Indian Journal of Plant Sciences ISSN: 2319–3824
Online, International Journal, Available at http://www.cibtech.org/jps.htm
2024 Vol.13, pp.23-29/Jayaraj
Research Article (Open Access)

A NOTE ON GENOME WALKING AND ITS IMPORTANCE IN THE STUDY OF GENOMIC DIVERSITY IN *EUCALYPTUS NITENS*, CONTROLLING WOOD TRAITS

***G. Jayaraj**
Centre for Research and Dept of Botany and Biotechnology, Andhra Loyola College (Autonomous), Vijayawada, Andhra Pradesh.
**Author for Correspondence: jgorantla3@gmail.com*

ABSTRACT
By using the Genome walking method in the present study, it is to find out the genomic diversity in *Eucalyptus nitens* We have isolated 5 important genes which are involved in wood traits in *E. nitens* is done using PCR based molecular technique. This is to find out the timing and location of gene expression in Promoter region of DNA. In the present study, we have isolated the promoter regions in *E. nitens* for EnicelK, Eni4CL, EniSusy1, EniUP17 and EniLIM genes. We have identified the TSS, TATA and CATT boxes. The genes need to be cloned for their expression by designing primers in order to genotype SNPs (Single Nucleotide Polymorphic sights that are associated in controlling the wood traits.

Keywords*: G. walking, Traits, PCR, DNA, Promoter, Primer and SNPs*

INTRODUCTION
This study involves isolation and analysis of promoter regions of 5 important genes involved in controlling wood traits in *E. nitens* by a method called Genome walking. Genome walking is a basic PCR-based molecular technique. This may give insights into the genomic diversity within *E. nitens*. The regulation of timing and location of gene expression is done by the promoter segments of DNA. The sequence elements (*cis*-elements) within the DNA of the promoter regulate transcriptional activity by determining the type, affinities and spatial arrangement of regulatory proteins (transcription factors). They are usually located upstream of the transcription start site, but regulatory elements can be present in 5' untranslated regions (UTR), within introns or in the 3' UTR of genes. For the analysis of upstream and down stream noncoding regions the isolation and characterization of unknown DNA sequences flanking the known segments is very vital.

In the past for obtaining unknown sequencers flanking the known sequences, it involved a long and time-consuming successive probing of libraries with clones obtained from prior probes. It also required the use of radioactive probes (Brown *et al.*, 1999). The genome walking does not require the use of the radioactive probes. It is very reliable and relatively fast method for sequencing the adjacent region to the known region of the DNA. The candidate genes of interest include *EniCelK1, EniSusy1, Eni LIM, Eni4CL* and *EniUP17* (Glucosyl transferase) genes. These are constitutive promotors which are used widely in plant genetic engineering these days to express genes throughout the plant.

MATERIALS AND METHODS
Selection of plant material:
The *E. nitens* used in this study is from the collection made at CSIRO forestry division, Yarralumla by Colleen.M. The collection includes multiple samples from Mount Gambier region of Victoria, Australia.

The genome walking:
We used a Genome walker kit (Unitech, USA). It Genomic DNA is isolated from 8 grams of leaf material of *E. nitens*. It is checked for high average molecular weight on a 0.8% agarose gel. The DNA is digested by blunt end cutting enzymes *DraI, PVuII, EcoRV* and *StuI*. The digested DNA is then purified and ligated overnight at 16°C with Adaptors I & II provided in the kit. These are now

called Genome walker libraries. In the next step we designed the Gene specific primers *GSPI* and *GSPII* . The *GSPI* is from coding region and *GSPII* from is further upstream.

The primer designing:

The Primer3 programme was used for designing GSPI and GSPII (Primer3 software (Steve Rozen and Skaletsky, 2000) code available at http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Forward and reverse gene-specific primers are listed in Table 4.

The primers are 26-30 nucleotides in length and have G/C content 40-60. Annealing and extension temperature of 67° C is ensured for effective annealing of primer to the template.

