Genetic Distance among Selected Cotton Genotypes and Its Relationship with $F_2$ Performance


ABSTRACT

Knowledge of genetic diversity and relationships among breeding materials has a significant impact on crop improvement. Association between parental divergence and progeny performance has not been well documented in cotton (Gossypium hirsutum L.). The objectives of this study were to estimate genetic diversity among selected cotton genotypes on the basis of simple sequence repeat (SSR) markers, and to investigate the relationship between genetic diversity and $F_2$-bulk population performance. Five U.S. and four Australian cultivars, and two day-neutral converted lines of G. hirsutum were genotyped by means of 90 SSR primer pairs providing 69 polymorphic marker loci. Genetic distance (GD) between genotypes ranged from 0.06 to 0.34 for the 11 parental genotypes. The highest GD (0.34) was observed between ST474 and the day-neutral converted line B1388. The lowest GD (0.06) was detected between cultivars FM832 and FM975. The GD between day-neutral converted lines and cultivars ranged from 0.26 to 0.34. Among the Australian cultivars, GD ranged from 0.06 to 0.19 while GD among U.S. cultivars varied from 0.10 to 0.22, indicating a narrow genetic base. Significant correlations between agronomic and fiber traits of $F_2$-bulk populations and GD ranged from negative to positive depending on the traits, genetic background, and environment. On the basis of SSR markers, GD revealed a lack of genetic diversity among all genotypes and it was a poor predictor of overall $F_2$ performance. However, when genotypes with maximum range of GD were present, it was a better predictor for some traits.

Information about the degree and distribution of genetic diversity and relationships among breeding materials has a significant effect on crop improvement. Selection of suitable parents is one of the most important criteria used to allocate resources to the most promising crosses and increase the efficiency of breeding programs. Molecular markers increasingly play an important role in crop improvement programs. They have been used to predict genetic variance among inbred lines (Manjarrez-Sandoval et al., 1997), estimate genetic diversity in crops (Wendel et al., 1992; Tatineni et al., 1996), predict plant cultivar rights (Smith and Smith, 1992), classify heterotic groups (Dudley et al., 1991; Senior et al., 1998), study phylogenetic relationships among crops and their wild relatives (Li et al., 2000), analyze pedigrees (Smith et al., 1997), and select desired traits (Young, 1999).

Simple sequence repeats, also called microsatellites, are tandem repetitive DNA sequences. The availability and abundance of SSR markers throughout the cotton genome, their polymorphic nature, codominance, and polymerase chain reaction (PCR)-based assay make SSRs useful in detecting genetic diversity (Reddy et al., 2001).

The usefulness of GD as a predictor of hybrid performance has been studied in several crops. In maize (Zea mays L.) and sunflower (Helianthus annuus L.), significant correlations between GD and hybrid performance were observed by Lee et al. (1989), Smith et al. (1990), Lanza et al. (1997), and Cheres et al. (2000). In contrast, Godshalk et al. (1990) and Dudley et al. (1991) observed weak correlations between marker genotype and hybrid performance in maize. Furthermore, Martin et al. (1995) working with wheat (Triticum aestivum L.) found no association between measures of diversity and hybrid performance. Meredith and Brown (1998) using restriction fragment length polymorphic (RFLP) markers reported that in cotton the correlations between yield of $F_2$ hybrids, heterosis, and GD were very low ($r = 0.07$).

The utilization of heterosis for lint yield and fiber quality in $F_2$ hybrids has eluded most researchers because of inconsistent expression of heterosis and lack of economically effective means of delivering $F_2$ seeds to growers. Nonetheless, it has been demonstrated that $F_2$ hybrids can exhibit superior performance for agronomic and fiber traits when compared with their parental lines. Meredith (1990) reported that $F_2$ hybrids yielded 8% higher than their parents and no differences in adaptation were detected; however, the $F_2$ hybrids had significantly shorter fiber. Moreover, Tang et al. (1992, 1993a,b) observed that almost all $F_2$ hybrids developed from crosses between selected pest-resistant germplasms and cultivars were equal to their highest parent for most fiber traits. Perhaps the greatest motivation to investigate prediction of $F_2$ performance is to spur cotton improvement. Presently, $F_2$-bulk population performance is used in about 24% of commercial cotton breeding programs in the USA for early generation testing and identification of populations upon which to focus selection (Bowman, 2000).

