

PEDIGREE- VS. RFLP-BASED GENETIC SIMILARITY ESTIMATES IN COTTON

Guillermo Van Becelaere¹, Edward L. Lubbers¹, Peng W. Chee¹,
and Andrew H. Paterson²

¹Dept. of Crop and Soil Sciences, UGA, Tifton

²Plant Genome Mapping Laboratory, UGA, Athens

Introduction

Genetic similarity estimates among different genotypes are helpful to select parental combinations for developing segregating populations. Crosses between genetically divergent parents are expected to have a larger genetic variance among progenies than crosses between closely related parents, increasing the opportunity for selecting superior progenies. Genetic similarity among genotypes can be estimated by different approaches, which include the use of pedigrees and DNA fingerprinting. The coefficient of parentage (COP) measures the degree of genetic relatedness among genotypes based on pedigree information. The analyses of molecular markers such as restriction fragment length polymorphisms (RFLPs) directly measure DNA sequence variation. It is important to determine whether pedigree- and DNA marker-based estimates of genetic similarity provide similar information about the genetic relationships among germplasm. The objective of this study was to determine the correspondence between pedigree- and RFLP-based genetic similarity (RFLP-GS) 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. As part of a more comprehensive study, the cultivars were assayed with 261 codominant RFLP markers. The markers were chosen to provide even coverage of the cotton genome. Genetic similarity based on RFLP markers was estimated for all possible pairs of cultivars using Nei's normalized identity (Nei, 1978). COPs for all possible pairs of cultivars were calculated by Bowman et al. (1997).

Cluster analyses of the COP and RFLP-GS matrices were performed to compare the relationships among cultivars as revealed by pedigrees and RFLP markers. Correlation analysis was conducted to determine the association between COP and RFLP-GS values. Spearman's rank correlation between COP and RFLP-GS values was calculated for the pairs of cultivars with $COP \geq 0.1$, following the assumption that cultivars with $COP < 0.1$ are not related by their pedigrees.

Results and Discussion

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. In general, cultivars that were closely related according to their

pedigrees also had an apparent relationship in the RFLP-based dendrogram.

A highly significant though moderate association ($r = 0.41$, $P < 0.001$) was detected between the COP and RFLP-GS matrices. Seventy-eight percent of the pairs of cultivars had a COP smaller than 0.1 and consequently these cultivars were considered unrelated. Rank correlation between COP and RFLP-GS for the pairs of related cultivars was highly significant ($r_s = 0.53$, $P < 0.001$). The exclusion of the most distant pedigree relationships from correlation analysis intensified the association between pedigree- and RFLP-based estimates of genetic similarity.

The moderate association between COP and RFLP-GS was not surprising since the COP is an estimate of the proportion of loci with alleles identical by descent, whereas RFLP-GS estimates the proportion of alleles alike in state. 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.

The moderate association between COP and RFLP-GS may have also resulted from violations to the assumptions made when computing them. Some of the assumptions underlying the calculation of COP are unrealistic for cotton breeding materials. By assuming that ancestors were unrelated, COP may have underestimated true genetic resemblance. In contrast, COP probably overestimated genetic relationships by assuming that all parents were homozygous and homogeneous. The COP also assumes that the progeny receives half its alleles from each parent, ignoring the effects of selection and genetic drift during cultivar development. This reduces the reliability of COP as a measure of true genetic resemblance.

Genetic similarity estimates such as COP and RFLP-GS are helpful to select parents that are less likely to possess similar genes in order to maximize the level of genetic variance in segregating populations. The moderate association between pedigree- and RFLP-based genetic similarity estimates indicated the potential weakness of reliance on either estimate. The estimation of genetic relationships among cultivars could be improved by combining COP and RFLP-GS into a composite index.

Acknowledgements

We thank Cotton Inc. and the USDA-IFAFS for financial support for this project.

References

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.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89:583-590.