

Molecular marker applications in sorghum

G. Ejeta, P.B. Goldsbrough, M.R. Tuinstra, E.M. Grote, A. Menkir, Y. Ibrahim, N. Cisse,
Y. Weerasuriya, A. Melake-Berhan, and C.A. Shaner

Purdue University, West Lafayette, IN 47907

Introduction

The use of DNA-based markers for the genetic analysis and manipulation of important agronomic traits has become an increasingly useful tool in plant breeding. DNA markers have the potential to enhance the operation of a plant breeding program through a number of ways, ranging from finger printing of elite genetic stocks, assessment of genetic diversity, increasing the efficiency of selection for difficult traits, to making environment-neutral selection possible. However, their greatest potential appears to be in accelerating the rate of gain from selection for desirable genotypes and in the manipulation of quantitative trait loci (QTL) that condition complex economic traits. DNA markers also permit plant breeders to correctly map or place the various interacting genes that condition complex agronomic traits. Genetic mapping is essential for effective manipulation of important genes. Effective use of marker-based selection or marker assisted introgression should permit genetic recombination beyond the range possible in traditional breeding. While the possibilities appear limitless, it should be pointed out that the application of the science is in its infancy and it may take some time before it becomes a routine operation of most plant breeding programs.

Molecular markers have been used to identify and characterize QTL associated with several different traits in sorghum including plant height and maturity (Pereira and Lee, 1995), characters concerned with plant domestication (Patterson et al., 1995), disease resistance (Gowda et al., 1995), and drought tolerance (Tuinstra et al., 1996, 1997, 1998). In addition several sorghum linkage maps (Hulbert et al., 1990; Melakeberhan et al., 1993; Xu et al., 1994; Chittenden et al., 1994; Pereira et al., 1994; Lin et al., 1995; Dufour et al., 1996; Boivin et al., 1999) have been generated, but they have not yet been properly integrated to produce a more global and functional map with 10 linkage groups. At Purdue University, we have used molecular markers in our program to both produce a linkage map and to address important research questions in sorghum. The following is a summary of activities in our sorghum

In: B.I.G. Haussmann, H.H. Geiger, D.E. Hess, C.T. Hash, and P. Bramel-Cox (eds.). 2000. Application of molecular markers in plant breeding. Training manual for a seminar held at IITA, Ibadan, Nigeria, from 16-17 August 1999. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, Andhra Pradesh, India.

research program at Purdue University, where we have employed DNA marker technology to address some important research questions.

Assessment of genetic diversity in sorghum using molecular markers

Analyses of the extent and distribution of genetic variation in a crop are essential in understanding the evolutionary relationships between accessions and to sample genetic resources in a more systematic fashion for breeding and conservation purposes. Traditionally, genetic resources in sorghum are classified by taxonomists based on morphological markers. However, these morphological traits used in classification of sorghum to different races are conditioned by a relatively small number of genes. On the other hand, important traits which are related to habitat adaptation and exhibit enormous variability among sorghum germplasm are complex and quantitatively inherited. Hence, classifying germplasm accessions based solely on a few discrete morphological characters may not provide an accurate indication of the genetic divergence among the cultivated genotypes of sorghum. In this study (Menkir et al.,1997), we used molecular markers to analyze genetic diversity in cultivated races of sorghum. We hypothesize that both natural and human selection efforts have contributed to current genetic differences in sorghum and hence landraces of the same race grown in different habitats may have greater genetic dissimilarity than those of different races from the same habitat.

We sampled 190 sorghum accessions from the five major cultivated races , namely *bicolor*, *guinea*, *caudatum*, *kafir*, and *durra*. Accessions representing each race and each of the major geographical centers of distribution were randomly selected from the world collection of sorghum maintained at the International Crop Institute of the Semi-Arid Tropics (ICRISAT). The 190 accessions were divided into 38 sets, in each of which five randomly selected accessions from all five main cultivated races of sorghum were included to ensure parallel comparison of races and minimize biases while scoring marker products. A total of 82 RAPD primers were used for DNA amplification but only 53 primers produced clearly scorable bands generating 220 bands across all sorghum accessions.

