Full Length Research Paper

Analysis of expressed sequence tags derived from inflorescence shoot of *Tectona grandis* (teak)

H. Mohd Rosli[1]*, N. M. Saleh[2], M. I. Mohd Noor[3], M. Norwati[1], B. Norlia[1] and A. Norwati[1]

[1]Genetic Unit, Forest Research Institute Malaysia (FRIM), 52109 Kepong, Selangor, Malaysia.
[2]Cell and Molecular Biology Department, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.
[3]Malaysia Genome Institute, Helix Emas Block, UKM-MTDC Technology Centre, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia.

Teak Inflorescence Shoot Stage 4 (TIS4) shoots bearing the floral meristems were used to construct a cDNA librariy with insert size range of 1500 - 5000 bp. The titer of the library was 7.5 x 10^5 pfu/ml (primary) and 4.5 x 10^9 pfu/ml (amplified). EST generation and analysis were performed using the cDNA library where a total of 1384 plaques were randomly picked and their inserts PCR-amplified using T3 and T7 universal primers. Only 1125 plaques generated single amplified fragments, each which were purified and sequenced using the SK universal primer. The generated raw 5' ESTs were filtered and clustered. A total of 674 nonredundants (69 consensus sequences and 605 singletons) were generated and their identities searched through BLASTX. Of the 674 nonredundants, 107 of them (15.9%) showed no hits or no identity. All the 567 nonredundants identified through BLASTX were categorized into their functional categories and were further analysed using InterProScan to detect their protein signatures and to assign their GO numbers. From all the sequences analysed, only 186 (32.8%) sequences were given the GO numbers and grouped into the three GO main categories namely biological process, cellular component and molecular function. Several important ESTs were highlighted based on their functional categories. There were five sequences found to be related to flowering and light induction.

Key words: ESTs, flowering, *Tectona grandis.*

INTRODUCTION

Teak (*Tectona grandis* Linn. f.) is a tropical tree species distributed naturally in countries including India, Myanmar, Thailand and Laos. Teak has been planted both in the countries where it is naturally distributed and where it is an exotic species such as in Africa, central and southern America and in the Pacific region (Rao, 1997). Teak is considered to be one of the most promising plantation tree species due to its qualities including resistance to termite, fungus and weather, lightness with strength, attractiveness, workability and seasoning capacity (Kadambi, 1972). Plantation grown teak trees have been reported to flower as early as two years of age

(Boonkird, 1966). It was reported that early flowering in teak is controlled by both genetic and environmental factors (Boonkird, 1966; Gram and Larsen, 1958). This early terminal flowering causes forking of the main axis during the first year of flowering and forking of other leading shoots in the subsequent flowering seasons (Boonkird, 1966; Nanda, 1962). The forking of the main axis at a very early stage could reduce the length of un-broken stump for quality timber. Flowering also reduces vegetative growth due to the energy utilization in the flowering processes. Therefore, it is important to apply genomics and molecular biology approaches in order to isolate and ultimately to understand the function and interaction of some important genes involved in the flowering processes in teak. However, to isolate those important floral genes, it is crucial to determine and select suitable shoots that bear floral meristem where those desired genes are highly expressed. Morphological and his-tological observations must first be carried out during the

---

*Corresponding author. E-mail: mrosli@frim.gov.my.

Abbreviations: ESTs, Expressed sequence tags; **GO,** gene ontology; **TIS4,** teak inflorescence shoot stage 4.

reproductive phase on several developmental stages of the terminal shoots, to observe the transition from the vegetative to the inflorescence stage. From these observations, the suitable developmental stage of the terminal shoots can be selected for the isolation of floral genes.

Flowering can be divided into four sequential steps; the activation of flowering time genes by both environmental and endogenous signals, the activation of meristem identity genes by signals from the various flowering time pathways which specify floral identity, the activation of floral organ identity genes by the meristem identity genes in which specify floral organs and the activation of organ building genes that are involved in the formation of the four floral organs (Jack, 2004). The flowering time genes function to control the activity of the meristem identity genes that are required to promote the floral meristem fate. The meristem identity genes can be divided into two subclasses: the shoot meristem identity genes and the floral meristem identity genes. The shoot meristem identity genes specify the inflorescence shoot apical meristem as being indeterminate and nonfloral, whereas, the floral meristem identity genes specify lateral meristems to develop into flowers rather than into leaves or shoots (Jack, 2004). Molecular genetic and genomic approaches have been applied to isolate the genes involved in flowering in many plants. In this study, genes expressed in teak terminal shoots that were observed to bear floral meristem were investigated through the construction of cDNA library and the generation of Expressed Sequence Tags (ESTs).

