

Pedigree- vs. DNA Marker-Based Genetic Similarity Estimates in Cotton

Guillermo Van Becelaere, Edward L. Lubbers, Andrew H. Paterson, and Peng W. Chee*

ABSTRACT

Knowledge of genetic diversity and relationships among breeding materials is essential to the improvement of crop species. Genetic similarity estimates among cultivars are helpful to select parental combinations for segregating populations so as to maintain genetic diversity in a breeding program. The objective of this study was to determine the correspondence between pedigree- and restriction fragment length polymorphism-based genetic similarity (RFLP-GS) estimates for a set of 36 Upland cotton (*Gossypium hirsutum* L.) cultivars. Coefficients of parentage (COPs) and genetic similarity estimates based on 261 codominant RFLP markers for all possible pairs of cultivars were compared. A significant though moderate association ($r = 0.41$, $P < 0.001$) was detected between the COP and RFLP-GS matrices. Spearman's rank correlation for the 142 pairs of related cultivars ($\text{COP} \geq 0.1$) was somewhat higher ($r_s = 0.53$, $P < 0.001$). There was a significant linear relationship between COP and RFLP-GS for the pairs of related cultivars; however, the coefficient of determination was low ($R^2 = 0.25$), indicating that the COP only explained a small portion of the variation observed for RFLP-GS. COP and RFLP-GS estimate different types of genetic resemblance; however, the moderate association may have also resulted from violations to the assumptions made when computing COP. RFLP-GS is a more accurate estimate of true genetic resemblance among cotton cultivars. Nevertheless, the pedigree- and RFLP-based dendrograms were somewhat similar, suggesting that pedigree information will continue to be useful to inexpensively identify diverse parents in a breeding program.

KNOWLEDGE of genetic diversity and relationships among breeding materials is essential to the plant breeder in the efficient improvement of crop species. Genetic similarity (or genetic distance) estimates among genotypes are helpful in selecting parental combinations for segregating populations so as to maintain genetic diversity in a breeding program. These estimates are, in turn, estimates for availability of alternate alleles for desirable traits, which is the basis for long-term selection gains. Crosses between genetically divergent parents are expected to have a larger genetic variance among progenies than crosses between closely related parents (Messmer et al., 1993), increasing the opportunity for selecting rare progenies that may be superior. Despite this, the repeated crossing of closely related elite parents to develop new elite cultivars is common in many crop species and has likely narrowed their genetic base, increasing their vulnerability to potentially widespread losses from diseases and pests.

Genetic similarity among genotypes can be estimated

by different approaches, which include the use of pedigrees and DNA fingerprinting. The COP measures the degree of genetic relatedness among genotypes based on pedigree information by estimating the probability that a random allele from one genotype is identical by descent to a random allele of another genotype at the same locus (Kempthorne, 1969). The accuracy of COP depends on the availability of reliable and detailed pedigree records; at times these records may be either unavailable or unreliable. The calculations also do not account for the effects of selection, mutation, and genetic drift and require several simplifying assumptions that are generally not met. Nevertheless, pedigree information offers plant breeders a simple and inexpensive way to estimate genetic relatedness among breeding materials.

DNA markers, such as RFLPs, amplified fragment length polymorphisms (AFLPs), random amplified polymorphic DNAs (RAPDs), and simple sequence repeats (SSRs), can directly assess DNA sequence variation. DNA-based genetic similarity measures the degree of genetic relatedness among genotypes by estimating the proportion of alleles that are alike in state, thus differing from COP in the type of genetic resemblance estimated. As such, they provide a more precise estimate of genetic similarity among genotypes and do not require the assumptions inherent to pedigree analysis. All the DNA markers mentioned above can provide a large number of polymorphic loci and have been utilized to estimate genetic similarity among genotypes in crop species (Graner et al., 1994; Kim and Ward, 1997; Messmer et al., 1993).

The level of association between pedigree- and DNA marker-based genetic similarity estimates may vary among different crop species. In maize (*Zea mays* L.), for example, most studies found a close association between COP and RFLP-based genetic similarity (RFLP-GS) (Messmer et al., 1993). However, in other species such as wheat (*Triticum aestivum* L.) (Kim and Ward, 1997), barley (*Hordeum vulgare* L.) (Graner et al., 1994), and oat (*Avena sativa* L.) (O'Donoghue et al., 1994) only moderate to low associations have been observed between both estimators. Therefore, it is important to determine within each species whether pedigree- and DNA marker-based estimates of genetic similarity provide similar information about the genetic relationships among germplasm.

