5-2001

The Characterization of a Novel Petite Phenotype in Transgenic Arabidopsis Thaliana and the Examination of the Expression Patterns of a Gamma Carbonic Anhydrase Homologue in Arabidopsis Thaliana

Sumeet K. Asrani

THE CHARACTERIZATION OF A NOVEL *PETITE* PHENOTYPE IN TRANSGENIC *ARABIDOPSIS THALIANA* AND THE EXAMINATION OF THE EXPRESSION PATTERNS OF A GAMMA CARBONIC ANHYDRASE HOMOLOGUE IN *ARABIDOPSIS THALIANA*

A Thesis

Submitted to the Honors College of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for
Upper Level Honors

in

Biochemistry

by
Sumeet Asrani
May 2001

THE CHARACTERIZATION OF A NOVEL PETITE PHENOTYPE IN TRANSGENIC ARABIDOPSIS THALIANA

Table of Figures

Abstract

A novel petite phenotype was observed using *Arabidopsis thaliana* transformed with T-DNA carrying a γCA-3 promoter-GUS fusion. Since this transgenic plant was the only one of 12 transgenic plants examined that had a phenotype different from wild type, we concluded that the phenotype resulted from random insertional mutagenesis using T-DNA. Using thermal asymmetric interlaced PCR (TAIL-PCR), the site of insertion of the right border of the T-DNA was amplified, cloned and sequenced. Sequence analysis revealed that the T-DNA had inserted between two genes (XMS 254 and XMS 255) expressing uncharacterized proteins on chromosome II. It was hypothesized that the expression of either of the two genes could have been disrupted due to the presence of the T-DNA insertion either in the promoter region or downstream activating sequence region of the two genes. An RT-PCR assay showed that there is no obvious difference in the expression of XMS 255 in wildtype and mutant plants. Phenotypic observations showed that the petite plant is considerably shorter, has a reduced hypocotyl, and tightly packed leaves. It also produces fewer siliques. Leaf cross sections revealed that the petite cells were smaller indicating a problem with cell expansion. Scanning electron micrographs revealed that the petite plants had fewer stomates and trichomes than wild type plants. In addition, many of the trichomes on the petite plants were bipartite rather than tripartite. Further localization studies using promoter/reporter fusions can be conducted to further understanding of interrupted gene function.

Background

Random Insertional Mutagenesis

In *Arabidopsis thaliana*, random insertional mutagenesis (RIM) serves as a molecular technique to discover functions of various genes. The disturbance in genotype, if not lethal, can be observed through phenotypic alterations. Dramatic alterations, such as the size of the plant, are often valuable indicators of gene knockouts or disruption of function. Insertional mutagenesis relies on the integration of foreign DNA into the plant genome. Transposable elements such as transferred DNA (T-DNA) can be used for RIM. The latter is preferred because such insertions will not transpose subsequent to integration and are therefore chemically and physically stable through multiple generations (Krysan et al. 1999). RIM is accomplished by infecting the plant with Agrobacterium-mediated transformation techniques. Common techniques include the floral dip method, where the flowering parts are immersed in an Agrobacterium suspension, and the vacuum infiltration method, where the whole plant is saturated with the suspension by use of a vacuum chamber. The seeds are then harvested and sown on selective media and observed for altered phenotype.

Characteristics of T-DNA

T-DNA is an important tool in insertional mutations. Studies have shown that insertions preferentially occur in potentially transcribed regions of plant genome (Koncz et al 1992). Using tissue culture methods, the insertion is in one locus only one third of the time, so multiple insertions are likely. However, with floral transformation, insertion is usually in one locus. There is also a chance of inversion and duplication of the T-DNA giving repeat sequences. However, the stability of the insertion provides a large window

of time for studies of disrupted genes. While T-DNA RIM is useful for studying loss of function phenotypes, lethal mutations cannot be studied.

Dwarf Mutants

In this project, RIM by a former graduate student resulted in a dwarf mutant. Dwarf mutants in *Arabidopsis* result from various alterations in cell pathway and function. They include disruptions in the hormones responsible for plant growth through cell division and cell expansion. For example, gibberellin is required for stem growth and leaf expansion and an absence of gibberellin leads to short stemmed plants with altered leaf growth patterns (Bowman, 1993).

Plant growth is accomplished by a combination of cell division and cell expansion. When new cells grow they mainly elongate. In a region of active growth, cells can expand to a volume as much as 50 times their original size. Expansion occurs when the cell wall yields to the turgor pressure of the cell. Acids secreted by the cell causes chemical changes that weaken the cross-links within the cell wall. The cell is hyperosmotic to the surrounding solution and it takes up water by osmosis and expands. Growth continues until the wall again becomes knit tightly enough to offset the cell's turgor pressure. Most of the water is packaged into the large vacuole (Campbell 1993).

TAIL PCR

Several methods are available to characterize the left and right border junction of the T-DNA insertion into the plant genome. They include plasmid rescue, inverse PCR, and TAIL-PCR. Thermal asymmetric interlaced (TAIL)-PCR, utilizes nested sequence-specific primers together with a shorter arbitrary degenerate primer so that the relative amplification efficiencies of specific and nonspecific products can be thermally

controlled. One low-stringency PCR cycle is carried out to create annealing site(s) adapted for the arbitrary primer within the unknown target sequence bordering the known segment. This sequence is then preferentially and geometrically amplified over non target ones by interspersion of high-stringency PCR cycles with reduced-stringency PCR cycles (Liu et al., 1995).

Linking phenotype and genotype

Once the junctions are identified, several cloning techniques are used to identify the knock-out gene. Once the gene is identified, there are several questions that need to be addressed. For example, is there a concrete connection between the phenotype and the deduced genotype? There is a possibility that there might have been other disruptions by T-DNA elsewhere that might have caused the expression or complemented the expression by the deduced gene. One should try to reconstitute the knocked out gene using transgenic techniques. The patterns of segregation should be followed over many generations. The T-DNA insertion could directly knock out the gene or influence the expression of the gene. T-DNA could insert within the promoter region or even disrupt upstream or downstream activating sequences.

PCR can serve as a powerful analytical tool in more ways than one. Using genomic DNA as template, the presence or absence of T-DNA insertions within certain genes can be established and reconfirmed. Reverse Transcriptase PCR (RT-PCR) can be used to track patterns of expression and production of gene product. A variability in the amount of gene product obtained from wild type and control plants can provide clues about the type of mutation. Finally, localization studies can be performed using a variety of reporter assays.

Materials and Methods

DNA extraction

Total DNA was extracted from 100 mg of plant leaves from the laboratory stock of G8-3 plants exhibiting the petite phenotype. The manufacturer's protocol (Qiagen) was followed for the extraction. The resulting pellet was dissolved in TE and a working dilution of 20 ng/ul was made.

Thermal Asymmetric InterLaced PCR (TAIL PCR)

Primers were designed to isolate the right border of the T-DNA insertion into the Arabidopsis genome (NCBI accession number- U12639). Three primers, NPT1, 2,and 3, were spaced so that NPT1 and NPT2 were separated by 300 base pairs; NPT-2 and NPT3 were separated by 75 base pairs. NPT3 was about 350 base pairs inside the T-DNA right border and was the closest primer to the border. Three degenerate primers (AD1, AD2, AD3) were designed according to Liu et al. (1995) to hybridize with the plant genome near the T-DNA right border. The steps of the tri-PCR reaction are given below. For all three reactions (primary, secondary and tertiary), 2 units of Taq polymerase (Promega), 200 uM of dNTPs, and 2 mM $MgCl_2$ were used. In the primary reaction, a combination of NPT1 (.2 uM) and AD1 (2 uM), AD2 (3 uM), and AD3 (4 uM) primer sets were used to amplify 20 ng of template DNA.

Primary Reaction	
TEMP	Time (min)
93	1
95	2
5 cycles ()	
(94	1
62	1
72	2.5
94	1)
25	3

ramp to 72 over 3 min	
72	2.5
15 cycles ()	
(94	0.5
68	1
72	2.5
94	0.5
68	1
72	2.5
94	0.5
44	1
72	2.5)
72	5

A 1:50 dilution of the PCR product was used as the template for the secondary reaction. Here NPT2 (.2 uM) was used in combination with AD1 (1.5 uM), AD2 (2 uM) and AD3 (2 uM).

Secondary Reaction	
TEMP	Time (min)
12 cycles ()	
(94	0.5
64	1
72	2.5
94	0.5
64	1
72	2.5
94	0.5
44	1
72	2.5)
72	5

A 1:10 dilution of the PCR product was used for the tertiary reaction. The total volume was 100 ul. Here NPT3 (.2 uM) was used in combination with AD1 (1.5 uM), AD2 (2 uM) and AD3 (2 uM).

Tertiary Reaction

TEMP	Time (min)
20 cycles ()	
(94	1
44	1
72	2.5)
72	5

Sequencing

The products from all three reactions were analyzed by electrophoresis on a 1.2% agarose gel in TBE (Figure 2).