The objectives of this study were to estimate GD on the basis of SSR markers among selected cotton genotypes, and to investigate the association between GD and fiber and agronomic characteristics of $F_2$-bulk populations.

MATERIALS AND METHODS

Plant Materials and Field Experiments

Eleven cotton genotypes, five U.S. cultivars, four cultivars developed by the CSIRO cotton breeding program in Austra-

*Corresponding author (ssaha@msa-msstate.ars.usda.gov).


Abbreviations: GD, genetic distance; SSR, simple sequence repeats.
lia, and two-day-neutral converted lines of *G. hirsutum* derived from photoperiodic wild accessions were selected for this study (Calhoun et al., 1994) (Table 1). The U.S. cultivars are a representative sample of those grown in the Delta region and some represent the most influential breeding programs in terms of genetic contributions to modern cultivars (Bowman et al., 1996). The Australian cultivars were obtained from Cotton Seed International. These four were used because they were offered for sale to growers in the USA and their GD from U.S. cultivars was unknown.

Day-neutral converted line A239 was derived from photoperiodic primitive accession T239 (PI 163693) *G. hirsutum* race *latifolium* from Guatemala. A239 was developed from a cross of ‘Deltapine 16’ by T239 and selected for the day-neutral flowering habit in the F1. A day-neutral selection was then backcrossed three times to T239 and selected for day neutrality after each backcross. Following the third backcross a single high fiber strength plant was selected in the BC2F3. Subsequently, a single high fiber strength plant was selected in the BC3F2. The BC3F2 plant was advanced by bulk increase to the BC3F6. Day-neutral converted line B1388 was derived from accession T1388 (PI 415112) of unknown race from Nicaragua. Line B1388 was developed from a cross of ‘DES 56’ by T1388. In the F1, a single high fiber strength day-neutral plant was selected. Subsequently, a single high fiber strength plant was selected in the F2 and F3. The F3 plant was advanced by bulk increase to the F5.

The mating design used for this experiment was an 11 parent half diallel. Crosses were made during 1998 at the R. R. Foil Plant Science Research Center, at Mississippi State University. The F1 generation was grown in the USDA-ARS winter nursery with the following profile: 7 min at 95°C/H11002 with 10 plants m-2. Plants were planted on 11 and 12 May 1999, and 11 and 15 May 2000. Trials were planted in 9- or 12-m single-row plots with 10 plants m-2 of row and a row spacing of 0.97 m. The soil type at the first location was a Leeper silty clay loam (fine, smectitic, nonacid, thermic, Vertic Epiaqupts) and at the second location was a Marietta loam (fine-loamy, siliceous, active, thermic, Fluvaquentic Eutrudepts) soil. Standard cultural practices were followed during the growing season.

Plots were harvested with a mechanical picker for yield determination on 7 and 8 Oct. 1999 and 20 Sept. and 10 Oct. 2000. Samples containing 50 bolls were hand-harvested both years from each plot before mechanical harvesting. The boll samples were weighed and ginned on a 10-saw laboratory gin for lint percentage and lint weight/seed cotton weight. A lint sample (20 g) from each boll sample represented a fiber sample. Fiber samples were sent to STARLAB Inc. (Knoxville, TN) for measurements of elongation (%), fiber strength (kN m kg⁻¹), 2.5% span length (mm), 50% span length (mm), and micronaire reading by single-instrument testing.

**SSR Markers**

Leaves were collected from field plots in summer 1999. Samples of 20 young leaves from each of the 11 parents were collected and freeze-dried (Saha et al., 1997). The freeze-dried tissue was then ground into dry powder and stored at −20°C. Freeze-dried, powdered tissue was ground further with liquid nitrogen just before extraction following the manufacturer’s protocol. DNA was isolated from 40 mg (dry weight) of cotton leaf tissue with the DNeasy Plant mini kit (Qiagen, Santa Clarita, CA) following the manufacturer’s protocol. The DNA concentration was measured with a spectrophotometer. The purity of the uncut DNA samples was also visually determined by means of agarose gel electrophoresis.

