

**Towards Identifying Cis and Trans Regulators of Expression of Xylem Cysteine Protease 1
(XCP1) in Arabidopsis**

By
William Jefferson Stroud Jr.

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in
partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In

Horticulture

Dr. Eric P. Beers, Chairman

Dr. Jerzy Nowak

Dr. Glenda Gillaspay

May 6, 2009

Blacksburg, Virginia

Keywords: *Arabidopsis thaliana*, cysteine protease, promoter deletion, tracheary element, yeast one-hybrid, southwestern blot

Towards Identifying Cis and Trans Regulators of Expression of Xylem Cysteine Protease 1 (XCP1) in *Arabidopsis thaliana*

By

William Jefferson Stroud Jr.

Dr. Eric P. Beers, Chairman

Department of Horticulture

(ABSTRACT)

Secondary xylem, commonly known as wood, is a valuable commercial commodity. Among the major components of wood are the elongated, thick-walled water-conducting cells known as tracheary elements. Understanding tracheary element differentiation and maturation is of scientific and commercial importance as it may lead to broad understanding of cellular differentiation processes as well as ways to increase both the quality and quantity of wood produced by economically important tree species. One way to begin to understand the regulation of tracheary element differentiation is to identify elements that control expression of genes associated with tracheary elements. In *Arabidopsis thaliana*, Xylem Cysteine Protease 1 (XCP1) is specifically expressed in tracheary elements where it catalyzes microautolysis. Thus XCP1 can serve as a useful model for identifying factors that regulate tracheary element-specific gene expression. A deletion analysis of the XCP1 promoter was conducted to identify promoter elements that are necessary and sufficient for tracheary element-restricted gene expression. Two regions required for tracheary element-specific gene expression were identified. One of these was assembled as a multimeric bait construct and used in yeast one-hybrid assays to identify candidate transcription factors that bind to the XCP1 promoter region. Subsequently, a southwestern blot analysis was used to identify transcription factors displaying specific binding to a previously reported cis-element, CTCAAAGCCA, found in the XCP1 promoter and other tracheary element-associated genes from multiple species.

ACKNOWLEDGEMENTS

I'd like to thank all the members of the Beers lab for their assistance in completing this degree, especially:

Dr. Eric Beers: Thank you for your time, knowledge, and patience. I deeply appreciate the opportunities you've opened for me.

Dr. Chengsong Zhao: Thank you for sharing your time, your technical expertise, and your unique perspective on many different issues. I've benefitted immensely from having access to your technical knowledge of molecular biology.

Earl Petzold: Thank you for supporting my work, your technical advice, and your friendship. Say hi to Eileen for me.

I'd also like to thank:

My parents Jeff and Sandra Stroud: Thank you for all your encouragement and support, and not letting me quit.

Martha Vaughan: Thanks for all the years of friendship. I hope we are colleagues for many long years to come.

Ellie Sturgis: Your timely help made all the difference. Thank you.

I dedicate this thesis to my dog Tobey, my very best friend.

TABLE OF CONTENTS

Acknowledgements.....	iii
Table of Contents	iv
List of Figures.....	v
Chapter 1: Literature Review.....	1
Chapter 2: Materials and Methods	14
2.1 Cloning and Transformation.....	14
2.2 Promoter Deletion Analysis	15
2.3 Yeast One-Hybrid Analysis.....	16
2.4 Southwestern Blot Analysis.....	16
2.4.1 3' End-Labeling of Probe and Probe Complement.....	16
2.4.2 Plaque Screening.....	16
Chapter 3: Results.....	18
3.1 Promoter Deletion Results.....	18
3.2 Yeast One-Hybrid Results.....	20
3.3 Southwestern Blot Results.....	25
Chapter 4: Discussion.....	27
References.....	32

LIST OF FIGURES

Figure 1 <i>XCP1</i> full sequence.....	5
Figure 2 Transcription factor-mediated activation/repression of transcription.....	6
Figure 3 Promoter deletion analysis, 3' deletion example.....	8
Figure 4 Promoter deletions and their ability drive GUS expression in tracheary elements.....	10
Figure 5 Primer locations used for <i>XCP1</i> promoter deletion analysis.....	11
Figure 6 Overview of positive yeast one-hybrid reaction	13
Figure 7 <i>XCP1</i> promoter from positions -268 to -168.....	21
Figure 8 Construction of pXCP1p ^{3'24} bait vector for use in Y1H analysis.....	22
Figure 9 Southwestern blot; chemiluminescent visualization.....	26
Figure 10 Primer sequences for <i>XCP1</i> promoter deletion analysis between bases -253 & -213..	31

LIST OF TABLES

Table 1 BLAST results for sequences generated by yeast one-hybrid.....	24
---	----

Chapter 1

Literature Review

Xylem, (pro)cambium and phloem are the tissues that comprise the plant vascular tissue system (Ye et al., 2002). Xylem is primarily responsible for the transport of water, while phloem is primarily responsible for the transport of sugars. Primary xylem and phloem are derived from procambium. Secondary xylem, also called wood, and secondary phloem are derived from the vascular cambium. Together with the cork cambium, the vascular cambium is responsible for lateral expansion characteristic of woody plants.

Secondary xylem in angiosperms is commonly composed of four cell types: tracheids, vessel elements, fibers and parenchyma (Turner et al., 2007). Collectively, tracheids and vessel elements are referred to as tracheary elements (TEs). For a review of TE differentiation and function, see Turner et al. (2007). TEs are elongated cells, arranged axially within the plant stem, and lack a protoplast at maturity. Their cell walls are thickened due to the deposit of a network of cellulose, hemicellulose and lignin polymers in layers that form the secondary cell wall. Water passes freely through vessel elements because they are perforated at their connecting ends. In contrast, water moves through tracheids less efficiently, as it must pass through modified areas of the cell wall known as pit membranes. When joined end-to-end, mature vessel elements form hollow tubes capable of efficiently transporting water. While mature tracheary elements are dead and incapable of maintaining themselves, adjacent living parenchyma cells help maintain tracheary element integrity and regulate the content of TEs.

In the final step of TE maturation the cell digests its own protoplast, leaving behind the thickened secondary cell wall. This process of cellular autolysis is a form of programmed cell death (PCD). Thus differentiation of TEs is achieved via the integration of pathways that regulate division, elongation and/or expansion, cell wall synthesis and PCD.

Increases in serine and/or cysteine protease activity (Minami and Fukuda, 1995; Beers and Freeman, 1997), gene expression (Moreau et al., 2005; Zhao et al., 2005), and protein levels (Funk et al., 2002) are correlated with TE death and autolysis. In *Arabidopsis* xylem, two paralogous C1A-family, papain-like cysteine proteases (Beers et al., 2004), XCP1 and XCP2, are specifically associated with late-stage TE differentiation (Funk et al., 2002). It was recently demonstrated that XCP1 and XCP2 cooperate to catalyze autolysis in differentiating TEs (Avci et al., 2008) and they may also be components of basal defense against plant vascular pathogens (van Esse et al., 2008). This potential for dual functionality makes these proteases very interesting subjects for investigations of both cis and trans components regulating their expression. See Figure 1 for the full length XCP1 sequence.

In a simplified analysis, the two important parts of a eukaryotic gene are the coding region and the promoter region. The coding region contains the information needed to create a protein through the combined processes of transcription and translation. The promoter region is a non-coding region that regulates transcription of the adjacent coding region. It does so partly by providing binding sites, known as cis-elements, for transcription factors (TFs), which regulate the process of transcription (see Figure 2). The identification of cis-element(s)

contained within the promoter and the corresponding binding transcription factors are crucial to understanding the regulation of a gene.

Several cis-elements have been empirically determined to have positive or negative regulatory effects on genes associated with secondary xylem formation (Leyva et al., 1992; Hauffe et al., 1993; Hatton et al., 1995). In these studies, promoter deletion analyses of phenylalanine ammonia-lyase (*PAL*) and 4-coumarate:CoA ligase (*4CL*), genes involved in phenylpropanoid metabolism, were conducted. Sequences for several putative cis-elements were identified, including TCCACCAACCCC (Leyva et al., 1992), AGTCAACT and TCTCACCAACCCC (Hauffe et al., 1993), and GTTAGGTC, GCCACGTCGCGC, CCACCAACCCCC, and CCCACCTACC (Hatton et al., 1995).

Work in the Fukuda lab resulted in the identification of the tracheary-element-regulating cis-element, or TERE (Pyo et al., 2007). Using a combination of promoter deletion analysis, mutational analysis, and gain-of-function analysis, an 11-bp cis-element was identified that “confers TE-specific expression to both genes related to secondary wall formation or modification and PCD” (Pyo et al., 2007). Five tandem copies of the sequence CTTGAAAGCAA from the of the *Zinnia cysteine protease 4* (*ZCP4*) promoter were cloned upstream of the CaMV 35S minimal promoter-reporter, -46CamVGUS, and introduced into Arabidopsis plants. The construct conferred high levels of differentiating TE-specific reporter expression. Based on similar gain-of-function experiments with conserved sequences from several other TE-specific genes in Arabidopsis (including *XCP1*, *XCP2*, and *XSP1*), the TERE sequence was identified as CTTNAAAGCNA. Based on this study and *XCP1* promoter deletion analysis data gathered by the

Beers lab, it is likely that the TERE cis-element is one of at least two elements in the *XCP1* promoter responsible for regulating *XCP1* expression.