All DNA manipulations including precipitation, ligation, transformation, and restriction digestions were performed using standard methods (Sambrook et al., 1989). The PCR product was precipitated with sodium acetate (10% of volume) and two volumes of 100% percent ethanol. The pellet was then washed in 70% ethanol and the final pellet was raised in TE. The PCR product was run on a low melt agarose gel and was ligated with pGEM-T Easy (Promega) according to the manufacturer's protocol. The ligated product was transformed into DH5α cells by electroporation and plated on media containing ampicillin and x-gal. Colonies were picked and alkaline lysis of overnight cultures was carried out. Restriction analysis was conducted using *EcoRI.* Secondary confirmation and the orientation of the product were revealed by digestion with *Pst I* and *NheI/Nsi I.* Cultures containing plasmids with the correct insert were grown overnight in liquid culture. The plasmid was extracted and purified using the Qiagen Plasmid Mini Kit according to the manufacturer's protocol. The DNA was quantitated by electrophoretic gel analysis and the sample was sent to Gene Lab, LSU, for sequencing.

Plant growth

The transgenic plants (generation T4) were grown for observation and RNA and DNA extraction. Seeds were suspended in 0.15% agar and were spread over soil. After staying at 16 C for two days, the pots were transferred to a growth chamber and were watered on alternate days. Fertilizer was added once every week. The plants were monitored and photographed when they flowered. Leaf samples were obtained to visualize a cross section. The cell surface was analyzed using a scanning electron microscope.

RNA extraction

RNA was extracted from plant samples using the Qiagen RNeasy Plant Mini Kit according to the manufacturer's protocol. The RNA was quantitated using a Spectrophotometer.

RT PCR assay

Primers were designed to span the open reading frames encoded by *XMS 255* and *XMS 254* that could possible be affected by the T-DNA insertion. An RT- PCR assay was used to determine expression levels of the two gene products. The reaction program was set up as per the manufacturer's protocol (Stratagene).

DNA analysis

Primers (2 uM) for *XMS 255* and *XMS 254* open reading frames of the two gene products were utilized to determine the presence of the two genes in transgenic plants (100 ng of template). 0.2 mM of dNTPs and Vent polymerase (New England BioLabs) were used. The parameters of the reaction were 4 minutes at 94 C, 35 cycles of 30 seconds at 94 C, 66 C for 30 seconds,72 C for 1 minute, and finally 3 minutes at 72 C.

Results and Discussion

TAIL PCR

In TAIL PCR three degenerate primers along with 3 primers within the right border of the T-DNA insert were used (Figure 1). As seen in Figure 2, a combination of primer NPT-1, NPT-2, NPT-3 with AD-3 showed a positive result. Products differing in length of about 70 bases and 300 bases were expected and observed. The smallest fragment that covered the border was recovered. When resolved, more bands were seen and the largest band was used to insert into pGEM-T Easy for sequencing.

Cloning into P-GEM

The presence of the insert in the sequencing vector was confirmed by digesting with *Eco RI* that flanks the insert. Further confirmation and orientation information was obtained by digesting separately with *PstI* and then with *NheI/NsiI* in the appropriate buffers (Figure 3). The DNA was sequenced by GeneLab, LSU. The sequence displayed the right border of the T-DNA and the junction where it inserted into the plant genome.

Sequence Analysis

The results of the sequence revealed the following: Homology was observed with a fragment of DNA between two genes, designated *254 XMS* and *255 XMS,* that express uncharacterized proteins on chromosome II of *Arabidopsis* (Figure 4 and 5). Both genes are expressed on the complement strand. *254 XMS* has a WD-40 repeat that is a Trp-Asp 40-amino acid long motif and found in the beta subunit of G-proteins. *255-XMS* has a PPR repeat motif that is characteristic of plant organellar proteins (BLAST conserved domain locator). Homozygous sequences are expressed in many higher plants and multiple homologues appear in the *Arabidopsis* genome. Due to the nature of T-DNA

insertions, it was hypothesized that the T-DNA insertion could have affected the promoter of 244 XMS or downstream activating sequences of 255 XMS.

RT-PCR assay and DNA analysis

Primers were designed to flank the two genes in order to determine if they were intact in the mutant plant (Figure 4b and 5b). Using genomic DNA from mutant and wild type plants as template, only the open reading frame of *254 XMS* was confirmed in both varieties. *255 XMS* was not amplified (Figure 6). An RT-PCR assay was also conducted to compare the levels of *254 XMS* and *255 XMS* mRNA in *Columbia* and *petite* leaves. Despite numerous tries, the expression of only *255 XMS* was confirmed. As seen in Figure 7, there is a small difference between expression in the mutant and in the wild-type plant. However, since no control reaction for the amount of input RNA from wildtype and mutant plants was carried out, it is not clear whether these differences are real.

Phenotype observations

The mutant plant was named *petite (XMS)* due to its short and stubby Christmas-tree like appearance. *Petite XMS* has a significantly shortened flower stalk and fewer siliques than wild type plants. The siliques are bunched up near the base of the plant In addition, *Petite XMS* has shorter internodes with tightly packed leaves (Figure 8). These phenotypic differences were not observed during early stages of vegetative growth, suggesting that the mutation is more pronounced when there is a change from circular to more elongated leaves.

Examination of Microscopy

Leaf cross sections revealed that the cells in the petite leaf were, on average, smaller than those in wild type plants (Figure 9). Hence, rather than fewer cells (indicating a problem early on in cell division), there might be a disruption in cell expansion after a certain stage in plant development.

Fewer stomates were noticed in cross sections of *XMS* leaves and this was confirmed by SEM microscopy (Figure 10). SEM also revealed certain *XMS* leaf samples with bipartite trichomes rather than the usual tripartite trichome (Figure 11).

Growth Experiments

Experiments were conducted to see whether the petite was due to a deficiency in gibberellin (a growth hormone). In the dark, Arabidopsis *gib* mutants would be short, due to a failure of the hypocotyl to elongate. However, dark grown *petite XMS* plants still produced a long hypocotyl. De-etiolated (*det*) mutants are also dwarf. However, dark grown *petite XMS* seedlings were etiolated. These experiments ruled out the possibility that the *petite XMS* is a *gib* or *det* mutant.

Future directions

Further experiments need to be done to confirm the expression of 254 XMS and 255 XMS. New primer pairs should be designed for optimal RT-PCR conditions. Since an internal control of ribosomal RNA was not conducted, there might be discrepancy in the expression levels of 255 XMS. Possibly, fewer rounds of RT-PCR would have shown a difference in expression levels in wild type and mutant plants. There is a possibility that one or both genes are temporally expressed. Hence DNA and RNA need to be extracted at various stages of growth. Localization studies can easily be carried out

by transforming *Arabidopsis* with promoter/reporter fusions to discern tissue and development expression patterns.

It has been suggested that the PPR motif has protein binding properties. Small and Peters (2000) also hint at interaction of the motif with RNA and a possible role in RNA editing. Since the PPR domain found in *255 XMS* can serve as a signaling motif for mitochondria and chloroplasts, a loss of function might hamper growth pattern and lead to a petite phenotype. A reduced gene expression could lead to defective plant organelles. Photosynthesis could be affected which would lead to less growth. Fewer stomates might develop if there is less demand for $C0_2$. If RNA editing is affected at any level, there is a chance of altered growth patterns. *254 XMS* has a WD-40 repeat motif. These repeats are characteristic of the β subunit of G-protein kinase complex, among other candidates. Any alteration in the signaling cascade could lead to phenotypic differences in growth patterns.

gtttacccgc caatatatcc tgtcaaacac tgatagttta aactgaaggc gggaaacgac
aatctgatca tgagcggaga attaagggag tcacgttatg acccccgccg atgacgcggg
acaagccgtt ttacgtttgg aactgacaga accgcaacgt tgaaggagcc actcagccgc
gggtttctgg agtttaatga gctaagcaca tacgtcagaa accattattg cgcgttcaaa
agtcgcctaa ggtcactatc agctagcaaa tatttcttgt caaaaatgct ccactgacgt
tccataaatt cccctcggta tccaattaga gtctcatatt cactctcaat ccaaataatc
tgcaccggat ctggatcgtt tcgc**atg**att gaacaagatg gattgcacgc aggttctccg
gccgcttggg tggagaggct attcggctat gactgggcac aacagacaat cggctgctct
gatgccgccg tgttccggct gtcagcgcag gggcgcccgg ttctttttgt caagaccgac
ctgtccggtg ccctgaatga actgcaggac gaggcagcgc ggctatcgtg gctggccacg
acgggcgttc cttgcgcagc tgtgctcgac gttgtcactg aagcgggaag ggactggctg
ctattgggcg aagtgccggg gcaggatc**tc ctgtcatctc accttgct**cc tgccgagaaa
gtatccatca tggctgatgc aatgcggcgg ctgcatacgc ttgatccggc tacctgccc**a**
ttcgaccacc aagcgaaaca tcgcatcgag cgagcacgta ctcggatgga agccggtctt
gtcgatcagg atgatctgga cgaagagcat caggggctcg cgccagccga actgttcgcc
aggctcaagg cgcgcatgcc cgacggcgat gatctcgtcg tgacccatgg cgatgcctgc
ttgccgaata tcatggtgga aaatggccgc ttttctggat tcatcgactg tggccggctg
ggtgtggcgg accgctatca ggacatagcg ttggctaccc gtgatattgc tgaagagctt
ggcggcgaat gggctgaccg **cttcctcgtg ctttacggta** tcgccgctcc cgattcgcag

Figure 1: A part of the NPT region of the T-DNA insert is shown. The translational start site of the NPT region is marked. The primers were made complementary to the sequences indicated above (in the order NPT-3, NPT-2, NPT-1) to serve as the 3' end primers. Three degenerate primers (AD1-NTCGASTWTSGWGTT; AD2-NGTCGA SWGANAWGAA ; AD3-WGTGNAGWANCANAGA) were also made as possible 5' end primers for the TAIL-PCR reaction.