Table 1

Gene	API GSPI	APII GSPII
EniCelK	5'–GTAATACGACTCACTATAGGGC–3' CGTTGAAGAACATGAGAGACTTGTGGA	5'–ACTATAGGGCACGCGTGGT–3' ATGCAGCCCAGGTCGACGTACTTCTT
EniSusy1	5'–GTAATACGACTCACTATAGGGC–3' CATGCGATCAGCCATGTCTAAATTCTT	5'–ACTATAGGGCACGCGTGGT–3' TTGGAAAACCAGGAAACAGAAAAGTGG
EniLIM	5'–GTAATACGACTCACTATAGGGC–3' GAGCTTGTCCACCAGATAGACTGTCTTC	5'–ACTATAGGGCACGCGTGGT–3' TGTCTGTGAGTGTAAGGAGAAGGAGAGA
Eni4CL	5'–GTAATACGACTCACTATAGGGC–3' ATCAGCTCGACCTCGGCATAGGTGTAG	5'–ACTATAGGGCACGCGTGGT–3' GGAGAGGTTGTCGGGAATGTAGATGTC
EniUP17	5'–GTAATACGACTCACTATAGGGC–3' AACAGCTGAGCAAAACAATAATTAAGGC	5'–ACTATAGGGCACGCGTGGT–3' AACTCTATCAAAGAAAAATGGAAGCTT

The primary PCR is performed with *API* as outer adapter primer and outer gene specific primer *(GSPI).* It gave multiple bands in each lane with a few exceptions with a general background smear. The primer PCR is diluted in 1: 9 ratio and it is used as a template for the secondary PCR reaction which uses *APII* and *GSPII* primers. It amplified the desired product in the subsequent nesting PCR resulting a single bright band. A positive control and as well as a negative control were used provided by the kit. The PCR products are Gel extracted using DNA kit. The DNA obtained is purified by doing Miniprep.

The PCR

The Cloning of PCR Products:

All purified DNA amplicons were ligated into the pGEM-T Easy vector using the pGEM-T Easy vector kits (Promega, USA). Ten microliters ligation mixtures contained 5 µl of 2xRapid Ligation Buffer, 1µl of pGEM-Teasy vector (50ng), 3 µl PCR products (25-30ng/µl), 1µl of T4 DNA Ligase (3U/µl) and 1 µl distilled water. A positive control was included for checking transformation and ligation efficiency, and used control insert DNA instead of PCR product as template. A no template negative control was also included. Two microliters of each ligation reaction were transformed into bacterial cells (JM109 and DH5α) by heat-shocking for 50 second at 42˚C water bath, and plated out in LB/ampicillin/IPTG/X-Gal medium, and then incubated at 37˚C for overnight. Twenty-four white colonies were picked and cultured in 5 ml LB/ampicillin medium for overnight at 37˚C and plasmid DNA isolated using the QIAprep Spin Miniprep kit (Hilden, Germany). A single positive colony was picked from each individual and cultured for plasmid DNA preparation. The sizes of all inserts were verified by digestion with *Not*I followed by gel electrophoresis. Each 10 µl digestion reaction was consisted of 1 µl of buffer D, 0.1 µl of BSA, 1 µl of plasmid DNA, 0.20 µl of *Not*I (Invitrogen, USA) and 7.70 µl of distilled water and was incubated at 37˚C in a water bath for 2-3 hours.