A high level of genetic variation was detected among sorghum accessions. The results of our study indicated that genetic diversity within a race was high for race *bicolor*, and *guinea* and low for race *kafir*. Partitioning the genetic variation further revealed that 86% of the total genetic variation occurred among accessions and 14% among races. Examination of the degree of association of accessions with their geographic areas of origin indicated that only 13% of the total genetic variation was attributable to divergence among regions. In spite of the limited differentiation among regions, the extent of genetic diversity within and among regions showed

some trends. Though represented by a large number of accessions, Southern African germplasm exhibited the least amount of genetic diversity suggesting a narrow genetic base of accessions from this region. By contrast, West Africa exhibited a high level of genetic diversity with a least number of accessions. Genetic diversity in Central and Eastern Africa as well as accessions from the Middle East was as high as that observed in accessions from West Africa. In conclusion, our data suggest that molecular markers are suitable to assess genetic diversity and to identify diverse sources in crop germplasm collections. In particular, genetic distance estimates determined by molecular markers help identify suitable germplasm for introgression into breeding stocks. Selecting the most divergent accessions for introgression may increase the chances for extracting suitable inbred lines from backcross populations. Such inbred lines may, in turn, become useful sources of favorable alleles to improve the productivity of varieties and hybrids.

QTL identification and genetic analysis of drought tolerance in sorghum

The development of molecular genetic markers and the use of these markers in quantitative trait loci (QTL) analysis is increasingly becoming a common approach for evaluating the inheritance and feasibility of accelerating gains from selection for complex quantitative traits in crop plants. Drought tolerance is one such trait for which QTL analysis holds great promise. The genetic and physiological mechanisms that condition the expression of drought tolerance in crops are poorly understood. Controlled by many genes and dependent on the timing and severity of moisture stress, drought is one of the more difficult traits to study and characterize. Sorghum is one of the most drought tolerant grain crops and its rich genetic diversity for stress tolerance makes it an excellent crop model and choice for studying the genetic and physiologic mechanisms of drought tolerance. Nonetheless, even in sorghum, direct selection for drought tolerance using conventional approaches has been slow and difficult. A number of physiological and biochemical traits have been implicated to enhance drought tolerance. Yet, only a few of these mechanisms has been demonstrated to be causally related to the expression of tolerance to drought under field conditions. We believe the use of molecular markers and QTL analysis based on carefully managed replicated tests has the potential to alleviate the problems associated with inconsistent and unpredictable onset of moisture stress or the confounding effect of other stresses such as heat. To this end, we conducted several experiments on both phenotypic selection for drought tolerance as well as QTL analysis of drought tolerance in sorghum. We summarize below the highlights of these findings:

Phenotypic selection for drought tolerance

We have made a slow but significant progress via empirical breeding of sorghum for drought tolerance by breaking the trait of drought tolerance into specific phenological stages. Our approach has been to breakdown the complex trait of drought tolerance into simpler components by studying drought stress expressions at specific stages of plant development. We have been particularly interested in midseason (pre-flowering) and late-season (post-flowering) drought expressions in sorghum germplasm. Our rationale is that if individual components associated with a complex trait can be identified, we can measure the contribution of each of the factors or mechanisms independently without the confounding effect of other factors. Using this approach, we have identified sorghum germplasm that are uniquely pre-flowering or post-flowering drought tolerant and few that combine tolerance at both stages. We have developed new improved drought tolerant sorghum lines in diverse and elite germplasm background. Some of these lines have been officially released and distributed to both public and private sorghum research concerns. Several more await release and distribution following further characterization and cataloguing to facilitate specific mode of utility. Our breeding and selection effort was based on reliable phenotypic markers associated with morphological and yield related symptoms that occur at pre-flowering and post-flowering stages of crop development. Some of these marker traits are simply inherited and others appear quantitative rendering them amenable to QTL marker analysis and introgression.