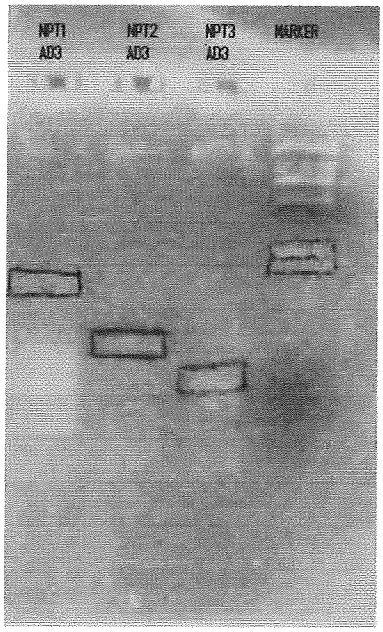

Figure 2: The results of the TAIL-PCR are shown above. Primer pairs with degenerate primer AD3 worked to produce products separated by 70 and 300 base pairs respectively. The PCR product from primer pair NPT3/AD3 was further resolved.

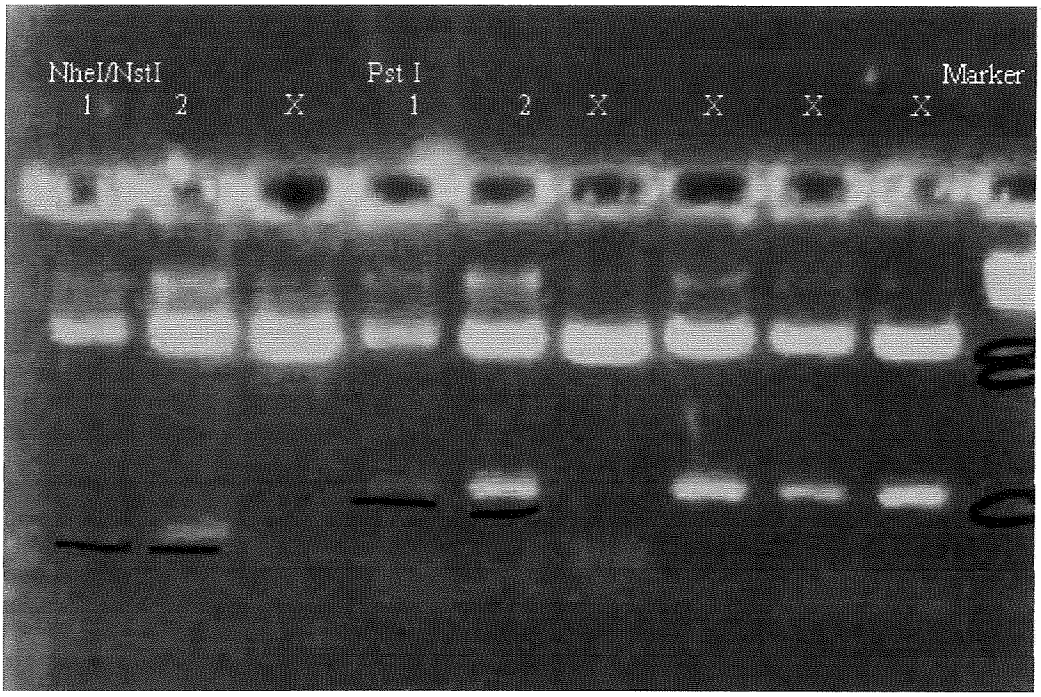

Figure 3: The presence of the TAIL PCR product within pGEM-T Easy was confirmed by digesting with *Nhe I/ Nsi I*. An expected 450 base pair fragment was observed. After digesting with *PstI*, an expected 600 base pair fragment was seen.

(A)

MEEELRVRLNDHQVSKVFPVKPKSTAKPVSESETPESRYWSSFKNHSTPNLVSSVAALAFSPVHPHSLAVA
HSATVSLFSSQSLSSSRRFSFRDVVSSVCFRSDGALFAACDLSGVVQVFDIKERMALRTLRSHSAPARFVK
YPVQDKLHLVSGGDDGVVKYWDVAGATVISDLLGHKDYVRCGDCSPVNDSMLVTGSYDHTVKVWDARVHTS
NWIAEINHGLPVEDVVYLPSGGLIATAGGNSVKVWDLIGGGKMVCSMESHNKTVTSLRVARMESAESRLVS
VALDGYMKVFDYGRAKVTYSMRFPAPLMSLGLSPDGSTRVIGGSNGMVFAGKKKVRDVVGGQKKSLNLWSL
ISDVDESRRRALRPTYFRYFQRGQSEKPSKDDYLVKEKKGLKLTRHDKLLKKFRHKEALVSVLEEKKPANV
VAVMEELVARRKLMKCVSNMEEGELGMLLGFLQRYCTVQRYSGLLMGLTKKVLETRAEDIKGKNEFKGLLR
NLKREVNQEIRIQQSLLEIQGVIAPLMRIAGRS

(B)

ATGGAGGAAGAGCTTCGTGTTCGGCTGAATGACCATCAAGTGTCA
AAGGTTTTCCCTGTGAAGCCCAAATCAACGGCGAAACCTGTTTCCGAATC
GGAGACTCCGGAATCGAGGTACTGGTCCTCCTTCAAAAATCATTCCACAC
CTAATCTCGTCTCCTCCGTAGCCGCTTTGGCTTTCTCTCCGGTCCACCCT
CACTCCTTGGCCGTCGCTCACTCCGCCACCGTCTCCCTCTTTTCCTCGCA
GTCGCTTTCCTCCTCCCGCCGTTTCTCCTTCCGCGATGTTGTCTCCTCCG
TCTGTTTCCGTTCCGACGGAGCTCTTTTCGCGGCATGCGATCTCTCCGGT
GTGGTCCAGGTCTTCGACATCAAGGAACGAATGGCTCTCCGTACTCTCCG
CTCCCACAGCGCCCCCGCTCGATTCGTCAAGTATCCAGTTCAGGACAAGC
TCCACTTGGTCTCCGGAGGCGACGATGGAGTCGTCAAGTACTGGGATGTC
GCCGGAGCAACAGTCATCTCAGATCTGCTAGGCCACAAGGATTATGTCCG
CTGCGGTGATTGCTCGCCCGTCAACGATTCAATGTTGGTAACAGGCTCTT
ACGATCACACGGTGAAGGTGTGGGACGCCAGAGTCCACACGTCCAATTGG
ATTGCCGAGATCAACCACGGACTCCCAGTGGAGGACGTCGTATACCTGCC
TTCCGGTGGATTGATCGCAACCGCTGGTGGCAACAGCGTCAAGGTATGGG
ATTTGATCGGAGGCGGGAAGATGGTTTGCTCCATGGAGAGTCACAACAAG
ACCGTCACGTCTCTCCGCGTTGCGAGAATGGAATCGGCTGAGAGTCGACT
TGTGAGCGTTGCTTTGGACGGATACATGAAAGTCTTCGATTACGGGAGAG
CAAAGGTGACTTACTCTATGAGGTTTCCTGCTCCTCTCATGTCTCTTGGC
CTTTCCCCCGATGGTTCCACTCGGGTAATCGGAGGGTCTAATGGCATGGT
GTTTGCCGGTAAGAAGAAGGTGAGAGATGTTGTTGGAGGACAGAAGAAAA
GTTTGAATCTCTGGAGTCTGATAAGCGATGTGGATGAGAGTAGGAGAAGG
GCACTGAGGCCTACGTATTTCAGGTACTTCCAGAGAGGGCAGAGCGAGAA
GCCATCCAAAGATGATTACTTGGTCAAGGAGAAGAAAGGATTGAAGCTGA
CGAGACACGATAAGCTCTTGAAGAAGTTCAGGCATAAGGAAGCTCTGGTC
TCGGTGTTGGAGGAGAAGAAACCGGCCAATGTTGTGGCAGTGATGGAAGA
ATTAGTGGCGAGGAGGAAGCTGATGAAGTGCGTGTCGAATATGGAGGAAG
GCGAGTTGGGAATGCTGCTAGGCTTCTTACAAAGGTACTGCACTGTGCAG
AGGTATTCAGGATTGTTGATGGGATTGACGAAGAAGGTTTTGGAGACGAG
AGCTGAAGACATCAAGGGTAAAAACGAGTTTAAAGGGCTTCTCAGGAACT
TGAAGAGGGAGGTTAATCAGGAGATTAGGATACAACAGTCTCTTCTAGAG
ATTCAGGGTGTTATTGCTC**CTTTGATGAGAATTGCGGGTAGAAGTTGATC**
TA

Figure 4: The deduced amino acid sequence of *254 XMS* protein, a product of a gene potentially affected by the T-DNA insertion is shown in 4a. *254 XMS* is expressed on the complementary strand (see NCBI Accession number ACO05309). The DNA sequence of the open reading frame, as well as the primers designed to amplify it, are shown in 4b.