The pedigrees of cotton cultivars have been assembled and published (Calhoun et al., 1997) and have been used in selecting parents for crossing blocks; whereas large-scale DNA marker data were not available until recently. Bowman et al. (1996) reported an average COP of 0.07 for 260 Upland cotton cultivars released

G. Van Becelaere, E.L. Lubbers, and P.W. Chee, Dep. of Crop and Soil Sciences, Univ. of Georgia, Tifton, GA 31793; A.H. Paterson, Plant Genome Mapping Lab., Univ. of Georgia, Athens, GA 30602. Received 7 Dec. 2004. *Corresponding author (pwchee@uga.edu).

Published in Crop Sci. 45:2281–2287 (2005).

Plant Genetic Resources

doi:10.2135/cropsci2004.0715

© Crop Science Society of America

677 S. Segoe Rd., Madison, WI 53711 USA

Abbreviations: COP, coefficient of parentage; RFLP-GS, restriction fragment length polymorphism-based genetic similarity.

in the USA between 1970 and 1990, which indicated a greater level of genetic diversity in cotton than in most crops. However, there are indications that the level of genetic diversity of cotton, as estimated by the average COP, may be overestimated (Van Esbroeck et al., 1999). In fact, studies using isozymes (Wendel et al., 1992) and DNA markers (Brubaker and Wendel, 1994) found that there is very little genetic diversity in modern Upland cotton cultivars. This inconsistency casts doubt in regard to the correspondence of COP and RFLP-GS in cotton, although this remains to be tested. The objective of this study was to determine the correspondence between pedigree- and RFLP-based genetic similarity estimates for a set of 36 Upland cotton cultivars.

MATERIALS AND METHODS

Thirty-six Upland cotton cultivars, selected on the basis of historical importance and pedigree availability, were included in this study. The cultivars contain representatives from the four primary production areas of commercial cotton grown in the USA: Western, High Plains, Midsouth, and Southeast. The pedigrees and growing regions of the cultivars are presented in Table 1. All cultivars were released in the USA between 1970 and 1990.

Coefficients of parentage for all possible pairs of cultivars were calculated by Bowman et al. (1997) based on pedigree information reported by Calhoun et al. (1997). The COPs were calculated under the assumptions described by Murphy et al. (1986), i.e., ancestors and cultivars with unknown pedigree were unrelated to each other, cultivars derived from crosses obtained half their genes from each parent, all parents

were homozygous and homogeneous, and COP between a cultivar and a reselection was 0.75 [a compromise between an outcross to an unknown (COP = 0.5) and a self pollination (COP = 1)].

The cultivars were assayed with 261 codominant RFLP markers previously mapped by Reinisch et al. (1994), chosen to provide even coverage of the cotton genome. DNA extraction, electrophoresis, blotting, and hybridization were performed as described by Reinisch et al. (1994). Genetic similarity based on RFLP markers was estimated using GDA software package (Lewis and Zaykin, 2001) for all possible pairs of cultivars using Nei (1978) unbiased identity:

$$\hat{I} = \hat{G}_{XY} / \sqrt{\hat{G}_X \hat{G}_Y}$$

where \hat{G}_X and \hat{G}_Y are the averages of $(2n_x J_x - 1)/(2n_x - 1)$ and $(2n_y J_y - 1)/(2n_y - 1)$ over the number of loci studied, respectively, and $\hat{G}_{XY} = \hat{J}_{XY}$ (Nei, 1978). J_x , J_y , and J_{XY} are the averages of $\sum x_i^2$, $\sum y_i^2$, and $\sum x_i y_i$ over all the loci studied, respectively, and x_i and y_i are the allele frequencies of the i th allele in cultivars X and Y , respectively (Nei, 1972). Nei's unbiased identity estimates can be greater than 1 on very rare occasions. This is caused by sampling error and may occur if the number of the sampled individuals is large. Nei suggests that these values should be replaced with 1. With this as given, both RFLP-GS and COP values range from 0 (completely unrelated) to 1 (identical).

Unweighted pair group method using arithmetic averages (UPGMA) cluster analyses of the COP and RFLP-GS matrices were performed to compare the relationships among cultivars as revealed by pedigrees and RFLP markers. Dendrograms were obtained using the SAHN clustering routine of NTSYS-pc (Rohlf, 2004). Simple correlation analysis was conducted to

Table 1. Pedigrees and growing regions of the 36 Upland cotton cultivars used in this study.