Using a computational biology approach linked with transcript profiling, Oh et al., (2003) identified several known cis-element sequences that showed increased frequency within putative promoters of genes exhibiting expression associated with *Arabidopsis* secondary growth, including Abscisic Acid Response Motif 3 (ABRE3, CAACGTG) and Extensin A (extA, AACGTGT) elements. Further investigation is needed to determine what roles, if any, these elements play in xylem formation.

Several recent papers have focused on illuminating the network of transcription factors regulating deposition of lignin, cellulose and hemicellulose in vascular tissue. Work in the Ye lab has described a broad network of transcription factor interactions leading to the process of the secondary wall deposition, starting with NST1 and SND1 acting as master switches in fibers, while the SND1 close homologs VND6 and VND7 trigger secondary wall biosynthesis in vessels (Zhong et al., 2008). MYB46 also regulates secondary wall biosynthesis. SND1 was shown to bind specifically to the MYB46 promoter (Zhong et al., 2007). In turn, MYB58 and MYB63 appear to be downstream targets of MYB46 and control lignin biosynthesis (Zhou et al., 2009).

gtaagttaaccacattttgcttttgctaataacaatcatctcaacttttgatatcaaaagtagccaccaacat
 ttatgcacataagagccactgactgcgtaaaagggtgtctcataaaatgcattactgtggttgcaactttgca
 gggtcgtctctggttcagtataattacattccgcaggaagctacgatagtaaagatcttcacacttacat
 ctaccaagaaaacaccattattctagtagacagagataaagcttttgaacgattacgattttggttttcggtg
 aggtatattcacattttgcttctgtattgattttttttccgctctcatgaaaccttctattctcatttgt
 acgataccaaaacttatcaagaatagttacaatttcagaaattttaagcctctacgtaatttttttataa
 aacacccctccggatatttttaataatcattaactgtctgaacgtgagctgtatgttgaaaattgcacgctt
 agaacaaaaggcttaaccaaaaatgatccaaccgtgaagactcggagaccggcgacttagtttttaaatt
 aatcattgtctttgcttcaaagccaatcctatgagactttgtcttctccaacttgtatataagctctat
 tcctctactctgtttcacaccatctcttcttcttcttcttctctcagtgaaacaaatgtggctATGGCT
 TTTTCTGCACCATCACTTTCCAAATTTCTCTCTTTTGGTTGGCATTTCAGCATCAGCTCTCCTGTGTGTC
 TTTTGGCCGTGATTTCTCCATTGTTGGATACACGCCGGAGCATTGACAAACACTGACAAGCTTCTAGAGC
 TCTTCGAGTCATGGATGTCAGAACACAGCAAGGCTTACAAAAGCGTGGAGGAGAAGGTGCACAGGTTTGGAG
 GTTTTCAGAGAGAATCTGATGCATATAGACCAGAGGAACAATGAGATCAACAGTTACTGGCTCGGTTTGAA
 CGAGTTTGGGATTTGACCCATGAAGAGTTCAAAGGAAGATATCTAGGACTTGCAAAGCCACAATTCTCTA
 GAAAGAGACAGCCCTCAGCTAACTTCAGGTACAGAGATATCACGGACTTGCCATAATCCGTAGACTGGAGA
 AAGAAAGGCGCTGTGGCTCTGTCAAGGACCAGGGTCAATGTGgtaacatttctctctttccatcttggtc
 ccaggtatataagaacagaactcagagattcgaatagctatccaacgatcttattttgttgctgcagGTAGC
 TGTTGGGCATTTTCAACAGTTGCAGCTGTGAGGGGATCAACCAGATCACAACAGGGAATCTGAGTTTCGCT
 TTCAGAGCAAGAACTCATAGACTGTGACACAATTTCAACAGTGGCTGCAATGGAGGTCTCATGGACTACG
 CATTCCAGTACATAATTTGACCGGTGGTCTCCACAAAGAAGATGATTACCCTTATCTCATGGAGGAAGGA
 ATTTGTCAAGAGCAGAAAgtacaacactagacacaaacataatcaaaacttaacactgggaaacagatt
 ttttataaacactgatcaatcatgaaccttacacagGAGGATGTGGAACGTGTGACAATCAGCGGCTACGA
 AGATGTCCCTGAAAATGATGACGAAAGCCTGGTGAAGGCTTTAGCTCATCAGCCAGTCAGTGTGGCTATTG
 AGGCTTCAGGAAGAGACTTCCAGTTCTACAAAGGGgtatataacagatgattatgatatcgaaattgtggt
 ctatgtatccaatctttcattctctcaaaaaccgatttggagatttttgtttgttttacagGGAGTGTTTAA
 TGGTAAATGTGGAACAGACCTAGACCACGGTGTGGCAGCGTTGGATATGGTTCATCAAAGGGATCTGACT
 ATGTTATTGTCAAGAACTCATGGGGACCAAGATGGGGAGAGAAAAGGTTTATTAGGATGAAGAGAAACT
 GGTAACCAGAGGGACTCTGTGGAATCAACAAGATGGCCTCATATCCTACCAAGACCAAGTGA

Figure 1: *XCP1* full sequence. Black: putative promoter region. Red: 5' UTR. Yellow: exons. Purple: introns.

Transcription Factor Activation/Repression of Transcription

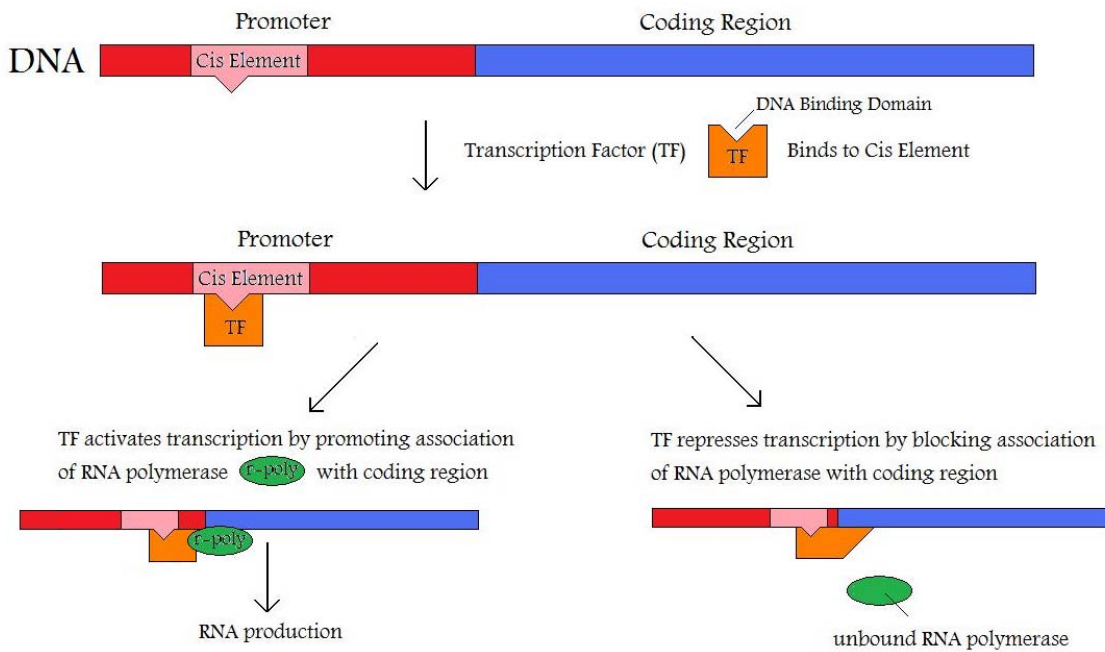


Figure 2: Transcription factor-mediated activation/repression of transcription. Cis-elements in a gene's promoter region recruit transcription factors that bind to specific sequences known as cis-elements. Once bound, transcription factors can either recruit RNA polymerase, which initiates transcription of the associated coding region, or block RNA polymerase recruitment by a variety of methods and repress transcription (Latchman, 1997).

One well documented approach to discovering functional cis-elements is called a promoter deletion analysis (de Bruijn et al., 1989). Starting with a 1-2 kb portion of the promoter, deletions are made from the 5' or 3' end. These shortened promoters are inserted directly upstream of the coding region of a reporter gene, sometimes in fusion with a minimal promoter as in the case of 3' deletions. By ligating this gene construct into a binary vector and using that vector to transform wild type *Arabidopsis* plants, the regions necessary for expression, and presumably containing the essential cis-elements, are identified by virtue of a loss, diminution or incorrect pattern of reporter expression or activity. Once these necessary regions are identified, other techniques such as predictive computational biology, tests of synthetic multimers of putative cis-elements, yeast one-hybrid assays, gel mobility shift assays, etc., can be used to specifically identify functional cis-elements contained in the region (Figure 3).

