MAPPING OF CANDIDATE TRICHOME GENES AND QUANTITATIVE TRAIT LOCI FOR FIBER YIELD AND QUALITY

Liping Wang
New Mexico State University
Las Cruces, NM
Xinjiang Agriculture University
Urumqi, China
Yanying Qu
Xinjiang Agriculture University
Urumqi, China
Las Cruces, NM
Sidney E. Hughs
USDA-ARS Southwestern Cotton Ginning Research Laboratory
Mesilla Park, NM
Jinfa Zhang
New Mexico State University

Abstract

Cotton fibers are epidermal trichomes on the seed surface which share similarities with leaf trichomes. The objective of this study was to identify homologous cotton genes based on known genes involved in the pathway of leaf trichome initiation in Arabidopsis. Sequence variations including single nucleotide polymorphisms (SNPs) of the homologous trichome genes were identified in cotton for primer designing and the single strand conformation polymorphism (SSCP) technique was then employed. Of a total of 69 primer pairs, 13 yielded polymorphisms between Upland and Pima cotton and six markers were mapped to linkage groups. The results are being used to identify markers and genes associated with fiber traits, i.e., quantitative trait loci.

Introduction

Cotton is the most important fiber crop worldwide and cotton fiber production represents a multi-billion dollar industry. The cotton fiber, called seed trichome, is derived from seed protodermal cells. Cotton fiber cell growth and development has received much attention in research, but its molecular basis is still unknown. It is highly likely that it shares some similarities with leaf trichome development in Arabidopsis (Lee et al., 2007). It is currently known that leaf trichome initiation in Arabidopsis is promoted by the positive transcription regulators GLABROUS1 (GL1), TRANSPARENT TESTA GLABRA1 (TTG1) and GLABRA3 (GL3) that are counteracted by the negative regulators TRIPTYCHON (TRY) and CAPRICE (CPC). To take advantage of the knowledge of trichome initiation in Arabidopsis, we searched the homologous gene sequences of these trichome genes in cotton to identify sequence variations and develop polymorphic markers.

Materials and Methods

Plant materials

The plant materials used to assess PCR amplification and polymorphisms included three Upland (*Gossypium hirsutum*)- Acala 1517-99, TM-1 and SG 747, two Pima cotton (*G. barbadense*)- DP 340 and 3-79, and their two ancestral species *G. arboretum* (A2) and *G. raimondii* (D5) represented by an accession each. Leaf tissues from each genotype or species were harvested for DNA extraction using a quick method (Zhang and Stewart, 2000). These samples were used to test primers designed from homologous trichome genes in cotton. Once a polymorphism between Upland and Pima cotton amplified by some of the primers was validated, a backcross inbred line (BIL) population of 96 lines was used to locate homologous genes in a linkage map.

Primer designing

Sequences of the trichome genes were first downloaded from http://arabidopsis.org and then used to search for homologous cotton gene sequences in the Cotton Gene Index database. When tentative consensus (TC) sequences (i.e., contigs) existed, individual genes or expressed sequence tags (ESTs) were further divided into two subgroups (subcontigs) based on ESTs from these of A2 and/or D5 if available. This allowed separation of allelic homologs from homoelogs between the two subgenomes (A_h and D_h) in the tetraploid cotton. Sequence variations including

one or more simple nucleotide polymorphisms (SNPs) within each subcontig were manually identified for primer designing. PRIMER3 were used to design primers that amplify a 200-300-bp region containing the SNPs. Primers were named after their corresponding sequence ID number.

Polymerase chain reactions (PCR) and marker analysis

PCR were performed in a volume of 10 µL, consisting of 5 µM of each primer, 25 mM $MgCl_2$, 5× buffer, 10 mM dNTPs, and 0.5 U of Taq Gold DNA polymerase (Applied Biosystems). PCR cycling conditions were 5 min of denaturation at 94°C followed by 30 cycles of 1 min denaturation at 94°C, 1 min annealing at 55°C, and 1 min extension at 72°C, with a final extension of 7 min at 72°C after the last cycle on a PE 2700 Thermocycler (Applied Biosystems). Single strand conformation polymorphism (SSCP) analysis was conducted as described by Lu et al. (2009). Briefly, 5 µL of PCR products was mixed with 5 µL SSCP loading dye and denatured at 95°C for 7 min; then, the samples were loaded on an 8.1% polyacrylamide gel for electrophoresis at a constant 15 W for 8 h at room temperature; and finally, the DNA fragments in the gels were observed by silver staining.

Data Analysis

The DNA fragments in the gels were scored as follows: "1" for presence and "0" for absence. The data matrix for the above seven genotypes or species was used to calculate Jaccard's similarity coefficients (Guo and Luo, 2006). The cluster analysis (UPGMA) and dendrogram construction were performed using the SHAN program in NTSYS-pc 2.10e software (Rohlf 2000). A linkage mapping on the BIL population was performed using JOINMAP4 software.

Results and Analysis

Screening of homologous trichome gene primers

A total of 69 primer pairs containing sequence variations including SNPs were designed, including 14 primer pairs for two TC sequences homologous to GL1 in Arabidopsis, 20 primer pairs for four TC sequences homologous to GL2 in Arabidopsis, 6 primer pairs for two cotton genes homologous to GL3, 23 primer pairs for eight gene/TC sequences homologous to TTG1, and one primer pair for an EST homologous to TRY and two primer pairs for homologous CPC and ZWI genes each.