Cultivar	U.S. region	Pedigree†
Acala 1517-70	western	B1413/Hopicala
Acala 1517-75	western	Acala 688/Acala 9608
Acala 1517-88	western	Acala 1517-77BR/DP 70
Acala Maxxa	western	T7538/S4959
Acala SJ-3	western	C6TE/NMB7378
Acala SJ-4	western	C6TE/NMB3080
Blightmaster A-5	High Plains	Stormmaster/Stoneville 20/Acala 5675/3/Stormmaster
Coker 310	Southeast	CKR 100 Staple/DP 15
Coker 312	Southeast	Selection CKR 310
Deltapine 50	Midsouth	DP 16//DP Smoothleaf/DP 45/3/DES 56
Deltapine 90	Midsouth	DP 6516/DP 6582
Dunn 1047	High Plains	Tamcot SP21/Dunn 219
Lankart LX571	High Plains	Lankart 57/Lankart 3840
McNair 220	Southeast	CKR 201/PD2165
McNair 235	Southeast	CKR 201/PD2165
Paymaster 111-A	High Plains	Selection PM 111
Paymaster 145	High Plains	Selection Tamcot SP21
Paymaster 266	High Plains	AZ 6024-11-1-2/DP 5540//PM 101-A/TAES B4
Paymaster HS26	High Plains	Acala SJ-4/5B9-184
Paymaster HS200	High Plains	107 × 329 123171-74/160 × 145 145521
PD-1	Southeast	PD4381/PD8623
PD-2	Southeast	FTA 266/Atlas//AC235/Dixie King
PD-3	Southeast	PD9363/PD9240
Ranger BB-53	High Plains	Stripper 31/PM 111-A
SC-1	Southeast	CKR 421/PD4398
Stripper 31A	High Plains	Complex cross of: Roldo Rowden #5 & 27, Empire WR, BBR 4-1-36 B2
Tamcot CAB-CS	High Plains	CAMD-21-S-78/BCUS-8-76
Tamcot CAMD-E	High Plains	MDR.SP7-67/17M2//SP46-67/17M2
Tamcot CD3H	High Plains	Tamcot SP37H/CDPS-1-77
Tamcot GCNH	High Plains	CAMD-21S-5-80/GN-8-76
Tamcot HQ95	High Plains	Tamcot CD3H/MAR-CABU ^c CS-2-1-83
Tamcot SP21	High Plains	K4808-5 (1&2)D//Blightmaster/39-11-20/3/K4808-5 (1&2)A/PM 54-M-105-3
Tamcot SP21S	High Plains	SP21F/SP33F//SP21V/SP37V
Tamcot SP23	High Plains	K4808-5 (1&2)D//Blightmaster/39-11-20/3/K4808-5 (1&2)A/PM 54-M-105-3
Tamcot SP37	High Plains	K4808-5 (1&2)D//Blightmaster/39-11-20/3/K4808-5 (1&2)A/PM 54-M-105-3
Tamcot SP37H	High Plains	66N, B.V.65/52o, B.V.65

† Calhoun et al. (1997).

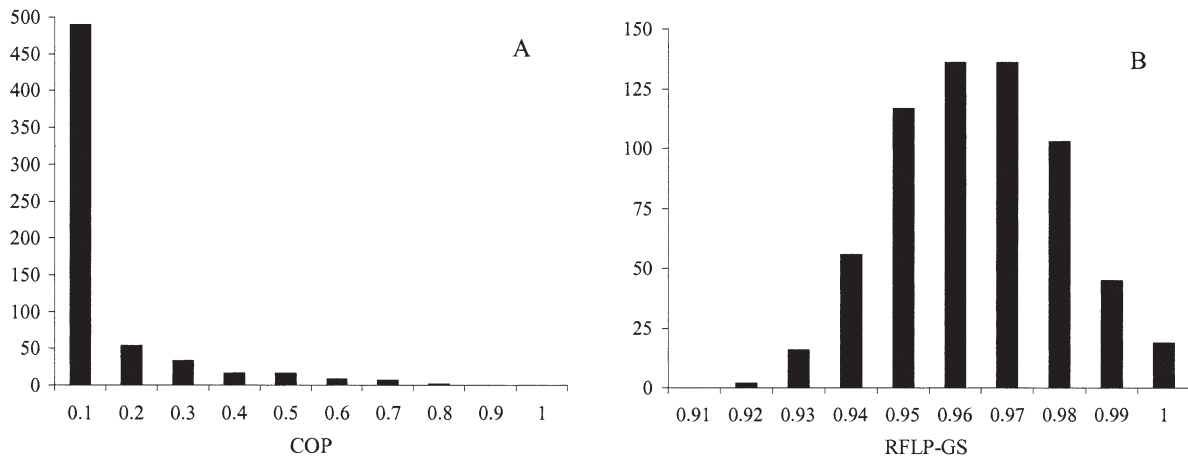


Fig. 1. Frequency distributions of (A) coefficients of parentage (COPs) and (B) RFLP-based genetic similarity (RFLP-GS) values for 630 pairs of Upland cotton cultivars.

determine the association between COP and RFLP-GS values. The normalized Mantel statistic Z (Mantel, 1967) was used to assess the significance of the correlation between the COP and RFLP-GS matrices. A linear regression of COP on RFLP-GS was conducted including only the pairs of cultivars with $COP \geq 0.1$, following the assumption that cultivars with $COP < 0.1$ are not related by their pedigrees (Melchinger et al., 1995). Spearman's rank correlation (Steel et al., 1997) between COP and RFLP-GS values was calculated for the pairs of related cultivars.

