



Characterisation of α -expansin genes in
Gossypium hirsutum

by

Sarah Elizabeth Harmer, B.Sc.(Hons)

A thesis submitted for the degree of Doctor of Philosophy

in

The University of Adelaide

School of Molecular and Biomedical Science

Discipline of Genetics

(Faculty of Sciences)

March, 2003

Table of contents

| | |
|---------------------------------------------------------------------------------------------------------|-----------|
| Abstract | vi |
| Declaration | ix |
| Acknowledgments | x |
| List of Abbreviations | xi |
| | |
| CHAPTER ONE | 1 |
| INTRODUCTION | 1 |
| 1.1 THE COTTON CROP AND INDUSTRY | 3 |
| 1.2 HISTORY AND EVOLUTION OF <i>GOSSYPIUM HIRSUTUM</i> | 3 |
| 1.3 THE CULTIVATED COTTON PLANT AND FLOWER | 4 |
| 1.4 THE QUALITY AND DEVELOPMENT OF COTTON FIBRES | 4 |
| 1.5 COTTON FIBRE GROWTH AND DEVELOPMENT | 5 |
| 1.5.1 <i>Initiation</i> | 8 |
| 1.5.2 <i>Elongation</i> | 9 |
| 1.5.3 <i>Secondary wall synthesis</i> | 11 |
| 1.5.4 <i>Maturation</i> | 12 |
| 1.6 PLANT CELL GROWTH | 12 |
| 1.6.1 <i>The primary cell wall</i> | 12 |
| 1.6.2 <i>The expansins</i> | 14 |
| 1.6.3 <i>A fibre-specific α-expansin</i> | 17 |
| 1.7 GENERATION OF IMPROVED COTTON VARIETIES | 17 |
| 1.7.1 <i>Recombinant DNA technology</i> | 18 |
| 1.8 TARGETING TRANSGENE EXPRESSION | 19 |
| 1.8.1 <i>Fibre-specific genes and promoters</i> | 19 |
| 1.8.2 <i>Cotton fibre-specific gene promoters and α-expansin transgene expression</i> | 20 |
| 1.9 AIMS OF THE PROJECT | 21 |
| CHAPTER TWO | 23 |
| MATERIALS AND METHODS | 23 |
| 2.1 MATERIALS | 25 |
| 2.1.1 <i>Samples of plant material</i> | 25 |
| 2.1.2 <i>Nucleic acids</i> | 25 |
| 2.1.3 <i>Chemical reagents</i> | 25 |
| 2.1.4 <i>Stains and dyes</i> | 25 |
| 2.1.5 <i>Enzymes</i> | 26 |
| 2.1.6 <i>Radioactive isotopes and nucleic acids</i> | 26 |
| 2.1.7 <i>Molecular weight markers</i> | 27 |
| 2.1.8 <i>Antibiotics and indicators</i> | 27 |
| 2.1.9 <i>Bacterial strains</i> | 27 |
| 2.1.10 <i>Cloning vectors</i> | 27 |
| 2.1.11 <i>Kits and miscellaneous materials</i> | 28 |
| 2.2 SOLUTIONS, BUFFERS AND MEDIA | 34 |
| 2.2.1 <i>Solutions and buffers</i> | 34 |
| 2.2.2 <i>Media</i> | 35 |
| 2.3 METHODS | 36 |

