TOWARDS GENETIC MANIPULATION OF FIBRE QUALITY IN AUSTRALIAN COTTON

Sharon J Orford, Sarah E Harmer, Sven K Delaney, John A Humphries,
Damien J Lightfoot and Jeremy N Timmis
Department of Molecular Biosciences, The University of Adelaide, Adelaide SA, 5005

Introduction

Cotton fibres, or lint, are very long single cells containing almost pure cellulose. The fibres develop in the weeks after flowering from single cells on the surface of the young seed. Each fibre cell is small at first but elongates and develops rapidly, eventually forming the mature cotton fibre. These processes require the ordered expression of genes which make enzymes, structural proteins and signaling molecules that together determine the properties of the fibre.

Domestication has modified fibre development to produce cotton varieties with greatly improved fibre length, strength and quality. However, the selection and breeding of plants with desirable fibre characteristics is slow and expensive. As such, future crop improvement is likely to depend upon genetic engineering and the cotton industry has been a leader in research and commercialisation of transgene technology, with momentous consequences for the agronomic properties of the crop such as insect and herbicide resistance.

Whilst fibre improvement has not been the initial focus of the application of DNA technology, research to this end has been vastly accelerated in the last decade, particularly in the USA. The recent formation of the International Cotton Genome Initiative between scientists in USA, France and Australia represents a global effort to develop a saturated and fully integrated genetic and physical map of the cotton genome. Gene discovery is also advancing rapidly, with the release of over 20,000 partial gene sequences (called ESTs) from elongating fibres of both diploid and tetraploid cotton, and the use of cDNA microarrays to facilitate global expression profiling.

Genetic modification of the cotton fibre has two major requirements. The first of these is an understanding of the genes that control commercially important fibre properties. Potentially useful fibre genes have been isolated in a number of laboratories including our own and fundamental studies on these genes have led to the creation of the first molecular and cellular model of fibre development (Wilkins and Jernstedt, 1999). The model has provided the basis for some early success in the alteration of fibre growth (Wilkins *et al.*, 2000), and this paves the way for near limitless opportunities to alter the fibre by manipulation of the biosynthetic and/or metabolic pathways that impact on physical or biochemical properties. The second requirement is the development of techniques for inserting genes of interest into cotton and regulating their expression. Both of these subjects are being addressed in industry-funded research at The University of Adelaide.

Genes which are expressed during cotton fibre development

Previous CRDC-funded research in Adelaide has focused on identifying genes that are expressed at high levels in fibre cells but not in other tissues, since these genes are likely to control fibre characteristics such as yield and quality. Our work has identified six genes which are only expressed in cotton fibres and not elsewhere in the plant, in tissues such as leaves, stems and roots. All are structural genes which encode proteins mostly involved with cell wall synthesis, a result which is not surprising given that cell wall deposition is the primary activity of the rapidly growing cotton fibre.

The most interesting isolate encodes an expansin (Orford and Timmis, 1998). Expansins are a large family of well-characterised proteins which loosen the components of rigid plant cell walls, thereby allowing a cell to expand. We have carried out a detailed analysis of the expansin gene family in cotton. Six members of the expansin gene family were characterised (Harmer *et al.*, in press), including one which is expressed at high levels during elongation of cotton fibre cells. This expansin may be pivotal to fibre cell initiation and elongation (Ruan *et al.*, 2001) and is therefore an attractive target in the design of strategies to alter fibre growth. To this end, we have created a number of transgenic cotton lines in which expansin gene expression in the fibres is prolonged, down-regulated or up-regulated. The whole-plant transformations (being performed at CSIRO Division of Plant Industry in Canberra) are progressing such that we will be able to analyse them late in 2002. These experiments provide the first example of alteration of the expression of an endogenous cotton gene, and offer exciting prospects for the genetic improvement of cotton fibre length.