RESULTS AND DISCUSSION

The frequency distributions of COP and RFLP-GS values for the 630 pairs of cultivars were distinctly different (Fig. 1). These COP values from Bowman et al. (1997) ranged from 0 to 0.75, their median being 0.051 due to skewness toward low COP values. The RFLP-GS values had a normal-shaped distribution, ranging from 0.914 to 1 with a median of 0.975, which reflects the low genetic diversity reported by Brubaker and Wendel (1994) and the narrow ranges of genetic distance and similarity reported by Pillay and Myers (1999), Iqbal et al. (2001), Linos et al. (2002), and Gutiérrez et al. (2002). The distributions of COP and RFLP-GS values were similar to those observed by Barrett et al. (1998) in wheat using AFLP markers.

A highly significant though moderate association ($r = 0.41$, $P < 0.001$) was detected between the COP and RFLP-GS matrices. This result agreed with those obtained in other autogamous crops such as wheat (Kim and Ward, 1997), barley (Graner et al., 1994), and oat (O'Donoghue et al., 1994) using RFLPs. In contrast, the correlation coefficient was much lower than that obtained by Messmer et al. (1993) in maize, which is an allogamous species. Seventy-eight percent (488) of the pairs of cultivars had a COP smaller than 0.1 and consequently these cultivars were considered unrelated (Melchinger et al., 1995). Rank correlation between COP and RFLP-GS for the remaining 142 pairs of related cultivars ($COP \geq 0.1$) was highly significant ($r_s = 0.53$, $P < 0.001$). Although the correlation coefficient was still moderate, the exclusion of the most distant pedigree relationships ($COP < 0.1$) from correlation analysis in-

creased the association between pedigree- and RFLP-based estimates of genetic similarity. This agreed with the results observed in barley by Tinker et al. (1993) and in wheat by Kim and Ward (1997) using RAPD- and RFLP-based estimates, respectively. In contrast, Barrett et al. (1998) found that the association between pedigree- and AFLP-based estimates was weak throughout the whole range of diversity detected within their wheat germplasm. The RFLP-GS estimates for the pairs of related cultivars were plotted against their COP values (Fig. 2). There was a significant linear relationship between COP and RFLP-GS; however, the coefficient of determination was low ($R^2 = 0.25$), indicating that the COP only explained a small portion of the variation observed for RFLP-GS.

RFLP markers were effective in estimating the level genetic similarity throughout the entire range of diversity present in this study. In contrast, the use of pedigree information had a lower genetic resolution due to skewness toward low values, making COP uninformative for pa-

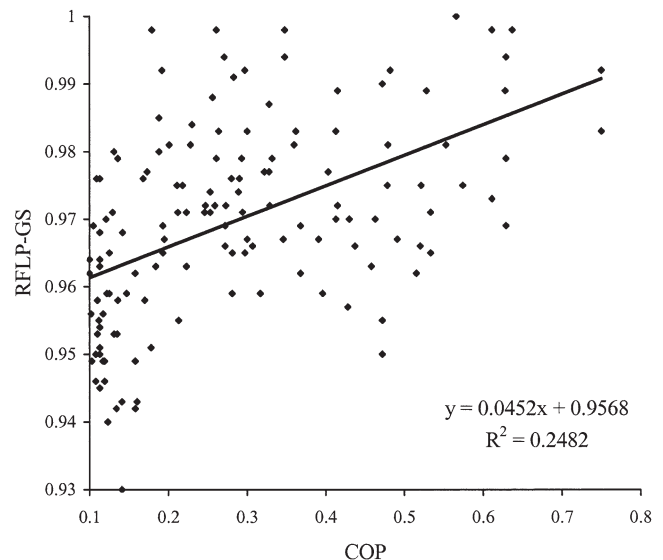


Fig. 2. Scatter plot of RFLP-based genetic similarity (RFLP-GS) versus coefficient of parentage (COP) for 142 pairs of related ($COP \geq 0.1$) Upland cotton cultivars.

rental combinations with low genetic similarity ($COP < 0.1$). Nevertheless, the dendrograms resulting from UPGMA cluster analyses of the COP and RFLP-GS matrices were somewhat similar (Fig. 3) since cluster analysis starts grouping the cultivars with the closest genetic relationships, for which the correlation between COP and RFLP-GS was highest. In general, cultivars that were closely related according to their pedigrees, such as 'Coker 310' and 'Coker 312' or 'McNair 220' and 'McNair 235', also had an apparent relationship in the RFLP-based dendrogram. However, some cultivars, such as 'Paymaster 145' and 'Tamcot SP21' or 'Paymaster 111-A' and 'Ranger BB-53', were farther apart in the RFLP-based dendrogram than might be expected based on their pedigrees. Some other relationships, such as those between 'Dunn 1047' and 'Stripper 31A' or 'Lankart LX571' and 'Paymaster 266', were close in the RFLP-based analysis but were not evident in the pedigree-based dendrogram. The similarity between both dendrograms suggests that pedigree information should still be useful to identify diverse parents in a cotton breeding program.