The 90 SSR primer pairs were received from Research Genetics Inc., Huntsville, AL. The PCR was performed with 80 ng DNA as template, 0.15 μM each fluorescent-labeled and nonlabeled specific SSR primer pairs labeled either HEX (4,7,2,’4,5,7,’hexachloro-6-carboxyfluorescein), NED (7’-benzo-5’fluoro-2’,4,7-trichloro-5’-carboxyfluorescein), or FAM (6-carboxyfluorescein), 0.2 mM each dNTP, 1× GeneAmp PCR Gold Buffer, 2.5 mM MgCl₂, and 0.5 units AmpliTaq Gold (Perkin-Elmer, Norwalk, CT) DNA polymerase in a 10-μL reaction solution. PCR was carried out in a thermal cycler with the following profile: 7 min at 95°C, followed by 15 s at 94°C for denaturation, 30 s at 55°C for annealing and 2 min at 72°C for synthesis. The PCR amplification was continued for 40 cycles with a final extension for 30 min at 72°C. Samples were stored at −20°C until needed.

Capillary electrophoretic analysis of fluorescently-labeled, amplified DNA markers were visualized as peaks on electropherograms with the automated ABI PRISM 310 Genetic Analyzer (Perkin-Elmer Corp., Norwalk, CT) equipped with genotyping and Genescan analysis software (PE Applied Biosystems, Foster City, CA, USA). Because of the high sensitivity of the system, a DNA marker was considered valid if it had a peak height of at least 100 fluorescent units and plus or minus one base size difference with the nearest DNA fragment peak. Fluorescently labeled PCR products (1:30 diluted in sterile water) were mixed in 10 μL formamide (AMRESCO Inc., Solon, OH) and 0.2 μL of an internal size standard DNA labeled with a ROX dye, denatured at 95°C for 5 min, kept on ice at least for 4 min and loaded on the automated ABI PRISM 310 Genetic Analyzer. Variation in the amplified products was compared with an internal size standard DNA labeled with the ROX dye different from that of the sample. Computer-assisted analysis of the data was performed with

---

**Table 1. Pedigrees and origin of cotton cultivars and day-neutral converted lines used to estimate genetic distance and its relationship with F₂-bulk population performance.**

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Origin</th>
<th>Pedigree</th>
</tr>
</thead>
<tbody>
<tr>
<td>A239</td>
<td>Guatemala</td>
<td>DP16/PI 163693/PI 163693×3</td>
</tr>
<tr>
<td>B1388</td>
<td>Nicaragua</td>
<td>DES56/PI 415112</td>
</tr>
<tr>
<td>DP90</td>
<td>Delta &amp; Pine Land Co. (US)</td>
<td>DP16/DP Smoothleaf/DP45/3/DES56</td>
</tr>
<tr>
<td>DP90</td>
<td>Delta &amp; Pine Land Co. (US)</td>
<td>DP6516/DP6582</td>
</tr>
<tr>
<td>FM832</td>
<td>Cotton Seed International (AU)</td>
<td>Confidential</td>
</tr>
<tr>
<td>FM963</td>
<td>Cotton Seed International (AU)</td>
<td>Confidential</td>
</tr>
<tr>
<td>FM975</td>
<td>Cotton Seed International (AU)</td>
<td>Confidential</td>
</tr>
<tr>
<td>FM999</td>
<td>Cotton Seed International (AU)</td>
<td>Confidential</td>
</tr>
<tr>
<td>PM1560</td>
<td>Paymaster Cottonseed (US)</td>
<td>DES119/La434-RKR</td>
</tr>
<tr>
<td>SG501</td>
<td>Sure Grow Seed, Inc. (US)</td>
<td>DES119/DES237-7</td>
</tr>
<tr>
<td>ST474</td>
<td>Stoneville Pedigreed Seed Co. (US)</td>
<td>DES119/ST453</td>
</tr>
</tbody>
</table>

---

The matching design used for this experiment was an 11 parent half diallel. Crosses were made during 1998 at the R. R. Foil Plant Science Research Center, at Mississippi State University. The F1 generation was grown in the USDA-ARS winter nursery at Tecoman, Colima, Mexico, to produce F2 seed. Seed from the 54 F2 hybrids (one cross was lost) and the eleven at 72°C for synthesis. The PCR amplification was continued for 40 cycles with a final extension for 30 min at 72°C. Samples were stored at −20°C until needed.