| | | |
|---------|---------------------------------------------------------------------------------|----|
| 2.3.1 | <i>Plant growth conditions</i> | 36 |
| 2.3.2 | <i>General techniques</i> | 36 |
| 2.3.2a | <i>Restriction digests</i> | 36 |
| 2.3.2b | <i>Electrophoretic separation of DNA</i> | 36 |
| 2.3.2c | <i>Recovery of DNA from agarose gels</i> | 37 |
| 2.3.2d | <i>Purification of DNA using phenol/chloroform extraction</i> | 37 |
| 2.3.2e | <i>Precipitation of DNA</i> | 37 |
| 2.3.2f | <i>Determination of DNA and RNA concentration</i> | 38 |
| 2.3.3 | PREPARATION OF DNA SAMPLES | 38 |
| 2.3.3a | <i>Genomic DNA from plants</i> | 38 |
| 2.3.3b | <i>Small-scale plasmid preparation</i> | 38 |
| 2.3.3c | <i>Large-scale plasmid preparations for transient assays</i> | 38 |
| 2.3.3d | <i>Preparation of plasmid DNA for automated sequencing</i> | 39 |
| 2.3.3e | <i>Preparation of λ bacteriophage DNA from plate lysates</i> | 39 |
| 2.3.4 | MANIPULATION OF BACTERIOPHAGE | 40 |
| 2.3.4a | <i>Preparation of host cells for phage propagation</i> | 40 |
| 2.3.4b | <i>Plating and isolating bacteriophage</i> | 40 |
| 2.3.5 | CLONING STRATEGIES | 40 |
| 2.3.5a | <i>Cloning of PCR products</i> | 40 |
| 2.3.5b | <i>Preparation of insert DNA and plasmid vector</i> | 41 |
| 2.3.5c | <i>Ligations</i> | 41 |
| 2.3.5d | <i>Preparation and transformation of competent <i>E. coli</i></i> | 41 |
| 2.3.6 | RNA MANIPULATION | 41 |
| 2.3.6a | <i>Total RNA preparation</i> | 41 |
| 2.3.6b | <i>Electrophoretic separation of RNA</i> | 42 |
| 2.3.7 | TRANSFER OF DNA TO NYLON MEMBRANES | 42 |
| 2.3.7a | <i>Southern blotting</i> | 42 |
| 2.3.7b | <i>Plaque lifting</i> | 42 |
| 2.3.7c | <i>Colony hybridisation</i> | 42 |
| 2.3.8 | RADIOACTIVE LABELLING OF DNA PROBES | 43 |
| 2.3.9 | PREHYBRIDISATION, HYBRIDISATION AND WASHING OF FILTERS | 43 |
| 2.3.9a | <i>Prehybridisation</i> | 43 |
| 2.3.9b | <i>Probe denaturing and hybridisation</i> | 44 |
| 2.3.9c | <i>Washing of filters</i> | 44 |
| 2.3.9d | <i>Autoradiography and stripping of filters</i> | 44 |
| 2.3.10 | SEQUENCING OF DOUBLE-STRANDED DNA TEMPLATES | 45 |
| 2.3.10a | <i>Preparation of template DNA</i> | 45 |
| 2.3.10b | <i>Automated sequencing reactions</i> | 45 |
| 2.3.11A | SEQUENCE ANALYSIS SOFTWARE | 45 |
| 2.3.11B | DOTPLOT COMPARISONS OF PROMOTER SEQUENCES | 45 |
| 2.3.12 | OLIGONUCLEOTIDE DESIGN | 46 |
| 2.3.12a | <i>Sequencing primers</i> | 46 |
| 2.3.12b | <i>Primers for Polymerase Chain Reaction (PCR)</i> | 46 |
| 2.3.13 | POLYMERASE CHAIN REACTION (PCR) | 46 |
| 2.3.13a | <i>PCR of plasmid and genomic DNA templates</i> | 46 |
| 2.3.13b | <i>Reverse Transcription Polymerase Chain Reaction (RT-PCR)</i> | 47 |
| 2.3.13c | <i>Rapid amplification of 5' cDNA ends (5'-RACE)</i> | 47 |
| 2.3.13d | <i>GenomeWalker™ PCR</i> | 47 |
| 2.3.14 | TRANSIENT ASSAYS USING THE PARTICLE INJECTION GUN | 48 |
| 2.3.14a | <i>Preparation of gold particles</i> | 48 |
| 2.3.14b | <i>Preparation of Swinney filters</i> | 48 |
| 2.3.14c | <i>Precipitation of DNA onto gold particles for fibre bombardments</i> | 48 |
| 2.3.14d | <i>Preparation of target tissue</i> | 49 |

