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

GENETIC RELATIONSHIPS OF HISTORICALLY IMPORTANT EASTERN US UPLAND COTTON

Edward L. Lubbers, Peng W. Chee and O. Lloyd May
University of Georgia
Tifton, GA
John R. Gannaway
Texas Agricultural Experiment Station
Lubbock, TX
Andrew H. Paterson
University of Georgia
Athens, GA

Abstract

Correct utilization of parental materials is an important aspect of a breeding program from which the entire breeding program is derived. Historically important cultivars and germplasm lines are leading candidates for parents due to their proven performance. The genetic relationships of these cultivars and germplasm lines provide insights in effective use of the available cotton germplasm in the United States. The eastern region of the United States includes the Southeastern and MidSouth subregions which, in turn, include some of the oldest production areas in the United States. This area includes three of the most recognized commercial cotton breeding companies as well as highly regarded public breeding programs. The 115 sampled cultivars and germplasm lines are from the early 1900s to the 1990s. Genetic distance was calculated using 261 linkage map-derived RFLP markers surveyed against the DNA of the samples in a Southern analysis. A cladogram was constructed to reveal the clustering relationships of the cultivars and germplasm lines. Comparisons of the clustering of individuals and breeding programs against their putative pedigrees were explored. Our results show that, for the most part, pedigrees corresponded well with genetic distance measured via molecular markers. Outcrossing, either inadvertent or on purpose, with selection is the likely explanation for any observed deviation.

Objective

The objective of this study is to evaluate the genetic relationships of the eastern region of the United States Upland cotton germplasm.

Materials and Methods

A collection of 115 cotton lines/cultivars most of them of historical importance from the eastern cotton belt of the United States (Table 1) were surveyed against a series of 261 RFLP loci (Reinisch et al., 1994) that were chosen to evenly cover the cotton genome. The identification of thirteen of these lines/cultivars were uncertain, unknown, or not applicable as to production region or development location and included one *Gossypium hirsutum* var. palmeri as a outgroup and two varieties outside the US; Colombia, collected in Trinidad & Tabago, and Sivon from Israel. Genetic Data Analysis (Lewis and Zaykin, 2001) was used to calculate the genetic distance matrix with allele data from mapping populations, PAUP* (Swofford, 2003) was used to develop the optimal cladogram via maximum parsimony using heuristic algorithms, and TREEVIEW (Page, 1996) was used to provide the high resolution graphic of the cladogram. Pedigrees came from Ware (1950) and Calhoun et al. (1997).

Table 1. The listing of the individuals in grouping cotton cultivars and germplasm lines for the eastern region of the United States Cotton Belt.

Groupings	n	Names of the Cultivars and Germplasm Lines
Pee Dee medium staple program USDA-ARS, Florence, SC	31	AC241, FJA, FTA, PD0109, PD0111, PD0113, PD0259, PD0695, PD0875, PD2164, PD2165, PD3246, PD3249, PD4381, PD8619, PD9232, PD9363, PD9364, PD4461Q, PD-1, PD-2, PD-3, SC-1, PD-3-14, PD93030, PD93034, PD93007, PD93043, PD93009, PD93019, PD93021
Pee Dee extra long staple program	9	Sealand 1, Sealand 2, Sealand 7-white flower, Sealand 7-yellow

USDA-ARS, Florence, SC		flower, Sealand 391, Sealand 472, Sealand 542, Sealand 883, Earlistaple 7
Tidewater	5	Tidewater 4, Tidewater 29, Tidewater E372-4, Tidewater Seabrooks, Ewing Long Staple x Tidewater
Coker Pedigreed Seed Co. Hartsville, SC (Now Emergent Genetics, Memphis, TN)	20	Hartsville, Hartsville 5, Hartsville (Tucson), Columbia, Deltatype Webber, Deltatype Webber 4, Deltatype Webber 253-1 T142-8, Coker's Deltatype Webber 7, Coker's Deltatype Webber 9, Lightning Express, Coker's Wilds 2, Coker's Wilds 4, Coker's Wilds 9, Wilds 5, Wilds 15, Wilds 18, Wilds 34-4 T82-2, Wilds 34-4 T85-2, Coker 310, Coker 312
Southeast misc. cultivars	6	Half and Half, Mexican Big Boll, Empire, Auburn 56, McNair 220, McNair 235
Delta & Pine Land Co. Scott, MS	10	Express, Express 121, Express 432, Ewing's Long Staple, Deltapine 14, Deltapine 15, Deltapine Smooth Leaf, Deltapine 5540, Deltapine 50, Deltapine 90
Stoneville Pedigreed Seed Co. (Now Emergent Genetics, Memphis, TN)	7	Jackson Round Boll, Lone Star, Stoneville 5, Stoneville 7, Stoneville 20, Stoneville 213, Stoneville BXN 47
MidSouth cultivars	4	Deltatype Webber 2139, Rex, MO-Del, Dixie King
Cultivars and germplasm lines developed in Arkansas with Rowden parentage	10	Arkansas 9 (Nucala X Rowden 11-8-?), Arkansas 17 (Nucala X Rowden 14-2), Arkansas 21 (Rowden 1-6-4-1), Arkansas 22 (Roldo Rowden), Arkansas 23 (Rowden 11-4-1), Arkansas 25 (Rowden 41B-100), Rowden 40, Rowden 40-80-3, Rowden 2088, Rowden 2088-2-10-1
Obsolete and miscellaneous cultivars/germplasm lines	13	TH458, TH 386-2758, McNair TH 149-20, Beasley's Hybrid 49-0-4, Tideland TPSA #1, Tideland TPSA #69, Stovepipe, Carter Long Staple, Spears Upland Early Long Staple, Jackson Heritage 216, Colombia, Sivon, *G. Hirsutum* var. palmeri

Results and Discussion

The cladogram (Figure 1, *contact author for 11 by 17 copy, cladogram is unlikely to be clearly reproducible in this 8.5 by 11 format*) shows an optimal tree of the 115 cultivars and germplasm lines from the eastern US cotton belt with seventeen clusters labeled. Other analytic techniques can be used and do show a somewhat different result (Lubbers et al., 2003) which is why PAUP software was used to present a more optimal tree. Many of the cluster pairs, which are the closest grouping, match expectations based on pedigree; Coker 310 with Coker 312 (a selection of Coker 310), Sealand 7 with white flowers and with yellow flowers, and McNair 220 and 235 which are sister lines.

