.65**a**	2.14**a**	3.00**a**
Weslaco	1.21**ab**	1.41**a**	0.60**b**
Halfway	2.07**a**	0.69**b**	1.81**a**

† F-values sharing the same letter across rows do not have statistically different discriminatory efficiencies.

Schumann-Bradley Test

For the data combined across locations, the Schumann-Bradley test failed to reject the null hypothesis that the F-values produced by each tester for testing the Male effect are equal (Table 18). This indicates that the three testers are considered to have the same efficiency in discriminating among R-lines. These results were expected, considering that the combined data reduces the role of the environment and therefore, differences between the testers based on adaptation regions.

The Schumann-Bradley test for College Station also failed to find significant differences between the testers in their efficiencies for discrimination among the lines. However, the F_1 tester did produce the numerically largest F-value of the three testers in College Station.

In Halfway, ATx2752 produced the largest F-value for the Male effect, but the F-value produced by the F_1 tester, though smaller, was not considered statistically different from ATx2752's F-value. Therefore both testers provide the same discrimination efficiency at Halfway per the Schumann-Bradley test. The smallest F-value for Male

effect was produced by ATx623, and it was declared statistically different from both the ATx2752 and the F_1 tester F-values. These results provide evidence that the discrimination efficiency of ATx623 is significantly less than ATx2752 and the F_1 tester at Halfway.

In Weslaco, the discrimination efficiencies of both ATx2752 and ATx623 were not different, but the F-value for Male effect produced by the F_1 tester was declared statistically different by the Schumann-Bradley test. The smaller F-value of the F_1 tester indicates that its discriminatory efficiency was significantly less than that of the other testers in Weslaco. Because Weslaco represents a subtropical to tropical environment, it was assumed that ATx623 would be a better tester than ATx2752 and therefore, should produce a larger F-value. The F_1 tester F-value was expected be between the values produced by the two inbred line testers. It may be that the 2006 Weslaco environment was constrained in some manner so as to prevent the production of large differences between the hybrids. In such a case, much of the discriminatory efficiency advantage of a subtropical-tropical adaptation tester over a temperately adapted tester would be reduced. This would also reduce some of the advantage provided by the combination of alleles from both adaptation types in the F_1 tester. In addition, the increased heterogeneity of the F_1 tester would further reduce the variance among lines, and lead to the production of a smaller F-value. Weslaco is the only location where the F_1 tester produced the smallest F-value, and is the only location where the F-value was declared statistically different from that of the tester with the largest F-value. Repeating this test in Weslaco over more years would help elucidate whether some of the observed results

can be attributed to the specific environment in 2006, or to the general performance of the testers in Weslaco.

Testcross Variances

In College Station, the F_1 tester produced the largest testcross variance, but ATx623 was similar (Table 19). In Weslaco, The F_1 tester produced a negative testcross variance, which can be assumed to be zero, indicating that the F_1 tester produced virtually no true differences among the R-lines. However, none of the three testers produced a significant effect for the differences among the R-lines in Weslaco. ATx2752 and ATx623 produced substantially different testcross variances, indicating that they differ in their efficiencies for discriminating among the R-lines. But again, none of the testers produced a significant effect for the R-lines. In addition, the Schumann-Bradley test failed to detect differences between the efficiencies of the two inbred line testers. The results suggest that, despite the differences in testcross variances, both inbred line testers provided similar discrimination efficiencies. As was previously mentioned, conditions in Weslaco may have reduced the expression of the differences among the R-lines, thereby reducing the ability to discern differences between the testers. And the increased heterogeneity of the F_1 tester would further reduce the small testcross variances that would be observed in these conditions.

In Halfway, ATx2752 produced the largest testcross variance, followed by the F_1 tester and then ATx623. Because ATx2752 represents the tester more adapted to the Halfway location, it was expected to allow for the maximum expression of the variance between lines. And because ATx623 was expected to donate alleles that would limit the

expression of variance among the lines, it was expected to produce the smallest testcross variance. And indeed, the testcross variance produced by ATx623 was negative, so it can be assumed to be zero. With a combination of both positive and negative alleles for Halfway, the F_1 tester yielded the expected testcross variance smaller than that of ATx2752, yet superior to that of ATx623.

At two of the locations, the testcross variances produced were sufficient to permit the observation of statistically significant differences for the variance among the R-lines (Male effect). Unexpectedly, in College Station the F_1 tester had the largest testcross variance as well as the highest degree of significance for the Male effect. ATx623 also detected a significant Male effect in College Station. In Halfway, ATx2752 produced the most significant Male effect and the F_1 tester also produced similar results, despite having a smaller testcross variance. At all the locations and across locations, the F_1 tester was consistent in its declaration of significance for the Male effect as compared to the inbred testers. This provides evidence that despite differences in testcross variances, the F_1 tester provided efficiency results similar to those of the inbred line testers and represents a viable option for discriminating among the R-lines.

Table 19. Testcross variances with their standard errors produced by each tester within each location in 2006 for the R-line test.

Environment	ATx2752	ATx623	ATx623/BTx2752
College Station	23451 +/- 28503	50613 +/- 58834	59712 +/- 53174
Weslaco	6846 +/- 31533	17486 +/- 41325	18611 +/- 35567
Halfway	33575 +/- 37605	-13000 +/- 31292	15388 +/- 20560

Conclusions

By identifying differences among the hybrid testcross progeny produced by the three testers, conclusions regarding the performance of each tester for evaluating the experimental R-lines can be made. Of particular interest, are inferences that can be made regarding the utility of the F_1 tester.

The non-significant Line x Tester analyses both at the individual location level and across locations provides evidence that the three testers evaluate the R-lines with a similar degree of accuracy. This implies that the F_1 tester ranked the R-lines in a manner consistent with the inbred line testers. The large correlations between the individual tester rank assignments and the ranks based on the average performance of the material also demonstrates the accuracy of the testers in evaluating the R-lines. In all the rank correlation analyses performed, the F_1 tester performed with a degree of accuracy similar to the inbred line testers. This permits the conclusion that the F_1 tester was similar to the inbred line testers in terms of accuracy.

All of the testers consistently identified the majority of the top performing R-lines in each location and across locations, implying they performed with similar accuracies.

With the Schumann-Bradley test, the F-value for the Male effect produced by the F_1 tester was declared statistically different from the both inbred line testers only once. This occurred in Weslaco where the environment may have reduced the variance among the lines and subsequently, the F-values of the testers. This may have been of particular importance in the more heterogeneous F_1 tester. Both ATx2752 and ATx623 yielded similar results as testers, indicating that the adaptation type of the testers may have had a reduced role. In such a case, the advantage of combing two adaptation types into an F_1

tester would be reduced. And the increased heterogeneity of the F_1 tester would become a liability by reducing its discrimination efficiency. For Weslaco, the F_1 tester was significantly less efficient than the inbred line testers per the Schumann-Bradley test. An accumulation of data from multiple years is needed to elucidate whether this is due to a location x year interaction, or to the effect of the Weslaco location in general. However, for the remaining locations and in the combined analysis, the Schumann-Bradley test consistently failed to reject the hypothesis that the F_1 tester F-value was equal to that of both the inbred line testers, or to the largest F-value produced. With the exception of Weslaco, the F_1 tester provided a discriminatory efficiency similar to that of the inbred line testers.

With the exception of Weslaco, the F_1 tester always produced the largest or, as was expected, the second largest testcross variance. If the testcross variance produced by a tester in one of the analyses was able to detect a significant effect for the R-lines, the F_1 tester was also able to detect that effect. The F_1 tester produced testcross variances large enough to be consistent with the tester generating the largest testcross variance in detecting significance among the R-lines.