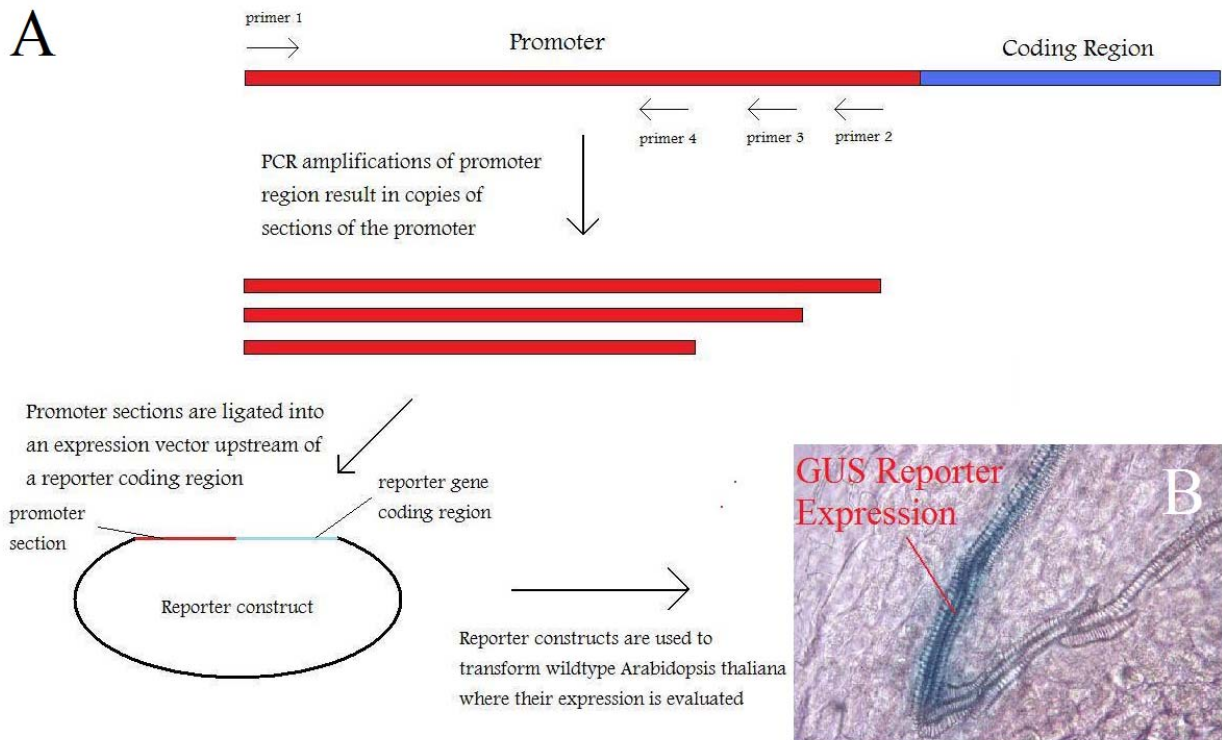


Figure 3: Promoter deletion analysis, 3' deletion example.

A. PCR is used to generate specific sections of the promoter region of interest. These sections are then ligated into an expression vector.

B. The construct's ability to drive reporter gene expression can be tested in transgenic plants. From reporter gene expression, regions of the promoter which are necessary for native gene expression can be inferred. Photo by C. Zhao, Beers lab, Virginia Tech.

A previous promoter deletion analysis was conducted in the Beers lab by Ihab Ismail. Sequential 5' deletions of the *XCP1* promoter (XCP1p) were fused to the *uidA* gene for the reporter enzyme β -glucuronidase (GUS). A 237-base pair (bp) region of XCP1p was identified as both necessary and sufficient for driving *GUS* expression in Arabidopsis TEs (I. Ismail and E. Beers, unpublished; see Figure 4). A follow-up promoter deletion series was created to analyze the 5' end of this 237-bp region, and the results of that work are discussed below (See Figure 5 for primer sequences and positions used). Promoter deletion and synthetic promoter analyses are capable of revealing which regions of a promoter are necessary and sufficient to promote native gene expression. While useful, such experiments cannot directly identify the trans factor that binds to a putative regulatory sequence. To answer this question, tests for protein-DNA interactions are needed. Two tests for revealing protein-DNA interactions are the yeast one-hybrid system and southwestern blotting.

XCP1 Promoter Deletion::GUS Fusion Results

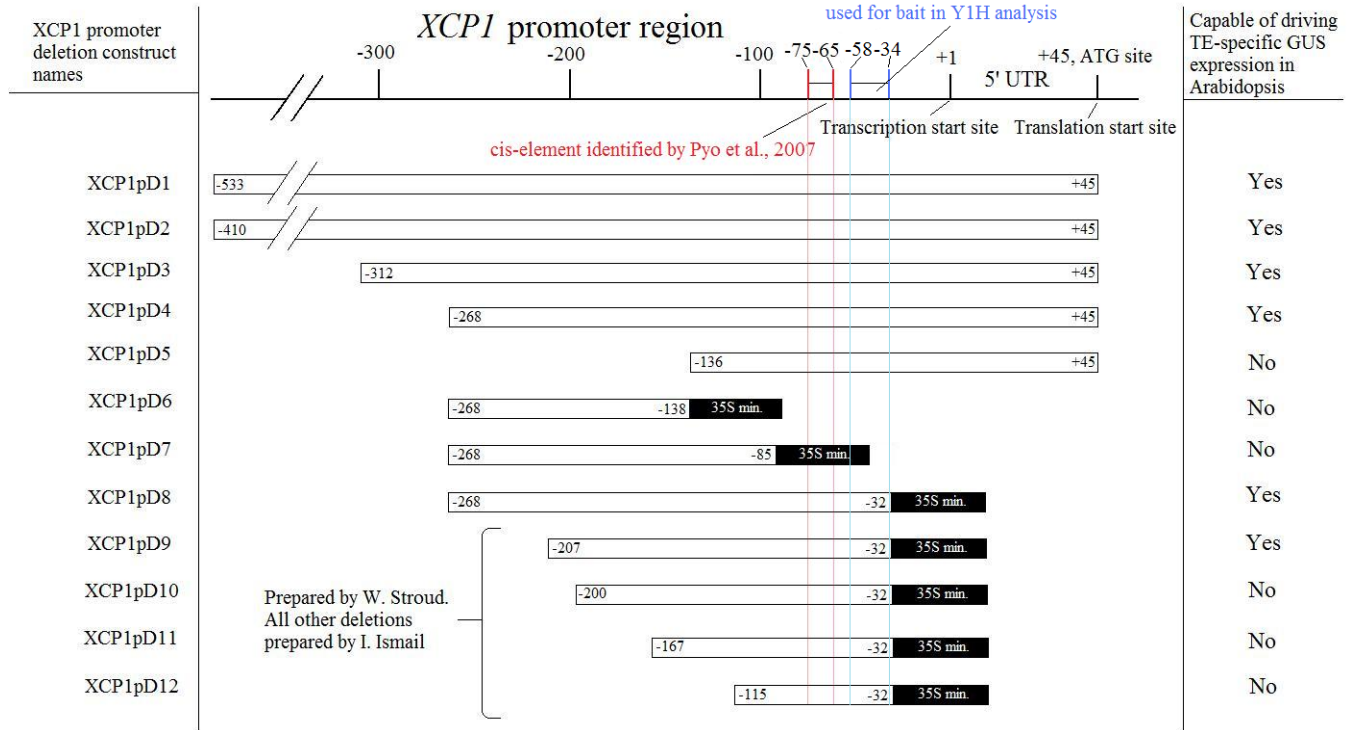


Figure 4: XCP1 promoter deletions and their ability drive GUS expression in tracheary elements. Numbers at the 5' and 3' ends of each deletion correspond to the positions of the respective 5' and 3' end bases, with +1 defined as the 5' end base of the XCP1 5' UTR.

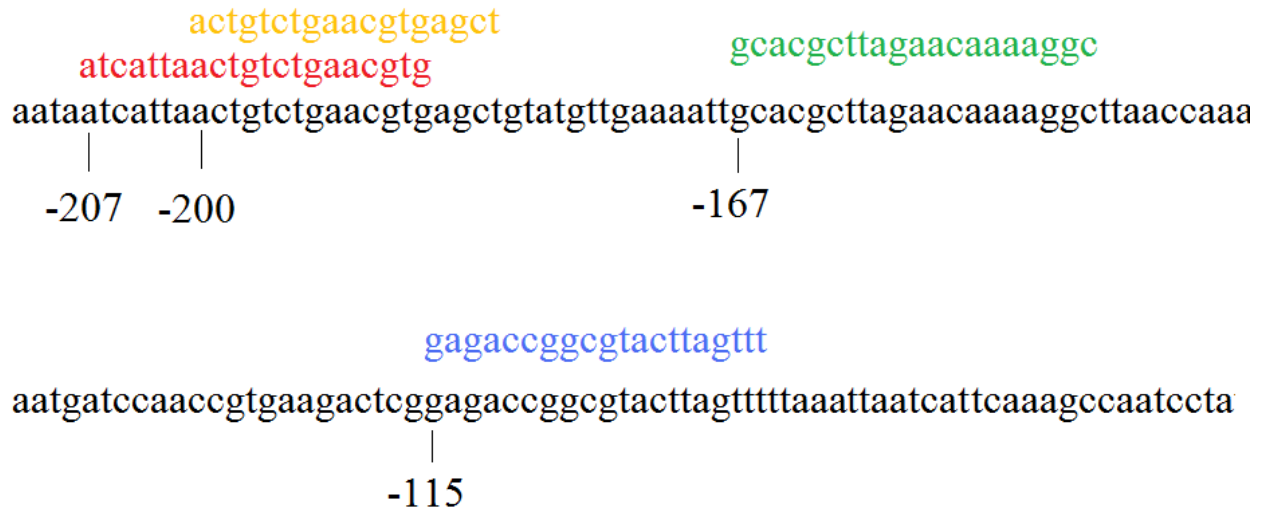


Figure 5: Primer locations used for *XCP1* promoter deletion analysis conducted by W. Stroud. *XCP1* promoter sequence and base locations shown in black, primer sequences shown in color. Red: PR-207, Orange: PR-200, Green: PR-167, Blue: PR-115. Primers correspond to deletion constructs XCP1pD 9-12 (PR-207 for XCP1pD9, PR-200 for XCP1pD10, PR-167 for XCP1pD11, and PR-115 for XCP1pD12). See Figure 6 for more information.