(A)

MKKRMWRISLISQISDLLCLSRGSSSTLKTLTPFCFTLSRSPFHQSGGDDDASGLKNQLLRFRNDSGKVAS
VLERNKIQGAAFVELLRQLRPWPVLSQLVFDWRRNKALCDGLPMTADEYAKGITISGRLKNVDLALSLFHE
SANKTTSVYNALMGAYLCNGLSHHCEQLFLDFNSQQDGPSSSTPSVSTYNILISLYGRLIMVERMESVFLQ
LQQLNILPDSSTYNNLIAGYIYAWDWDKMEATFHSMKNGLVKPTLATYLLMLRGYANSGNLLRMEDMYQAV
KRHVDRNEIKLIESMICAYYRSCHKDRIRKIKTLSKLIPKKSYKPWLYLLLMQVYAKDDNLHAMENFIDQA
ITKGLQIETDGIMRSIVASYFRCNAVDKLAKFVQRANSAGWKMSRSMFHGLMIMYGSQKRFKEMENVLSEM
ESFKISRSKKTLCILLRVYAATHGQEHKVNQVAGMMLKHGHDFQRPEASKRVMGK

(B)

ATGAAGAAGCGCATGTGGAG

AATCTCACTGATTTCCCAGATATCTGATCTGCTCTGCCTTTCCCGTGGCT
CTTCTTCAACCTTGAAGACTCTGACCCCATTCTGCTTTACCCTCTCCAGG
TCTCCCTTCCACCAATCCGGGGGCGATGATGATGCTTCGGGATTGAAAAA
CCAATTGTTGCGCTTCCGTAATGATTCCGGTAAAGTTGCCTCCGTTTTGG
AGAGGAACAAAATCCAAGGAGCTGCGTTTGTTGAGCTCCTTAGGCAGTTA
CGTCCATGGCCCGTTTTGTCCCAACTGGTATACATACATCTTTTTTGTCT
TCGAATTTCACAACCACAAAAGAAAAAGCCTTTTTTCCGTCATGTTTGAT
GTTGTTGCTTTGCAGGTGTTCGACTGGAGAAGAAACAAGGCACTATGCGA
TGGGTTGCCTATGACTGCTGATGAATACGCTAAGGGCATTACCATTTCTG
GTAGACTGAAGAATGTCGATTTGGCCCTTTCCTTGTTTCACGAGTCTGCA
AACAAGACCACCTCCGTCTACAATGCTTTGATGGGTGCTTACCTCTGCAA
CGGTTTGTCTCATCATTGCGAGCAGCTCTTCCTTGACTTCAATTCTCAAC
AAGACGGACCCTCCTCTTCTACTCCATCTGTCTCTACTTACAACATCCTC
ATTTCTCTTTATGGCCGTTTGATTATGGTTGAGCGGATGGAATCAGTATT
CCTACAGCTTCAACAACTCAATATCCTTCCCGATTCCAGCACCTACAACA
ACTTGATTGCTGGGTACATCTATGCTTGGGATTGGGATAAGATGGAAGCA
ACTTTCCACTCCATGAAGAACGGCCTTGTTAAGCCTACGCTTGCCACCTA
TTTGCTGATGCTCAGAGGGTATGCAAACTCGGGAAATCTGCTACGGATGG
AGGACATGTATCAAGCCGTCAAGCGTCACGTTGACAGGAATGAGATTAAG
TTAATCGAGTCCATGATTTGCGCATACTACAGGAGTTGTCACAAGGATAG
GATCAGAAAAATCAAGACTTTATCCAAACTTATCCCCAAGAAAAGTTACA
AGCCTTGGTTGTATTTGCTCTTGATGCAGGTCTATGCTAAGGATGATAAT
CTACATGCCATGGAGAATTTCATAGACCAGGCTATTACAAAAGGACTCCA
GATAGAGACAGATGGTATTATGAGATCAATAGTGGCTAGTTACTTCAGGT
GCAATGCGGTTGATAAGCTTGCCAAATTCGTGCAGCGTGCAAACTCTGCT
GGATGGAAAATGAGCAGGTCTATGTTCCATGGCTTAATGATCATGTATGG
ATCTCAAAAGCGTTTTAAAGAGATGGAAAACGTCCTCTCTGAAATGGAAA
GCTTTAAAATAAGCCGGTCCAAGAAGACGCTTTGTATTCTGCTTAGAGTA
TACGCAGCAACGCATGGACAGGAGCATAAGGTTAATCAAGTCGCAGGAAT
GATGTTAAAGCATGGACATGATTTCCAACGACCTGAAGCGAGTAAAAGAG
TTATGGGAAAA**TAG**GATACTAGAGAGA**GAGATATGAATAAATACAAGCAG**
CA

Figure 5: The deduced amino acid sequence of *255 XMS* protein, a product of a gene potentially affected by the T-DNA insertion is shown in 4a. *255 XMS* is expressed on the complementary strand (see NCBI Accession number ACO06072). The DNA sequence of the open reading frame, as well as the primers designed to amplify it, are shown in 4b.

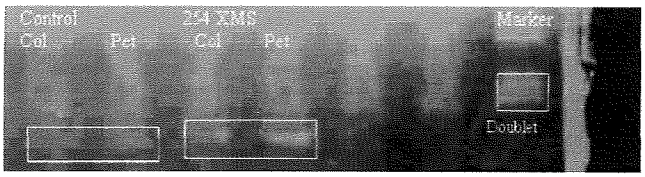

Figure 6: Primers were designed to amplify the 254 XMS and 255 XMS coding sequences. However, only the presence of 254 XMS was confirmed in the control Columbia plant as well as the mutant petite plant. The control primers amplified endogenous carbonic anhydrase.

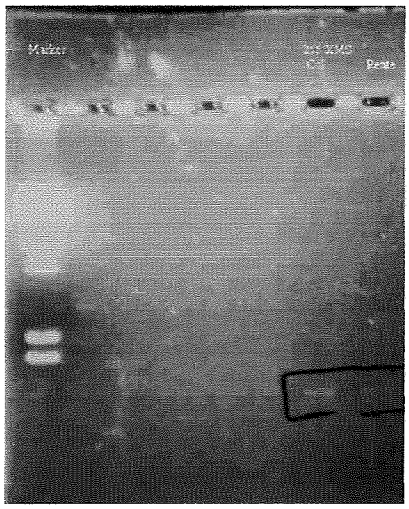

Figure 7: An RT-PCR assay was conducted to compare expression of the 254 XMS and 255 XMS in wildtype and mutant *petite* plants. However, only the expression of 255 XMS was obtained. As seen above, the expression levels appear to be slightly different in both control and mutant plant.