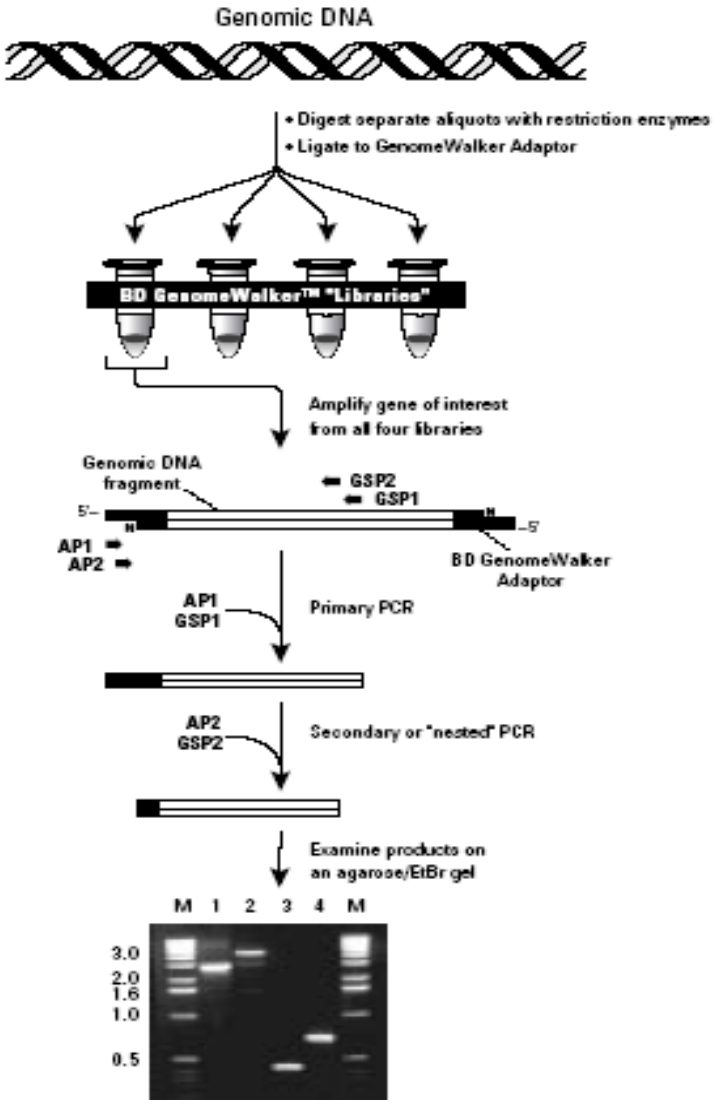

Figure 1: Cloning of PCR

Sequencing

Twelve different amplicons of each fragment of *Eni CelK1*, *EniSusy1*, *Eni LIM*, *Eni4CL* and *EniUP17* products were purified following a BigDye reaction using forward and reverse M13 primers the DNA is sequenced. The was precipitated using ethanol, dried down under vacuum and sent to the John Curtin School of Medical Research (JCMSR) for sequencing.

Sequencing was carried out using BigDye Terminator version 3.1 reagents and an ABI PRISM sequence analyzer using 1/8 reaction volume. Plasmid DNA (0.8 µl) was added to 14.2 µl of distilled water, 1µl of BigDye version 3.1 mix, 3.5 µl of 5x sequencing buffer and 0.5 µl of each primer (10µM). Cycle sequencing used an initial step at 94˚C for 5 min, then 30 cycles of denaturation at 96˚C for 10 s, annealing at 50˚C for 5 s, and extension at 60˚C for 4 min.

Sequencing analysis:

These promoter sequences are analysed in combination with databases like PLACE (http://www.dna.affrc.go.JP/htdocs/PLACE/fasta.html) and PlantCARE/http://oberon.fvms.ugent.be:8080/PlantCARE/index.html).

It provides insights into regulation of these genes involved in wood traits in eucalypts. DNA motifs are identified in promoter regions. This information can be used for SNP discovery and for further comparative analysis of these genes by genotyping the SNPs.