The yeast one-hybrid system is designed to reveal specific binding between DNA and proteins. *Saccharomyces cerevisiae* strain Y187 was used in this test because it does not produce histidine, an amino acid essential for normal growth. In a yeast one-hybrid screen, Y187 cells are co-transformed with bait and prey plasmid. Bait plasmid consists of the pHIS2 plasmid with DNA containing known or suspected cis-elements ligated into the promoter region of *HIS3*, a reporter gene capable of complementing the *HIS3* deficiency of Y187. Prey plasmid consists of cDNAs (in this case prepared from RNA isolated from Arabidopsis xylem) ligated directly upstream of the coding region of the GAL4 activation domain. Transcription and translation of the prey vector should result in a fusion protein containing a library protein fused to a functional GAL4 activation domain. When Y187 cells are co-transformed with bait and prey, and the prey cDNA codes for a protein capable of binding to bait elements with proper positioning, the GAL4 activation domain fused to the prey protein initiates *HIS3* expression and hence histidine production, allowing yeast growth on histidine-deficient media. (See Figure 6) In the research discussed below, a region of the *XCP1* promoter from positions -58 to -32 was used as bait in a yeast one-hybrid analysis with the objective of identifying transcription factor(s) that potentially act as regulators of *XCP1* expression.

Yeast One-hybrid Overview

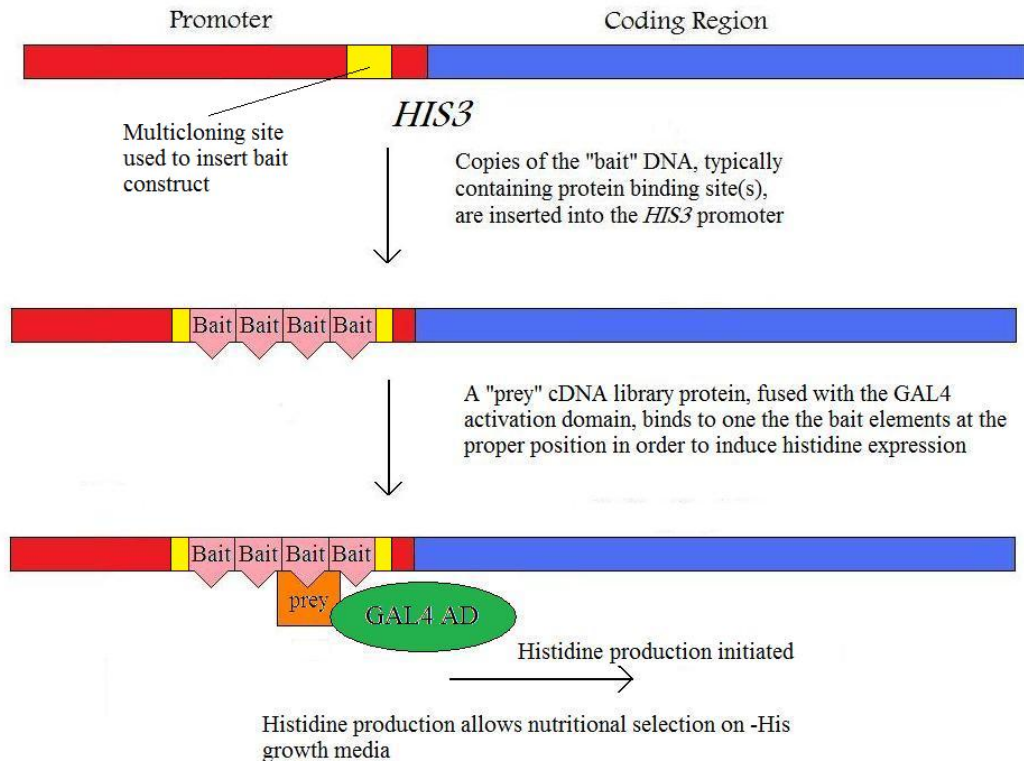


Figure 6: Overview of positive yeast one-hybrid reaction.

Southwestern blot analysis, as the name suggests, involves a combination of Southern blot technique and western blot technique. Southwestern blotting is used to detect a specific interaction between a protein and a DNA sequence. Either DNA or protein can be bound to a membrane, and then probed to detect specific binding interactions: Membrane-bound DNA to protein probe or membrane-bound protein to DNA probe. In the research discussed below, a PVDF-membrane was used to bind a broad range of proteins, which were then probed for specific binding to a sequence of DNA with the objective of identifying transcription factors that potentially act as regulators of *XCP1* expression.

Chapter 2: Methods and Materials

2.1 Cloning and transformation

Polymerase Chain Reaction (PCR) was used to amplify desired sequences from template genomic DNA. PCR product was ligated into pGEM-T Easy vector, (Promega Corp. Madison, WI) according to the Promega protocol. Recombinant vector was then used to transform *Escherichia coli* strain DH5 α cells via heat shock. Transformed *E. coli* was selected on LB plates containing 100 $\mu\text{g/ml}$ ampicillin. Plasmid DNA was purified using the Qiagen DNA Mini Kit (Qiagen Corp. Germantown, MD) and inserts were sequenced to confirm the absence of mutations.

Upon identity confirmation, if recombinant DNA was to be used for plant transformation, the desired sequence was removed from pGEM-T Easy via restriction digest using *Bam*HI and *Hind*III endonucleases, isolated by agarose gel electrophoresis, collected by agarose gel purification using the QIAEX II Gel Extraction Kit (Qiagen Corp. Germantown, MD), and ligated into the binary vector pBI121.3 constructed by Zhao (2005). Recombinant pBI121.3 was then used to transform *Agrobacterium tumefaciens* strain GV3101 following the procedure outlined by (Chen et al., 1994), except that the liquid nitrogen freezing step was replaced by a heat shock method. The heat shock consisted of thawing *A. tumefaciens* GV3101 cells to room temperature, adding recombinant pBI121.3, placing cells in a 42°C water bath for 45 seconds, and then immediately placing them on ice for 5 minutes. 1 ml of 2XTY broth was added and cells were incubated at 28°C for 2 hours with shaking at 250 rpm. Cells were pelleted by centrifugation at 2000 rpm for 5 minutes, resuspended in 100 μl of 2XTY and spread on LB

plates containing 34 µg/ml rifampicin, 50 µg/ml gentamicin, and 50 µg/ml kanamycin for transformant selection. Plates were grown at 28°C for 24-36 hours until visible colonies formed. Single colonies were used to inoculate liquid cultures as described by Clough and Bent, (1998).

For transformation of *Arabidopsis thaliana*, the floral dip method was used as described by Clough and Bent, (1998). Finale® Herbicide (0.03% final concentration of active ingredient Glufosinate-ammonium) was sprayed on seedlings at 1 and 2 weeks post germination to select for transformants.

2.2 Promoter Deletion Analysis

To construct XCP1pD9, XCP1pD10, XCP1pD11, and XCP1pD12, relevant sequences were amplified by PCR using *Arabidopsis* genomic DNA as template. All reactions used the primer sequence GGATCCCAAGTTGGAGACAAGACA (PR67 from Ismail, 2004) for the reverse primer. Forward primer sequences for XCP1pD9-12 were AAGCCTATCATTAAGTGTCTGAACGT (PR-207 for XCP1pD9), AAGCTTACTGTCTGAACGTGAAGCTG (PR-200 for XCP1pD10), AAGCTTGACGCTTAGAACAAAAGGC (PR-167 for XCP1pD11), and AAGCTTGAGACCGGCGTACTTAGTTT (PR-115 for XCP1pD12).

All plasmids used consisted of modified p1058 (Ismail, 2004), which contains a 35S minimal promoter fused to the reporter gene β-glucuronidase (GUS). p1058 underwent *Bam*HI and *Hind*III digestion to remove unneeded DNA, and PCR products were then ligated into the linearized plasmid. Recombinant plasmids were used to transform *A. tumefaciens*, and transgenic *A. tumefaciens* was then used to transform *A. thaliana* using floral dip as described

above. After plant transformation and selection, GUS activity was assayed according to Petzold (2007).