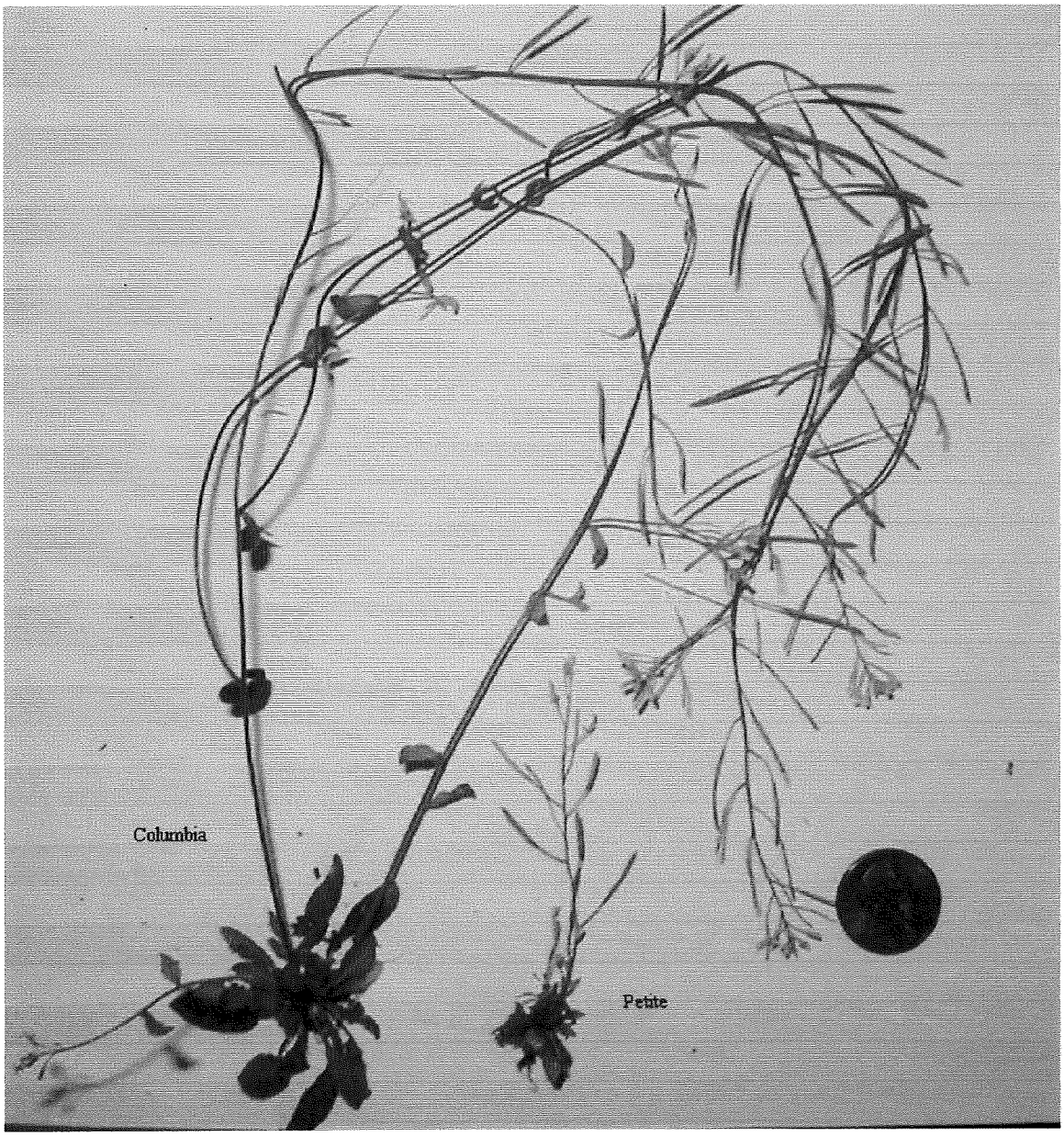

Figure 8: As seen above, the phenotypic difference between wild type Columbia and mutant petite plant is striking.

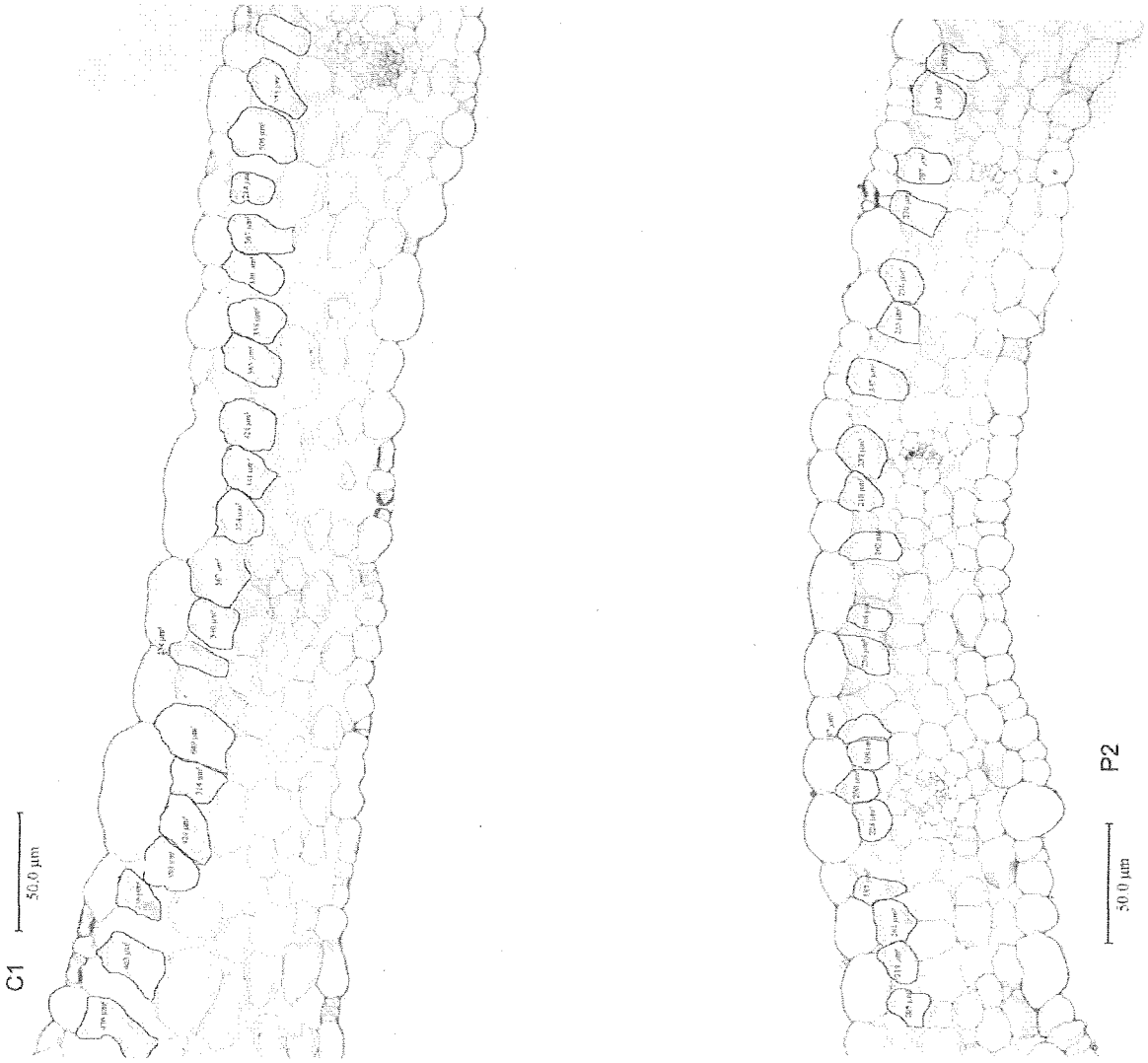

Figure 9: Cross sections of the leaf revealed that the mesophyll cells in the *petite* leaf (right) were, on average, smaller than those in wildtype leaf (left).

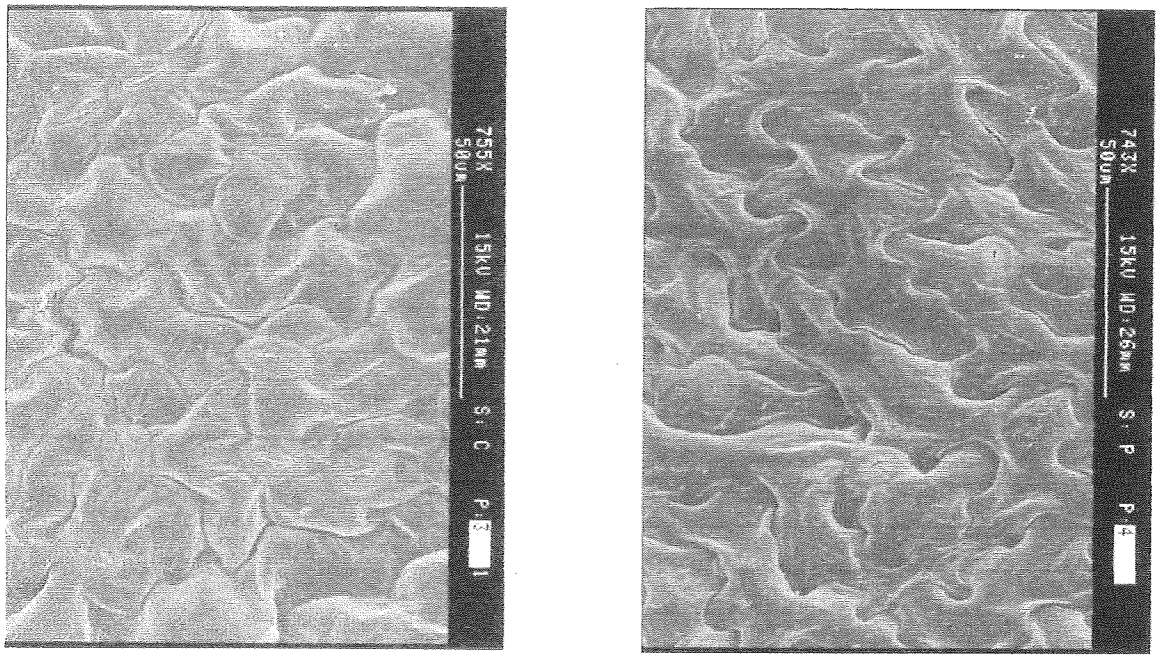

Figure 10: As seen above, wildtype (left) exhibits a greater number of stomates than the *petite* (right).