| | | |
|------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|------------|
| 2.3.14e | <i>The particle injection gun</i> | 49 |
| 2.3.14f | <i>Particle bombardment</i> | 51 |
| 2.3.14g | <i>Histochemical staining of bombarded tissues</i> | 51 |
| 2.3.15 | PHYLOGENETIC ANALYSIS | 52 |
| CHAPTER THREE | | 53 |
| CHARACTERISATION OF AN α-EXPANSIN GENE FAMILY IN COTTON (<i>GOSSYPIUM HIRSUTUM</i>) | | 53 |
| 3.1 | INTRODUCTION | 55 |
| 3.2 | ISOLATION AND PRELIMINARY ANALYSIS OF α -EXPANSIN GENOMIC CLONES | 55 |
| 3.2.1 | <i>Subcloning of genomic fragments with homology to pFS14</i> | 60 |
| 3.2.2 | <i>Mapping of genomic subclones</i> | 63 |
| 3.3 | PRELIMINARY SEQUENCE ANALYSIS OF PUTATIVE α -EXPANSIN GENOMIC REGIONS | 63 |
| 3.3.1 | <i>GhExp10 sequence</i> | 63 |
| 3.3.2 | <i>GhExp12 sequence</i> | 77 |
| 3.3.3 | <i>GhExp14 sequence</i> | 77 |
| 3.3.4 | <i>GhExp6 sequence</i> | 77 |
| 3.3.5 | <i>GhExp5 sequence</i> | 85 |
| 3.4 | PRELIMINARY ANALYSIS OF FIVE COTTON EXPANSIN GENES: A SUMMARY | 85 |
| 3.5 | A SEARCH FOR THE FIBRE-SPECIFIC α -EXPANSIN GENE | 91 |
| 3.5.1 | <i>Obtaining the genes which correspond to gw14.2 and gw14.4</i> | 91 |
| 3.5.2 | <i>The GhExp14.2 gene sequence</i> | 96 |
| 3.5.3 | <i>The GhExp14.4 gene sequence</i> | 101 |
| 3.6 | OBTAINING THE FULL-LENGTH GHEXP5 GENE | 117 |
| 3.6.1 | <i>Genome Walker™ PCR of GhExp5</i> | 117 |
| 3.6.2 | <i>The GhExp5 gene sequence</i> | 117 |
| 3.7 | DETAILED SEQUENCE ANALYSIS OF SIX α -EXPANSIN GENES FROM COTTON | 123 |
| 3.7.1 | <i>The coding region of GhExp10</i> | 123 |
| 3.7.2 | <i>Sequence analysis of the GhExp10 promoter</i> | 135 |
| 3.7.3 | <i>The structure and analysis of five other cotton α-expansin genes</i> | 137 |
| 3.7.4 | <i>Analysis of available promoter sequence of cotton α-expansin genes</i> | 140 |
| 3.8 | EXPRESSION OF α -EXPANSIN GENES IN COTTON | 141 |
| 3.8.1 | <i>Designing gene-specific primers for each cotton α-expansin gene</i> | 141 |
| 3.8.2 | <i>Testing the gene-specific primers</i> | 142 |
| 3.8.3 | <i>RT-PCR transcription analysis of cotton α-expansin genes</i> | 142 |
| 3.9 | α -EXPANSIN GENE EXPRESSION IN COTTON: A DISCUSSION | 143 |
| 3.10 | SUMMARY AND CONCLUSIONS | 149 |
| CHAPTER FOUR | | 151 |
| EVOLUTIONARY ANALYSIS OF SIX COTTON α-EXPANSIN GENES | | 151 |
| 4.1 | INTRODUCTION | 153 |
| 4.2 | GENE REDUNDANCY AND THE EXPANSINS | 154 |
| 4.3 | THE α -EXPANSIN FAMILY: MAXIMUM PARSIMONY ANALYSIS | 155 |
| 4.4 | THE α -EXPANSIN PHYLOGENETIC TREE: GENERAL OBSERVATIONS OF THE FOUR CLADES | 157 |
| 4.5 | RELATIONSHIP BETWEEN α -EXPANSIN SUBCLADES AND A SPECIES TREE | 160 |
| 4.5.1 | <i>The α-expansin A clade</i> | 168 |
| 4.5.2 | <i>The α-expansin B clade</i> | 169 |
| 4.5.3 | <i>The α-expansin C clade</i> | 170 |
| 4.5.4 | <i>The α-expansin D clade</i> | 171 |
| 4.5.5 | <i>Database searches in Arabidopsis</i> | 171 |