A valid tester for the evaluation of experimental lines should rank lines with a high degree of accuracy and efficiently discriminate among the lines. The overall results of the R-line test indicate that ATx623/BTx2752 provides a degree of accuracy similar to that of ATx2752 and ATx623. The results also indicate that across locations and in every individual location except Weslaco, ATx623/BTx2752 was measured to be sufficiently efficient to produce results comparable to the inbred line testers. While not necessarily a superior tester, ATx623/BTx2752 represents a valid tester for the evaluation of experimental R-lines in College Station, Halfway, and across locations. Further testing in Weslaco would be needed to make reliable conclusions regarding the performance of ATx623/BTx2752 as a tester in that location.

In addition to being a valid tester, the F_1 tester permits the evaluation of R-lines against alleles from two genetic backgrounds, representing two adaptation types. In this manner, the probabilities of incorrectly discarding material that may have superior performance with alleles from ATx623 or ATx2752 are reduced in a single testcross. R-lines that are identified as having a superior GCA can subsequently be tested with the individual inbred lines for the identification of specific combining abilities.

CHAPTER IV

B-LINE TEST

Introduction

In developing new sorghum inbred lines, it is essential to evaluate those inbred lines for performance in hybrid combination. This is done using a testcross, in which experimental lines are crossed to a common line, producing hybrids which are evaluated for performance. New inbred lines will either be advanced or discarded, based on the performance of the hybrid produced with the tester.

To ensure the most accurate evaluation of experimental lines, it is essential that an appropriate tester be selected. Tester options range from broad genetic base populations to elite inbred lines; most sorghum breeding programs will have their own criteria for tester selection, based on their needs and goals. However, any tester selected must meet certain requirements. An effective tester should correctly rank inbred lines for performance in hybrid combination, and it should maximize the variance between testcross progeny for efficient discrimination (Rawlings and Thompson 1962).

Most sorghum breeding programs utilize elite inbred line testers to evaluate new lines. Elite inbred testers provide breeders with valuable information regarding the performance of new lines with specific genetic backgrounds. However, utilizing an elite inbred tester only allows for the evaluation of new lines with alleles from a single genetic background. An alternative would be to use two inbreds, but this increases the numbers and is usually not economically or logically feasible. The use of a single inbred line tester increases the risk of incorrectly discarding material that may combine well with germplasm adapted to a different region or genetic background than that of the tester.

Most sorghum hybrids are adapted to either temperate or tropical environments. Few, if any, are well adapted to both types of environments. Because the state of Texas produces sorghum in both temperate and tropical environments, it is possible to evaluate experimental lines in both environments in a single state.

A potential solution for evaluating experimental lines against multiple groups of germplasm, would be to use a single-cross (F_1) tester. An F_1 tester derived from the cross of an elite temperately adapted line to an elite subtropical-tropical line could reduce the probabilities of incorrectly discarding material.

The purpose of the B-line test is to determine whether the A3 cytoplasm F_1 tester of two commonly used male inbred lines, RTx430 and RTx436, is a valid tester for evaluating sorghum B-lines versus its inbred line parents alone. A3Tx436/RTx430 combines the genetic background of A3Tx430 and RTx436. And thus, A3Tx436/RTx430 may represent a viable tester option for estimating the GCA of sorghum B-lines against two genetic backgrounds, in a single testcross.

Evaluation of hybrid testcrosses derived from A3Tx430, A3Tx436, and A3Tx436/RTx430 allows for the detection of differences between their performance as testers. By comparing parameters that quantify their accuracy in ranking the B-lines, conclusions can be drawn regarding the validity of these testers for obtaining correct GCA estimates of new B-lines. And parameters that measure tester efficiency can be used to establish differences between the testers in maximizing the variance between experimental lines, for the efficient discrimination between those lines.

The correlation between the ranking of experimental lines by a tester and the ranks produced by the average performance of the same material in multiple hybrid

combinations is an indicator of a tester's accuracy in ranking the experimental lines (Castellanos *et al.* 1998). A hypothesis of this test is that all three testers will rank the experimental lines with sufficient accuracy to produce strong correlations between their rank assignments and the ranks produced by the average performance of all testcross hybrids. Using an unadapted tester may mask and confound performance and ranking of the material. A3Tx436 and A3Tx430 are sometimes considered to have different environmental adaptations, the distinction is not as clear as with the testers in the R-line test. Therefore, obvious differences in the performance of the testers based on environmental adaptations regions are not expected to be observed, but are plausible.

When considering the data combined across locations, the differences in rank correlations between the testers should be reduced. The combined data will contain information from multiple locations, thereby reducing the effect of the environment in the tester rank assignments. In this manner, the advantage of one tester in a location may be offset by the advantage of the opposite tester in another environment, thereby reducing differences in rank correlations across environments.

Line x Tester analyses provide an additional parameter for evaluating the accuracy of a group of testers. A Line x Tester analysis is the interaction between the experimental lines and the testers for the dependant variable in a statistical model. A significant line x tester interaction provides evidence that the ranking of experimental lines will differ depending on the tester used. In such cases, testcross evaluations made with one tester will not be comparable to those made with one of the other testers. In this test, the hypothesis is that the three testers will not differ sufficiently in their assessment of the experimental B-lines to produce significant line x tester interactions. Such results

would provide evidence that the testers have similar accuracies in ranking the experimental lines, and that testcross evaluations derived from the different testers, including the F_1, are comparable. But because A3Tx430 and A3Tx436 represent different genetic backgrounds, a significant line x tester interaction is plausible. If this occurs, the previously mentioned tester rank correlations can provide insight regarding the differences amongst the testers in assigning ranks.

The Schumann-Bradley test provides a statistical method to elucidate differences in the discriminatory efficiencies of the testers. The Schumann-Bradley test compares the F-ratios of similar experiments to determine whether the experiments have significantly different efficiencies. This test has previously been used to compare tester discriminatory efficiencies (Sharma et. al. 1967). Within the B-line test, significant differences in efficiencies are likely to be observed between the testers when evaluated with the Schumann-Bradley test. In particular, A3Tx430 and A3Tx436 may demonstrate significantly different efficiencies based on environmental adaptation. In these environments, an unadapted tester will mask some of the differences between the experimental lines and reduce the testcross variance and efficiency that otherwise would be seen with an adapted tester. Because the F_1 tester combines alleles from both genetic backgrounds, this reduction in efficiency should not be as pronounced, and the F_1 tester is not expected to differ significantly from the superior inbred line tester in efficiency, as measured by the Schumann-Bradley test.

Another method for evaluating tester efficiencies is to compare the testcross variances produced by each tester. A larger testcross variance implies an increase in the variance between experimental lines and therefore, greater efficiency in the

discrimination among those lines. Testcross variances are maximized when the genetic potential of each experimental line is allowed to be fully expressed with minimal interference from the tester. In environments with reduced interaction with the testers, the testcross variances of the inbred lines should be similar. In such cases, the increased heterogeneity of the F_1 tester will partition more variance within the B-lines rather than among the B-lines, and reduce testcross variances. However, the reduction in testcross variances should not be sufficient to invalidate its use as a tester. In environments producing a strong interaction with the testers, the two inbred lines are likely to produce testcross variances that differ greatly in size.