The moderate association between COP and RFLP-GS was not surprising since the estimation of genetic relationships among genotypes based on pedigrees and RFLP markers are fundamentally different approaches; in fact, they estimate different types of genetic resemblance. The COP is an estimate of the proportion of loci with alleles identical by descent, that is, copies of the same allele from a common ancestor, whereas RFLP-GS estimates the proportion of alleles alike in state, that is, indistinguishable by their effects. The COP ignores alleles that are alike in state but not identical by descent, assuming that genotypes not related by pedigree do not carry homologous fragments. This difference should be especially important in a crop such as cotton, which has a narrow genetic base and thus possesses numerous alleles alike in state but not identical by descent. The RFLP-GS is more informative to the plant breeder since the proportion of alleles alike in state determines the amount of genetic variance among progeny (Helms et al., 1997). That these measures estimate different types of genetic resemblance suggest that the estimation of genetic relationships among cultivars might be improved by combining COP and RFLP-GS into a composite index, as proposed by Cox et al. (1985). The composite index would be expected to decrease the effect of the independent inaccuracies of both estimates (Schut et al., 1997).

The moderate association between COP and RFLP-GS may have also resulted from violations to the assumptions made when computing the COP. Some of the assumptions underlying the calculation of COP are unrealistic for cotton breeding materials. The COP expresses the degree of relatedness between two genotypes relative to a conceptual population of original ancestors that are assumed to be unrelated. For instance, ancestral cultivars Kekchi, Half and Half, Young's Acala, and Jackson Round Boll were considered unrelated ($COP = 0$) (Van Esbroeck et al., 1999), yet the RFLP-GS values among them ranged from 0.914 to 0.933, indicating not only that ancestors carry alleles that are alike in state but

also that some pairs of ancestors are more related than others. By erroneously assuming that ancestors were unrelated, COP underestimated true genetic resemblance; however, this underestimation has little impact on the relative genetic similarity estimates among cultivars (Van Esbroeck et al., 1999).

In contrast, COP probably overestimated genetic relationships by assuming that all parents were homozygous and homogeneous. Cotton is normally an autogamous species; however, cross-pollination can occur by insect vectors. The COP does not reflect outcrossing nor residual segregation; however, the RFLP scores showed that cultivars were not always completely homozygous or homogeneous.

The COP also assumes that progeny receive half of their alleles from each parent, ignoring the effects of selection and genetic drift during cultivar development. However, selection and/or genetic drift during inbreeding may favor the recovery of alleles from one parent, biasing the allelic contribution of the parents to the progeny. For example, in the absence of selection pressure, the RFLP-GS between 'Ranger BB-53' and each of its parents are expected to be the same; however, Ranger BB-53 was closer to its parent 'Stripper 31A' (0.976) than to its other parent, 'Paymaster 111-A' (0.966). In addition, even though 'McNair 220' and 'McNair 235' were derived from the same parents, the RFLP-GS between the parent 'PD2165' and McNair 220 (0.974) was higher than that between 'PD2165' and McNair 235 (0.964). Hence, the assumption of equal allelic contribution of the parents to the progeny reduces the reliability of COP as a measure of true genetic resemblance.

Estimates based on DNA markers are a more accurate measure of overall genetic similarity than COP as they reflect the resemblance among genotypes at the DNA level by direct sampling of the genome. Most authors (Graner et al., 1994; Kim and Ward, 1997; Messmer et al., 1993) recommend the use of DNA marker-based genetic similarity to estimate true genetic resemblance among cultivars rather than using pedigree information alone. However, the extent of the utility of DNA markers for estimating genetic similarity may depend on the nature of the markers and the genotypes under study (Soleimani et al., 2002), while the accuracy of the estimates depends on the number of markers employed, the uniformity of their distribution in the genome, and the independence of the information they provide (Messmer et al., 1993).