QTL mapping and analysis of drought tolerance

Molecular markers linked to QTL for drought tolerance could be used in increasing efficiency of breeding efforts to select sorghum germplasm with enhanced drought tolerance once these markers are identified through carefully monitored characterization of appropriate germplasm under stress conditions. Such an approach provides a more systematic mode for identifying specific traits that contribute to drought tolerance. Further analysis of these traits could lead to better understanding of the biological basis of drought tolerance. In the last several years, we undertook a number of studies toward this goal using a set of recombinant inbred (RI) sorghum lines especially developed for an array of interdisciplinary evaluation of the genetics and physiology of drought tolerance in sorghum. First, the RI lines were carefully evaluated for response to drought in a series of pre-flowering and post-flowering stress environments. Drought tolerance was estimated in several ways: evaluation of grain yield under drought, stability of yield, rate and duration of grain fill, seed weight, stay green and associated traits. Evaluation of the RI lines indicated segregation of drought tolerance during both developmental stages affirming its genetic basis and suggesting complementary interaction of loci from both parental sources. Second, the RI population was scored for the segregation of

RAPD, RFLP, and SSR markers and these markers were ordered into a genetic map by linkage analysis and used to determine the contribution of the parental genotypes to each of the RI lines. Single factor analysis was used to identify QTL associated with yield and other measures of agronomic performance under drought and non-drought conditions. Several regions of the genome were associated with the expression of yield or yield components under pre-flowering and post-flowering drought, and under fully irrigated conditions. In most cases, the marker allele associated with higher yield under fully irrigated condition was also associated with improved tolerance or agronomic performance under drought. On the other hand, two regions on two separate linkage groups were strongly associated with agronomic performance under pre-flowering drought but not under full irrigation. Similarly, two other regions of the genome on yet two other linkage groups were found to be associated with agronomic performance under post-flowering drought but not under full irrigation. These findings suggest that these loci mediate the expression of pre-flowering (Tuinstra et al., 1996) or post-flowering drought (Tuinstra et al., 1997a) tolerance independent of mechanisms that control yield. Several QTL for staygreen were identified on five linkage groups, however, QTL on three of these linkage groups were also positively associated with grain yield under non-drought conditions. This indicates that there may be a physiological link between the expression of staygreen under post-flowering drought and grain yield under non-drought conditions.

Development of near-isogenic lines (NILs) that differ for drought QTL

Although our QTL analysis identified regions of the sorghum genome that condition the expression of drought tolerance, it provided little information concerning the expression of individual QTL. Analysis of near-isogenic lines that differ at QTL can be an effective approach for the detailed mapping and characterization of individual loci. However, the use of NILs in analysis of important agronomic traits has been limited perhaps because of the time and effort required to develop these lines. We, therefore, developed a procedure for drawing NILs for any region of the genome that can be analyzed with molecular or other genetic factors (Tuinstra et al., 1997b). The procedure utilizes molecular markers to identify heterogeneous inbred families that are isogenic at most loci in the genome from NILs that differ for markers linked to QTL of interest. Using this procedure we developed NILs for several QTL associated with yield under drought environments and other morphological traits associated with drought tolerance.

Evaluation of NILs that differ for drought QTL

The process of identifying linkage between markers and traits in a mapping population followed by test of marker effects in NILs can be powerful and useful to resolve several issues. First, marker linkage to a QTL can be confirmed by examining the phenotype on NILs that only differ for individual QTL. Initial QTL analysis indicates regions of the genome that may contain QTL but the particular phenotypic effects of these loci need to be confirmed. Second, NILs can be used for fine mapping of QTL. Evaluation of a series of NILs that contrast at a specific locus can be used to narrow the genetic interval known to contain the QTL. Third, NILs that differ at a QTL can be used to characterize the expression and function of a specific locus. In our case, we reasoned that NILs differing for QTL associated with drought tolerance can be used to identify the specific mechanism of drought tolerance controlled by each QTL. We focussed on the analysis of NILs contrasting at three loci and evaluated differences in the size of the genomic region differentiating each set of NILs by testing markers flanking each target QTL. Agronomic evaluation of these NILs indicated large differences in yield and seed weight associated with each QTL marker. In most cases, NILs contrasting for as specific locus differed in phenotype as predicted by QTL analysis (Tuinstra et al., 1998). Further analyses indicated that differences in agronomic performance may be associated with effects of heat tolerance, water status, and expression of staygreen suggesting that these loci mediate the expression of drought tolerance via different biological mechanisms. This can be corroborated with careful physiological studies that can be more readily undertaken using NILs than random and unrelated genotypes. We plan to conduct these studies to identify and define the specific mechanisms of drought tolerance mediated by these loci. We believe that the approach of narrowly focussing on specific genomic regions associated with drought tolerance holds promise for developing a clearer understanding of the specific biological basis of this complex trait.