Dissection of cotton fibre-specific promoters

The expression of genes is controlled by regions known as promoters, which determine the timing, location and level of gene expression in response to signals from inside and outside the cell. We have identified six genes which are specifically expressed in cotton fibre cells (see above), and have isolated the corresponding promoters. These promoters provide important tools in the genetic engineering of cotton fibres in several respects. Firstly, expression of transgenes only in the fibre cells avoids any potential problems from gene expression in non-fibre tissues. Secondly, they regulate the level and timing of transgene expression and so control the effects of transgenes on fibre characteristics. Finally, the promoters may be used with any transgene, and could therefore be used to produce, for example, naturally pigmented cotton fibres or fibres with enhanced insect resistance. In addition, desired changes in fibre characteristics could be achieved by manipulating gene expression by altering the promoters themselves.

The aim of this part of the project is to determine how the promoters direct the timing, level and location of gene activity, and to identify the signals to which they respond. Sequence analysis of our six fibre-specific promoters has identified a number of elements, such as TATA and CAAT boxes (necessary for basal gene expression), elements that may respond to MYB transcription factors and elements that mediate responses to plant hormones. However, the precise sequence

elements that regulate fibre-specific gene expression have not been isolated and cannot yet be identified by computer comparisons.

Identification of such elements is therefore being approached by functional analysis of the promoters. In the first instance, we are testing our bank of candidate fibre-specific promoters for their ability to direct expression of a reporter gene, in four ways:

1. Transient assays in cotton fibres and other tissues by microprojectile bombardment;
2. Stable transformation of cotton plants;
3. Stable transformation of tobacco plants;
4. Stable transformation of *Arabidopsis* plants.

All six promoters have been shown to be fibre-specific in transient assays in cotton (1 above). That is, the reporter protein was detected only in the fibres and not in bombarded leaf or petal tissue. The whole-plant transformations (2 above; also being carried out at CSIRO Division of Plant Industry in Canberra) are progressing such that we will be able to analyse them late in 2002. Transgenic lines will be tested for expression of the reporter gene only in fibres and not elsewhere in the plant. Transgenic tobacco and *Arabidopsis* plants with stable genomic integration of each reporter construct will also be generated (3 and 4 above). Tobacco and *Arabidopsis* trichomes (hairs) are an attractive model for cotton fibre growth because of the developmental similarities between cotton fibres and trichomes, and the cost and time factors involved in transforming tobacco and *Arabidopsis* (approx 3 months) as opposed to cotton (at least 12 months). In these cases the reporter protein is expected to be detectable only in the leaf trichomes of transgenic plants.

Identification of fibre-specific promoter elements (FSEs) has been approached by a strategy in which progressive 5' deletions of the promoter are tested using the reporter gene system as above. The experiment has been completed for our strongest promoter, with the FSEs which are absolutely required for fibre-specific gene expression being confined to a region of approximately 100 bp. Two further promoters will be tested in a similar way, and a "minimal" fibre-specific promoter, containing the minimal elements necessary for fibre-specific gene expression, can be defined. The information gained will be used to generate novel promoter constructs, by "cutting and pasting" a combination of different promoter elements, which will then be tested for their ability to control gene expression in a defined and predictable manner. The promoter combinations that we intend to investigate are unique to this research effort and (to our knowledge) have not been studied in other laboratories.

The promoters developed during this project could be used to control the expression of any gene of interest in the fibre cells, and could enhance the function of specific fibre-modifying genes, such as those for length and uniformity. Such genes have been isolated in previous and current research at The University of Adelaide. In addition, the information generated would provide a foundation for developing improved systems for the control of gene expression in plants. The results of this part of the project are therefore expected to be useful in other cotton research and development programs as well as in the genetic engineering of cotton fibre properties.

Gene expression in the cotton fruit

The *Bt* toxin from *Bacillus thuringiensis* is a transgene which has been engineered into a number of commercially grown cotton varieties, including INGARD® cotton in Australia. In these transgenic crops, the toxin is expressed throughout all tissues of the plant, at a high metabolic expense to the cotton plant. Since more than 80% of crop losses attributable to cotton insect pests are the result of species that attack fruit (Reynolds *et al.*, 1982), and the most commercially important product of the cotton crop is the fibres, it would be advantageous to restrict expression of the toxin to the cotton boll wall. This would form a physical barrier of toxin-expressing tissue to protect the fruit and fibres from pest attack.