RESULTS AND DISCUSSION

In the last few years there has been found an interest in studying the genes responsible for plant secondary wall formation, lignin and cellulose biosynthesis and xylem development. (Arioli et al., 1998; Fukuda, 1997; Ye, 2002). Wood is important to us as timber, fuel wood, wood-pulp for paper manufacturing as well as a renewable source of energy that is environmentally friendly. The control and expression of the genes involved in wood formation in eucalypts is of particular interest because some of the species like *E.nitens, E. globulus, E. grandis* and *E. camaldulensis* are a good source of hard wood. This study has been undertaken for isolating the promoter regions and identifying the Transcription start site (TSS), the TATA and CAAT boxes of *Eni SUSY1, Eni4CL, EniLIM. EniCelK* and *EniUP17* genes in *Eucalyptus nitens*. The promoter regions size varies from 900 to 1200bps. These might give us insights into the transcriptional regulation of secondary growth. The expression patterns of the genes encoding transcription factors are of vital importance in this approach.

Identifying the regulatory *cis*-elements from the promoter regions of the genes that are upregulated in wood formation in eucalypts is important in the present study. It has been reported that *4CL* may be involved in lignin biosynthesis in eucalypts (Paux *et al.*, 2005). There is a reduced expression of the 4CL gene in transgenic *poplar* that has correlation to the reduction in lignin content (Hu et al., 1999). Gion *et al.*, (2000) reported that *PAL, 4CL, COMT, CCoAOMT, CAD*, and *CCR* are constituent of structural genes of lignin biosynthesis in eucalypts. 4-Coumarate: COA ligases (4CLs) are enzymes necessary for maintaining a continuous metabolic flux for the biosynthesis of phenylpropanoids such as lignin and flavonoids. 4CL is reported to be involved in the development of xylem tissue in *Populus tremuloides* (Hu *et al.*, 1999).

The rosette terminal complexes (as part of plasma membrane in making cellulose) are formed by glycosyltransferases (Burn *et al.*, 2002). The Korrigan cellulase was suggested to be involved in cellulose biosynthesis. It removes the Sitosterol from the primer for biosynthesis reaction of cellulose in *Arabidopsis* (Kimura S and Kondo T 2002). The cellulose biosynthesis apart from the *CesA* proteins, also requires genes like Korrigan (KOR). If there is a mutation in KOR gene, it leads to reduced cellulose production in plants (Mutwil *et al.*, 2008). KOR1 and related membrane-anchored endoglucanases may be involved in chain termination during cellulose biosynthesis or in the degradation of ß-D-glucan chain that have not been properly incorporated into cellulose microfibrils. A role for the KOR endoglucanase in cellulose synthesis has been established by several groups. This work is limited to isolation the promoter regions in E. nitens for *EniCelK*, *Eni4CL*, *EniSUSY*1, *EniUP*17 and *EniLIM* genes by genome walking and for the identification of TSS, TATA and CAAT boxes. These can then be cloned and further analysed for *GUS* expression in transgenic tobacco seedlings. These can be further analysed after designing primers, for SNP discovery in this region. The genotyping of SNPs in these segments of the genes might give us insights into association of SNPs and wood traits.