Capillary electrophoretic analysis of fluorescently-labeled, amplified DNA markers were visualized as peaks on electropherograms with the automated ABI PRISM 310 Genetic Analyzer (Perkin-Elmer Corp., Norwalk, CT) equipped with genotyping and Genescan analysis software (PE Applied Biosystems, Foster City, CA, USA). Because of the high sensitivity of the system, a DNA marker was considered valid if it had a peak height of at least 100 fluorescent units and plus or minus one base size difference with the nearest DNA fragment peak. Fluorescently labeled PCR products (1:30 diluted in sterile water) were mixed in 10 μL formamide (AMRESCO Inc., Solon, OH) and 0.2 μL of an internal size standard DNA labeled with a ROX dye, denatured at 95°C for 5 min, kept on ice at least for 4 min and loaded on the automated ABI PRISM 310 Genetic Analyzer. Variation in the amplified products was compared with an internal size standard DNA labeled with the ROX dye different from that of the sample. Computer-assisted analysis of the data was performed with...
Fig. 1. Dendogram presenting the association of 11 cotton genotypes determined by Unweighted Paired Group Method using Arithmetic Averages (UPGMA) cluster analysis of 69 polymorphic SSR marker loci. GeneScan software using local southern method. Additional parameters in the ABI 310 system were injection time 10 to 15 s, electrophoresis voltage 13 kV, injection voltage 15 kV, collection time 26 min, run temperature 60°C, and syringe pump time 180 s.

**Data Analysis**

Final analyses were performed on a total of 90 primer pairs. Peaks, representative of bands in the electropherograms were then coded as 1 or 0 for presence or absence of the band, respectively. Genetic distances between all pairs of parents were calculated with the PAUP* 4.0b5 software (Swofford, 2000) as GD = 1 - $S_{xy}$ according to the method developed by Nei and Li (1979):

$$S_{xy} = 2 \frac{n_{xy}}{n_x + n_y},$$

where $S_{xy}$ is a measure of genetic similarity between pairs of parents, $n_{xy}$ is number of bands common in parents $X$ and $Y$, $n_x$ is number of bands in parents $X$, and $n_y$ is number of bands in parents $Y$.

A phylogenetic tree (Fig. 1) and matrix of GD (Table 2) were constructed by the Unweighted Pair Group Mean Average (UPGMA) method of Saitou and Nei (1987) by PAUP* 4.0b5 software (Swofford, 2000). Heterosis of F2-bulk population for all traits was calculated by comparing each F2-bulk population with the respective midparent (F2-MP) mean per replication at each location.

Environments were considered as a combination of years and locations. Combined analysis of variance over environments revealed significant genotype × environment interactions. Therefore, analyses of variance by individual environments were performed with the MIXED procedure of SAS (SAS Institute, 2000) for each set of data to test for significant differences ($P < 0.05$) among the parents and F2-bulk populations. Replications were considered random effects. Entries (Parents and F2s) were considered fixed effects. Least Significant Differences (LSD) were calculated for pair-wise comparisons among entries (i.e., parents and F2 bulk) in each environment by means of the restricted maximum likelihood (REML) estimator of the corresponding standard error of the mean differences (Littell et al., 1996). Entry means over replications within environments were used in correlation analyses with GD. Correlations between GD and the agronomic and fiber traits were estimated by PROC CORR of SAS (SAS Institute, 2000). For this analysis, data was divided into six subsets as follows: Set 1—U.S. cultivars × U.S. cultivars (US × US), Set 2—Australian cultivars × Australian cultivars (AU × AU), Set 3—Australian cultivars × U.S. cultivars (AU × US), Set 4—U.S. cultivars × day-neutral converted lines (US × day neutral), Set 5—Australian cultivars × day-neutral converted lines (AU × day neutral), and Set 6—all crosses.