The public breeding programs have most of their material fall into distinct clusters. For example, 4 of 6 Univ. of Arkansas Rowden-derived cultivars in cluster #3, 4 out of 5 of the Tidewater material in cluster #7, and 6 of 9 of the Pee Dee extra long staple program in cluster #15. In the Pee Dee medium staple length program 24 out of 31 cultivars and germplasm lines are in 5 clusters, #8 to #11 and #17 with each cluster mostly from this Pee Dee program. Cluster #8 has 12 Pee Dee lines out of 16 total in the cluster, cluster #9 has 5 of 7, cluster #10 has 1 of 1, cluster #11 has 5 of 6, and cluster #17 has 4 of 5. Besides this, one of the parents of the McNair cultivars in cluster #8 is PD2165. However it is interesting that PD2165 isn't in any of these clusters. Also, Earlistaple 7 in cluster #17 is a parent to many PD lines.

Unlike the public breeding efforts mentioned in the previous paragraph, the commercial programs do not cluster as cleanly. Coker material, for instance, is scattered throughout the cladogram with only some that group together. Coker does appear to have a wider assortment of material than Deltapine or Stoneville (Table 1). The three commercial programs have interchanged cultivars from one and another as breeding material which, in turn, has time and time again provided new, greatly improved cultivars for the cotton producer. A very interesting cluster, #12 in Figure 1, has a strong mix of the cultivars of these three commercial programs. Six of 10 Deltapine cultivars are in cluster #12 along with 3 of the 7 Stoneville cultivars and 4 Coker cultivars. These cultivars are also the more modern ones of our samples from the three commercial companies. This leads to the question of whether the even newer cultivars are equally or more closely related, thus further increasing field genetic uniformity of United States cotton.

The progression of a breeding program can be monitored over time. Looking at our examples from the Stoneville program; Jackson Round Boll, ST 5, and ST 20 are all in cluster #16 and follow our expectations based on the pedigree. However, Lone Star, which came from Jackson Round Boll and whose progeny include both ST 5 and 20, is in cluster #7. Next in the progression come ST 7, ST 213, and ST BXN 47 which are in cluster #12 (not with the other Stoneville cultivars). In the Deltapine program, Express 432 follows Express and is in cluster #16 but Express 121, a selection of Express 432, is in cluster #6. Coker's Lightning Express, from the Express group, is not in cluster #16 with Express and Express 432 but it is in cluster #15 and part of a higher order cluster. All of the rest of the Deltapine cultivars are in cluster #12 as was pointed out above. An unexpected clustering shows Jackson Round Boll, ST 5, and ST 20 in cluster #16 with Express and Express 432. Establishing that many of the foundational materials are related heightens our need to determine whether recent field genetic diversity in the United States is decreasing.

Genetic relationships revealed by molecular markers corresponded well, for the most part, with relationships expected from the pedigrees. Cultivar identity that may have been mishandled over time is an obvious possibility for why the pedigrees don't match the clustering but a more likely reason maybe from outcrossing to unknown parents along with selection that is skewed to the other parent. With so many of the early cultivars being the product of reselection from existing cultivars, outcrossing may have been an important source of variability for continued improvement.

Acknowledgements

The authors would like to thank the USDA-IFAFS for funding this research project and Cotton Incorporated for funding the Cotton Incorporated Fellowship of the senior author, Ed Lubbers.

References

Calhoun, D. S., D. T. Bowman, and O. L. May. 1997. Pedigrees of Upland and Pima cotton cultivars released between 1970 and 1995. Mississippi Agricultural and Forestry Experiment Station Bulletin 1069.

Lewis, P. O., and D. Zaykin. 2001. Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0 (d16c). Free program distributed by the authors over the internet from http://lewis.eeb.uconn.edu/lewishome/software.html

Lubbers, E. L., P. W. Chee, K. M. El-Zik, J. R. Gannaway, O.L. May, R.J. Wright, and A.H. Paterson. 2003. Genetic Diversity of Upland Cotton. *In* 2003 ASA-CSSA-SSSA Annual Meetings Abstracts. ASA, Madison, WI.

Page, R. D. M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* 12: 357-358.

Reinisch, A., Dong, J., Brubaker, C.L., Stelly, D.M., Wendel, J.F., and Paterson, A.H. 1994. A detailed RFLP map of cotton, Gossypium hirsutum x Gossypium barbadense: chromosome organization and evolution in a disomic polyploid genome. Genetics 138: 829-847.

Saitous, Naruya and Masatoshi Nei. 1987. The Neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4:406-425.

Swofford, D. L. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0b10. Sinauer Associates, Sunderland, Massachusetts.

Ware, J. O. 1950. Origin, rise, and development of American Upland cotton varieties and their status at present. Mimeo Publication, University of Arkansas, College of Agriculture, Agriculture Experiment Station, Fayetteville, AR.