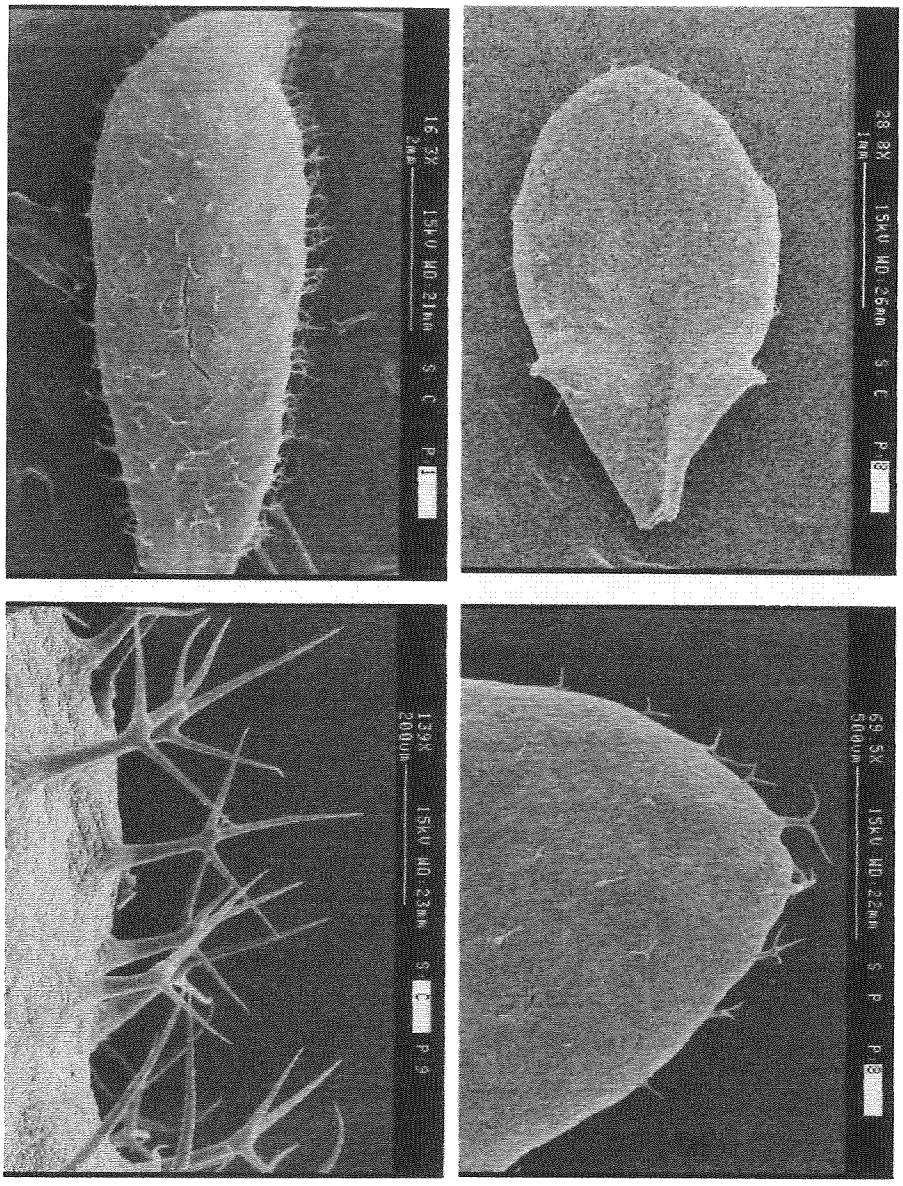

Figure 11: The scanning electron microscope shows that the wild type plant (left) has more trichomes that are tripartite. However, the *petite* (right) displays fewer trichomes and many are bipartite.

Works Cited

Bowman, J. (1993) *Arabidopsis-An Atlas of Morphology and Development.* Springer-Verlag, New York.

Campbell, N., (1993) *Biology* 3rd Ed. Benjamin Cummings Publishing, Redwood City, CA.

Koncz, C., Chua, N., Schell, J. (1992) *Methods in Arabidopsis Research.* World Scientific Publishing. Singapore.

Krysan, P., Young, J. (1999) T-DNA as an insertional mutagen in *Arabidopsis. Plant Cell*, **11**, 2283-2290.

Liu Y., Mitsukawa, N., Oosumi, T., Whittier, R. (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant Journal*, **8**, 457-463.

Liu Y., Whittier, R. (1995) Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics*, **25**, 674-81.

Sambrook, J., Fritsch, E., Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor. New York.

Small, I., and Peters, N. (2000) The PPR motif-a TPR related motif prevalent in plant organellar proteins. *Trends in Biochemical Sciences*, **25**, 46-47.

THE EXAMINATION OF THE EXPRESSION PATTERNS OF A GAMMA CARBONIC ANHYDRASE HOMOLOGUE IN *ARABIDOPSIS THALIANA*

Table of Figures

Abstract

Carbonic anhydrase is a zinc metalloenzyme that catalyzes the interconversion of CO_2 and HCO_3^-. Carbonic anhydrase (CA) is important in many physiological functions that involve carboxylation or decarboxylation reactions, including both photosynthesis and respiration. The purpose of the research project was to characterize the expression of a gamma carbonic anhydrase (γCA-1) in *Arabidopsis thaliana*. The promoter and cDNA of γCA-1 were isolated and sequenced. The promoter was introduced into pBI101, a GUS reporter vector system. The construct (pBI101/γCA-1/GUS) was transformed into *Arabidopsis* using *Agrobacterium* and the flower dip method. It was observed that γCA-1 expression followed the vasculature of small leaves and the molecule was possibly expressed in the cells of the procambium in small leaves. Keeping previous expression patterns of other γCA's in mind, it was hypothesized that different isoforms of this carbonic anhydrase are expressed spatially and temporally to have carbonic anhydrase present in cells that have the potential to divide throughout the plant. An antisense construct was also created and the cDNA was inserted backwards behind the promoter in pBI101/γCA-1/GUS. Antisense plants need to be created to examine the physiological role of γCA-1. Immunoprecipitation experiments can be conducted to study the interactions of γCA-1 with other molecules.

Background

Carbonic anhydrases

Carbonic anhydrase (CA) is a zinc metalloenzyme that catalyzes the interconversion of CO_2 and HCO_3^- (Khalifah, 1971). The enzyme has been identified in most organisms including animals, plants, archaebacteria, and eubacteria (Hewett-Emmett and Tashian, 1996). CA is important physiologically as it is involved in carboxylation and decarboxylation reactions, including both photosynthesis and respiration. CA also participates in the transport of inorganic carbon to actively photosynthesizing cells or away from actively respiring cells (Henry, 1996). CA's are classified into three families called α-CA, β-CA, and γ-CA (Hewett-Emmett and Tashian, 1996). *Arabidopsis thaliana* expresses all three types of CA's.

Mechanism

While the primary sequences of the CA families are different, all three types of carbonic anhydrases are Zn^{+2} metalloenzymes and all appear to share a similar catalytic mechanism (Lindskog, 1997). A $Zn\text{-}OH^-$ attacks a CO_2 molecule residing in a hydrophobic pocket, generating a Zn-bound HCO_3^-. HCO_3^- is released after a water molecule replaces the bicarbonate. HCO_3^- in solution can gain a H^+ to form H_2CO_3 or can lose an additional H^+ to form CO_3^{-2}.

The interconversion is shown below:

$$CO_2 + H_2O \leftrightarrow H_2CO_3 \leftrightarrow HCO_3^- + H^+ \leftrightarrow CO_3^{-2} + 2H^+$$

It is possible that the three forms of carbonic anhydrase represent an example of convergent evolution of catalytic function (Hewett-Emmett and Tashian 1996).

γ-Carbonic anhydrases

The first γ-CA was discovered in the archaebacterium *Methanosarcina thermophila* (Alber and Ferry, 1994). The γ-CA functions as a trimer of identical subunits and the structure of each monomer is dominated by a left-handed β-helix (Kisker et al. 1996). The trimer contains three Zn atoms, one at each of the three subunit interface. In contrast to α-CA, the histidines coordinating the Zn atoms are provided by two separate subunits. However, the active site of γ-CA is similar to that of α-CA (Kisker et al., 1996).

The three γ-Carbonic anhydrases

Several *Arabidopsis* EST's in the databases have homology with the γ-CA from *Methanosarcina thermophila.* Two of them, γ-CA2 from chromosome I, and γ-CA3 from chromosome V, have very similar sequences around the active site (described below). Third, γ-CA3 (described in thesis), initially identified from a shotgun clone of unknown chromosomal location is also on chromosome I. It is less similar to the others but still retains high identity and the histidines at the active site.