Molecular mapping of striga resistance genes

Field resistance to striga is a complex quantitative trait that has been difficult to address via conventional plant breeding approaches. Successful manipulation of such a trait requires possible breakdown to component parts that lend themselves to genetic manipulation. We have had a striga resistance breeding program at Purdue University for several years and have employed a mix of biotechnological approaches to address this problem. One such approach focussed on mapping and identification of genetic loci associated with resistance to striga. This work is closely integrated with development of laboratory assays that would allow identification of host variants that are resistant to striga because they disrupt essential developmental interaction with the parasite. Our objective is to develop a genetic linkage map

of sorghum with polymorphic markers dispersed throughout the genome so that we would be able to manipulate striga resistance genes on the map with reasonable ease.

We genotyped over 230 DNA markers and generated a fairly dense linkage map of sorghum. The estimated map size is 1628 centiMorgans (cM), with an average interval of 9.5 cM between adjacent loci. The locus for one of the better characterized mechanisms of resistance to striga, namely the low germination stimulant (lgs) production is now mapped at 11.8 cM from an RFLP marker PIO200725 BamH1, and 13.5 cM from SSR17g marker. We also used the linkage map to place putative QTL for striga resistance using phenotypic data from field evaluation of our mapping population against *Striga hermonthica* and *Striga asiatica*. Single marker analyses detected six QTL for resistance to *S. hermonthica* and five QTL for resistance to *S. asiatica*. The QTL detected for resistance to *S. hermonthica* accounted for 37% of the variation in resistance and QTL detected for resistance to *S. asiatica* accounted for 49% of the variation in resistance. Two of these QTL were on the same linkage group as the *lgs* locus. Interval mapping confirmed most of the QTL detected by single marker analysis.

Genetics of cold tolerance in sorghum

Tolerance to early season cold temperature is a major trait that is needed in much of the sorghum production areas of the United States. The trait is needed for early season stand establishment to take advantage of the long crop season that would ensue if the crop is planted and established early. Important advantages have been attributed to seedling vigor, resulting in greater biomass and grain yield in cold and dry environments. We conducted QTL analyses for seedling vigor and cold tolerance using a set of RI lines in our mapping population. The RI lines were evaluated for germination, emergence, and growth under cold temperature and genotyped using RAPD markers. Markers located on two linkage groups were found to be significantly associated with seedling cold tolerance as estimated through visual scores of seedling vigor. Germination at low temperature and emergence and growth under optimum temperature were mostly under separate genetic control. QTL on one linkage group explained 75-80% of the variation for seedling growth. All markers with significant effects on shoot dry weight were associated with seedling growth as measured by height of seedlings three weeks after sowing and with visual scores. The visual score estimates used an estimate expected to integrate germination, emergence, and seedling height, only explained some 25% of the total variation. This suggests the presence of undetected QTL and the need for exhaustively detecting all relevant QTL by generating a series of well controlled and characterized phenotypic data. QTL determination and their potential utility in breeding programs is only as good as the quality of the phenotypic data employed in their generation.

QTL analysis for grain yield in sorghum lines and hybrids

We also employed DNA markers to determine if QTL affecting yield in inbred lines and hybrids are similar. It is generally believed that good inbreds make good hybrids, though the correlation between agronomic performance of lines and testcrosses can be poor. These differences are typically explained by differences in combining ability. Early reports on maize (Jenkins, 1924; Jorgenson and Brewbaker, 1927; Gama and Hallauer, 1977) gave positive association between performance in inbreds and hybrids for phenological traits, but correlations for yield were weak. We evaluated performance of recombinant inbred (RI) sorghum lines and their hybrids in five different environments. The correlation for grain yield between RI lines and hybrids were generally low, but significant associations were identified in high-yield environments. Genetic variation among RI lines and testcrosses was greatest in high yield environments and lowest in stress environments. We evaluated the genetic basis for the inbred-hybrid yield correlation through QTL analysis. We identified two QTL with similar effects on yield in both lines and hybrids. Several other QTL were identified for line or hybrid specific effects explaining perhaps differences in specific combining ability effects among inbreds. The results of this study indicated that QTL for grain yield of sorghum hybrids can be identified in testcross populations; however, these QTL may differ from those associated with performance of lines per se. It also suggests that the power for QTL detection for yield in grain sorghum is the most favorable under high-yielding or low- stress environments.