Figure 2: *EniLIM* gene. Promoter region

TGTCTGTGAGTGTAAGGAGAAGGAGAGAGGAGAATGGAGGATAAGAAAGGGAAGCCAAAAAAACAAGTGTTGG
TTGGAAAGTTGATGATTAAAGTGGTGGCATTGGACAATCATTTTGGGTGAAGGGTCTAACTAAACCTTTTATGACA
TGTTTTACCTTTTCCCCCCCTTCTTCTAGGTTTGTTAATGACCAAACACAAGATGGGCAGCATGAAGCTTAAGTTG
AGTAAGAATTAGAGAAAAATAATTCGGCCATTCCCTTTCTTTTTTTTTATCTATCATGATCTTTTAGTCGTCTATTTT
AACTAATATTGAAG**CAAT**GATATTGGTGAGAAATTCCGAGAAATCAAAAGCGGAACTTCTAAAAAATTCATTAGT
AGATTTCATCTTAGATTGCTCTAATTAAACTGGACCAGCCCGGTTACTATGATCAACTTTTCCAAAGGGCACTACT
ATGCTTTTAGCTGCATTCCAGAAAAATATCTGGCAGATGTGACCCTAGTAGCAAAGGGGCCATGGAATAGAAACT
TAGCTGGATTGCGCATGGAAAGGAGGAGAGGACCAGAGAGGGAAA**TATA**ATTGGAGTCACTGAATCGAGCTGGA
GC**T**GCATCGTCCTCCAACTGTTTGTTGTAAGGACTTCATGTGAGTTGGCCCAATGATTGTCCATCAAATGCTCTCC
CCCATCATGTCGCTTTGCATACACGTTCTTCAACTTTACTTTCCTCTCTCTTACATCACATTACCTTTCTTTTTTTCTT
TTTTGGGCTGAGAGCTCCAGTCCCATCGACGATCCTTGTTTTGAGAAAAAATTTCCTATCTGAAAATAGTTTCTTTA
CTTTTTGGCTTTGGATTGTTTGCAAAAATAAGTTTACAGAAAATGTTTTCATAGTAAATGAAAAGCATGCGTTTAA
ATTCAAGAAAATAATTATGATAAAAAAATCAGAAAATATTTCTTCTTTTTGAAATGAAAATCATTTTTAAAATACG
ACTAACATGAATCCCGACCTGTACTGATCATTAGTGATTAGGCTCGGGTCCCAAAGGCTTAGTCCAAAACTTAATG
TCTATGCTAGGGTCTCTAGTTGCCATGCTCGGGTACCTAGCTACTGACCTAAATCACTAGCGCCTTGGGCGCCGAA
CTAGTAAGGTTGGGTCTCATTCCTTG**ATG**GTTGAGCTTGGATCCATAGAGGCTTGGCCCAAATATCTTGAGGACCA
GCCCGGGCCGTCG

ATG
T - transcription start site 566 - 616 90 % cinfidence
AGAGGGAAATATAATTGGAGTCACTGAATCGAGCTGGAGCTGCATCGTCC
TATA-box
CAAT-box

Figure 3
***EniSUSY1* promoter region**
TTGGAAAACCAGGAAACAGAAAAGTGGAGCGAACGGAACCCTCGTGCCGAATTCCCA
TCATGTTTTTTGCTTTGCTAGTTTGCTTGAGAAAGCAGGAAAAATAGTAGAGGGAGAT
CAAATGGACATATAATCTTCTCCTCAATGATTGAAATGTAAGTTCTTTCCATGTTTGCA
GCAGAAAGCAGACCCTTTTATGCTATTGAAAATGATCTTTCTCGCACGAACTGTTTTTC
GGTTACACTCGAACCCCCTTCATGTCGTGCTGTGCATTAAAAAGAATGCATTAGAAGA
CTTGTTACAGCAGTCCTTTACAGGTTCTGCGTCACTGCCGATGGCTTGTATAAGACACC
TTTCGCCACGGTGGCCCAACAGCTAATCTTCTACACCTATCGTCTGGTTAATTTGATAT
TACTCATTACACACACCCACGCACACAGAGTGCGATAA**CATAAAT**GACAGTCATCAGT
AAGATTCATAAT**CAAT**CATCTCAGCATCATCTGTCCATCAC**TATA**ACAACATTCGCCTT
TCCATTTCTGATGGCTGGCAAACACGCGATGGTTTAATCATGTTGATCATAACCTGCTT
CCTAAGCTAAAAAAAACCAGAGCTTACAGCAGTGACTGT**G**TGAATCCAGTTCGTAGG
GATCTGGATACCAGCCCGGGCCGTCGACCACGCGTGCCCTATAGTAATCACTAGTGAA
TTCGCGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCCACGCGTGGAGCATAGCT
GAGTATTCTATAGGTCACCTAAATGACCACGCGTGCCCTATAGTGCCCTATAGTGAGT
CGTATTAC

ATG
G- transcription start site 585-635 92% confidence
TCCTAAGCTAAAAAAAACCAGAGCTTACAGCAGTGACTGTGTGAATCCAG
TATA-box
CAAT-box