With the combination of alleles from two genetic backgrounds, A3Tx436/RTx430 is expected to produce a testcross variance between that of the two inbred line testers under such conditions. Although the F_1 tester is not expected to produce the largest testcross variances, it should provide sufficiently large testcross variances to allow the detection of significant differences between the experimental lines, if significant differences are to be had. If the tester with the largest testcross variance produces a significant entry effect in a statistical model, the F_1 tester should produce a smaller, but sufficiently large, testcross variance to also detect a significant entry effect. Confirmation of this would lend validity to the use of F_1 testers for the efficient discrimination of experimental lines.

The B-line test objectives are to compare the accuracies of A3Tx430, A3Tx436, and A3Tx436/RTx430 in evaluating the experimental lines using rank correlations and line x tester analyses, as well as comparing their discriminatory efficiencies with the Schumann-Bradley test and comparisons of testcross variances. The results of these

analyses will be interpreted to assess the validity of A3Tx436/RTx430as a tester for sorghum B-lines.

Materials and Methods

Testcross Development

Thirty experimental sorghum B-lines were randomly selected from a set of F_5 breeding lines in the Texas A & M sorghum breeding program.

In the initial phases of testcrossing, male-sterile versions of experimental B-lines are not available. To evaluate their performance in crosses with R-lines, male sterile versions of common R-line parents are used as testers. For this test, two common R-line parents representing distinct genetic backgrounds and sterilized in male-sterile A3 cytoplasm were used as testers for the experimental B-lines. These A3 R-lines are isocytoplasmic to the original versions of these R-lines, and were pollinated by the experimental B-lines to yield hybrids. These B-lines do not restore male fertility, resulting in male-sterile testcross hybrids. Previous research has shown that testcrosses made with A3 R-lines are accurate predictors of A/B line performance when the reciprocal cross is made with standard fertility restoring A1 R-lines (Lee et al. 1992). During the summer of 2005, each experimental B-line was hybridized to A3Tx430, A3Tx436, and their F_1 hybrid, A3Tx436/RTx430 to yield a total of 90 hybrid testcross entries. Each experimental B-line was represented by three entries, each derived from a cross to one of the three testers.

Experimental Design

The 90 entries were arranged using the "sets and rep" design with 10 sets and 9 entries/set and three replications. The three entries representing each experimental B-line were assigned together into the same set. During the summer of 2006, the B-line test was grown in two locations: College Station, TX and Halfway, TX. In each environment, standard agronomic management practices were followed for fertilization. Supplemental irrigation was used at both locations. In each environment, a plot was defined as two rows 6 meters in length spaced 76 cm apart. Each row within a plot was planted with four grams of seed pre-treated with fungicides. Because the B-line test entries were all male-sterile, pollinator rows were planted on every seventh and eighth row within the test. In addition, all the border rows and ranges surrounding the test consisted of the same pollinator mix found in the pollinator rows. To provide a consistent source of pollen over the potential flowering dates for the experimental plots, the pollinator mix used consisted of 5 different male fertile hybrids: ATx3197/RTx7000, ATxARG-1/RTx430, ATx2752/RTx2908, ATx631/RTx436, and ATx635/R9623.

Data Collection

Standard agronomic notes were taken for all the plots at every location. These include height and days to mid-anthesis. Plant height was measured in inches from the base of the plant to the tip of the panicle as an average for the plot. Days to mid-anthesis was recorded as the Julian date that 50% of the plot reached 50% anthesis. Data for days to mid-anthesis were not collected for Halfway, TX.

Two heads per plot were covered with bags before flowering, to confirm the male sterility of the entries. Prior to combine harvest, two randomly selected panicle samples were harvested from each single cross hybrid plot while four panicles were harvested from the three way hybrids (due to segregation in these hybrids). These samples were used to obtain average panicle length and 100 seed size. Panicle length was measured in centimeters from the from the bottom pedicel to the tip of the panicle as an average for the plot. One hundred seed size is measured as the weight of 100 seeds in grams as an average for the plot.

Grain yield was measured using a modified John Deere 3300 plot combine, with plot grain weight, grain moisture, and test weight collected by the HarvestMaster HM-1000 weigh system onboard the combine. Total plot yield (kg/ha) was obtained by adjusting the plot weight following formula:

[(((100 - % Moisture)/87) * Plot Weight) * 385] * 1.115 = Total Plot Yield (kg/ha)

The plot yield obtained with this formula was further adjusted by multiplying it by the stand rating for each plot to account for missing plants. The adjusted yield data for each location was organized into datasets using Microsoft Excel, and subsequently imported into SAS 9.1 statistical analysis software as text files for analysis.

Statistical Analysis

The initial analyses of the B-line test were performed on an individual location basis using the PROC GLM procedure available in SAS 9.1 using the following statistical model:

$$\text{Yield} = \mu + \text{Tester} + \text{Female} + \text{Rep} + \text{Set(Rep)} + \text{Tester*Female} + \text{Error}$$

The Female effect represents the proportion of the variance attributable to the differences among the experimental B-lines. All the effects were analyzed utilizing the SAS 9.1 defaults, and were considered fixed effects. The size and significance value of the Tester*Female interaction term represents the Line x Tester interaction, and was used in evaluating the accuracy of the testers in each location

Rank correlations were based on average performances in each environment. The ranks based on average performances were then compared to the rank assignments generated by each tester to derive rank correlations using the CORR function in Microsoft Excel spreadsheets. The ranks produced by each individual tester were also compared to each other using the Microsoft Excel CORR function to determine whether any of the testers, particularly the F_1 tester, assigned ranks more like one tester than to another.

Data for each location of the B-line test was analyzed by tester with the following statistical model:

$$\text{Yield} = \mu + \text{Female} + \text{Rep} + \text{Set(Rep)} + \text{Error}$$

In this model, an F-ratio for the effect due to the variance between the B-lines (Female effect) was obtained for each tester. These F-ratios were then compared against each other via the Schumann-Bradley test to determine whether they were statistically different from each other (Sharma *et al*. 1967).

The Schumann-Bradley test compares the efficiencies of similar experiments by deriving a *w* statistic using the F-ratios of two experiments having the same degrees of freedom and number of replications (Bradley and Schumann 1957).

$$w = \text{Exp. 1 F-ratio/Exp. 2 F-ratio}$$

This *w* statistic is compared against tabulated significance values using transformed degrees of freedom and parameters derived from the two experiments being compared. The first step in conducting a Schumann-Bradley test is to calculate two parameters, **a** and **b**. The **a** parameter is equal to half the degrees of freedom of the effect being tested by the F-ratios of the experiments. In the case of the B-line test, this is equal to half the degrees of freedom for the Female effect. The **b** parameter is equal to half the error degrees of freedom shared by both experiments.

$$\mathbf{a} = ½(\text{Male effect d.f.})$$

$$\mathbf{b} = ½(\text{Error d.f.})$$

The **a** and **b** parameters are used to calculate a third parameter, λ_i, for each experiment using the following formula:

$$\lambda_i = \mathbf{a}(\text{F-ratio}_i - 1)$$

For the B-line test, a λ_i was calculated for each tester at each location. When comparing two experiments, the λ_is' for both experiments are added together to produce a total λ.

$$\lambda = \lambda_1 + \lambda_2$$

The total λ and the individual λ_is' form the basis for the hypothesis tests of the Schumann-Bradley test.

$$H_0: \lambda_1 = \lambda_2 = \lambda$$

$$H_A: \lambda_1 > \lambda_2$$

In more basic terms, the Schumann-Bradley test is a one-sided test of significance with the null hypothesis that both experiments have equal efficiencies and with the alternative hypothesis that Experiment 1 has a greater efficiency than Experiment 2. Experiment 1 represents the F-ratio used as the numerator in calculating the *w* statistic.