ESTs represent a simple but powerful sequence-based tool to probe or monitor expressed genes in a particular tissue. ESTs were generated by partially sequencing randomly chosen gene transcripts from a cDNA library. These partial sequences or tags were then used to infer the putative function of the gene by searching for homology to functionally related gene products from phylogenetically distant organisms. Clones could be isolated and partially sequenced by the thousands with the improvements in DNA sequencing technology. High-throughput DNA sequencing has greatly reduced both the cost and time involved in obtaining large ESTs data sets. There are several applications of ESTs including as a tool for gene discovery, as a tool to understand gene families through homology and motifs recognition and also to provide an indication of gene expression patterns in a particular tissue. In this study, the genes expressed in the teak terminal shoots that were observed to bear floral meristem were investigated through the construction of a cDNA library and the generation of Expressed Sequence Tags (ESTs).

MATERIALS AND METHODS

Plant materials

Morphological changes of terminal shoots from vegetative to the

inflorescence stages were monitored at FRIM Research Station, Mata Ayer, Perlis, which is located at the northern part of Peninsular Malaysia. These morphological changes were classified into several stages named as TIS1-TIS4 (Teak Inflorescence Shoot Stage 1-4). Histological observations showed that floral meristems were only observed in TIS4 shoots. For cDNA library construction, the TIS4 shoots were collected from planted teak stands at the research station. To avoid RNA degradation, the collected tissues were immediately wrapped in labeled aluminium foil and dipped into liquid nitrogen.

cDNA library construction

Total RNA was isolated from TIS4 tissues using a method outlined by Schultz et al. (1994) with slight modifications. mRNA was purified from the total RNA using the µMACS mRNA Isolation Kit (Miltenyi Biotec, Germany). cDNA library was constructed using the ZAP-cDNA Synthesis Kit and ZAP-cDNA Gigapack III Gold Cloning Kit (Stratagene, USA). The primary cDNA library generated was amplified following the method described in the supplied user manual of the kit.

ESTs generation and analysis

The amplified library was plated onto LB agar (1% (w/v) of tryptone, 0.5% (w/v) of yeast extract, 1% (w/v) of NaCl, 15 g/L of bacto-agar, pH 7.5 adjusted with 5N NaOH). Plates were incubated at 37°C for 8 h. Using 1 ml pipette tips, randomly selected single plaques were cored and transferred into microfuge tubes containing 500 µl of SM buffer (50 mM Tris-HCl pH 7.5, 5.8 g/L of NaCl, 2.0 g/L of $MgSO_4$, 0.001% (w/v) of gelatin). The tubes were vortexed and incubated overnight at 4°C to let the bacteriophages to be suspended in the SM buffer.

PCR amplification of insert DNAs of all the selected plaques was carried out using the plaques suspended in SM buffer directly as template and a universal T3-T7 primer pair (T3 5' AATTAACCCTCA CTAAAGGG 3'; T7 5' TAATACGACTCACTATAGGG 3'). For each plaque, a 100 µl PCR reaction was prepared containing 20 µl of plaque suspension, 10 µl of 10X PCR Buffer with $MgCl_2$, 1 µl of 10 mM dNTP mix, 1 µl each of 25 µM T3 and T7 universal primer and 1 U of *Taq* DNA polymerase. PCR amplification was performed for 35 cycles of denaturation at 94°C for 30 s, annealing at 55°C for 30 s and extension at 72°C for 2 min. After amplification, 10 µl of the PCR reaction was analysed in 1.0% (w/v) of agarose gel. Only reactions that yielded a single amplified fragment each were purified and sequenced.

Purification of PCR amplified DNA was carried out using the GeniSpin PCR Purification Kit (BST Techlab, Malaysia). 30 ng of each purified DNA fragment were sequenced with universal SK primer (5' CGCTCTAGAACTAGTGGATC 3') using ABI Prism Big Dye Terminator Cycle Sequencing Kit Version 2.1 (PE Biosystems, USA). Purified sequencing products were analysed using ABI Prism 377 DNA Sequencer (PE Biosystems, USA). Raw sequence files were trimmed and filtered using PHRED software (Ewing et al., 1998; Ewing and Green, 1998). The quality sequences generated were then clustered and assembled using STACK_PACK software package (Hide et al., 1999) were clustered and assembled based on overlapping regions to generate longer contigs or consensus sequences. Sequences which did not overlap with any other sequences were left as singletons. These contigs and singletons made up the nonredundants, which were then searched for their identities using BLASTX (http://www.ncbi.nlm.nih.gov). All the non-redundants that have been identified were then classified based on their biological functions. Further characterization of all the non-redundants was carried out using InterProScan software (http://www.ebi.ac.uk/InterProScan). InterProScan is a tool that combines

Table 1. Classification of the 567 nonredundants based on their functional categories where proteins involved in the respective activities. No hit; no sequence similarity found among known amino acid sequences. Unknown /Unclassified Function; amino acid sequences similar to known sequences whose function is unclear.