The study of γ-CA2 by Nielsen (2001) has shown that the pattern of expression is centralized at the immature leaves with limited expression radiating out a short distance into the petioles of the mature leaves and cotyledons and down the hypocotyl. However, microscopic analysis of the transgenic plants containing the first 21 amino acids of γ-CA2 fused to GUS revealed that the GUS product is associated with organelles, perhaps plastids (Nielsen 2001).

γ-CA3 appears to be expressed in regions undergoing active cell division and expansion, such as the root tip, leaf primordia and pollen mother cells (Xi 1999). The

result of RT-PCR showed that γ-CA3 expression is much higher in root tips than in mature regions of the root or mature leaves (Xi 1999).

Materials and Methods

RT-PCR

Total RNA extracted from seedlings of *Arabidopsis thaliana* was obtained from laboratory stocks. Expressed sequence tags AV549680 and AV564979 were used as a guide for the design of the primers flanking the open reading frame of the mRNA (5'-CACTTTGGGTAATCGGAGAATCA ; and 3'-TCTCTCTTGACCTTCCAGAGAAT) . Reverse transcriptase PCR (RT-PCR) was performed using the specifications provided by the manufacturer (Stratagene). The product was extracted with an equal volume of phenol/choloroform. The aqueous phase was suspended in two volumes of isopropanol, kept on ice for a few minutes and centrifuged to obtain a pellet. The pellet was dissolved in TE and the product was visualized on a low melt agarose gel.

cDNA ligated into pGEM-T Easy

All DNA manipulations including precipitation, ligation, transformation, and restriction digestions were performed using standard methods (Sambrook et al., 1989). The PCR product (which was run on a low melt gel) was ligated with pGEM-T Easy (Promega) according to the manufacturer's protocol. The ligated product was transformed into DH5α cells by electroporation, and the cells were plated on media containing ampicillin (50ug/ml) and x-gal. Colonies were picked and alkaline lysis of overnight cultures was carried out. The plasmid product was subjected to restriction enzyme analysis. The presence of the cDNA insert was established by digesting with *EcoRI*. Further confirmation and orientation was obtained by digesting with *AvrII/PstI*. One culture was selected and grown overnight in LB broth. The plasmid was extracted

and purified using the Qiagen Plasmid Mini Kit according to the manufacturer's protocol. The amount of DNA was quantitated by electrophoretic gel analysis, and the sample was sent to Genelab, LSU, for sequencing.

Promoter of γCA1

Total DNA from *Arabidopsis thaliana* ecotype Columbia was obtained from laboratory stocks. Primers (5'-TGAAGCCACACCACCGATATA G; 3'-TGATTC TCCGATTACCCAAAG TG) were designed to flank a 1 kb region upstream of the translation start site for the putative γ-CA1. PCR was performed (100 ng of template, 2 uM Primer concentration, .2 mM of dNTPs, 2 units Vent polymerase). The parameters of the reaction were 4 minutes at 94 C, 35 cycles of 30 seconds at 94 C, 66 C for 30 seconds,72 C for 1 minute, and finally 3 minutes at 72 C.

The PCR product was subject to A-tailing after being cleaned and precipitated (see above). The product was raised in TE and 2.5 mM of $MgCl_2$ and 0.2 mM of dATP was added. One unit of Taq polymerase was added and the reaction was kept at 72 degrees for 20 minutes. The product was precipitated with sodium acetate (10% of volume) and two volumes of 100% percent ethanol. The pellet was then resuspended in 70% ethanol and the final pellet was raised in TE. The PCR product was run on a low melt agarose gel and was ligated with pGEM-T Easy (Promega) according to the manufacturer's protocol. The ligated product was transformed into DH5α cells by electroporation, and the cells were plated on media containing ampicillin (50ug/ml) and x-gal. Colonies were picked and alkaline lysis of overnight LB/amp cultures was carried out. Restriction analysis was conducted using *Eco RI* and *Bam HI*. Secondary confirmation and the orientation of the product were revealed by digestion with *XhoI* and

Bam HI. Cells harboring positive clones were grown overnight in a LB/amp culture. The plasmid was extracted and purified using Qiagen Plasmid Mini Kit according to the manufacturer's protocol. The amount of DNA was quantitated by electrophoretic gel analysis and the sample was sent to Genelab, LSU, for sequencing.

Introduction of the γCA1 promoter into pBI-101

A Kpn I site was introduced into the 3' primer (CCTTGGTACCTGATTCTCCGATTA CCCAAAGTG). The PCR protocol described previously was used to amplify the new product. The product was precipitated, A-tailed, reprecipitated, gel purified, and was once again reintroduced into pGEM-T Easy. The insert was isolated by digesting with *Kpn I* and *Sal I* and ligated into pBI101 and digested with the same enzymes. The ligated product was transformed into DH5α cells by electroporation. Colonies were picked from plates containing kanamycin (30 ug/ml). Alkaline lysis of overnight LB/kan cultures was carried out. Restriction analysis was conducted using *Kpn I /Sal I* to confirm construction of pBI101/γCA1/GUS.

Introduction in Agrobacterium

PBI101/γCA-1/GUS was transformed into AgroGU 3101 using electroporation. The culture was plated in medium containing kanamycin (30 ug/ml) and gentamycin (50 ug/ml). The plates were maintained at 28-30 degrees for optimal growth of Agrobacterium. Colonies were selected and subject to alkaline lysis. The insert was confirmed by the use of PCR with the original primers used to amplify the promoter sequence.

Transformation into plants

Arabidopsis seeds were suspended in 0.15% agar and directly spread over soil in a 4" pot covered with a nylon sieve. The plants were covered with a plastic dome and placed in the dark for two days at 4 C and then transferred to continuous light at 23 C. After the seeds sprouted, the cover was removed. The plants were watered three times a week and were fertilized once a week. Plants were thinned to 9 plants per pot and were ready for transformation after a month. There were a few dozen open flowers. Plants were watered well one day before infiltration. One liter of infiltration media was prepared containing 2.3g Murashige and Skoog salts with vitamins, 50 g of sucrose, 0.5 g MES, pH to 5.7 with KOH, 44 nM Benzylaminopurine, and 200 ul of Stilwet L-77. Overnight cultures of the selected Agrobacterium (containing the insert) were spread on four kan/gen plates. The colonies were scraped into the media and mixed thoroughly. The plants were inverted and immersed into the infiltration media using the floral dip method. The plants were kept in that position for five minutes and then returned to the incubation room on its side. The pots were uncovered from plastic wrap and set upright the next day. The transformed plants were grown at 23 C until the seeds were harvested.

Selection of transgenic plants

Seeds were sterilized by washing in 100% Ethanol (1 min), 5 minutes rocking in 50% bleach/0.02% Triton X-100, and three rinses with sterile water. The seeds (approximately 1000 per plate) were plated onto selection plates containing vancomycin and kanamycin. The plates were sealed with paper surgical tape, placed in the dark for two days and then transferred to a lighted growth chamber at 23 C.

GUS Assay

The growth of the plants was monitored and the plants that survived after 15 days were slowly pulled out. The plants were transferred to microfuge tubes containing 0.5 ml of GUS staining solution (100 mM sodium phosphate pH 7, 10 mM EDTA, 0.1% triton X-100, 1 mg/ml X-Gluc and 100 ug/ml chloramphenicol). The tubes were placed in a vacuum dessicator and vacuum was drawn for 10 minutes. The tubes were wrapped in aluminum foil and incubated at 37 C overnight. The plants were destained with 70% ethanol for several hours.

Construction of Antisense construct

The cDNA construct was reamplified using the original cDNA primers with an added *Acc 65 I* site on both 5' and 3' ends. The PCR conditions were the same as listed above. The product was precipitated, cleaned, A tailed and run on a low melt agarose gel. It was ligated with pGEM-T Easy. The resulting construct was digested with *Acc65 I* to produce the insert. γCA-1 promoter/pBI101 was digested with *Acc65 I* and was treated with alkaline phosphatase at 37 C (1hr) to prevent self-ligation. The insert was ligated with γCA-1 promoter. The ligated product was transformed into DH5α cells by electroporation. Colonies were picked from plates containing kanamycin (30 ug/ml). Alkaline lysis of overnight LB/kan cultures was carried out. The product was subject to restriction enzyme analysis. The presence of the cDNA and its required orientation in the backwards direction was established by digesting with *EcoRI* and *Bam HI*.

Results and Discussion

cDNA

Primers designed to amplify the cDNA produced an approximately 700 base pair product using RT-PCR (Figure 1). The cDNA was isolated and cloned into pGEM-T Easy. To confirm the insertion, restriction analysis was carried out using *Eco RI*. As seen in Figure 2, the sites flanking the insert as well as an internal *Eco RI* site at the 3' end produced three fragments (approximately 3kb, 500bp, 200bp). Further digestions were made with *Avr II* and *Pst I* and it was concluded that the cDNA ligated into the plasmid in reverse orientation with respect to the promoter. The sequence received from GeneLab, LSU confirmed the expected results (Figure 3)

Promoter

Primers designed to amplify the promoter region produced an approximately 1000 base pair product (Figure 4). The promoter was isolated and cloned into pGEM-T Easy. To confirm the insertion, restriction analysis was carried out using *EcoRI* and *Bam HI*. As seen in Figure 5, the sites flanking the insert as well as an internal *Bam HI* site produced three fragments. Digestions with *BamHI* and *Xho I* confirmed the presence of the insert. The sequence received from GeneLab, LSU, confirmed the expected results.

