9th January, 1996

Cotton Research and Development Corporation
Project UA1C: Final Report

Analysis of gene expression during cotton fibre development

Jeremy N Timmis, supervisor
Sharon J Orford, PhD student

Cotton fibres are differentiated from single ovule epidermal cells and grow synchronously, initiating at about the time of anthesis. In view of the commercial desirability of long fibres, the study of factors involved in controlling the extent of fibre growth is important. The aims of this project were to:

A. Isolate and characterise cDNA clones of mRNAs which are specific to or important in cotton fibre development;

B. Elucidate ribosomal gene expression and its role in fibre development.

The majority of the results obtained from part A. of this project are contained in two papers, both of which have been submitted for publication as journal articles and which are included with this report. A copy of the PhD thesis also forms part of this report and expected date of completion is 31st March 1996, after which time a copy will be forwarded to the CRDC.

The differential screening of a cotton fibre cDNA library and resultant isolation of five fibre-specific messages is described in the paper entitled "Abundant mRNAs specific to the developing cotton fibre". Sequences of two of the clones, pFS5 and pFS19, showed significant similarity to known plant proteins and a detailed characterisation of pFS5 and pFS19 is also described in this paper. Screening of the cDNA library with each of the fibre-specific cDNA clones showed that pFS19 transcripts are the most abundant fibre-specific transcripts in 13 day old fibre cells. Two genomic clones with sequence similarity to pFS19 have been isolated and mapped. The clones were found to be identical and the detailed analysis of one of them, λFS19(B), is described in the paper entitled "Selective expression of a lipid transfer protein gene in developing fibre cells".

The remaining three clones, pFS3, pFS17 and pFS18, did not show any similarity to sequences in the nucleotide databases, possibly due to incomplete sequence information. Rapid amplification of cDNA ends (5'-RACE) was used to obtain full-length cDNA clones for each of the three clones of unknown identity:

- Three plasmid clones from the pFS18 RACE experiment were characterised by an Honours student and sequencing results thus far are unclear;
- Four plasmid clones from the pFS3 RACE experiment were obtained and their identity verified by Southern analysis. Screening of the nucleotide databases with the full-length pFS3 sequence derived from the RACE clones was uninformative, implying that pFS3

may encode a novel protein product. The longest open reading frame encoded a protein of only 103 residues which showed no significant sequence similarity to any known protein. The size of the full-length sequence did not correlate well with the estimated pFS3 transcript length and suggests that the full-length pFS3 cDNA has yet to be isolated;

- Three plasmid clones were obtained from the pFS17 RACE experiment and their identity verified by Southern analysis. The full-length pFS17 sequence was constructed from the original pFS17 clone sequence and a consensus sequence derived from the three RACE clones. The full-length sequence of approximately 1100 basepairs correlated well with an estimated transcript size of 1200 nucleotides. The nucleotide sequence contained a long open reading frame which encodes a putative protein of 299 residues. The pFS17 sequence showed significant homology to a family of proline-rich proteins (PRPs) from a variety of plants, at both the nucleotide and amino acid levels. PRPs play a structural role in the walls of plant cells are thought to have a cross-linking or defense-related function. PRP genes are tightly controlled and often exhibit cell type specificity, such as the one isolated from cotton fibres. Southern analysis showed that PRPs in *Gossypium hirsutum* may form a small gene family, typical of PRP genes in other plants.

Results from the work outlined clearly have commercial potential. Genes which correspond to the five fibre-specific cDNAs contain elements which may control the expression of any transgenes to which they are attached. Foreign genes, potentially from any source, may be transferred into cotton plants and expressed specifically within the fibre, allowing for the alteration, and possible improvement, of cotton fibre morphology.

Studies have shown that ribosomal RNA metabolism may be related to cotton fibre development. The nucleolar size in fibres from three *G. hirsutum* varieties, differing in their final fibre lengths, was measured at early stages of growth. The mean nucleolar diameter reached a maximum at different times after anthesis and declined gradually, confirming previous work. One of several levels at which accumulation of rRNA may be controlled is by quantitative variation in the number of copies of the rRNA gene. The rDNA repeat unit from *G. hirsutum* var. Deltapine 90 was cloned and used to estimate the number of rRNA gene copies in six cotton varieties. Significant differences were observed between some of the varieties, but these did not correlate with final fibre length.

Jeffery N. Timms

S. J. Orford
Sharon Orford