2005 Beltwide Cotton Conferences, New Orleans, Louisiana - January 4 - 7, 2005

ASSOCIATION BETWEEN PEDIGREE- AND RFLP-BASED GENETIC SIMILARITY ESTIMATES

Guillermo Van Becelaere, Edward L. Lubbers and Peng W. Chee
Dep. of Crop and Soil Sciences, University of Georgia
Tifton, GA
Andrew H. Paterson
Dep. of Crop and Soil Sciences, Botany, and Genetics, University of Georgia
Athens, GA

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 developing segregating populations. The objective of this study was to determine the agreement between pedigree- and restriction fragment length polymorphism (RFLP)-based genetic similarity 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 (RFLP-GS) 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 ($COP \geq 0.1$) was somewhat higher ($rs = 0.53$, $P < 0.001$). 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 them. The estimation of genetic relationships among cultivars could be improved by combining COP and RFLP-GS in a composite index.

Introduction

Knowledge of genetic diversity and relationships among breeding materials is essential to the plant breeder in the improvement of crop species. 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 effectiveness of selection. Despite this, the use of reselection and the repeated crossing of closely related cultivars are common in cotton breeding and have likely narrowed the genetic base of Upland cotton.

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 agreement 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. The cultivars contain representatives from the four primary categories of commercial cotton grown in the U.S.: Acala, Plains, Delta, and Eastern. All cultivars were released in the U.S. between 1970 and 1990. As part of a more comprehensive study to survey the genetic diversity of cotton germplasm, 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). 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. 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 (Melchinger et al., 1995).

Results and Discussion

The frequency distributions of COP and RFLP-GS values for the 630 pairs of cultivars were distinctly different; while RFLP-GS values had a normal-shaped distribution, COP values exhibited a "cliff-like" distribution. Nevertheless, the pedigree- and RFLP-based dendrograms were somewhat similar. In general, cultivars that were closely related according to their pedigrees also had an apparent relationship in the RFLP-based dendrogram. However, some cultivars were farther apart in the RFLP-based dendrogram than might be expected based on their pedigrees, whereas some other relationships were close in the RFLP-based analysis but were not evident in the pedigree-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 (488) 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 ($rs = 0.53$, $P < 0.001$). 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. However, selection during inbreeding favors the recovery of alleles from one parent, biasing the allelic contribution of the parents to the progeny. 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 in a composite index, as proposed by Cox et al. (1985). The composite index is expected to decrease the inaccuracies of both estimates.

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.

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.

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.

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