Functional categories	Number of sequences	Percentage (%)
No hit	107	15.9
Unknown/Unclassified Function	277	41.1
Energy	38	5.6
Metabolism	36	5.3
Disease and Defense	29	4.3
Signal Transduction	18	2.7
Transcription and Post-Transcription	28	4.2
Protein Synthesis	51	7.6
Protein Destination and Storage	38	5.3
Cell Structure	12	1.8
Cell Growth/Division	0	0.0
Secondary Metabolism	4	0.6
Transporters	23	3.4
Intracellular Traffic	14	2.1
Transposons	1	0.1
Total	674	100

different protein signature recognition methods based on the InterPro member databases for sequence function and annotation (Zdobnov and Apweiler, 2001).

RESULTS

cDNA library

Total RNA was successfully extracted from teak shoots with good quality based on optical density (OD) readings using the method described by Schultz et al. (1994). From the purified mRNA, a cDNA library was constructed and amplified. The titer of the library was 7.5×10^5 pfu/ml (primary) and increased to 4.5×10^9 pfu/ml (amplified).

Processing of ESTs

EST generation and analysis were performed using the cDNA library with the insert size range of 1500 - 5000 bp. In total, 1384 plaques were randomly picked and their inserts were PCR-amplified using the T3 and T7 universal primers. Only 1125 plaques generated single amplified fragments and were purified. DNA sequencing was carried out on the 5' end of the inserts using SK primer. All the 1125 purified inserts successfully generated raw 5' ESTs which were then filtered using the PHRED software. Low quality sequences in both ends were trimmed. All the 1125 filtered and trimmed ESTs by PHRED software were analysed by STACK_PACK software for clustering. From the total 1125 ESTs, 520 of them were grouped into 69 clusters to generate 69 contigs or con-

sensus sequences. The other 605 ESTs were not clustered and remain as singletons. Therefore, in total of 674 nonredundants; 69 consensus sequences and 605 singletons, were generated.

Identity search and functional classification

All the 674 nonredundants generated were searched for their identities in several databases namely SwissProt, PDB (Protein Data Bank), PIR (Protein Information Resource), PRF (Protein Research Foundation) and nr (non-redundant GenBank CDS translations) through BLASTX. From all the 674 nonredundants, 107 of them (15.9%) showed no hits or no identity. Most of them were short sequences, of less then 50 bases, generated after the cleaning up and trimming processes by PHRED. All the remaining 567 nonredundants that found their identities through BLASTX were categorized into functional groups as shown in Table 1. Majority of the sequences either did not hit any identity (15.9%), or found their significant identities but could not be grouped into any known functional categories (41.1%). The remaining 43% of the nonredundants were categorized as follows: energy (5.6%), metabolism (5.3%), disease and defense (4.3%), signal transduction (2.7%), transcription and post-transcription (4.2%), protein synthesis (7.6%), protein destination and storage (5.3%), cell structure (1.8%), cell growth/division (0.0%), secondary metabolism (0.6%), transporters (3.4%), intracellular traffic (2.1%) and transposons (0.1%).