2.3 Yeast one-hybrid analysis

Yeast cultures, yeast transformation, recombinant plasmid construction, and preparation of growth media were conducted according to the user manual for BD Matchmaker™ Library Construction and Screening Kit (BD Biosciences, Palo Alto, CA). The Arabidopsis root hypocotyl cDNA library was constructed using total RNA isolated from 8-week-old Arabidopsis root hypocotyls (Zhao et al., 2000).

Since plasmid could not be isolated directly from putative positive yeast cells, colony PCR was conducted on putative positive yeast colonies to amplify plasmid inserts. PCR products were then sequenced and analyzed using a Basic Local Alignment Search Tool (BLAST) search (Altschul et al., 1990).

2.4 Southwestern blot analysis

2.4.1 3' end-labeling of probe and probe complement

The Pierce Biotin 3' End DNA Labeling Kit (Thermo Fisher Scientific, Rockford, IL) was used to label all probe DNA. The Pierce Biotin 3' End DNA Labeling Kit user manual was followed for this procedure (Thermo Fisher Scientific, Rockford, IL).

2.4.2 Plaque Screening

LB agar plates and top agar were prepared with the addition of 1 mM IPTG. LB agar plates were prepared 24 hours before use. LB agar plates were preincubated for 1 hour at 37°C immediately before use. LB top agar was prepared as per standard LB agar plates, but only half

the amount of agar was used (i.e. 7.5 g instead of 15 g for 1L of media). Top agar was stored in 15 ml tubes in 10 ml aliquots.

E. coli XL-1 Blue cells were grown overnight in 25 ml liquid LB media supplemented with 10mM MgSO₄ and 0.2% (v/v) maltose at 37°C with shaking at 250 rpm. Bacteria were pelleted by centrifugation at 2000rpm for 10 minutes, gently resuspended in half the original volume of sterile 10mM MgSO₄. OD600 was adjusted to 0.5. Top agar tubes were placed in a 50°C water bath for ~10 minutes (until top agar was melted). Phage (pretitered) was added to bacterial cells and incubated at 37°C for 15 minutes. Phage/bacteria mixture was then added to liquid top agar, mixed by inverting the 15 ml tubes several times, and poured directly onto a pre-warmed LB agar plate. The plate was swirled several times for even distribution of top agar, and the plates were allowed to cool for ~30 minutes until top agar was completely solidified. Plates were incubated at 42°C until visible plaques formed. Plaques were typically visible in 8-12 hours, but times varied up to 36 hours.

PVDF membranes were prepared by placing marks with a pencil to confirm identification and orientation. PVDF membranes were wet in methanol, rinsed with sterile water to remove methanol, and placed onto the agar plates in the proper orientation. Forceps were used to gently guide air bubbles to the perimeter. Once air bubbles were eliminated, plates were incubated at 37°C for 3.5 hours. The membrane was removed using forceps to gently peel away membrane from the top agar. If significant amounts of top agar clung to the membrane, plates were cooled to 4°C to ease top agar/membrane separation. Membrane was then rinsed in sterile water.

100 μ l of Salmon sperm DNA (SS DNA) at 1 μ g/ μ l was boiled for 3 minutes, then added to TE 10:1 pH 8.0 for a final volume of 1ml. TE/SS DNA mixture was then pipetted onto the membrane (protein side up), and membrane was incubated at room temperature for 5 minutes with gentle shaking. 4 μ l of biotinylated DNA probe (prepared according to the Pierce Biotin 3' End DNA Labeling Kit user manual) was added to 100 μ g SS DNA and TE 10:1 pH 8.0 for a final volume of 1ml. The membrane was washed with sterile water, then the probe/SS DNA/TE mixture was pipetted on top of the membrane (protein side up). The membrane was incubated at room temperature with gentle shaking for 5 minutes. The membrane was washed with sterile water and placed in a UV crosslinker (protein side up). The UV crosslinker was set to 1 minute at "optimum crosslink" setting.

The Pierce Chemiluminescent Nucleic Acid Detection Module (Thermo Fisher Scientific, Rockford, IL) was used for membrane-bound probe detection, and all procedures were followed according to the corresponding user manual.

Chapter 3: Results

3.1 Promoter deletion results

Previous research conducted in the Beers lab by I. Ismail established a region of the *XCP1* promoter 237-bp in length as necessary and sufficient to drive reporter gene expression (shown as XCP1pD8 in Figure 4). To increase the resolution of this region, four additional 5' deletions of XCP1pD8 (shown in Figure 4 as XCP1pD9, XCP1pD10, XCP1pD11, and XCP1pD12)

were fused to the GUS gene within a modified pBI121 vector (Ismail I., 2004; Zhao C., 2005) and used to transform *Agrobacterium tumefaciens*. These transgenic *Agrobacterium* strains were then used to transform wild type Arabidopsis. After selection with herbicide, all T₁ survivors showed no GUS activity except XCP1pD9, which showed weak GUS activity when compared to XCP1pD8 transformants. XCP1pD9-transformant GUS activity required light microscopy for visualization, while XCP1pD8-transformant GUS was easily visualized with the naked eye (data not shown). This result suggests that the sequence ATCATTA, located between 61 and 68 bp downstream of the 5' end of the XCP1pD8, is essential for GUS expression (see Figure 7B). Considering that XCP1pD9 transformants showed weak GUS activity, other sequences immediately upstream of this element may also be necessary to induce full native expression. Figure 7 shows an expanded view of this region and predicted cis-elements contained within it (PLACE search, <http://www.dna.affrc.go.jp/PLACE/>). GUS activity results were confirmed by a second complete set of transformations and evaluations.

Resolution of the XCP1p region necessary and sufficient to control reporter gene expression was improved from 237-bp (XCP1pD8 in Figure 4) to 175-bp (XCP1pD9 in Figure 4). A higher resolution scan of -268 to -200 is needed to identify the cis-element apparently located within this region. These results suggest that either multiple cis-elements or enhancer and cis-elements exist in the XCP1p, possibly at both the 5' and 3' ends of XCP1pD9, complicating the identification of a single necessary and sufficient region using a promoter deletion analysis.

3.2 Yeast One-Hybrid Results

Previous work by I. Ismail demonstrated the importance of the -32 to -85 region, and this work combined with computer predicted cis-element analysis formed the basis for our selection of the -34 to -58 region, here termed XCP1p²⁴, for Y1H analysis (Figure 8). These bases, located near the 3' end of the *XCP1* promoter, contain the inverse repeat sequence `gagactttgtcttgtctc`. They also contain a putative sulfur deficiency response element and an ARF1 binding site motif (see Figure 8F). Since this area had been shown by promoter deletion analysis to be necessary to drive native gene expression and contain multiple elements of possible importance, XCP1p²⁴ was selected to test if it contained a cis-element capable of specifically binding a transcription factor.

A tetrameric repeat of XCP1p²⁴ was constructed in pGEM (Fig. 8 A-D), digested out of pGEM, and ligated into pHIS2. pHIS2 is a yeast expression vector containing genes for kanamycin resistance (*Kan^r*), tryptophan production (*TRP1*), and histidine production (*HIS3*). The tetrameric repeat of XCP1p²⁴ was inserted directly upstream of *HIS3* and was designed to facilitate activation of *HIS3* expression in the event of XCP1p²⁴ binding to a TF fused to the GAL4 transactivation domain. See Figure 2 for an overview of the yeast one-hybrid system.