Upon the calculation of a total λ, it can then be used to calculate the a' parameter that is used for interpreting the tabulated w_0 values in the Schumann-Bradley table. The a' is calculated in the following manner:

$$a' = (\mathbf{a} + \lambda)^2/[\mathbf{a} + 2(\lambda)]$$

In conjunction with **b**, a' is used to identify the cutoff w_0 in the Schumann-Bradley table that is used for determining the significance of the observed *w*. If the observed *w* exceeds w_0, the null hypothesis is rejected and the alternative hypothesis that the experiment providing the numerator F-ratio in calculating *w* is more efficient than the experiment providing the denominator F-ratio is accepted. In the B-line test, the F-ratios for the Female effect produced by each tester were compared against each other per location using the Schumann-Bradley test to identify significant differences in their efficiencies.

Testcross variances were estimated using the same model. The testcross variances were obtained by subtracting the Error mean squares from the Female mean squares produced by each tester within a location and then dividing that number by the number of replications and sets. Standard errors for the testcross variances were derived by calculating a 95% confidence interval around the variance estimate (Bernardo, 2002) and dividing it by two. After the derivation of the testcross variances, they were compared against each other to make inferences regarding the discriminatory efficiencies of the testers within a location.

Upon completing the analyses on a per location basis, the yield data for both locations was combined to analyze the testers across locations. This was deemed appropriate using the HOVTEST command within the PROC GLM procedure of SAS 9.1 with the combined yield data. The HOVTEST command performs a Levene's test for the homogeneity of variances, with the null hypothesis that the variances at both locations are

equal. To further investigate the appropriateness of combining the data, the combined yield data was segregated by tester to determine whether the variances for each tester were equal across locations. This was also done using the HOVTEST command within the SAS PROC GLM procedure, but on a per tester basis. The results in (Table 20) indicate that the Levene's test failed to reject the null hypothesis that the variances for each tester are equal across locations. In addition, the data was normally distributed with no obvious outliers. With these results, analyses evaluating the testers with the data combined across both locations were performed.

Table 20. Probability values of the Levene's test for the homogeneity of location variances for each tester and the combined tester data in the B-line test.

Analysis	F-Value	Pr < F
Combined Tester Data	2.31	0.13
A3Tx430	0.81	0.37
A3Tx436	0.03	0.87
A3Tx436/RTx430	0.87	0.35

The combined data was analyzed in a manner similar to that used on a per location basis. Using the PROC GLM procedure of SAS 9.1, the following model was used:

$$\text{Yield} = \mu + \text{Location} + \text{Tester} + \text{Female} + \text{Rep} + \text{Set(Rep)} + \text{Tester*Female} + \text{Location*Female} + \text{Error}$$

As before, all the effects were considered fixed per the SAS defaults, and the Female effect represents the variance between the B-lines. The Line x Tester analysis and rank correlations were completed using the same methodology as previously described. As in the individual environment analysis, the combined data was analyzed by tester using the following model:

$$\text{Yield} = \mu + \text{Location} + \text{Female} + \text{Rep} + \text{Set(Rep)} + \text{Location*Female} + \text{Error}$$

The reduced model for the combined data produces F-ratios for testing the effect due to the variance between the B-lines (Female effect) for each tester. These F-ratios were compared against each other via the Schumann-Bradley test (as previously described) to determine whether they were statistically different from each other and elucidate differences in efficiencies among the testers. If the Schumann-Bradley test declared two F-ratios statistically different, the tester providing the F-ratio used as the numerator in calculating the *w* statistic was considered as being more efficient in discriminating amongst the B-lines across the locations.

The reduced combined model was used to estimate testcross variances across locations using a fixed effect model. The testcross variances for each tester were obtained by subtracting the Error mean square from the Female mean square produced by each tester across locations and then dividing that number by the number of replications, locations, and sets. Estimates of testcross variances were compared to make inferences regarding the discriminatory efficiencies of the testers across locations. Testers with larger testcross variances were considered superior in maximizing the variance between the lines and therefore, more efficient in discriminating among the lines.

Results and Discussion

Individual and Combined Analysis

Significant variation was detected for most traits in the combined analysis and in most environments (Tables 21-23). While there is merit in pursuing the detailed analysis of agronomically important traits such as plant height and days to mid-anthesis, for the purposes of this study and subsequent breeding efforts, emphasis and further analysis will be primarily focused on grain yield. Yield data was relatively consistent as indicated by C.V. and R-square values (Tables 21-23).

In the BY tester analysis, significant variation was detected for most traits in the combined analysis and in most environments (Tables 24-25). While there is merit in pursuing the detailed analysis of agronomically important traits such as plant height and days to mid-anthesis, for the purposes of this study and subsequent breeding efforts, emphasis and further analysis will be primarily focused on grain yield. Yield data was relatively consistent as indicated by their C.V values (Tables 24-25).

Table 21. Yield, Height, and Days to Mid-Anthesis analysis of variance results for hybrid testcrosses of sorghum B-lines to A3Tx430, A3Tx436, and A3Tx430/RTx436 across locations in 2006.

	Yield		Height		Days/Mid-Anth.	
Source of Variance	df	Mean Square	df	Mean Square	df	Mean Square
Location	1	11173556.8**	1	75.1	--	--
Tester	2	40499290.2**	2	87.1	2	26.8**
Female	29	3618169.7**	29	283.8**	29	14.7**
Rep	2	24495511.1**	2	39.2	2	89**
Set(Rep)	27	710793.5	27	104.2	27	5.1**
Tester*Female	58	942698.0	58	61.3	58	4.9**
Location*Female	29	1672772.9	29	76.4	--	--
Error	374	1177852.2	372	91.1	150	2.6
	C.V. = 27.8 R^2 = 0.51		C.V. = 8.6 R^2 = 0.38		C.V. = 2.4 R^2 = 0.76	

*,** Significant at $p < .05$ and .01, respectively

† Days to Mid-Anthesis data was not collected for Halfway, TX

Table 22. Yield, Height, and Days to Mid-Anthesis analysis of variance results for hybrid testcrosses of sorghum B-lines to A3Tx430, A3Tx436, and A3Tx430/RTx436 in College Station, TX in 2006.

	Yield		Height		Days/Mid-Anth.	
Source of Variance	df	Mean Square	df	Mean Square	df	Mean Square
Tester	2	41534078.7**	2	6.4	2	26.8**
Female	29	3234663.7**	29	266.8**	29	14.7**
Rep	2	30668906.3**	2	2555.9**	2	89**
Set(Rep)	27	1987856.2**	27	173.5*	27	5.1**
Tester*Female	58	1479404.8**	58	96.7	58	4.9**
Error		910301.6	150	101.6	15	2.6
	C.V. = 23.8 R^2 = 0.76		C.V. = 9.2 R^2 = 0.64		C.V. = 2.4 R^2 = 0.76	

*,** Significant at $p < .05$ and .01, respectively

Table 23. Yield and Height analysis of variance results for hybrid testcrosses of sorghum B-lines to A3Tx430, A3Tx436, and A3Tx430/RTx436 in Halfway, TX in 2006.

	Yield		Height	
Source of Variance	df	Mean Square	df	Mean Square
Tester	2	6137899.9**	2	47.6
Female	29	2792135.3**	29	179**
Rep	2	33995057.2**	2	1888.2**
Set(Rep)	27	1190863.4	27	105.9*
Tester*Female	56	710872.9	55	48.4
Error	141	807827.1	141	61.5
	C.V. = 24.1 R^2 = 0.71		C.V. = 7.09 R^2 = 0.65	

*,** Significant at $p < .05$ and .01, respectively

Table 24. Yield, Height, and Days to Mid-Anthesis analysis of variance results by tester (A3Tx430, A3Tx436, and A3Tx430/RTx436) for hybrid testcrosses of B-lines in College Station, TX in 2006.