Figure 4
***EniUP*17 (glucosyl transferase) promoter region**
AACTCTATCAAAGAAAAATGGAAGCTTTTTCTCTGCCGCTGCCCTCAAACGAAGACG
GAAGCAGAAACGAGAAACGAATGAACAAGGGGGGAGAAGAGAGGAAATGAGAACG
AAGACGAAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCAACCTACTGCAGCACA
AAACAGAGTGAGGGAGGCTTGAGATATATACGAAGTGACGCTCTCAACAACCATCAG
ACGGGAATGAGAAGGCACTGTTAGGTTAAAACCCTCCGGAGCTTTCTACATTACAACA
TCTGCTTTTCCCAATTTGCGTGGTTCGCAGAAACTAGTTCTTTTCAAGTCTCAACCATA
ATTTCAGAAGCGTGATTCGAGCTCGTTTACTCTTCTCTGAACCTACAACAAAAACCCTC
GTCTGGCCACTCGGTATTTCACCGACATACACAGCATTTGACGGGACGGTCCCTCCAG
CTGCCCCGCAGACGCCGGTCCCGCCATCACCTCATCGAGCAGTGCCGCACATGTTCTC
GACCGACGAATCATTTCAATACGAAACTTGCCATGATTCGATCCATTTCCATAGGAAG
AAAATACCCTTCACGATGGAGAGTGAATACTTCAAATTTCAATCGATGATTTTTCGATT
TCTTCGCGAGTAGAAAATTTGTTTGTTGAGATGTCCGCGTAAACTTGTCGTTCCCTAGT
TCAACCTTAATTCAGACAAAAATGATTGG**CAAT**GGTCCTTTGGACAAACATCTCGTGG
AGTGCATATGGTCTATCCCAACTACGGTTCATTCCATCTGAGACCTTATGCTAGCGTCT
CACTACTTTTGTCAAGACTCTAGGAAACATAACCGCGTTACCATAAATAGGTGAACTC
CGATTTCATGA**TATA**TATACTCCACAATCTCGAAGGCCGA**T**AGACGCCAAGCCAAAAG
CCATTAATTCTTA**ATG**ATTTGATGAGATCCTGACCTAAGCCAAGATCGGCAAATCCAT
CTAAGTAATTAAGTTTAGAGGATCACTTTAACTTGGGTTAATCACTGCCATCACTGAGT
GACGATGATCAATCGTAACCTTATCGGCTGTTGTGGCCTCGATACCAGCCCGGGCGGC
G

ATG
T- transcription start site 873- 923 98% confidence
GATTTCATGATATATATACTCCACAATCTCGAAGGCCGATAGACGCCAAG