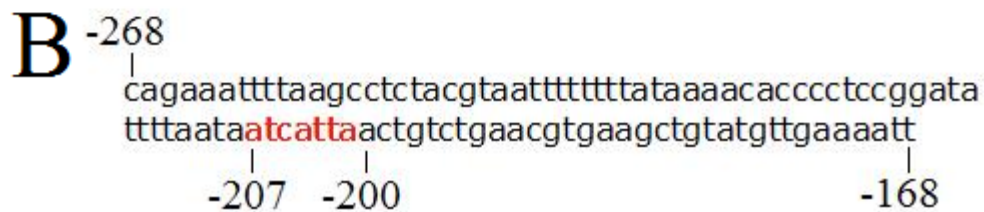
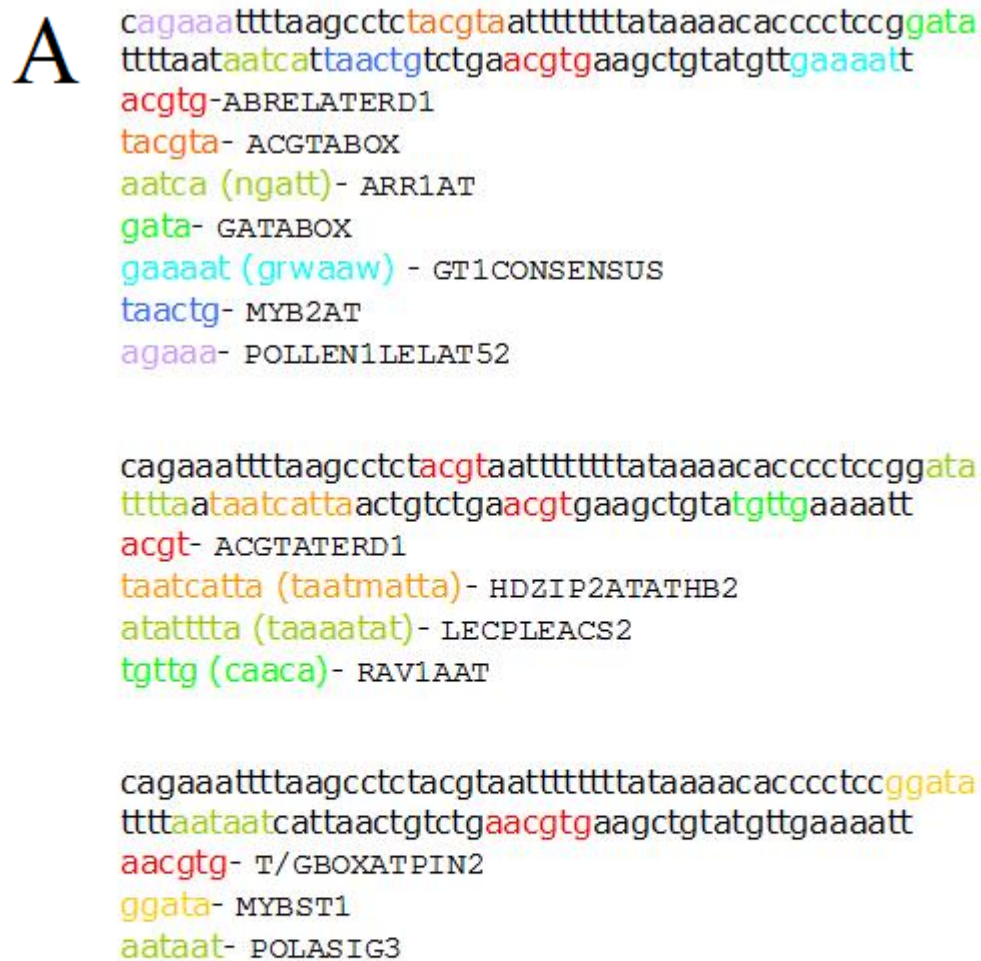


Figure 7: *XCP1* promoter from positions -268 to -168

A: cis-elements predicted by PLACE. (<http://www.dna.affrc.go.jp/PLACE/>). Multiple copies of the sequence are shown to avoid overlap between some predicted cis-elements. Symbols used: M = A or C; N = A, C, G or T; R = A or G; W = A or T.

B: Key base positions shown, with the partial ATHB2 binding site bases between -207 and -200 highlighted in red. Full ATHB2 site shown in part A, second copy, orange.

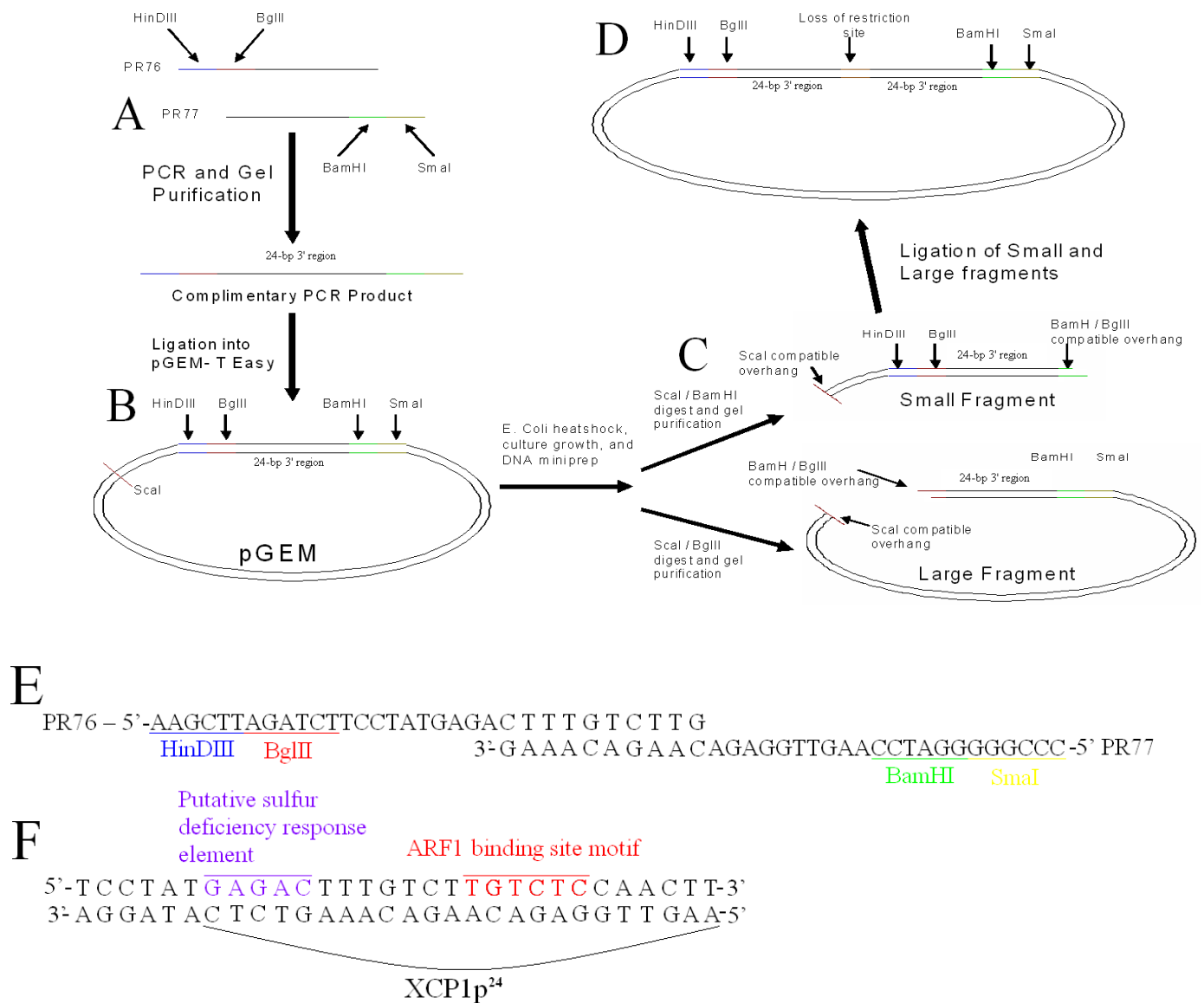


Figure 8: Construction of pXCP1p³²⁴ bait vector for use in Y1H analysis.

A-D shows the construction of a single repeat of XCP1p²⁴. The same technique was used, starting with a single XCP1p²⁴ repeat, to create an XCP1p²⁴ tetrameric repeat. The tetrameric construct was subcloned into the pHIS2 yeast bait vector creating pXCP1p³²⁴. A. PCR used to generate double stranded product containing XCP1p²⁴. B. Double stranded product generated in A is ligated into pGEM T-Easy vector. C. pGEM T-Easy containing XCP1p²⁴ is digested with either *Scal/BamHI* or *Scal/BglII*. Two fragments are generated by each digest, a small fragment of ~1kb, and a large fragment of ~2kb. D. Fragments generated in C are ligated forming a new plasmid containing two copies of XCP1p²⁴. E. Primer

complementation and restriction sites. For convenience of cloning, six bps upstream of XCP1p²⁴ (TCCTAT) were included in PR76. Note PR77 is shown in reverse complementary orientation to illustrate hybridization site for priming the polymerase reaction. F. Double stranded XCP1p²⁴ insert minus restriction sites. The putative ARF1 binding site motif is highlighted in red. Putative sulfur deficiency response element is highlighted in purple. XCP1p²⁴ is indicated by horizontal bracket. Note that the sequence shown in F is 12 bases downstream of the TERE element predicted by Pyo et al., (2007).

By co-transforming competent Y187 with pXCP1p^{3'24} (the bait vector) and prey cDNAs from an Arabidopsis cDNA library recombined with pGADT7-Rec[2], then growing the transformants on histidine dropout media (-HIS media), it was possible to screen for transformants that contained a putative trans-factor binding to the bait promoter. Since *HIS3* is known to have leaky expression, 3-amino-1,2,4-triazole (3-AT) was used as a competitive inhibitor of the *HIS3* enzyme activity to titrate leaky background *HIS3* expression. Three diagnostic experiments were conducted to minimize the effect of leaky *HIS3* expression. By transforming competent Y187 with only bait vector and plating on -HIS media containing 3-AT concentrations ranging from 10 to 100 mM (as recommended by the supplier), the minimum concentration of 3-AT needed to inhibit low levels of leaky *HIS3* expression generated by pXCP1p^{3'24} was determined to be 90 mM.

An Arabidopsis xylem cDNA library was generated using Arabidopsis root-hypocotyl tissue (Zhao et al., 2000). The double stranded cDNAs from this library were ligated into pGADT7-Rec[2] and used in conjunction with pXCP1p^{3'24} to complete the Y1H analysis as described above.