		Dependent Variables		
A3Tx430	df	Yield	Height	Days/Anth.
Female	29	2259813.6	239.8	5.6*
Rep	2	15590722.7**	1480.1**	38.6**
Set(Rep)	21	1313207.5	153.1	3.1
Error	36	1509875.0	155.4	2.7
C.V.		25.8	11.5	2.4
A3Tx436	df	Yield	Height	Days/Anth.
Female	29	1786410.3**	78.8*	9.5**
Rep	2	16097846.6**	483.2**	28.9**
Set(Rep)	21	1680604.2**	95.3*	4.6
Error	37	620994.6	40.6	3.7
C.V.		20.9	5.9	2.8
A3Tx436/RTx430	df	Yield	Height	Days/Anth.
Female	29	1572236.8**	167.2	5.4**
Rep	2	370432.2	499.6*	10.1*
Set(Rep)	17	1232157.0**	117.9	1.5
Error	41	413467.8	110.3	2.2
C.V.		18.4	9.3	2.2

*,** Significant at $p < .05$ and .01, respectively

Table 25. Yield, and Height analysis of variance results by tester (A3Tx430, A3Tx436, and A3Tx430/RTx436) for hybrid testcrosses of B-lines in Halfway, TX in 2006.

		Dependent Variables	
A3Tx430	Df	Yield	Height
Female	28	990446.0	100.2**
Rep	2	12348998.0**	416.3**
Set(Rep)	21	463328	70.9
Error	35	1079411.5	40.2
C.V.		25.1	5.7
A3Tx436	Df	Yield	Height
Female	28	1575312.0*	105.6
Rep	2	14588985.0**	445.7*
Set(Rep)	21	1052029.1	82.6
Error	29	721759.7	83.3
C.V.		23.6	8.3
A3Tx436/RTx430	Df	Yield	Height
Female	29	1400120.0*	126.2*
Rep	2	226023.1	449.4**
Set(Rep)	17	1144427.7	83.2
Error	41	728121.3	65.4
C.V.		24.7	7.2

*,** Significant at $p < .05$ and .01, respectively

Line x Tester Analysis

Statistical analysis of the College Station data revealed a highly significant Line x Tester interaction (Table 22). These results indicate that the relative performance of the B-line hybrids differed in College Station depending on the tester used. In addition to the highly significant Line x Tester interaction, the remaining effects; Tester, Female, Rep, and Set(Rep), were all highly significant in the College Station B-line test.

However, the validity of the College Station B-line Line x Tester analysis is suspect because of a lack of randomization of the plots and disease. An incorrect randomization mistakenly placed entries representing crosses to the F_1 tester together into large blocks. In addition, entries within sets were not consistent across the replications. The College Station location also developed a high degree of anthracnose (*Colletotrichum graminicola*) infection in certain areas of the field. A3Tx436 is highly resistant to anthracnose while A3Tx430 is not. Positive or negative interactions for anthracnose resistance combined with a reduction in randomization of the entries combined may have contributed to the highly significant Line x Tester analysis.

These factors may have caused violations of the ANOVA assumptions that the residuals of a dataset should be randomly, independently and normally distributed. Shapiro-Wilk normality tests failed to reject the hypothesis that the residuals were normally distributed. However, plots of the standardized residuals performed for each tester were not random in nature, with the presence of clear outliers. Due to these issues, inferences regarding the accuracy of the testers were not made with the College Station Line x Tester analysis. But rank trends were similar for all testers in College Station despite some rank shifts (Table 26).

The B-line test in Halfway did not share the errors of the College Station location, and a significant Line x Tester interaction was not detected. This indicates that the testers were consistent in ranking the experimental lines. While there were some rank shifts, rank trends were similar for all testers in Halfway (Table 27). They were consistently reliable in identifying poor performing lines, which is of particular importance.

Consistent with the results in Halfway, the Tester*Female interaction for the data combined across locations was also non-significant. These results provide evidence that the testers assigned ranks to the experimental lines in a consistent manner (Table 28).

However, some important rank shifts between the two inbred line testers were seen. For example, B05146 was the top performing line with A3Tx436 in College Station, but was ranked twenty-first by A3Tx430. A3Tx436/RTx430 ranked the same line thirteenth. Use of the F_1 tester in this case would reduce the probability of discarding B05146 as compared to using A3Tx430 alone as the tester. Other similar examples can be found in tables 26-28.

Table 26. Ranks, based on the grain yield of 30 experimental B-lines in testcross combination with A3Tx430, A3Tx436, and A3Tx436/RTx430 when evaluated in College Station, TX in 2006.

FEMALE	AVERAGE	A3Tx430	A3Tx436	A3Tx436/RTx430
B05137	1	1	4	2
B05186	2	2	8	3
B05154	3	12	2	9
B05159	4	3	6	18
B05275	5	4	7	19
B05146	6	21	1	13
B05167	7	11	3	23
B05242	8	25	20	1
B05265	9	23	5	6
B05240	10	5	22	14
B05219	11	7	18	15
B05148	12	14	10	16
B05273	13	16	14	5
BTx2752	14	17	21	4
B05151	15	8	9	25
B05267	16	9	24	7
B05257	17	20	19	8
B05209	18	10	13	24
(B9701*SC1251)-CS6-CA1	19	18	25	10
B05129	20	26	12	21
BTx2928	21	13	26	22
(B9701*SC1251)-CS6-CA3	22	22	27	12
B05274	23	28	15	11
B05128	24	27	11	27
(B9202*97CA2258)-CS2-CA2	25	19	23	26
B05130	26	15	16	29
B05266	27	6	28	28
B05179	28	29	17	17
B05224	29	24	29	20
B05208	30	30	30	30

Table 27. Ranks, based on the grain yield of 30 experimental B-lines in testcross combination with A3Tx430, A3Tx436, and A3Tx436/RTx430 when evaluated in Halfway, TX in 2006.

FEMALE	AVERAGE	A3Tx430	A3Tx436	A3Tx436/RTx430
B05275	1	2	1	11
(B9202*97CA2258)-CS2-CA2	2	1	3	2
B05257	3	5	2	1
B05273	4	7	6	4
B05266	5	6	10	5
B05130	6	4	16	14
B05209	7	8	12	23
B05154	8	9	20	12
B05137	9	17	18	10
B05274	10	19	21	3
(B9701*SC1251)-CS6-CA1	11	12	8	24
B05240	12	11	17	20
B05267	13	23	5	21
B05265	14	3	25	7
B05148	15	27	11	6
B05242	16	21	4	25
BTx292	17	25	7	19
B05179	18	20	9	22
B05167	19	16	22	16
B05159	20	24	15	15
B05146	21	15	23	13
(B9701*SC1251)-CS6-CA3	22	22	14	26
B05186	23	29	13	8
B05151	24	14	27	9
B05219	25	13	19	27
B05128	26	18	26	18
BTX275	27	10	29	28
B05224	28	28	28	17
B05129	29	26	24	29
B05208	30	30	30	30

Table 28. Ranks, based on the grain yield of 30 experimental B-lines in testcross combination with A3Tx430, A3Tx436, and A3Tx436/RTx430 when evaluated in College Station, TX and Halfway, TX in 2006.