Current research is aimed at the isolation of genes which are expressed only in the outer wall (pericarp) of the cotton boll. Our approach is based on a differential screening technique, which is well-established in the laboratory and has proven successful for the isolation of fibre-specific genes (see above). The isolation of pericarp-specific mRNAs will allow identification of the corresponding genes and promoter regions, which could be used to express the toxin (or other transgene) only in the outer wall of the cotton boll in a transgenic cultivar. The other tissues of the plant, such as leaves and seeds, would be free of *Bt* toxin (or other transgene) such that, for example, seeds would be more acceptable for oil production or animal feed. An additional advantage is that this would lead to reduced mortality among non-target species due to the reduced *Bt* toxin manufacture, and the plants would, in essence, contain an in-built refuge crop.

The molecular switch to fibre cell differentiation

An intriguing aspect of cotton fibre growth is the process by which fibre cells are determined. All ovule epidermal cells are able to differentiate into fibre cells, but only about 10% elongate to become fibres. Although the anatomy of young fibres has been well-studied by microscopy (Ryser, 1999), nothing is known about the molecular basis of cotton fibre initiation, except that the process appears to be controlled by plant growth factors such as auxins and gibberellins (Beasley and Ting, 1974).

A second type of differentiated epidermal cell, "fuzz", initiates growth up to several days after the lint fibres and, being very short, these structures have little or no commercial value. If the components of the developmental pathway can be elucidated, it may be possible to switch development of "fuzz" fibres into lint fibres. Added to this exciting prospect is the clear indication that metabolism in the cotton boll is able to sustain greater cellulose synthesis in additional fibres if "fuzz" formation can be abolished or reduced.

We have approached this complex question using the trichomes of *Arabidopsis thaliana* as a model system. *Arabidopsis* trichomes are branched hair-like structures which are found on all parts of the plant, and, like cotton fibres, they originate from single epidermal cells. The simplicity, visibility and dispensability of trichomes in *Arabidopsis* make them an ideal system for studying plant cell differentiation and morphogenesis, with the result that at least 20 genes have been implicated in the process. A model of the control of *Arabidopsis* trichome initiation, into which all the available

genetic and molecular data have been incorporated, suggests that normal trichome initiation requires only three or four genes (Szymanski *et al.*, 2000).

We have isolated and characterised four cotton homologues of one of the pivotal *Arabidopsis* trichome genes, TRANSPARENT TESTA GLABRA (*AtTTG1*). *AtTTG1* is required for both trichome initiation and pigment production in *Arabidopsis*, with mutants having no trichomes and colourless seeds. AtTTG1 encodes a protein of 341 amino acid residues with four WD-40 repeats (Walker *et al.*, 1999). WD-40 proteins have been implicated a diverse range of cellular processes such as signal transduction, RNA processing, gene regulation, vesicular traffic and regulation of the cell cycle (Yu *et al.*, 2000). The identification of TTG1 as a WD40 repeat protein indicated that TTG1 probably interacts with other proteins to control trichome initiation, rather than acting directly as a transcription factor.

Four cotton genes were isolated using a combination of PCR-based and library screening techniques. The sequences fall into two distinct pairs, with one pair (*GhTTG1* and *GhTTG3*) having 80% amino acid identity to *AtTTG1*, and the other pair around 62%. All four genes were functionally tested using *Matthiola incana* plants which have a mutation in the TTG1 gene, such that their normally purple petals are white. White *Matthiola* flowers were particle bombarded separately with the four constructs, and *GhTTG1* and *GhTTG3* were able to "rescue" the pigment pathway, producing purple spots on the bombarded petals. This significant result indicates that two cotton TTG1-like genes were able to function *in planta* like *Matthiola TTG1*.

Given the sequence and structural similarity between *AtTTG1* and cotton *GhTTG1* and *GhTTG3*, and the result with *Matthiola*, we hypothesise that either of the two cotton genes might be able to substitute for the function of the *Arabidopsis* trichome regulator. This will be tested by stable transformation of *TTG1* mutant *Arabidopsis* plants with the two cotton genes. Progeny from the transformed plants will be inspected for restoration of trichome formation and coloured seed coats. It is possible that proteins produced by the two cotton genes act in a similar way to *Arabidopsis TTG1* and may therefore play a pivotal role in trichome, or cotton fibre, initiation. As no regulators of cotton fibre cell growth have yet been isolated, further characterisation of these genes may lead to significant advances in the understanding of cotton fibre initiation.