Polymerase chain reaction (PCR)-based markers such as AFLPs, RAPDs, and SSRs have also been used in cotton to determine genetic relationships (Abdalla et al., 2001; Liu et al., 2000; Tatineni et al., 1996). In cotton, these markers are generally more polymorphic than RFLPs and, therefore, genetic similarity estimates based on them may show a closer association with COP. However, genetic similarity estimates using PCR-based markers may not accurately estimate functional genomic similarity, which is more relevant for plant breeding, since they generally analyze sequences from nonexpressed portions of the genome. Whereas, most of the RFLP mark-

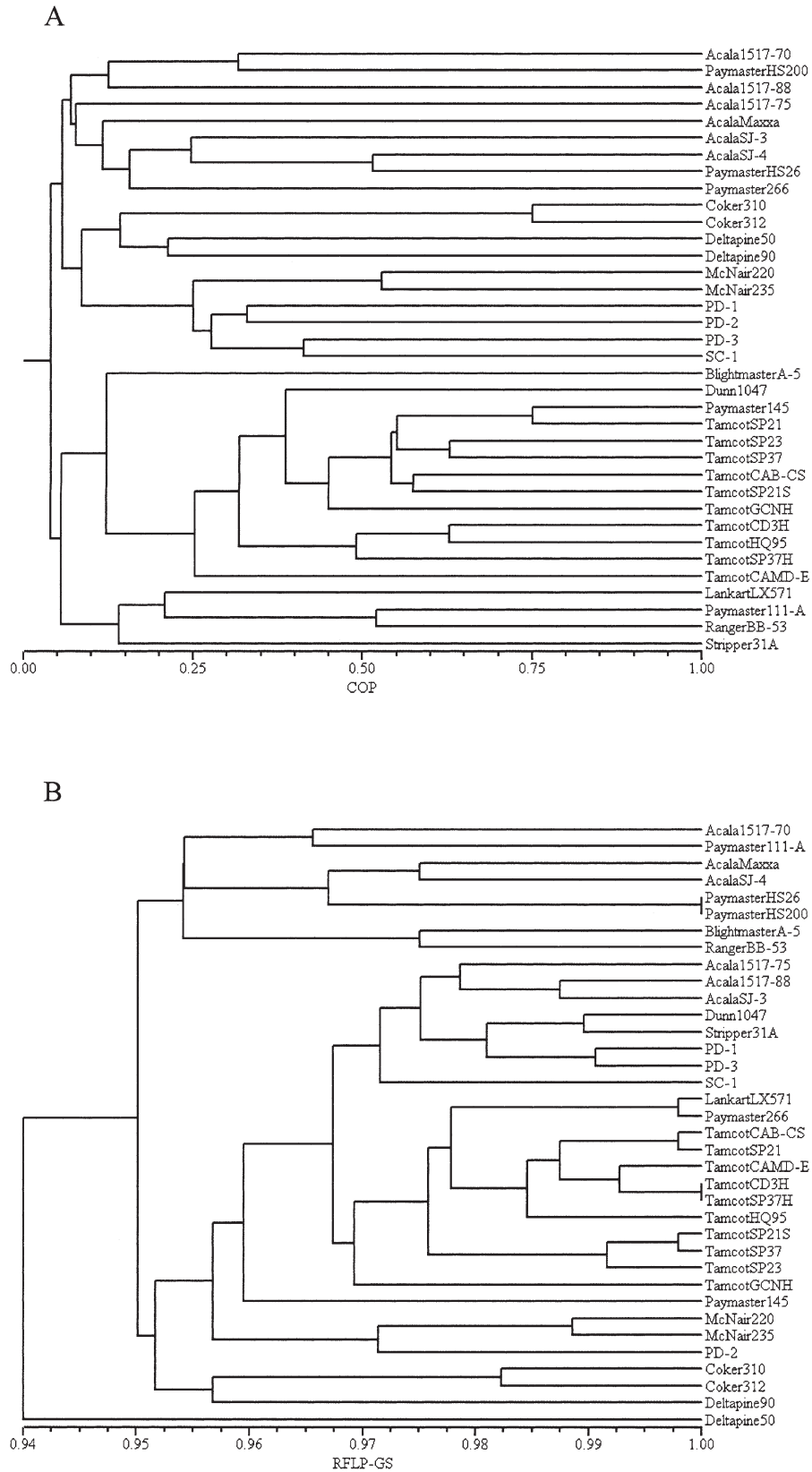


Fig. 3. Dendrograms resulting from UPGMA cluster analyses of (A) coefficients of parentage (COPs) and (B) RFLP-based genetic similarity (RFLP-GS) values among 36 Upland cotton cultivars.

ers used in this study were from either low copy or cDNA clones (Reinisch et al., 1994), which are likely to represent expressed regions of the genome. Further studies are needed to assess the relationship between genetic similarity estimates calculated using DNA markers versus genetic variance for economically importance traits.