TATA-box
CAAT-box

Figure 5
***EniCelK*3 – promoter region –allele 1**

ACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCTGCTTGTTTCTTTTGCCCCCCCTCCCCCCGCCCTA
AAAAATCTTGGATTGAAAACAATGGCCAAACCACGTGGGTTTTAGACATTGTTCAAGTGTCCATAACCAGG
GTATCGTATCATGTGATTTATACCTAAGATGTCGAGTGGCGGTTGTTCTTTCCCGTTATTACCAGTTCATCA
GAAATGGCATCTACGATTCATGAGCTTACATCCCATGTGCAATATCGATAGATAAATGCTGAGTACAACCT
GATCTTGTTTGCTCGAACGGCTAGTGAAACGTTGCTCTTACGGCGTGGATTATTACGCGTCGAGTCGGgGGG
AGACTTGGTAAATCGGTGTGGAAAAGTGGAGGACATTGCGGACAGAGGGTCGCTGTCGGTTAGTGGACTG
GGCACTGTCATCCACCCGACGCTC**CAAACAT**CAGTCCAACCATCCATGATATAATACTATTCTAATTCATTC
ATCTTAA**CAAT**AAAAAAAGAAC**TATA**TATGATTAAAAAAACTCTGGTAACCTCCACATGATC**A**TCAATCAT
CTTCTTGACTCTCGGCGAAGCAAAGCATCTTCCAAACGATCCATCCATTCCTTTCCATCCAGCCATCCATCA
TCACCATCAGCATCAGCATCAGCATCTTCTTCCTCCATCAAGAAACACAGAGCAGCCACCTCCAAACTCCA
TTATTGGTGAGGGAGAAACGCCACCCACCCCACCCCTCCATTTTTCTATGCTTTCCTCCCTGACCAATCTTT
GACGCAGATCTATCTCCCTCTCCCTCTCCCTCTCCCTCTCTCTCACCCACTGTCGGTGC**tgctt**CTTCTTCTTCTT
CTTCTTCTCCCTCTTCCTCCTCCTCACTTTCCCTGTTCTTGATTTTCTTGATATTACTATTGCTGGGAGCTGAA
GCAGACCCAGCAGTTCCTCTCTCTG**c**CGTTCTGCATCCACTGTGGTGGAGTGGTCGAGTGAGGGAAAAGGG
GAGCGACCCAGTTCGCGTCCGAGCGAGG**A**t**G**AGCATGTACGGACGGGACCCCTGGGGGGGACCCCTGGAG
ATCAACGCCGCCGACTCCGCCACCGAGGACGAGCGCAGCCGCAACCTCAACGACTACGACCGCGCCGCCC
TCTCCAACTCCCGCCCCCTCGACGAGACCCAGCAGAGCTGGCTCCTCGGCCAGGGGGAGCAGAAGAAGAA
G[AAGAAGTACGTCGACCTGGGCTGCAT]AATCG
N – SNP
tgctt – indel
[AAGAAGTACGTCGACCTGGGCTGCAT = AP2]
AtG – start codon
A – transcription start site 523- 573 99% confidence
TATATATGATTAAAAAAACTCTGGTAACCTCCACATGATCATCAATCATC
TATA- box
CAT- box

Figure 6
Eni4CL promoter region
GGAGAGGTTGTCGGGAATGTAGATGTCGGGGAGCTTCGACCGGAAGATGAACTCGCGGGGCTGCTCCG
ACGGCTTCGCCTCCATTGCTCGTGGTGGGCAGAATCGAGAGCGAGAAGTAGACAGAGAGGAGACAGAGCA
AGGCTGATGATAGAAGTAGCGTTGTGGTGGCCCTTCGTCCTTCCTCTGAGGTTGGTTGAGCTTTGGTCGCGG
GTAATATACACGAGCTGTGGCGGTAGAGGAGCTGGTGGAGTCGAGTTCGATGGGGGGGCATTTCCGTCTTC
GTATCGCCTGCCAGGGACGGCGGTGGAGGAAGTCCGCCCGGGCTGTCCCTTTGGGCGGGGACCAATCCGA
CCCGCCACTGCCGCCTCGGGCTTAGCCTTGACGGGAGTTCGCGGAGCGTCTCCGCTTGTCTTGTTGCGCACG
TGGCGCGCTGCGATTGGAGGGAGGGGAGGGGAGATTGGGTCGTCTCGGGGCGTCACCTACCACGCCGGGG
AGGGAGCCATCGGAGGTTGGTGAGGGACATTTGGTGGCATAAACAAAAGTCCATCGTTTTCTTTTCGTTTTC
CATTTTTTTTTTGGGGTGCCCTTTCGGGATTATTTTAGTGCATGCCACGTTTTGACCTTGGCCAG**CAAT**GCTG
TGCCACGTAGGATTTCACCTACCAAAGCGGGCCCTCCTTCCCGATTAAACGGCACCCCCACATCGCCCGAG
CTGGACATTCGAACGTGGCATCGAAATTATTCATGCTCTCTATTAGTCAGATTCCACTAGTTAATTCTGA**TA**
TATGTAAAAAGCAGCGCCTGAAATACCCATGTC**A**TTTAGGCTCGACAATCTAATATTTTCATTTCCACAAA
ATAATAATAAGGATAATAATAAGTCCAATCCTTATCTACACATTCACTTTGAGCACCTAAATAATGTCAAG
ACCCATCAATCAATGTACAAAAATGAATCGGCCCAGCCATTCATCTCACATGGCTTGAGGTAGACATTTGA
ACATGGCATCGGAGCCAGATTTTGCCCCTTTTCATGCATGCCTCTACCTTATCGTCAAATCCCACATACCAC
ACTTTCATCCGTCCTCC**ATG**TTAGAGAGAGGGAGTTGAAATTTCTCCATATTGCTCGGACCGAATAATTTAC
CAGCCCGGGCCGTCG