FEMALE	Average	A3Tx430	A3Tx436	A3Tx436/RTx430
B05275	1	1	1	11
B05137	2	2	4	1
B05257	3	12	7	2
B05273	4	11	9	4
B05154	5	10	2	9
B05186	6	20	6	3
(B9202*97CA2258)-CS2-CA2	7	5	14	12
B05159	8	8	8	15
B05240	9	4	21	14
B05265	10	14	22	6
B05146	11	21	5	10
B05242	12	25	11	5
B05167	13	16	3	22
B05148	14	22	10	8
B05209	15	6	12	27
B05267	16	18	15	13
B05266	17	3	24	26
B05219	18	7	18	25
(B9701*SC1251)-CS6-CA1	19	17	19	17
B05151	20	13	26	16
B05274	21	28	17	7
B05130	22	9	16	28
BTX2928	23	19	20	20
(B9701*SC1251)-CS6-CA3	24	23	23	23
B05179	25	29	13	21
BTX2752	26	15	28	19
B05128	27	24	27	24
B05129	28	26	25	29
B05224	29	27	29	18
B05208	30	30	30	30

Correlations

Rank correlations by location and combined across locations were significant (Table 29), including in College Station, despite the aforementioned errors. In College Station, A3Tx436 produced the largest rank correlation to the average performance ranks ($r = 0.72$). Both A3Tx430 and the F_1 tester had significant, but smaller, rank correlations ($r = 0.59$ and $r = 0.58$, respectively).

In Halfway, A3Tx430 and A3Tx436 produced identical rank correlations ($r = 0.70$), indicating that both testers provided similar accuracy. The F_1 tester produced a significant, yet somewhat smaller rank correlation than the inbred line testers ($r = 60$). Based on the size and significance of the rank correlations, all three testers represent valid testers for B-lines in Halfway.

Across both locations, A3Tx436 had the largest rank correlation to the average performance ranks, and based on this method, it would be considered the most accurate tester for evaluating B-lines ($r = 0.81$). However, the F_1 tester produced a rank correlation only slightly smaller than A3Tx436 ($r = 0.74$). ATx430 produced the smallest rank correlation, yet it was still significant ($r = 0.64$). The large rank correlations produced by all the testers indicate that they accurately ranked the experimental lines for GCA across locations. In addition, the similarity of the rank correlations gives credence to the interchangeable use of the testers to rank B-lines across locations.

Table 29. Correlations between the B-line rank assignments of each tester with their average performance in 2006 across locations, and in College Station, TX and in Halfway, TX.

Locations Combined	Average	A3Tx430	A3Tx436	A3Tx436/RTx430
Average	1.00	0.64**	0.81**	0.74**
A3Tx430		1.00	0.35	0.14
A3Tx436			1.00	0.57**
A3Tx436/RTx430				1.00
College Station	Average	A3Tx430	A3Tx436	A3Tx436/RTx430
Average	1.00	0.59**	0.72**	0.58**
A3Tx430		1.00	0.29	0.09
A3Tx436			1.00	0.19
A3Tx436/RTx430				1.00
Halfway	Average	A3Tx430	A3Tx436	A3Tx436/RTx430
Average	1.00	0.70**	0.70**	0.60**
A3Tx430		1.00	0.26	0.39*
A3Tx436			1.00	0.28
A3Tx436/RTx430				1.00

*,** Significant at $p < .05$ and .01, respectively

Table 30. Inclusion of the fifteen top performing B-lines (based on average performance) for College Station, TX in 2006 in the top fifteen selections of each tester.

Female	A3Tx430	A3Tx436	A3Tx436/RTx430
B05137	x	x	x
B05186	x	x	x
B05154	x	x	x
B05159	x	x	
B05275	x	x	
B05146		x	x
B05167	x	x	
B05242			x
B05265		x	x
B05240	x		x
B05219	x		x
B05148	x	x	
B05273		x	x
BTx2752			x
B05151	x	x	
TOTAL	10	11	10

Table 31. Inclusion of the fifteen top performing B-lines (based on average performance) for Halfway, TX in 2006 in the top fifteen selections of each tester.

Female	A3Tx430	A3Tx436	A3Tx436/RTx430
B05275	x	x	x
(B9202*97CA2258)-CS2-CA2	x	x	x
B05257	x	x	x
B05273	x	x	x
B05266	x	x	x
B05130	x		x
B05209	x	x	
B05154	x		x
B05137			x
B05274			x
(B9701*SC1251)-CS6-CA1	x	x	
B05240	x		
B05267		x	
B05265	x		x
B05148		x	x
TOTAL	11	9	11

Table 32. Inclusion of the fifteen top performing B-lines (based on average performance) across locations in 2006 in the top fifteen selections of each tester.

Female	A3Tx430	A3Tx436	A3Tx436/RTx430
B05275	x	x	x
B05137	x	x	x
B05257	x	x	x
B05273	x	x	x
B05154	x	x	x
B05186		x	x
(B9202*97CA2258)-CS2-CA2	x	x	x
B05159	x	x	x
B05240	x		x
B05265	x		x
B05146		x	x
B05242		x	x
B05167		x	
B05148		x	x
B05209	x	x	
TOTAL	10	13	13

All testers were consistent in the identification of the top performing lines. Of the top fifteen B-lines (averaged across all testers) in College Station, both A3Tx430 and A3Tx436/RTx430 identified ten, and A3Tx436 identified eleven (Table 30). In Halfway, both A3Tx430 and A3Tx436/RTx430 identified eleven, and A3Tx436 identified nine (Table 31). Across all locations, A3Tx430 identified ten, and both A3Tx436 and A3Tx436/RTx430 identified thirteen (Table 32). These results indicate that despite some rank shifts seen in Tables 24-26, as well as differences in rank correlations, all three of the testers identified the majority of the top performing B-lines for the individual locations and across locations. In addition, there were no large differences between the three testers, implying that they performed with similar degrees of accuracy.

Schumann-Bradley Test

For the data combined across locations, the Schumann-Bradley test failed to reject the null hypothesis that the F-values produced by each tester for testing the Female effect are equal (Table 33). This indicates that the three testers have the same efficiencies in discriminating among the B-lines. These results were expected, considering that the combined data reduces the role of the environment and therefore, differences between the testers based on environmental interactions.

The Schumann-Bradley test for College Station did find significant differences between the testers in their discrimination efficiencies. In College Station, the F_1 tester produced the largest F-value for the Female effect, but was not declared statistically different from the smaller A3Tx436 F-value. The A3Tx430 F-value was declared different, and therefore less efficient, from the F_1 F-value by the Schumann-Bradley test.

The F-values for the two inbred testers were not declared statistically different from each other. These results indicate that the discriminatory efficiencies of A3Tx436 and the F_1 tester can be consider similar and the discriminatory efficiency of A3Tx430 in College Station is less than that of the F_1 tester. The presence of anthracnose may have reduced the discriminatory efficiency of A3Tx430 in College Station.

In Halfway, A3Tx436 produced the largest F-value for the Female effect, but the smaller F-value produced by the F_1 tester was not considered statistically different. Therefore, both testers provide the same discrimination efficiency at Halfway per the Schumann-Bradley test. A3Tx430 produced the smallest F-value for Female effect was and was declared statistically different from the ATx436 F-value. These results lend support to the validity of the F_1 tester and provide evidence that the discriminatory efficiency of A3Tx430 was less than that of A3Tx436 in Halfway.

Table 33. Results of the B-line Schumann-Bradley test for statistically testing differences in the discriminatory efficiencies of the testers by comparing the Entry (Female) effect F-values they produced in an analysis of variance.