| | |
|--------------------------------------------------------------------------------------------------------------------------------------------|------------|
| 4.6 DISCUSSION: EVOLUTIONARY ASPECTS OF EXPANSINS | 172 |
| 4.7 CONCLUSIONS | 175 |
| CHAPTER FIVE | 177 |
| DEVELOPMENT, OPTIMISATION AND UTILISATION OF A TRANSIENT ASSAY IN COTTON FIBRES | 177 |
| 5.1 INTRODUCTION | 179 |
| 5.2 THE PARTICLE INJECTION GUN | 180 |
| 5.3 OPTIMISATION OF THE TRANSIENT ASSAY IN TOBACCO LEAVES | 183 |
| 5.4 OPTIMISATION OF THE TRANSIENT ASSAY FOR COTTON TISSUES | 185 |
| 5.5 UTILISATION OF THE TRANSIENT ASSAY SYSTEM TO TEST FIBRE-SPECIFIC PROMOTERS | 197 |
| 5.5.1 <i>The promoter::reporter constructs</i> | 197 |
| 5.5.2 <i>Transient assay results using the pGUSEXP14.4 construct</i> | 202 |
| 5.5.3 <i>Transient assay results using the pGUSEXP14.2 construct</i> | 202 |
| 5.6 DISCUSSION REGARDING TRANSIENT ASSAY RESULTS USING PGUSEXP14.2 AND PGUSEXP14.4 | 203 |
| 5.7 SUMMARY AND CONCLUSIONS | 208 |
| CHAPTER SIX | 211 |
| FUNCTIONAL DISSECTION OF THE <i>GHEXP14.2</i> α-EXPANSIN FIBRE-SPECIFIC PROMOTER USING A TRANSIENT ASSAY | 211 |
| 6.1 INTRODUCTION | 213 |
| 6.2 A SEARCH OF THE <i>GHEXP14.2</i> PROMOTER SEQUENCE FOR KNOWN <i>CIS</i> -ACTING ELEMENTS | 214 |
| 6.2.1 <i>Dof core binding site</i> | 215 |
| 6.2.2 <i>Gibberellin response elements</i> | 219 |
| 6.2.3 <i>The prolamin box</i> | 221 |
| 6.2.4 <i>MYB transcription factor elements</i> | 221 |
| 6.2.5 <i>The E-box</i> | 222 |
| 6.2.6 <i>Summary</i> | 223 |
| 6.3 SEQUENCE ANALYSIS OF COTTON FIBRE-SPECIFIC PROMOTERS | 223 |
| 6.3.1 <i>A search for known elements in the fibre-specific promoters</i> | 225 |
| 6.3.2 <i>Comparisons of nine fibre-specific promoters</i> | 225 |
| 6.4 ISOLATION OF THE FULL-LENGTH PFS14 CDNA | 232 |
| 6.4.1 <i>Primer design and 5'-RACE PCR of pFS14</i> | 232 |
| 6.4.2 <i>Sequence analysis of pFS14 5'-RACE products</i> | 238 |
| 6.5 FUNCTIONAL ANALYSIS OF A FIBRE-SPECIFIC PROMOTER | 239 |
| 6.5.1 <i>Primer design for creating gw14.2 promoter deletions</i> | 239 |
| 6.5.2 <i>PCR amplification and cloning of gw14.2 promoter deletions</i> | 242 |
| 6.5.3 <i>Functional analysis of a promoter using transient assays</i> | 242 |
| <i>pEXP14.2prom1 (848 bp full-length)</i> | 243 |
| <i>pEXP14.2prom2 (604 bp)</i> | 243 |
| <i>pEXP14.2prom3 (384 bp)</i> | 243 |
| <i>pEXP14.2prom4 (149 bp)</i> | 243 |
| 6.5.4 <i>Summary of GhExp14.2 promoter deletion transient assays</i> | 251 |
| 6.6 SUMMARY AND CONCLUSIONS | 252 |
| CHAPTER SEVEN | 255 |
| GENERATION OF α-EXPANSIN (PFS14) EXPRESSION AND MISEXPRESSION CONSTRUCTS FOR ANALYSIS IN TRANSGENIC COTTON PLANTS | 255 |
| 7.1 INTRODUCTION | 257 |
| 7.2 PROLONGED α -EXPANSIN EXPRESSION IN COTTON FIBRES | 259 |

| | |
|---------------------------------------------------------------------------------|------------|
| 7.3 INCREASED α -EXPANSIN GENE EXPRESSION IN COTTON FIBRES | 274 |
| 7.4 DECREASED α -EXPANSIN EXPRESSION IN COTTON FIBRES | 275 |
| 7.5 INCREASED AND PROLONGED α -EXPANSIN GENE EXPRESSION IN COTTON FIBRES | 282 |
| 7.6 TRANSFORMATION OF COTTON | 283 |
| 7.7 CONCLUSIONS AND CONSIDERATIONS | 287 |
| APPENDICES | 289 |
| REFERENCES | 295 |