After identifying the ideal concentration of 3-AT to inhibit leaky histidine expression and still permit yeast growth, numerous putative positives were detected. Direct recovery of

plasmids from putative-positive yeast colonies was not possible due to technical difficulties. Polymerase chain reaction (PCR) amplification of yeast plasmid inserts was used to obtain copies of prey cDNA from putative-positive yeast colonies. Of forty putative-positive yeast colonies screened using PCR amplification, sequencing, and BLAST search comparison, none were identified as known transcription factors, and none were found more than once.

Gene	Identity	Description/TAIR database	Function/TAIR database
AT3G33076.1	99/123	similar to Gypsy_Ty3-element GB:AAD22153 polyprotein	transposable element gene
AT2G09790.1	238/238	similar to Gypsy_Ty3-family F6H8 reverse transcriptase	transposable element gene
AT3G40640.1	206/206	similar to putative AP endonuclease/ reverse transcriptase	transposable element gene
AT1G45221.1	67/77	ECA1 gametogenesis related protein	Bifunctional inhibitor/plant lipid transfer protein/seed storage
AT3G46613.1	690/692	member of an angiosperm specific gene family	Putative plant development gene
AT2G16990.2	271/271	similar to Major facilitator superfamily MFS 1	general substrate transporter
AT1G10950.1	255/255	endomembrane protein 70; putative	endosomal integral membrane protein
AT1G03810	105/110	similar to DNA-binding protein-related [Arabidopsis thaliana] (TAIR:AT4G28440.1)	nucleic acid-binding

Table 1: BLAST results for sequences generated by yeast one-hybrid

3.3 Southwestern blot results

After the prediction of the sequence CTTC AAGCCA as a cis-element both necessary and sufficient to control XCP1 native expression by Pyo et al. (2007), an attempt was made to identify the TFs binding to this sequence. Competent *E. coli* cells were transformed using an *Arabidopsis thaliana* xylem library created by Zhao et al., (2005) using the Strategene ZAP cDNA synthesis kit. This kit utilizes the *lacZ* promoter to induce protein expression in transformants in the presence of IPTG. These transformed cells were grown to saturation density on growth media plates, while expressing *Arabidopsis* xylem library proteins due to the presence of 1 mM IPTG in the growth media. PVDF-membranes were placed on top of the bacteria lawn where they theoretically bound proteins representing a large portion of the *Arabidopsis thaliana* xylem protein library. The PVDF-membranes were then washed with biotinylated DNA probe containing the cis-element of interest. Membranes were then washed to remove non-specific bound DNA-probe. Theoretically, the only DNA-probe now left on the membranes would remain due to specific binding interactions between the cis-element of interest and a TF expressed in the xylem library and bound to the PVDF-membrane. Next, biotinylated DNA probe was detected by washing the membrane with streptavidin-bound horseradish peroxidase and induction of chemiluminescence.

Preliminary results showed numerous weak positive reactions (See Figure 9). Rescreening of these weak positives involved gathering phage from positive locations and using this phage to infect new *E. coli* cultures. These cultures were then grown to saturation density on growth media plates and screened using a second round of southwestern blot analysis. In

theory, multiple rounds of selection, growth, and screening would result in a population enhanced for production of proteins capable of binding DNA probe, making it less challenging to identify true positive interactions.

This strategy was not successful in identifying TFs. While numerous positive interactions were detected, these results proved unreproducible after three rounds of rescreening. Because all putative positive interactions proved unreproducible, it is very likely they were the result of non-specific binding of DNA probe to membrane-bound protein.

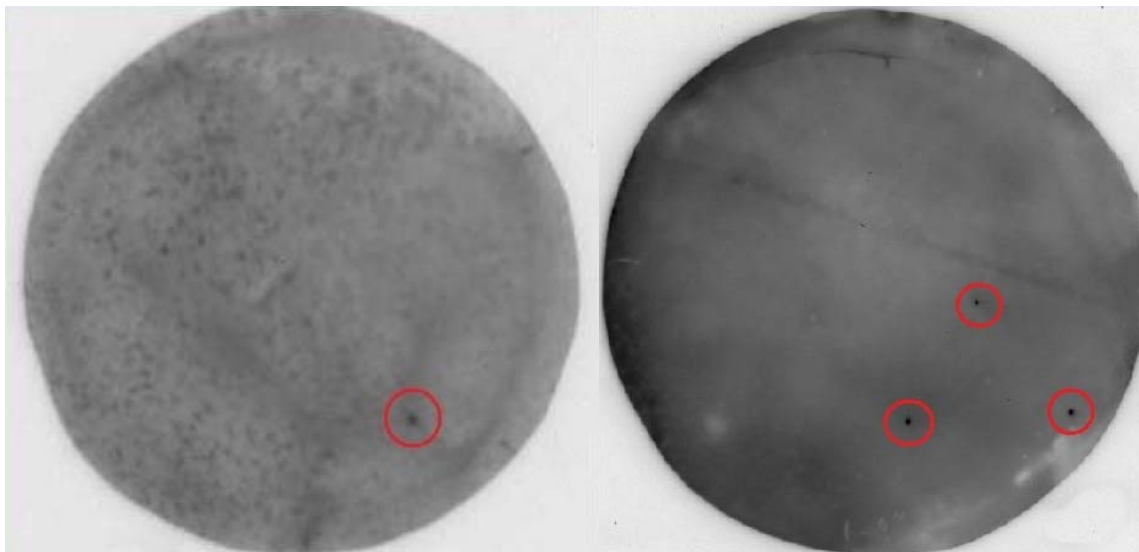


Figure 9: Southwestern blot; chemiluminescent visualization. Putative positives are circled in red. Photo by author, 2008.

Chapter 4: Discussion

Promoter deletion analysis resulted in increased resolution of the XCP1 promoter region both necessary and sufficient to drive native gene expression. However, yeast one-hybrid analysis and southwestern blot analysis were not successful at identifying an XCP1p cis-element necessary and sufficient to promote native gene expression or a TF capable of specific binding to such a cis-element. Promoter deletion analysis results, while partially consistent with results generated by Pyo et al. (2007), also revealed the existence of an enhancer or second cis-element located at the 5' end of XCP1pD9. XCP1pD9 successfully drove weak TE-specific GUS expression, while XCP1pD10 – XCP1pD9 with 7 bases deleted from its 5' end – was completely unable to drive GUS expression. This suggests that the sequence ATCATTA, the 7 deleted bases, is essential for GUS expression. ATCATTA is part of the cis-element TAATCATTA, the ATHB2 binding site motif (Sessa et al., 1993). *ATHB2* appears to be expressed independently of XCP1 based on data from over 2100 microarray experiments, which are available at the Genevestigator website (<https://www.genevestigator.com/gv/index.jsp>). However this does not rule out the possibility of XCP1 regulation by *ATHB2*, and investigation of this possibility may prove interesting.

During construction of XCP1pD9, PCR introduced a random point mutation a base -65, the 3' end of the TERE element. The A normally at this position was replaced with G. This mutant version of XCP1pD9 was used to transform *Arabidopsis* plants, and failed to drive GUS expression in all transformants. These results are consistent with data generated by Pyo et al.

(2007). This result independently supports the importance of the TERE cis-element as necessary to drive TE-specific expression of XCP1.

Yeast one-hybrid analysis resulted in so many false positives that determining which (if any) putative-positive yeast colonies were true positives became challenging. Since screening of all putative positive colonies was not practical, the 20 largest yeast colonies were selected for screening using PCR amplification, sequencing, and BLAST search comparison. Colony PCR may have been complicated by the presence of multiple prey plasmids in transformed yeast cells, resulting in poor quality or non-existent PCR products. Restreaking colonies aided PCR resolution in most cases. Restreaking colonies multiple times over several weeks may have further aided resolution.

None of the screens resulted in known transcription factors, though some putative DNA-binding proteins of unknown function were identified (see Table 1). Because of this putative positive result, 20 additional colonies were screened. No known transcription factors were identified, and no putative DNA-binding proteins were identified a second time. In general, true positive interactions generated by a yeast one-hybrid screen should be detected in multiple putative positive colonies. Since none of the putative DNA-binding proteins were detected a second time, they were most likely false positives.

XCP1p²⁴, containing bases -58 to -34, was selected as bait for yeast one-hybrid analysis based on previous data collected by promoter deletion analysis and computer predictive biology (see Figure 4). Unfortunately, this region did not yield informative results, and before

we could test additional elements in the -32 to -85 region, Pyo et al. (2007) demonstrated that the relevant cis-element, named TERE, was located at -65 to -75.

Southwestern blot analysis may have been undermined by the fact that the strength of binding specificity between the XCP1p cis-element and its potential TFs is unknown. The southwestern blot analysis was conducted using the assumption that binding between the XCP1p cis-element and its TFs would produce a positive result that stood out well above background levels. As background was significant in all trials, this assumption may have been faulty. Also, since *Arabidopsis thaliana* xylem proteins were expressed by the prokaryote *E. coli*, there is no guarantee these proteins were expressed at correct length and with proper folding. As a result, binding sites of potential binding TFs may have been masked or nonexistent. This experiment would also have been strengthened by the use of a true positive control.