A *KpnI* site was introduced at the 3' end to facilitate cloning in to pBI-101. The PCR product was reintroduced into pGEM-T Easy. The presence of the insert (and the *Kpn I* site) was confirmed by digesting with *Kpn I* and *Sac I*. The insert and pBI-101 were digested with *Kpn I* and *Sal I* and then ligated. The presence of the insert into the vector was confirmed by digesting with the same enzymes (Figure 6). The fused product

was transformed into *Agrobacterium* cells. Primers used to amplify the promoter were used to confirm the presence of the product within *Agrobacterium* (Figure 7).

Arabidopsis plants were transformed using the floral dip method. The seeds were harvested and grown on kanamycin/vancamycin selective plates. Possible transformed candidates were picked after 10 days and introduced into the GUS staining solution. γCA1appears to be expressed in young leaves (Figure 8). As was the case for γCA3 and γCA2, γCA1 is expressed in the vasculature of the young leaves, possibly in the procambium. The expression is intermediate in size between those expressing γCA3 and γCA2 (Figure 8).

Antisense construct

An *Acc65 I* (isochizmer of *Kpn I*) site was introduced into the cDNA 5' and 3' ends. The product was introduced into pGEM-T Easy and then cut out using *Acc 65I.* The γCa-1 promoter-pBI101 fusion vector was cut with the same enzyme and treated with Alkaline phosphatase. The vector and insert were ligated and the presence of the antisense construct (with the cDNA in reverse orientation behind the promoter) was confirmed by cutting with *Eco RI* and *Bam HI.* The presence of a 210 base pair *Eco RI* fragment confirmed the orientation.

Future directions

In this study, I showed that γCA1 is expressed in young leaves. However, the EST database contains clones identical to γCA1 that were obtained from roots and siliques. Clearly a developmental analysis of expression of the gCA1 will be required to identify all the organs in which the gene is expressed. Antisense plants need to be constructed in order to study function better. Mutations can be introduced into the

promoter element to examine the importance of γCA-1 in cells that express it. Immunopreciptation results can be conducted to study the interactions of the CA with other molecules and hence elucidate function better.

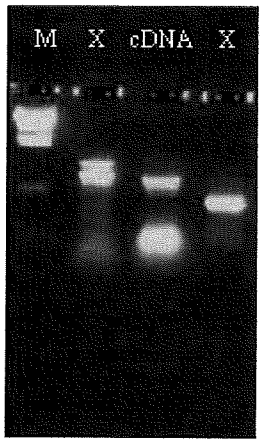

Figure 1: The results of the RT-PCR for γCA-1 are shown. An approximately 700 base pair product was amplified from extracted RNA. The last column is the control.

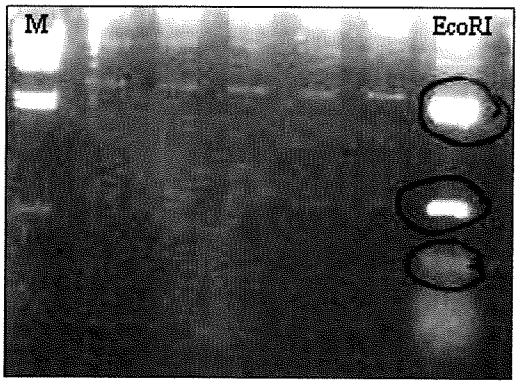

Figure 2: A diagnostic check was conducted with *Eco RI* to check the insertion of the cDNA into pGEM-T Easy. As expected, 3 bands were seen (approximately 3kb, 500bp, 200bp).

MGTLGRAIYTVGNWIRGTGQALDRVGSLLQGSHRIEEHLSRHRT
LMNVFDKSPLVDKDVFVAPSASVIGDVQIGKGSSIWYGCVLRGD
VNNISVGSGTNIQDNTLVHVAKTNISGKVLPTLIGDNVTVGHSA
VIHGCTVEDDAFVGMGATLLDGVVVEKHAMVAAGSLVKQNTRIP
SGEVWGGNPAKFMRKLTDEEIVYISQSAKNYINLAQIHASENSK
SFEQIEVERALRKKYARKDEDYDSMLGITRETPPELILPDNVLP
GGKPVAKVPSTQYF

Figure 3: Sequences sent to GeneLab, LSU, confirmed the sequence reported in NCBI (Accession number ACO079677). The deduced protein sequence of γCA-1 is shown above.

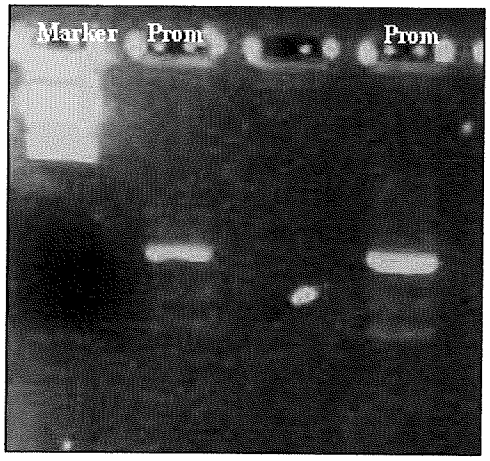

Figure 4: Primers were designed to amplify the 1kb promoter region upstream of the start site to γCA-1. The product is shown with different concentrations of template (25 ng and 100 ng respectively).

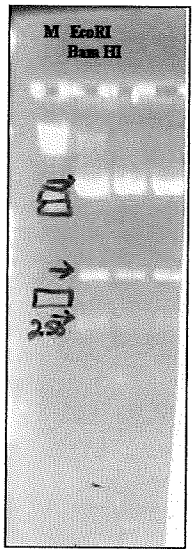

Figure 5: *EcoRI* and *Bam HI* were used as diagnostic tools to confirm the insertion of the promoter fragment into pGEM-T Easy. As expected, 3 bands were produced.

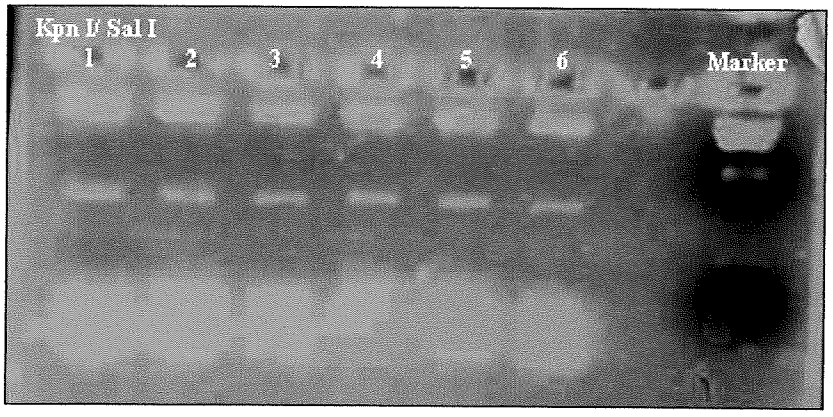

Figure 6: The presence of the promoter fragment in pBI 101 was confirmed by digesting with *Kpn I / Sal I.*

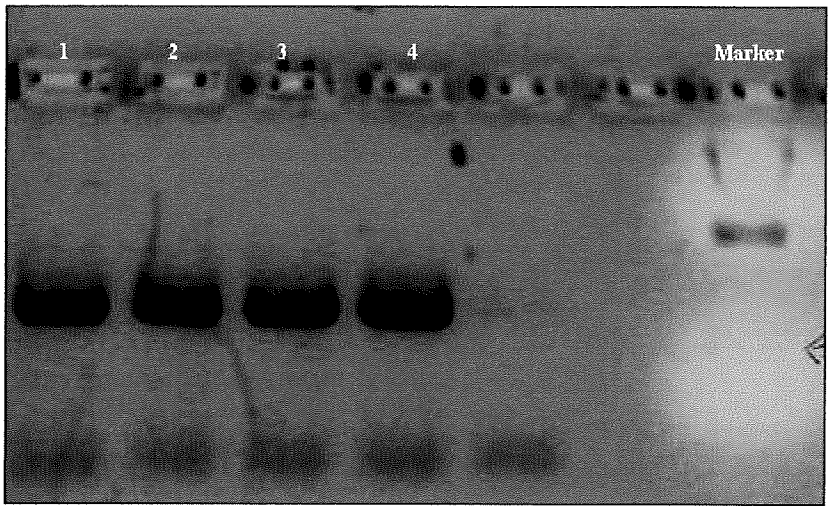

Figure 7: The presence of the promoter-pBI101 fusion vector in Agrobacterium was confirmed by PCR using primers that amplified the promoter. The approximately 1 kb fragment is seen.