### RESULTS AND DISCUSSION

**Genetic Diversity among the Parental Lines**

Sixty of the 90 SSR primer pairs amplified 69 polymorphic marker loci among 11 cotton genotypes of diverse genetic background (Table 1). Nine primer pairs (CML66, BNL530, BNL1417, BNL1551, BNL1897, BNL2440, BNL2448, BNL2646, and BNL3599) amplified two loci each and 51 primer pairs (CML43, CML63, CML68, BNL119, BNL169, BNL193, BNL256, BNL285, BNL409, BNL597, BNL632, BNL840, BNL1053,

---

Table 2. Genetic distance coefficients calculated for 10 cotton genotypes from 69 SSR marker loci.

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>A239</th>
<th>B1388</th>
<th>DP50</th>
<th>DP90</th>
<th>FM832</th>
<th>FM963</th>
<th>FM975</th>
<th>FM989</th>
<th>PM1560</th>
<th>SG501</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1388</td>
<td>0.21†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DP50</td>
<td>0.27</td>
<td>0.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DP90</td>
<td>0.31</td>
<td>0.26</td>
<td>0.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FM832</td>
<td>0.30</td>
<td>0.30</td>
<td>0.12</td>
<td>0.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FM963</td>
<td>0.31</td>
<td>0.31</td>
<td>0.18</td>
<td>0.14</td>
<td>0.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FM975</td>
<td>0.30</td>
<td>0.28</td>
<td>0.12</td>
<td>0.09</td>
<td>0.06</td>
<td>0.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FM989</td>
<td>0.29</td>
<td>0.28</td>
<td>0.10</td>
<td>0.11</td>
<td>0.07</td>
<td>0.17</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PM1560</td>
<td>0.28</td>
<td>0.26</td>
<td>0.10</td>
<td>0.17</td>
<td>0.17</td>
<td>0.19</td>
<td>0.15</td>
<td>0.15</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>SG501</td>
<td>0.33</td>
<td>0.30</td>
<td>0.14</td>
<td>0.16</td>
<td>0.16</td>
<td>0.18</td>
<td>0.13</td>
<td>0.14</td>
<td>0.12</td>
<td>0.16</td>
</tr>
<tr>
<td>ST474</td>
<td>0.31</td>
<td>0.34</td>
<td>0.16</td>
<td>0.22</td>
<td>0.18</td>
<td>0.19</td>
<td>0.17</td>
<td>0.16</td>
<td>0.15</td>
<td>0.16</td>
</tr>
</tbody>
</table>

† Genetic Distance coefficients calculated according to the formula GD = 1 - $S_{xy}$. 

---
BNL1059, BNL1064, BNL1162, BNL1231, BNL1414, BNL1423, BNL1440, BNL1495, BNL1672, BNL1681, BNL1694, BNL1707, BNL2499, BNL2550, BNL2590, BNL2597, BNL2895, BNL2986, BNL3065, BNL3084, BNL3090, BNL3264, BNL3279, BNL3280, BNL3368, BNL3383, BNL3400, BNL3441, BNL3442, BNL3479, BNL3482, BNL3649, BNL3792, BNL3816, BNL3895, BNL3955, BNL3971, and BNL3994) amplified one locus each. Similar results were previously obtained by Liu et al. (2000a, b) who mapped many of these SSR loci to different chromosomes. A total of 139 different alleles were amplified by 60 SSR primer pairs yielding an average of two SSR alleles per marker locus. The SSR markers used in this study were reported to be present on at least 13 of the 26 chromosomes of the cotton genome. Approximately 55% of the SSR markers have been assigned to the A genome and 46% to the D genome of tetraploid cotton.