Promoter deletion data suggests the presence of two elements necessary for native gene expression in the *XCP1* promoter. One is located near the 3' end of XCP1pD9 and is likely the TERE cis-element identified by Pyo et al., (2007). The other is located at or slightly upstream of the 5' end of XCP1pD9 and certainly contains some, if not all, of the bases between -200 and -207. This may be another cis-element, or possibly an enhancer or other element. A more detailed promoter deletion analysis is being conducted in the Beers lab to identify this 5' element. The *XCP1* promoter is being analyzed in 10 bp segments from position -253 to -213 (see Figure 10). Using results from this analysis, the resulting putative 5' element could subsequently be more finely mapped using two base substitutions within the identified relevant sequence as done by Pyo et al., (2007). Sequences containing the suspected 5' element could

then be tested for their ability to drive TE-specific GUS expression in combination with and separate from TERE. The 5' element could also be tested in single copy and as a multimeric repeat using yeast one-hybrid analysis to reveal potential binding transcription factors.

While research in *Arabidopsis thaliana* is interesting and necessary to promote general understanding of the fundamental biology of vascular tissue differentiation, studies conducted on the economically important genus *Populus* may prove more relevant to the forest products industry. Ongoing research in the Beers lab has identified a poplar *XCP1* ortholog, which may be important in controlling xylem development. Improvements in the methods of yeast one-hybrid analysis have yielded the Gateway-compatible yeast one-hybrid system which incorporates bait DNA into the yeast chromosome, creating a more gene-centered system compared to traditional Y1H screens (Deplancke et al., 2004). Conducting a Gateway-compatible yeast one-hybrid analysis using the promoter of the poplar *XCP1* ortholog may identify binding transcription factor(s) and help increase understanding of the transcription regulatory network associated with this gene and with wood formation in general. This information could prove to be an economically important result relevant to the regulation of wood development in commercially important tree species.

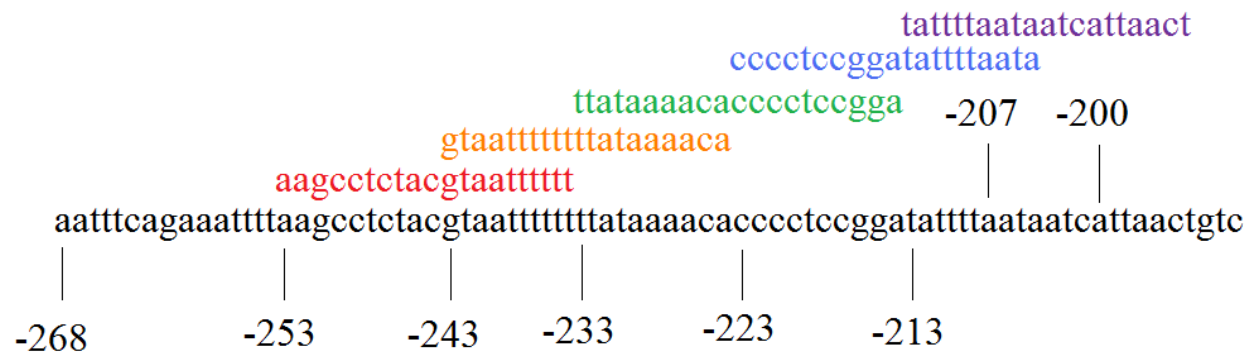


Figure 10: Primer sequences for *XCP1* promoter deletion analysis between bases -253 and -213. *XCP1* promoter sequence shown in black, primer sequences shown in color. Bases -268, -207 and -200 labeled for reference.

References:

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ** (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**: 403-410
- Avci U, Petzold HE, Ismail IO, Beers EP, Haigler CH** (2008) Cysteine proteases XCP1 and XCP2 aid micro-autolysis within the intact central vacuole during xylogenesis in *Arabidopsis* roots. *Plant J* **56**: 303-315
- Beers EP, Freeman TB** (1997) Proteinase activity during tracheary element differentiation in *Zinnia* mesophyll cultures. *Plant Physiol* **113**: 873-880
- Beers EP, Jones AM, Dickerman AW** (2004) The S8 serine, C1A cysteine and A1 aspartic protease families in *Arabidopsis*. *Phytochemistry* **65**: 43-58
- Clough SJ, Bent AF** (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**: 735-743
- de Bruijn FJ, Felix G, Grunenberg B, Hoffmann HJ, Metz B, Ratet P, Simons-Schreier A, Szabados L, Welters P, Schell J** (1989) Regulation of plant genes specifically induced in nitrogen-fixing nodules: role of cis-acting elements and trans-acting factors in leghemoglobin gene expression. *Plant Mol Biol.* **13**: 319-325.
- Deplancke B, Dupuy D, Vidal M, Walhout AJ** (2004) A gateway-compatible yeast one-hybrid system. *Genome Res.* **14**: 2093-2101
- Funk V, Kositsup B, Zhao C, Beers EP** (2002) The *Arabidopsis* xylem peptidase XCP1 is a tracheary element vacuolar protein that may be a papain ortholog. *Plant Physiol* **128**: 84-94
- Hatton D, Sablowski R, Yung MH, Smith C, Schuch W, Bevan M** (1995) Two classes of cis sequences contribute to tissue-specific expression of a PAL2 promoter in transgenic tobacco. *Plant J* **7**: 859-876
- Hauffe KD, Lee SP, Subramaniam R, Douglas CJ** (1993) Combinatorial interactions between positive and negative cis-acting elements control spatial patterns of 4CL-1 expression in transgenic tobacco. *Plant J* **4**: 235-253

- Ismail I** (2004) Function and regulation of xylem cysteine protease 1 and xylem cysteine protease 2 in Arabidopsis. PhD Dissertation, Virginia Polytechnic Institute and State University.
- Latchman DS** (1997) Transcription Factors: An Overview. *Int. J. Biochem. Cell Biol.* **29**: 1305-1312
- Leyva A, Liang X, Pintor-Toro JA, Dixon RA, Lamb CJ** (1992) cis-element combinations determine phenylalanine ammonia-lyase gene tissue-specific expression patterns. *Plant Cell* **4**: 263-271
- Minami A, Fukuda H** (1995) Transient and specific expression of a cysteine endopeptidase associated with autolysis during differentiation of Zinnia mesophyll cells into tracheary elements. *Plant Cell Physiol* **36**: 1599-1606
- Moreau C, Aksenov N, Lorenzo MG, Segerman B, Funk C, Nilsson P, Jansson S, Tuominen H** (2005) A genomic approach to investigate developmental cell death in woody tissues of Populus trees. *Genome Biol* **6**: R34
- Oh S, Park S, Han KH** (2003) Transcriptional regulation of secondary growth in Arabidopsis thaliana. *J Exp Bot* **54**: 2709-2722
- Petzold HE** (2007) Promoter Deletion Analysis of Xylem Cysteine Protease 2 (XCP2) in Arabidopsis thaliana. Masters Thesis, Virginia Polytechnic Institute and State University.
- Pyo H, Demura T, Fukuda H** (2007) TERE; a novel cis-element responsible for a coordinated expression of genes related to programmed cell death and secondary wall formation during differentiation of tracheary elements. *Plant J* **51**: 955–965
- Sessa G, Morelli G, Ruberti I** (1993) The Athb-1 and -2 HD-Zip domains homodimerize forming complexes of different DNA binding specificities. *Embo J* **12**: 3507-3517
- Turner S, Gallois P, Brown D.** (2007) Tracheary element differentiation. *Annu Rev Plant Biol* **58**: 407-433
- van Esse HP, Van't Klooster JW, Bolton MD, Yadeta KA, van Baarlen P, Boeren S, Vervoort J, de Wit PJ, Thomma BP** (2008) The Cladosporium fulvum virulence protein Avr2 inhibits host proteases required for basal defense. *Plant Cell* **20**: 1948-1963
- Ye ZH, Freshour G, Hahn MG, Burk DH, Zhong R** (2002) Vascular development in Arabidopsis. *Int Rev Cytol* **220**: 225-256

- Zhao C, Johnson BJ, Kositsup B, Beers EP** (2000) Exploiting secondary growth in Arabidopsis. Construction of xylem and bark cDNA libraries and cloning of three xylem endopeptidases. *Plant Physiol.* **123**: 1185-1196
- Zhao C** (2005) Identifying and characterizing genes that regulate vascular tissue-specific functions. PhD Dissertation, Virginia Polytechnic Institute and State University.
- Zhao C, Craig JC, Petzold HE, Dickerman AW, Beers EP** (2005) The xylem and phloem transcriptomes from secondary tissues of the Arabidopsis root-hypocotyl. *Plant Physiol* **138**: 803-818
- Zhong R, Richardson EA, Ye ZH** (2007) The MYB46 Transcription Factor Is a Direct Target of SND1 and Regulates Secondary Wall Biosynthesis in Arabidopsis. *Plant Cell* **19**: 2776–2792
- Zhong R, Lee C, Zhou J, McCarthy RL, Ye ZH** (2008) A Battery of Transcription Factors Involved in the Regulation of Secondary Cell Wall Biosynthesis in Arabidopsis. *Plant Cell* **20**: 2763–2782
- Zhou J, Lee C, Zhong R, Ye ZH** (2009) MYB58 and MYB63 Are Transcriptional Activators of the Lignin Biosynthetic Pathway during Secondary Cell Wall Formation in Arabidopsis. *Plant Cell* **21**: 248–266