	Entry (Female) Effect F-value		
Analysis	A3Tx430	A3Tx436	A3Tx436/RTx430
Combined Locations	1.53**a**	1.35**a**	1.90**a**
College Station	1.48**a**	2.49**ab**	3.80**b**
Halfway	1.12**a**	2.55**b**	1.92**ab**

† F-values sharing the same letter across rows do not have statistically different discriminatory efficiencies.

Testcross Variances

In College Station, the A3Tx436 produced the largest testcross variance, closely followed by the F_1 tester, and A3Tx430 producing the smallest testcross variance (Table 34)**.** As with the other analyses, the randomization errors and the presence of anthracnose in College Station may be influencing these results. In Halfway, A3Tx436 had the largest testcross variance, followed by the F_1 tester (Table 34). A3Tx430 produced a negative testcross variance Halfway, which can be assumed to be zero. This implies that the F_1 tester and A3Tx436 were much more efficient in discriminating among the lines in Halfway than A3Tx430 (Table 34).

In the data combined across locations as well as in College Station and Halfway, significance for the variance among the B-lines (Female effect) was detected. In all three of these statistical analyses, the F_1 tester produced a testcross variance large enough to detect significant differences among the B-lines. In the combined data and in College Station, the F_1 tester produced the largest degree of significance out of the three testers. In the combined locations data, both A3Tx430 and A3Tx436 failed to detect significance for the Female effect as the F_1 tester did. And in College Station and Halfway, A3Tx430 was the only tester that failed to detect a significant Female effect.

Table 34. Testcross variances with their standard errors produced by each tester within each location in 2006 for the B-line test.

Analysis	A3Tx430	A3Tx436	A3Tx436/RTx430
College Station	24997 +/- 44621	38847 +/- 29876	38625 +/- 25501
Halfway	-2965 +/- 30644	28451 +/- 28808	22399 +/- 25418

Conclusions

By identifying differences among the hybrid testcross progenies produced by the three testers, conclusions regarding the performance of each tester for evaluating the experimental B-lines can be made. Of particular interest, are inferences that can be made regarding the utility of the F_1 tester.

The non-significant Line x Tester analyses in Halfway and in the combined data provides evidence that the three testers evaluated the B-lines with similar degrees of accuracy. This implies that the F_1 tester ranked the B-lines in a manner consistent with the inbred line testers, and gives credence to its validity as a tester. While a significant line x tester interaction was detected in College Station, errors in the plot arrangements and a high degree of anthracnose infection call these results into question. Conclusions regarding the accuracy of the testers in College Station will be withheld until further data is available in 2007.

The large correlations between the individual tester rank assignments and the ranks based on the average performance of the material also demonstrates the accuracy of the testers in evaluating the B-lines. In all the rank correlation analyses performed, the F_1 tester had substantial rank correlations and performed with a degree of accuracy similar to the inbred line testers. This permits the conclusion that the F_1 tester provided a degree of accuracy similar to the inbred line testers.

In addition to the results of the Line x Tester analyses and rank correlations, all three of the testers consistently identified the majority of the top performing B-lines in each location and across locations, despite some important rank shifts. All three of the

testers identified approximately the same number of top performing lines, implying they performed with similar accuracies.

The F_1 tester produced the largest F-ratio for the Female effect in both College Station and in the combined data. When the F_1 tester did not provide the largest F-ratio, it was not declared statistically different from the largest F-ratio produced. And in both individual location analyses, the F_1 tester ratio was statistically different per the Schumann-Bradley test from the smallest F-ratio. While firm conclusions cannot be drawn with the College Station data, the balance of the Schumann-Bradley test results imply that the F_1 tester represents an efficient and valid tester compared to the inbred line testers.

As with the F-ratios, the F_1 tester produced a large largest testcross variance in College Station, very similar to the largest testcross variance. In Halfway, the F_1 tester produced a testcross variance between that of the two inbred line testers. Even when omitting the College Station results, the testcross variances produced by the F_1 tester support its use as a valid tester for evaluating the B-lines. In addition, the testcross variances produced by the F_1 tester consistently allowed for the detection of significance for the variance among the B-lines, providing further support to the validity of A3Tx436RTx430 as a tester.

A valid tester for the evaluation of experimental lines should rank lines with a high degree of accuracy and efficiently discriminate among the lines. The results of the B-line test indicate that A3Tx436/RTx430 provides a degree of accuracy similar to that of A3Tx430 and A3Tx436. The results also indicate that across locations and in both individual locations, A3Tx436/RTx430 was measured to be the most efficient tester or

similar to the inbred line tester with the greatest efficiency. Despite some inconclusive data from College Station, the results of the B-line test indicate that A3Tx436/RTx430 represents a valid tester for the evaluation of experimental B-lines versus A3Tx430 and A3Tx436.

In addition to being a valid tester, the F_1 tester permits the simultaneous evaluation of B-lines against alleles from two genetic backgrounds. In this manner, the probabilities of incorrectly discarding material that may have superior performance with alleles from A3Tx430 or A3Tx436 are reduced in a single testcross. B-lines that are identified as having a superior GCA can subsequently be tested with the individual inbred lines for the identification of specific combining abilities.

CHAPTER V

SUMMARY

For reliably evaluating experimental inbred lines in testcrosses, it is essential that an appropriate tester be selected. A valid tester is one that correctly identifies the relative performance of experimental lines and maximizes the differences between those lines to allow for efficient discrimination. While there are many tester choices available, most sorghum breeding programs use elite inbred lines as testers, so as to evaluate experimental lines in realistic hybrid combinations.

Inbred line testers only allow for the evaluation of experimental lines against one genetic background, and may increase the probability of incorrectly discarding material that may perform well in a different hybrid combination. A solution would be to test experimental lines with multiple inbred line testers in individual testcrosses. However, this is typically not feasible due to limitations in time and resources. Another potential solution would be to use an F_1 tester for evaluating experimental lines. An F_1 tester combines alleles from two genetic backgrounds, and may reduce the probability of incorrectly discarding material.

The purpose of this research was to compare the performance of F_1 testers versus that of their inbred line parents for evaluating experimental sorghum B and R lines. Sorghum B lines are the male-sterile female parents used in producing hybrid sorghum. Sorghum R lines are the male-fertile pollinators that are crossed to B lines to produce hybrid sorghum. R lines have male fertility restoration capabilities, thus a cross between a B line and an R line yields fully male fertile hybrids. The comparison of the F1 testers

to their respective inbred lines is based on data gathered in one year over two locations for the B lines and three locations for the R lines.

Line x tester analyses were performed for both the R-line test and the B-line test to asses the consistency of the testers in assigning ranks. A non-significant line x tester analysis provides evidence that the performance of the experimental lines was consistent regardless of the tester used. With the exception of one location in the B-line test, all of the line x tester analyses performed were statistically non-significant. While not providing definitive proof, these results indicate that both the inbred line testers and the F_1 testers provided similar assessments of the experimental lines for performance. The one significant line x tester analysis was in the College Station B-line test, and may have been due to the confounding effects of planting errors and anthracnose infection. To clarify these results, the College Station B-line test is being repeated for 2007.

Correlations between the ranking of the experimental lines by their average performance and the ranks produced by each tester were measured to asses the accuracy of the testers. This was done in both tests for each location, as well as with the data combined across locations. In all cases, each tester produced significant rank correlations to the ranks based on average performances. And the rank correlations produced by the F_1 testers were always similar to those produced by the inbred line testers. These results imply that, despite some differences, all of the testers evaluated the inbred lines with informative levels of accuracy.