References

- Boivin, K., M. Deu, J-F. Rami, G. Trouche, and P. Hamon. 1990. Towards a saturated sorghum map using RFLP and AFLP markers. *Theor. Appl. Genet.* 98:320-328.
- Chittenden, M.L., K.F. Schertz, Y-R. Lin, R.A. Wing, and A.H. Paterson. 1994. A detailed RFLP map of *Sorghum bicolor* × *S. Propinquum*, suitable for high density mapping, suggests ancestral duplication of sorghum chromosomes or chromosomal segments. *Theor. Appl. Genet.* 87:925-933.
- Dufour, P., M. Deu, L. Grivet, A. Dhont, F. Paulet, A. Bouet, C. Lanaud, J.C. Glaszmann, and P. Hamon. 1997. Construction of a composite sorghum genome and comparasion with sugarcane a related complex polyploid. *Theor. Appl. Genet.* 94:409-418.
- Gama, E.E.G., and A.R. Hallauer. 1977. Relation between inbred and hybrid traits in maize. *Crop Sci.* 17:703-706.
- Gowda, P.S.B., G.W. Xu, R.A. Frederiksen, and C.W. Magill. 1995. DNA markers for downey mildew resistance genes in sorghum. *Genome* 38:823-826.

- Hulbert, S.H., T.E. Richter, J.D. Axtell, and J.L. Bennetzen. 1990. Genetic mapping and characterization of sorghum and related crops by means of maize DNA probes. *Proc. Natl. Acad. Sci.* 87:4251-4255.
- Jenkins, M.T. 1929. Correlation studies with inbred and crossbred strains of maize. *J. Agric. Res.* 39:677-721.
- Jorgenson, L., and H.E. Brewbaker. 1927. A comparison of selfed lines of corn and first generation crosses between them. *J. Am. Soc. Agron.* 19:819-30.
- Lin, Y-R., K.F. Shertz, and A.H. Paterson. 1995. Comparative analysis of QTL affecting plant height and maturity across the poaceae in reference to an interspecific sorghum population. *Genetics* 141:391-411.
- Melake Berhan, A., S.H. Hulbert, L.G. Butler, and J.L. Bennetzen. 1993. Structure and evolution of the genomes of *Sorghum bicolor* and *Zea mays*. *Theor. Appl. Genet.* 86:598-604.
- Menkir, A., P. Goldsbrough, and G. Ejeta. 1997. RAPD based assessment of genetic diversity in cultivated races of sorghum. *Crop Science* 37:564-569.
- Patterson, A.H., Y. Lin, Z. Li, K.F. Schertz, J.F. Doebley, S.R.M. Pinson, S. Liu, J.W. Stanzel, J.E. Irvine. 1995. Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. *Science* 269:1714-1718.
- Pereira, M.G., and M. Lee. 1995. Identification of genomic regions affecting plant height in sorghum and maize. *Theoretical and Applied Genetics* 90:380-388.
- Pereira, M.G., M. Lee, P. Bramel-Cox, W. Woodman, J. Doebly, and R. Whitkus. 1994. Construction of a RFLP map in sorghum and comparative mapping in maize. *Genome* 37:236-243.
- Tuinstra, M.R., G. Ejeta, and P.B. Goldsbrough. 1997. Heterogeneous inbred family (HIF) analysis: An approach for developing near-isogenic lines that differ at quantitative trait loci. *Theoretical and Applied Genetics* 95:1005-1011.
- Tuinstra, M.R., G. Ejeta, and P.B. Goldsbrough. 1998. Evaluation of near-isogenic sorghum lines contrasting for QTL markers associated with drought tolerance. *Crop Science* 38:835-842.
- Tuinstra, M.R., E.M. Grote, P.B. Goldsbrough, and G. Ejeta. 1996. Identification of quantitative trait loci associated with pre-flowering drought tolerance in sorghum. *Crop Science* 36:1337-1344.
- Tuinstra, M.R., E.M. Grote, P.B. Goldsbrough, and G. Ejeta. 1997. Genetic analysis of post-flowering drought tolerance and components of grain development in sorghum. *Molecular Breeding* 3:439-448.
- Xu, G.W., C.W. Magill, K.F. Schertz, and G.E. Hart. 1994. A RFLP linkage map of *Sorghum bicolor* (L.) Moench. *Theor. Appl. Genet.* 89:139-145.