Genetic distance coefficients based on the polymorphic SSR marker loci ranged from 0.06 to 0.34 for the 11 parental lines (Table 2). The highest GD (0.34) was detected between the cultivar ST474 and the day-neutral converted line B1388. The lowest GD (0.06) was observed between Australian cultivars, FM832 and FM975. The GD between cultivars and day-neutral converted lines ranged from 0.26 to 0.34. Among the U.S. cultivars, GD ranged from 0.10 to 0.22 while GD of Australian cultivars varied from 0.06 to 0.19, indicating the presence of a narrow genetic base among the Australian and U.S. cultivars. Multani and Lyon (1995) observed GD of 0.01 to 0.08 among nine Australian cultivars which also showed a lack of genetic diversity. Iqbal et al. (1997) also found very high genetic similarity of 0.82 to 0.93 among 17 G. hirsutum cultivars on the basis of random amplified polymorphic DNA (RAPD) markers. On the basis of SSR markers, Ulloa et al. (1999) observed a GD of 0.18 within Acala and Delta cottons while the GD within the Pima PS series was approximately 0.16. The monoculture of a few successful cotton cultivars and the extensive use of them as parents in breeding programs has limited genetic diversity, possibly contributing to yield stagnation in the 1990s (Van Esbroeck et al., 1998). Furthermore, Bowman (2000) suggested that the present situation of high genetic uniformity in cotton is unlikely to change. He reported that genetic uniformity of cotton cultivars reflects reselection within cultivars, disregard for coefficient of parentage in commercial breeding programs, and reduced efforts in germplasm enhancement. Low GD values indicate that it is necessary to introgress new alleles into the U.S. cotton germplasm base to enhance its genetic diversity.

The GD between day-neutral converted lines of wild race stocks, A239 and B1388, was 0.21. This is somewhat surprising because their recurrent parental accessions were collected in different countries, and they have different donor parents for the day-neutral flowering trait (DP16 and DES56, respectively). These results are in agreement with the findings of Liu et al. (2000a), who reported a very narrow genetic base in a subset of day-neutral selections from crosses of race-stock accessions and upland cultivars compared with the G. hirsutum genetic standard TM-1. The lack of diversity between A239 and B1388 could be due to the fact that these accessions were selected for the same traits of insensitivity to photoperiod and high fiber strength. During this process, large chromosomal blocks surrounding genes for fiber strength, earliness, and day neutrality may be introgressed as common alleles in both lines. Lack of enough recombination in portions of the cotton genome could be also responsible for the small GD.

The UPGMA tree generated from genetic distance coefficients, grouped the 11 cotton genotypes into two major clusters (Fig. 1). Cluster A (denoted as Subclusters A1-A4 in Fig. 1) represents the cultivars, and Cluster B the day-neutral converted lines. Within Cluster A, there were four subclusters, A1 (ST474), A2 (‘SG501’, ‘PM1560’, and ‘DP50’), A3 (FM975, FM832, ‘FM989’, and ‘DP90’), and A4 (‘FM963’) (Fig. 1). Although PM1560, SG501, and ST474 have one parent in common, ‘DES119’ (Calhoun et al., 1994) (Table 1), ST474 was in a separate subcluster. DP 90 was included with the Australian cultivars in the A3 cluster. DP 90 was more closely related to the Australian cultivars compared with the U.S. cultivars indicating that the Australian cultivars and DP 90 may have a similar genetic background, or DP 90 is a frequent parent in the pedigrees of Australian cultivars. Subcluster A4 consists only of the Australian cultivar, FM 963. The Australian cultivars were distinguished by clustering close together compared with the U.S. cultivars, indicating their origin from similar genetic backgrounds. Long-term selection for similar agronomically desirable traits may have caused the cultivars to be genetically more uniform and thus to cluster close together. Even though pedigree information on the Australian cultivars was not available, the majority of the groups resulting from cluster analysis were as expected on the basis of their pedigree and geographic information which further supports the strength of SSR markers for the detection of GD.

Relationship of Genetic Distance with F2-Bulk Performance

There were significant differences (P < 0.001) among entries (i.e., parents and F2 bulk) for all agronomic and fiber traits in each environment. There were changes in the rank of the crosses from environment to environment. The number of F2-bulk population means differing from environment to environment demonstrated the strong impact of environment on fiber properties, yield, and yield components (Table 3). The effects of severe dry conditions in Environments 3 and 4 during 2000 perhaps affected the performance of F2-bulk populations, since several of them performed better than their mid-parents in agronomic and fiber traits.

Correlations between agronomic and fiber traits of F2-bulk populations and GD across environments varied in magnitude and direction from −0.99 to 0.90 (Table 4). Midparent heterosis for lint percentage and fiber elongation were significantly correlated with GD in all environments for Set 1. Lint percentage and mid-parent heterosis for seed cotton yield and lint yield were signifi-
Table 3. Number of F2-bulk population means out of 54 significantly higher than the mid-parent (±LSD) for yield components and fiber properties grown at four environments in Mississippi during 1999 and 2000.