In addition to the results of the Line x Tester analyses and rank correlations, all of the testers in both the R-line test and B-line test consistently identified the majority of the top performing experimental lines in each location and across locations, despite some

important rank shifts. All of the testers within the two tests identified approximately the same number of top performing lines, implying they performed with similar accuracies.

The F-ratios for the effect due to the variance among the experimental lines (entry effect) was compared for each tester within the individual locations as well as across locations. These comparisons were made with the Schumann-Bradley statistical test, which compares the efficiencies of similar experiments using their F-ratios. With the exception of one location in the R-line test, the F_1 testers always produced the largest or second largest Entry effect F-ratio. In cases where the F_1 testers produced the second largest F-ratio, it was not declared statistically different from the largest F-ratio by the Schumann-Bradley test. In the analysis where the F_1 tester produced the smallest F-ratio (Weslaco R-line test), differences between the inbred line testers were not clear. In such environments, the advantages of combining alleles may be reduced, and the increased heterogeneity of the F_1 tester may decrease discriminatory efficiencies. But the overall balance of the results indicates that the F_1 testers were effective in efficiently discriminating among the experimental lines as compared to the inbred line testers.

The testcrosses variances produced by each tester for every location and across locations were measured to compare the discriminatory efficiencies of the testers. A larger testcross variance allows for more efficient discrimination among experimental lines. With the exception of the Weslaco R-line test, the F_1 testers consistently produced the largest or second largest testcross variances. This provides evidence that, when compared to the inbred line testers, the F_1 testers were effective in maximizing the differences among the experimental lines for their efficient discrimination. As stated

before, the combing of alleles in the F_1 tester may not have been advantageous in Weslaco, and led to a decrease in the testcross variance.

The overall results of this research permit the conclusion that, while not necessarily a superior tester, the F_1 testers represent a valid option for the evaluation of experimental sorghum lines in testcrosses as compared to their respective inbred line testers alone. With only a few exceptions, the F_1 testers provided levels of accuracy and efficiency similar to their counterpart inbred line testers. In addition, the F_1 testers have the additional advantage of allowing the evaluation of experimental lines against two genetic backgrounds or adaptation types, in a single testcross. In this manner, the probability of incorrectly discarding material can be reduced while minimizing resource expenditures. Experimental lines identified as having superior combining abilities with the F_1 tester can then be advanced for testing with individual inbred line testers to identify specific combing abilities.

REFERENCES

Allison J. C. S., and R. N. Curnow. (1966) On the choice of tester parents for the breeding of synthetic varieties of maize (*Zea mays* L.) Crop Sci. 6: 541-544.

Bartel, A. T., (1949) Hybrid vigor in sorghums. Agron. J. 41: 147-152.

Bernardo, R., (2002) Breeding for quantitative traits in plants. Stemma Press, Woodbury, MN. p. 136.

Bradley R. A., D.E.W. Schumann (1957) The comparison of the sensitivities of similar experiments: Applications. Biometrics 13: 496-510.

Castellanos J. S., A. R. Hallauer, H. S. Cordova. (1998) Relative performance of testers to identify elite lines of corn (*Zea mays* L.) Maydica 43: 217-226.

Cowan J.R. (1943) The value of double cross hybrids involving inbreds of similar and diverse genetic origin. Sci. Agr. 23: 281-296.

Gebrekidan B., and D. C. Rasmusson. (1970) Evaluating parental cultivars for use in hybrids and heterosis in barley. Crop Sci. 10: 500-502.

Gilbert M. L., D. T. Rosenow, H. T. Nguyen. (1996) Utilization of alternate cytoplasms to evaluate combing abilities in a diverse group of sorghum B-lines. Texas J. Agric. Nat. Resour. 9: 51-60.

Hallauer A. R., and E. Lopez-Perez. (1979) Comparisons among testers for evaluating lines of corn. Proc. Annu. Corn & Sorghum Res. Conf. 34: 57-75.

Jenkins M. T. and A. M. Brunson. (1932) Methods of testing inbred lines of maize in crossbred combinations. J. Am. Soc. of Agron. 24: 523-530.

Johnson I.J., and H. K. Hayes. (1936) The combining ability of inbred lines of Golden Bantam sweet corn. J. Am. Soc. of Agron. 28: 246-252.

Kambal A. E., and O. J. Webster. (1965) Estimates of general and specific combining ability in grain sorghum, *Sorghum vulgare* Pers. Crop Sci. 5: 521-523.

Karper R. E., and J. R. Quinby. (1937) Hybrid vigor in sorghum. J. of Heredity 28: 82-91.

Kimber, C.T., (2000) Origins of domesticated sorghum and its early diffusion to India and China. *In* C. W. Smith and R. A. Frederiksen (ed.) Sorghum: origin, history, technology, and production. John Wiley and Sons, Inc., New York.

Lee, R. D., B. E. Johnson, K. M. Eskridge, J. F. Pedersen. (1992) Selection of superior female parents in sorghum utilizing A3 cytoplasm. Crop Sci. 32: 918-921.

Matzinger D. F. (1953) Comparisons of three types of testers for the evaluation of inbred lines of corn. Agron. J. 45: 493-495.

Quinby, J. R. (1963) Manifestations of hybrid vigor in sorghum. Crop Sci. 3: 288-291.

Rawlings J. O., and D. L. Thompson. (1962) Performance level as criterion for the choice of maize testers. Crop Sci. 2: 217-220.

Rooney, W. L., (2004) Sorghum improvement: Integrating traditional and new technology to produce improved genotypes. Adv. in Agronomy 83: 37-109.

Ross W. M., (1969) Performance of three-way grain sorghum hybrids. Proc. Annu. Corn & Sorghum Res. Conf. 24: 129-134.

Ross W.M., and K. D. Kofoid. (1978) Evaluation of grain sorghum R-lines with a single-cross vs. inbred line testers. Crop Sci. 18: 670-672.

Schertz, K. F., and J. W. Johnson. (1984) A method for selecting female parents of grain sorghum hybrids. Crop Sci. 24: 492-494.

Sharma D., J. P. Tandon, N. J. Batra. (1967) Effect of testers on combining ability estimates of maize germplasm complexes. Euphytica 16: 370-376.

Shebeski, L. H. (1966) Quality and yield in hybrid wheat. Can. J. Genet. Cytol. 8: 375-386.

Stephens, J. C., and R. F. Holland. (1954) Cytoplasmic male-sterility for hybrid sorghum seed production. Agron J. 46: 20-23.

Stephens, J. C., and J. R. Quinby. (1952) Yield of a hand produced hybrid sorghum. Agron J. 44: 231-233.

USDA, (2007a) [Online] June 29 Acreage report. National Agricultural Statisitics Service. Available at: http:\\usda.mannlib.cornell.edu/usda/current/Acre/Acre-06-29-2007.pdf

USDA, (2007b) [Online] Grain: World markets and trade. Circular series F6 06-07. Available at: www.fas.usda.gov/psdonline/circulars

Zambezi B. T., E. S. Horner, F. G. Martin. (1986) Inbred lines as testers for general combining ability in maize. Crop Sci. 26: 908-910.

VITA

Daniel Jacob Packer was born in Nuevo Casas Grandes, Chihuahua, Mexico in February of 1978. He graduated from Judson High School in Converse, TX in 1996 and received a B.S. in Plant Biology from Brigham Young University in 2005.

Address:

Department of Soil and Crop Sciences
c/o Dr. William Rooney
TAMU 2474
College Station, TX
77843