In recent years, the average genetic gain for lint yield has been on a downward trend (Meredith et al., 1997). Some argue (Helms, 2000; Meredith, 2000) that cotton may have reached a plateau for productivity and fiber quality traits due possibly to the narrow genetic base of modern Upland germplasm. Genetic similarity estimates such as COP and RFLP-GS are helpful for plant breeders in selecting parents that are more likely to possess dissimilar genes to increase the level of genetic variance in segregating populations. The mating of genetically diverse parents broadens the genetic base of Upland cotton; however, some cotton breeders think that genetically diverse parents tend to come from different production regions and are reluctant to use unadapted germplasm. Due to rigorous standards for fiber quality and the need to maintain or increase lint yield, breeders tend to use adapted cultivars with outstanding performance as parents. Nevertheless, selecting parents with outstanding performance does not necessarily mean that the parents must be closely related.

CONCLUSIONS

Pedigree- and RFLP-based genetic similarity estimates among 36 Upland cotton cultivars showed a significant though moderate association. The moderate level of association was likely due, in part, to the fact that both approaches estimate different types of genetic resemblance, while violations of unrealistic assumptions involved in the calculation of COPs may have also contributed to the observed results. The exclusion of the most distant pedigree relationships from correlation analysis intensified the association between both estimators. Regression analysis showed that pedigree relatedness was not an accurate measure of genetic similarity as revealed by RFLP markers. We recommend the use of RFLP-GS as a more direct estimate of true genetic resemblance among cultivars. Nevertheless, the pedigree- and RFLP-based dendrograms were somewhat similar, indicating that pedigree information will continue to be useful to inexpensively identify diverse parents in a breeding program. The combination of COP and RFLP-GS estimates would help plant breeders to maximize the level of genetic variance in segregating populations and focus their resources on the most promising crosses.

ACKNOWLEDGMENTS

We thank Dr. D.T. Bowman and Dr. O.L. May for many helpful discussions. This work was supported by Cotton Inc. and the USDA-IFAFS.

REFERENCES

Abdalla, A.M., O.U.K. Reddy, K.M. El-Zik, and A.E. Pepper. 2001. Genetic diversity and relationships of diploid and tetraploid cottons revealed using AFLP. *Theor. Appl. Genet.* 102:222–229.

- Barrett, B.A., K.K. Kidwell, and P.N. Fox. 1998. Comparison of AFLP and pedigree-based genetic diversity assessment methods using wheat cultivars from the Pacific Northwest. *Crop Sci.* 38:1271–1278.
- Bowman, D.T., O.L. May, and D.S. Calhoun. 1996. Genetic base of upland cotton cultivars released between 1970 and 1990. *Crop Sci.* 36:577–581.
- Bowman, D.T., O.L. May, and D.S. Calhoun. 1997. Coefficients of parentage for 260 cotton cultivars released between 1970 and 1990. U.S. Department of Agriculture, Tech. Bull. No. 1852.
- Brubaker, C.L., and J.F. Wendel. 1994. Reevaluating the origin of domesticated cotton (*Gossypium hirsutum*; Malvaceae) using nuclear restriction fragment length polymorphisms (RFLPs). *Am. J. Bot.* 81:1309–1326.
- Calhoun, D.S., D.T. Bowman, and O.L. May. 1997. Pedigrees of upland and pima cotton cultivars released between 1970 and 1995. *Miss. Agric. For. Exp. Stn. Bull.* 1069. Mississippi State Univ., Mississippi State, MS.
- Cox, T.S., Y.T. Kiang, M.B. Gorman, and D.M. Rodgers. 1985. Relationship between the coefficient of parentage and genetic similarity indices in the soybean. *Crop Sci.* 25:529–532.
- Graner, A., W.F. Ludwig, and A.E. Melchinger. 1994. Relationships among European barley germplasm: II. Comparison of RFLP and pedigree data. *Crop Sci.* 34:1199–1205.
- Gutiérrez, O.A., S. Basu, S. Saha, J.N. Jenkins, D.B. Shoemaker, C.L. Cheatham, and J.C. McCarty, Jr. 2002. Genetic distance among selected cotton genotypes and its relationship with F₂ performance. *Crop Sci.* 42:1841–1847.
- Helms, A.B. 2000. Yield study report. In P. Dugger and D. Richter (ed.) *Proc. Beltwide Cotton Prod. Conf.*, San Antonio, TX [CD-ROM]. 4–9 Jan. 2000. Natl. Cotton Council Am., Memphis, TN.
- Helms, T., J. Orf, G. Vallad, and P. McClean. 1997. Genetic variance, coefficient of parentage, and genetic distance of six soybean populations. *Theor. Appl. Genet.* 94:20–26.
- Iqbal, M.J., O.U.K. Reddy, K.M. El-Zik, and A.E. Pepper. 2001. A genetic bottleneck in the 'evolution under domestication' of upland cotton *Gossypium hirsutum* L. examined using DNA fingerprinting. *Theor. Appl. Genet.* 103:547–554.
- Kempthorne, O. 1969. An introduction to genetic statistics. Iowa State University Press, Ames.
- Kim, H.S., and R.W. Ward. 1997. Genetic diversity in Eastern U.S. soft winter wheat (*Triticum aestivum* L. em. Thell.) based on RFLPs and coefficient of parentage. *Theor. Appl. Genet.* 94:472–479.
- Lewis, P.O., and D. Zaykin. 2001. Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0 (d16c) [Online]. Available at hydrodictyon.eeb.uconn.edu/people/plewis/software.php (verified 14 June 2005).
- Linos, A.A., P.J. Bebeli, and P.J. Kaltsikes. 2002. Cultivar identification in upland cotton using RAPD markers. *Aust. J. Agric. Res.* 53:637–642.
- Liu, S., R.G. Cantrell, J.C. McCarty, Jr., and J.McD. Stewart. 2000. Simple sequence repeat-based assessment of genetic diversity in cotton race stock accessions. *Crop Sci.* 40:1459–1469.
- Mantel, N.A. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27:209–220.
- Melchinger, A.E., A. Graner, M. Singh, and N.M. Messmer. 1995. Relationships among European barley germplasm. I. Genetic diversity among winter and spring cultivars revealed by RFLPs. *Crop Sci.* 34:1191–1199.
- Meredith, W.R., Jr. 2000. Cotton yield progress—Why has it reached a plateau? *Better Crops* 84:6–9.
- Meredith, W.R., Jr., J.J. Heitholt, W.T. Pettigrew, and S.T. Rayburn, Jr. 1997. Comparison of obsolete and modern cotton cultivars at two nitrogen levels. *Crop Sci.* 37:1453–1457.
- Messmer, M.M., A.E. Melchinger, R.G. Herrmann, and Jürgen Bopenmaier. 1993. Relationships among early European maize inbreds: II. Comparison of pedigree and RFLP data. *Crop Sci.* 33:944–950.
- Murphy, J.P., T.S. Cox, and D.M. Rogers. 1986. Cluster analysis of red winter wheat cultivars based on coefficients of parentage. *Crop Sci.* 26:672–676.
- Nei, M. 1972. Genetic distance between populations. *Am. Nat.* 106:283–292.
- Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89:583–590.