Based on PCR and SSCP genotyping of seven different cotton genotypes, 14 primer pairs did not yield any specific amplification; 42 primer pairs did not produce polymorphisms between Upland and Pima cotton; and 13 primer pairs detected polymorphisms between the two cultivated tetraploid species. Primers designed from one TC for GL1, another TC for GL2, three TCs for TTG1 and an EST for TRY did not produce polymorphisms between the two species.

Cluster analysis

A total of 16 primer pairs (18.8% of all primer pairs) generated 108 polymorphic loci with an average of 6.75 loci per primer combination among the seven genotypes used. The cluster analysis showed that these polymorphic loci clearly distinguished between A2 and D5 and also separated Upland from Pima cotton (Figure 1). However, the two Pima genotypes 3-79 and DP 340 were highly identical (with a similarity coefficient of 0.95) from each other based on the trichome gene markers (Figure 1). However, the two Pima genotypes were unexpectedly closer to D5 than to A2 which was the most distant from other genotypes. This may indicate that most of the polymorphic trichome gene markers are located on the D subgenome. As expected, TM-1 and SG 747 were closer to each other than to Acala 1517-99 which had germplasm introgression from *G. barbadense*.

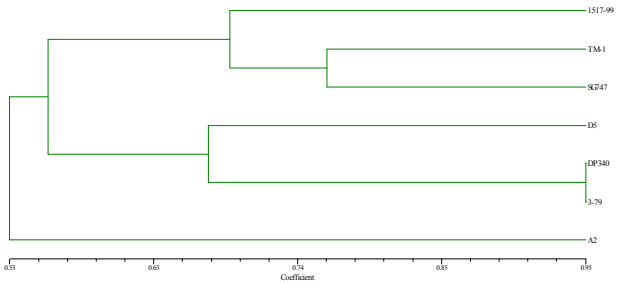

Figure 1. The UPGMA dendrogram of seven genotypes/four species based on homologous trichome gene markers.

Linkage mapping

We used the 16 informative primer pairs to genotype the BIL population with a subset of 93 individual lines. Of those, 3 primer pairs (TC231334-1, ES815673, and AF530912) did not produce polymorphism in the population. Of the remaining 13 primer pairs, six polymorphic SSCP markers were mapped onto five chromosomes (Figure 2). Interestingly, two trichome gene markers from two different genes were closely linked. Unfortunately, the chromosome origin of two linkage groups was currently unknown due to no chromosome-anchored SSR markers were mapped.

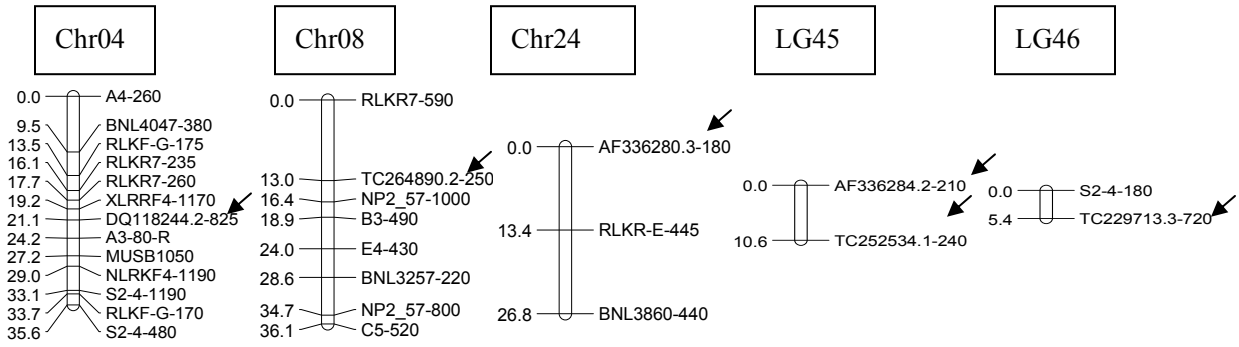

Figure 2. A linkage map containing 6 homologous trichome gene markers.

References

Guo, D.L., and Z.R.Luo. 2006. Genetic relationships of some PCNA Persimmons (*Diospyros kaki* Thunb.) from China and Japan revealed by SRAP analysis. Genet. Resour. Crop Evol. 53:1597–1603.

Lee, J.J., A.W. Woodward, and Z.J. Chen. 2007. Gene expression changes and early events in cotton fibre development. Ann. Bot. 100:1391–1401.

Lu, Y., J. Curtiss, R.G. Percy, S.E. Hughs, S.X.Yu, and J.F. Zhang. 2009. DNA polymorphisms of genes involved in fiber development in a selected set of cultivated tetraploid cotton. Crop Sci. 49:1695–1704.

Rohlf, F.J. 2000. NTSYS-pc: Numerical taxonomy and multivariate analysis system, Version 2.1, Exeter Software, New York

Zhang, J.F., and J.M. Stewart. 2000. Economical and rapid method for extracting cotton genomic DNA. J. Cotton Sci. 4:193-201.