<table>
<thead>
<tr>
<th>Traits</th>
<th>Environment 1</th>
<th>Environment 2</th>
<th>Environment 3</th>
<th>Environment 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seed cotton yield</td>
<td>10</td>
<td>18</td>
<td>42</td>
<td>42</td>
</tr>
<tr>
<td>Lint yield</td>
<td>9</td>
<td>12</td>
<td>38</td>
<td>40</td>
</tr>
<tr>
<td>Lint Percentage</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Boll Size</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Fiber Micronaire</td>
<td>6</td>
<td>3</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>Fiber Elongation</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Fiber strength</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Fiber 50% span length</td>
<td>7</td>
<td>5</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Fiber 2.5% span length</td>
<td>12</td>
<td>14</td>
<td>13</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 4. Correlations between agronomic traits of F2-bulk populations and GD for six sets of parental cotton genotypes.

<table>
<thead>
<tr>
<th>Environment</th>
<th>Trait</th>
<th>Set 1†</th>
<th>Set 2</th>
<th>Set 3</th>
<th>Set 4</th>
<th>Set 5</th>
<th>Set 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Seed cotton yield</td>
<td>0.08</td>
<td>0.39</td>
<td>0.23</td>
<td>-0.25</td>
<td>-0.04</td>
<td>-0.57***</td>
</tr>
<tr>
<td>2</td>
<td>MPHseed cotton yield</td>
<td>-0.32</td>
<td>-0.12</td>
<td>0.27</td>
<td>-0.10</td>
<td>0.03</td>
<td>-0.23</td>
</tr>
<tr>
<td>3</td>
<td>MPH lint yield</td>
<td>-0.26</td>
<td>0.04</td>
<td>0.58</td>
<td>0.90**</td>
<td>0.62***</td>
<td>0.59***</td>
</tr>
<tr>
<td>4</td>
<td>Lint yield</td>
<td>-0.46</td>
<td>-0.21</td>
<td>0.09</td>
<td>0.36</td>
<td>-0.59</td>
<td>-0.61***</td>
</tr>
<tr>
<td>5</td>
<td>Fiber Micronaire</td>
<td>0.34</td>
<td>0.79</td>
<td>0.68***</td>
<td>-0.07</td>
<td>0.17</td>
<td>0.36***</td>
</tr>
<tr>
<td>6</td>
<td>Fiber Elongation</td>
<td>0.14</td>
<td>0.05</td>
<td>0.60**</td>
<td>0.47</td>
<td>0.23</td>
<td>0.48***</td>
</tr>
<tr>
<td>7</td>
<td>Fiber strength</td>
<td>-0.18</td>
<td>1.14</td>
<td>-0.24</td>
<td>0.30</td>
<td>0.87**</td>
<td>0.30***</td>
</tr>
<tr>
<td>8</td>
<td>Fiber 50% span length</td>
<td>0.41</td>
<td>0.40</td>
<td>0.54*</td>
<td>0.53</td>
<td>0.65</td>
<td>0.26</td>
</tr>
<tr>
<td>9</td>
<td>Fiber 2.5% span length</td>
<td>0.18</td>
<td>0.34</td>
<td>0.29</td>
<td>-0.14</td>
<td>-0.01</td>
<td>0.67***</td>
</tr>
</tbody>
</table>

* Indicates significance at P = 0.05.
** Indicates significance at P = 0.01.
*** Indicates significance at P = 0.001.
† Set 1, US × US; Set 2, AU × AU; Set 3, AU × US; Set 4, US × day-neutral; Set 5, AU × day-neutral; Set 6, all 54 crosses.
were also significantly correlated with GD in most of the environments in Set 6.