The way forward

Since 1992 we have been conducting CRDC-funded research on the genes which are expressed during cotton fibre development. Several genes have been analysed in detail such that the work has recently entered a new and exciting phase. We are now in a position to carry out functional analysis of developmental genes and cotton fibre-specific promoters, and to test the effects of perturbing the expression of a particular gene on cotton fibre characters. Cotton transformation experiments are coming to fruition and a large number of transformant lines will be available for analysis in 2002.

Findings from our research will contribute substantially to our knowledge of the molecular control mechanisms which underlie fibre cell initiation, development and agronomic properties. More

generally, the project will contribute to our knowledge of cell fate determination and control of gene expression in plants by transcription factors. The research will significantly advance gene discovery in cotton, particularly key genes that control growth and development of the cotton fibre, and hence, determine economically important properties. Such information will provide a clear basis for fibre quality improvement in Australian cotton cultivars through genetic engineering, an outcome which can only result in economic benefits for the Australian cotton industry.

Acknowledgements

S Orford is grateful for the continued support for this research by the Cotton Research and Development Corporation, Narrabri, New South Wales. S Harmer is supported by an Australian Postgraduate Award (Industry) from the Australian Research Council and our industry partner is Cotton Seed Distributors, Wee Waa, New South Wales. S Delaney, J Humphries and D Lightfoot are all generously supported by Postgraduate Scholarships from the Cotton Research and Development Corporation. These experiments comply with the current Australian statutory requirements

References

Beasley CA and Ting IP (1974) Effects of plant growth substances on *in vitro* fibre development from unfertilised cotton ovules. *Amer J Bot* 61: 188-194

Harmer SE, Orford SJ and Timmis JN (2002) Characterisation of six α-expansin genes in *Gossypium hirsutum* (upland cotton). *Mol Gen Genom*, in press

Orford SJ and Timmis JN (1998) Specific expression of an expansin gene during elongation of cotton fibres. *Biochim Biophys Acta* 1398: 342-346

Reynolds HT, Adkisson PL, Smith RF and Frisbie RE (1982) Cotton insect pest management. *In* Metcalf RL and Luckman WH (eds) *Introduction to insect pest management*. Wiley and Sons, 2nd edn, pp 375-442

Ruan Y-L, Llewellyn DJ and Furbank RT (2001) The control of single-celled cotton fiber elongation by developmentally reversible gating of plasmodesmata and coordinated expression of sucrose and K^+ transporters and expansin. *Plant Cell* 13:47-60.

Ryser U (1999) Cotton fibre initiation and histodifferentiation. In Basra AS (ed) *Cotton fibres: developmental biology, quality improvement, and textile processing*. The Haworth Press, Inc, NY, pp 1-45.

Szymanski D, Lloyd A and Marks M (2000) Progress in the molecular genetic analysis of trichome initiation and morphogenesis in *Arabidopsis*. *Trends Plant Sci* 5: 214-219.

Walker A, Davison P, Bolognesi-Winfield A, James C, Srinivasan N, Blundell T, Esch J, Marks M and Gray J (1999) The TRANSPARENT TESTA GLABRA1 locus, which regulates trichome differentiation and anthocyanin biosynthesis in *Arabidopsis*, encodes a WD40 repeat protein. *Plant Cell* 11: 1337-1349.

Wilkins TA and Jernstedt JA (1999) Molecular genetics of developing cotton fibres. In Basra AS (ed) *Cotton fibres: developmental biology, quality improvement, and textile processing*. The Haworth Press, Inc, NY, pp 231-269.

Wilkins TA, Rajasekaran K and Anderson DM (2000) Cotton biotechnology. *Crit Rev Plant Sci* 19: 511-550.

Yu L, Gaitatzes C, Neer E and Smith TF (2000) Thirty-plus functional families from a single motif. *Prot Sci* 9: 2470-2476.