Abstract

Cotton fibres arise from single epidermal cells of the ovule and initiate as protrusions on the day of flowering (anthesis). Elongation continues for approximately 20 days, resulting in fibres which are 2-3.5 cm in length. As fibres are single cells of homogenous tissue and their development is synchronous, they are an excellent model system for the analysis of the molecular processes involved in cell growth.

One type of cell wall protein, the α -expansins, has been implicated in cell wall modification. α -expansins have been characterised in a number of plant species and are encoded by large gene families (31 members in *Arabidopsis*), with family members often showing heterogenous patterns of expression (Wu *et al.*, 2001). In some plant species, α -expansins have specifically been implicated in cell elongation, such that a cotton fibre-specific α -expansin is an attractive target for cotton biotechnology. Differential screening of a cotton fibre cDNA library allowed the identification of a fibre-specific α -expansin mRNA (pFS14) of which accumulated at high levels during fibre cell elongation (Orford and Timmis, 1998). The aims of this project focused on the isolation, characterisation and analysis of the fibre-specific α -expansin gene and other cotton α -expansin family members.

Using, a combination of genomic library screening and PCR-based approaches six different cotton α -expansin gene family members were isolated (Chapter Three). One of the genes, *GhExp14.2*, was 99% similar to the pFS14 cDNA, and therefore it was considered to be the genomic representative of pFS14 cDNA.

RT-PCR analysis (Chapter Three) showed that five of the six α -expansins have distinctive expression patterns in different cotton tissues, and that the transcripts of four genes were detected in elongating fibres. *GhExp14.2* transcripts were detected at high levels in elongating fibres, consistent with the results described by Orford and Timmis (1998), suggesting that *GhExp14.2* plays an important role in fibre development.

Evolutionary analysis of α -expansin genes from cotton and other plant species (Chapter Four) confirmed previous studies that demonstrate the α -expansin family is divided into four ancient, distinct subclades. Each subclade was further divided into groups on the basis of gene orthology. In some cases 'orthologous groups' within clades contain multiple genes that appeared to be expressed in a similar manner to each other and the evolutionary analysis provides a foundation by which specific orthologous α -expansin genes could be identified and characterised.

The development of a transient assay for use in cotton fibres was undertaken (Chapter Five). Various parameters and conditions were tested on tobacco and cotton tissues until an appropriate method was refined, with fibres aged approximately 4 DPA providing the best results. The validity of the transient assay technique was tested using the *GhExp14.2* promoter, with the full available sequence (848 bp up to and including the ATG initiation codon) directing GUS reporter gene expression only in developing fibre cells and not other cotton tissues tested. This work describes an efficient, simple method for the functional dissection of cotton fibre-specific promoters.

Known *cis*-acting consensus elements in the *GhExp14.2* fibre-specific promoter were accurately mapped using transcription factor binding site websites (Chapter Six). Detailed comparisons of multiple fibre-specific promoters allowed identification of four new putative fibre-specific promoter elements. A deletion analysis was undertaken to dissect the functional regions of, and putative consensus binding sites within, the *GhExp14.2* promoter. All *GhExp14.2* promoter deletions maintained fibre-specific expression of the GUS reporter gene, including the smallest, which consisted of only 149 bp. Within the *GhExp14.2* 149 bp minimal promoter five previously characterised consensus-binding sites (ACGT box, Q-element, RY pyrimidine repeat, an E-box and a GCN4 motif) and one novel binding site (PFE4) were present, and each was suggested to be involved in directing fibre-specific transcription.

In Chapter Seven, preparation of four constructs using the fibre-specific cDNA pFS14, under the control of different fibre-specific promoters, is described. These constructs were designed with a view of modifying fibre morphology by altering the temporal expression pattern of the α -expansin gene in transgenic cotton. Results from the transgenic cotton, although not available for incorporation into this thesis, will be used to define the important role that *GhExp14.2* plays in fibre elongation, and may also facilitate the production of cotton varieties with improved fibre morphology.