ATG
A- transcription start site 775- 825 99% confidence
TCTGATATATGTAAAAAGCAGCGCCTGAAATACCCATGTCATTTAGGCTC

TATA-box
CAAT-box

REFERENCES

Arioli T, JE Burn *et al.* (1998). A hot mutant for cellulose synthesis ... response: How many cellulose synthase-like gene products actually make cellulose? Trends in Plant Science **3**(5) 165-166..

Arioli T, LC Peng, *et al.* (1998). Molecular analysis of cellulose biosynthesis in Arabidopsis." Science **279**(5351) 717-720.

Brown, AJH., Perry, S.J., Saunders, S.E. and Burke JF (1999). Genome walking in pea: an approach to clone unknown flanking sequences. Biotechniques **26** 804-806.

Burn, J. E., C. H. Hocart, *et al.* (2002). Functional analysis of the cellulose synthase genes CesA1, CesA2, and CesA3 in Arabidopsis. Plant Physiology **129**(2) 797-807.

Burn, J. E., U. A. Hurley, *et al.* (2002). The cellulose-deficient Arabidopsis mutant rsw3 is defective in a gene encoding a putative glucosidase II, an enzyme processing N-glycans during ER quality control." The Plant Journal **32**(6) 949-960.

Fukuda, H. (1996). Xylogenesis: Initiation, progression, and cell death." Annual Review of Plant Physiology and Plant Molecular Biology **47** 299-325.

Fukuda, H. (1996). XYLOGENESIS: INITIATION, PROGRESSION, AND CELL DEATH." Annual Review of Plant Physiology and Plant Molecular Biology **47**(1) 299-325.

Hu, W.-J., S. A. Harding, *et al.* (1999). Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. Nat Biotech **17**(8): 808-812.

Gion, J.-M., P. Rech, *et al.* (2000). Mapping candidate genes in Eucalyptus with emphasis on lignification genes." Molecular Breeding **6**(5) 441-449.

Kimura, S. and T. Kondo (2002). Recent progress in cellulose biosynthesis." Journal of Plant Research **115**(4): 297-302.

Mutwil, M., S. Debolt, *et al.* (2008). Cellulose synthesis: a complex complex." Current Opinion in Plant Biology **11**(3) 252-257

Paux, E., V. Carocha, *et al.* (2005). Transcript profiling of Eucalyptus xylem genes during tension wood formation." New Phytologist **167**(1): 89-100.

Paux, E., V. Carocha, *et al.* (2005). Transcript profiling of Eucalyptus xylem genes during tension wood formation." New Phytologist **167**(1): 89-100.

Paux, E., V. Carocha, *et al.* (2005). Transcript profiling of Eucalyptus xylem genes during tension wood formation." New Phytologist **167**(1) 89-100.

Paux, E., V. Carocha, *et al.* (2005). Transcript profiling of Eucalyptus xylem genes during tension wood formation." New Phytologist **167**(1) 89-100.

Ye, Z H. (2002). Vascular tissue differentiation and pattern formation in plants. *Annaul Review of Plant Biology.* **53**(1) 183-202.