Correlations between agronomic and fiber traits of F2-bulk populations and GD were not consistent across environments and sets. Therefore, they were not generally useful for F2 performance prediction. However, it was a better predictor of traits such as mid parent-heterosis for seed cotton yield and fiber strength in Set 6 whose genotypes exhibited a maximum range of GD values. Our findings in cotton agree with Melchinger (1993) who concluded that GD estimates based on markers randomly arranged across the maize genome were of no value in predicting hybrid performance. Additionally, Bernardo (1992) stated that molecular marker heterozygosity would be most valuable for predicting hybrid performance in crop species under conditions such as strong dominance effects, and high trait heritability, while, Charcosset et al. (1991) anticipated the need for linkage disequilibrium between marker loci and quantitative traits (QTL) for marker heterozygosity and heterosis to be associated. In cotton, some of these conditions may not be present, thus affecting the ability of markers to predict F2-bulk population performance. For instance, we did not apparently have sufficient linkage between markers and QTLs to predict population performance. Furthermore, correlations of GD based on marker heterozygosity with performance and heterosis differed from one trait to another and depended on the genetic background of the germplasm. Lee et al., 1989; Godshalk et al., 1990; Melchinger et al., 1990; Boppenmaier et al., 1993, Zhang et al., 1996; Xiao et al., 1996; Saghai Maroof et al., 1997). In this study, the direction of the correlations often changed among crosses and environments. In Set 6, seed cotton yield, lint yield, lint percentage, fiber strength, and 2.5% fiber span length were the only traits for which the direction of the correlation between GD and F2 population performance was the same in all environments. Midparent heterosis for seed cotton yield was positively correlated with GD in Set 3 and 6; however, lint percentage was negatively correlated with GD in Set 6 and positively correlated in Set 3. A negative association between GD and mid parent heterosis for lint percentage was observed in Set 1 in all environments. Perhaps heterosis was expressed in some traits at some environments for certain genetic backgrounds.

In regards to fiber properties, significant negative correlations for micronaire reading were observed only in Set 5 in two environments indicating that day-neutral converted line A239, which has a low micronaire reading, could be used to develop lower micronaire reading germplasm. This also agrees with Cheatham (2001) who observed that predicted micronaire values for crosses with A239 decreased in advanced generations of selection. A negative correlation was observed between fiber elongation and GD in Set 1 despite parents with desirable fiber elongation.

In summary, commercial cultivars exhibited little genetic diversity with GD ranging from 0.06 to 0.22. The GD between the day-neutral converted lines A239 and B1388 was 0.21, which was similar to the greatest GD between commercial cultivars. The day-neutral converted lines, A239 and B1388, had lower yield, lower lint percentage, and shorter and stronger fibers than the cultivars, but their fiber strength were nearly as high as that of FM 832. Crosses of A239 and B1388 with the cultivars should gain alleles for higher fiber strength, but to use the strength genes from A239 and B1388, careful attention should be given to fiber length, lint percentage, and lint yield to avoid negatively affecting these traits.

An assessment of the usefulness of SSR markers in breeding cotton for yield improvement and fiber quality needs further consideration. We must acknowledge that a limited number of SSR markers were used to make this experiment cost effective. Perhaps more SSR markers covering all 26 chromosomes and at higher density are needed to better predict F2 population performance with GD. Improving the association between SSR marker diversity and F2-bulk population performance also would require SSR markers linked to QTLs for agronomic traits and fiber properties. Our results also demonstrated that the performance of F2-bulk populations does not always depend on the GD, but on the genetic background of the parental germplasm. Genetic diversity has to be incorporated in any breeding program with a proper plan, so that maximum potential can be achieved by incorporating favorable QTLs in a crossing program and removing the unfavorable ones from parental lines (Liu and Wu, 1998). Finally, our study also revealed very low genetic diversity among improved commercial cultivars and, thus the need for new germplasm introgression.

ACKNOWLEDGMENTS

The authors express thanks to Mr. Douglas Dollar, Dr. Mehmet Karaca, Ms. Chethana Bezwada, and Mr. Russ Hayes for their technical assistance. We also thank Drs. Roy G. Cantrell, Michael Lee, James B. Holland, Mauricio Ulloa, Lloyd May, and three anonymous reviewers for their critical evaluation of the manuscript and helpful comments. Finally, we are grateful to Dr. Clarence E. Watson, Jr., for his valuable assistance with the statistical analyses.

REFERENCES

Cheatham, C.L. 2001. Genetics and combining ability of yield and fiber properties associated with selected American and Australian
cotton genotypes. Ph.D. diss. (Diss. Abstr. Int. 02B:608), Mississippi State Univ., Mississippi State, MS.