- O'Donoghue, L.S., E. Souza, S.D. Tanksley, and M.E. Sorrells. 1994. Relationships among North American oat cultivars based on restriction fragment length polymorphisms. *Crop Sci.* 34:1251–1258.
- Pillay, M., and G.O. Myers. 1999. Genetic diversity in cotton assessed by variation in ribosomal RNA genes and AFLP markers. *Crop Sci.* 39:1881–1886.
- Reinisch, A.J., J.M. Dong, C.L. Brubaker, D.M. Stelly, J.F. Wendel, and A.H. Paterson. 1994. A detailed RFLP map of cotton, *Gossypium hirsutum* × *Gossypium barbadense*: Chromosome organization and evolution in a disomic polyploid genome. *Genetics* 138:829–847.
- Rohlf, F.J. 2004. NTSYSpc: Numerical Taxonomy System, ver. 2.11V. Exeter Publishing, Ltd., Setauket, NY.
- Schut, J.W., X. Qi, and P. Stam. 1997. Association between relationship measures based on AFLP markers, pedigree data and morphological traits in barley. *Theor. Appl. Genet.* 95:1161–1168.
- Soleimani, V.D., B.R. Baum, and D.A. Johnson. 2002. [*Triticum turgidum* L. subsp. *durum* (Desf.) Husn.] AFLP and pedigree-based genetic diversity estimates in modern cultivars of durum wheat. *Theor. Appl. Genet.* 104:350–357.
- Steel, R.G.D., J.H. Torrie, and D.A. Dickey. 1997. Principles and procedures of statistics: A biometrical approach. 3rd ed. McGraw-Hill Inc., New York.
- Tatineni, V., R.G. Cantrell, and D.D. Davis. 1996. Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPDs. *Crop Sci.* 36:186–192.
- Tinker, N.A., M.G. Fortin, and D.E. Mather. 1993. Random amplified polymorphic DNA and pedigree relationships in spring barley. *Theor. Appl. Genet.* 85:976–984.
- Van Esbroeck, G.A., D.T. Bowman, O.L. May, and D.S. Calhoun. 1999. Genetic similarity indices for ancestral cotton cultivars and their impact on genetic diversity estimates of modern cultivars. *Crop Sci.* 39:323–328.
- Wendel, J.F., C.L. Brubaker, and A.E. Percival. 1992. Genetic diversity in *Gossypium hirsutum* and the origin of Upland cotton. *Am. J. Bot.* 79:1291–1310.