All the 567 nonredundants identified through BLASTX were further analyzed using InterProScan to detect their
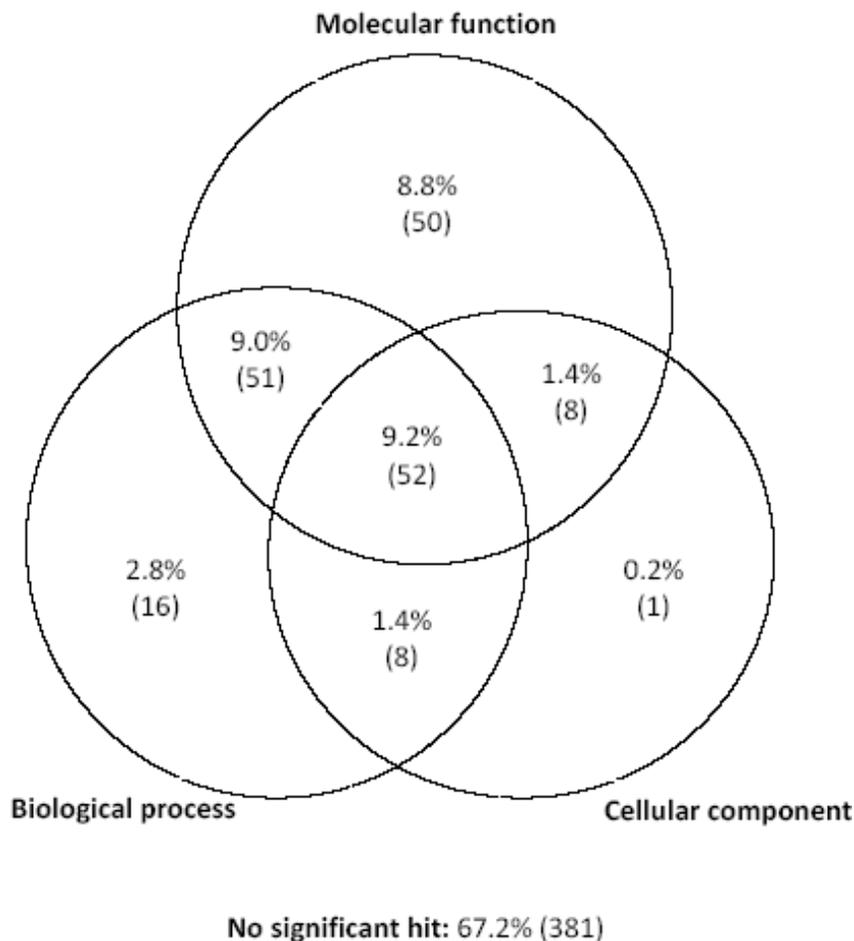
Figure 1. Diagram representing the classification of all the identified 567 nonredundants analyzed using InterProScan based on the main GO categories. No significant hit 67.2% (381); Biological process (GO:0008150) 2.8% (16); Cellular component (GO:0005575) 0.2% (1); Molecular function (GO:0003674) 8.8% (50); both the Biological process (GO:0008150) and Molecular function (GO:0003674) 9.0% (51), both the Biological process (GO:0008150) and Cellular component (GO:0005575) 1.4% (8), both the Cellular component (GO:0005575) and Molecular function (GO:0003674) 1.4% (8); all the three GO categories 9.2% (51).

protein signatures. From all the sequences analysed, only 186 sequences (32.8%) were given the gene ontology (GO) numbers and grouped into the three GO main categories, namely biological process (GO:0008150), cellular component (GO:0005575) and molecular function (GO:0003674). Figure 1 shows the classification of all the analyzed sequences based on the GO main categories. From the analysis, 67.2% (381) of the non-redundants were not grouped in any of the GO categories or have no significant hit, 2.8% (16) in the Biological process (GO:0008150), 0.2% (1) in the Cellular component (GO:0005575) and 8.8% (50) in the Molecular function (GO:0003674). There were sequences that were categorized into more than one GO categories; 9.0% (51) in both the Biological process (GO:0008150) and Molecular function (GO:0003674), 1.4% (8) in both the Biological

process (GO:0008150) and Cellular component (GO:000 5575), 1.4% (8) in both the Cellular component (GO:000 5575) and Molecular function (GO:0003674) and 9.2% (51) in all the three GO categories.

Similarity to known genes

There were five sequences found to be related to flowering and light induction as shown in Table 2. They were: (1) Clone ID = 286; 47% identity to putative negatively light-regulated protein (*Oryza sativa*); (2) Clone ID = 507; 76% identity to light-inducible protein AT-LS1 (*Arabidopsis thaliana*) (3) Clone ID = 509; 58% identity to photolyase/blue light photoreceptor PHR2 (*Arabidopsis thaliana*) (4) Clone ID = 900; 43% identity to early flower-

Table 2. Sequences related to flowering and light induction. Identity score; High, BLASTX score > 80% identity; Moderate, BLASTX score 60 - 80% identity; Low, BLASTX score 40 - 60% identity.

Clone ID	Putative identity	Identity score
286	Putative negatively light-regulated protein (*O. sativa*)	Low
507	Light-inducible protein AT-LS1 (*A. thaliana*)	Moderate
509	Photolyase/blue light photoreceptor PHR2 (*A. thaliana*)	Low
900	Early flowering 3 (*M. crystallinum*)	Low
978	Early light-induced protein (*G. max*)	Low

ing 3 (*Mesembryanthemum crystallinum*) and (5) Clone ID = 978; 49% identity to early light-induced protein (*Glycine max*). However, this identity search was based on the sequence tags generated and sequencing of the full length clones shall be carried out to confirm the identities and also to further characterize the genes. Other sequences that showed significant identities were shown in Table 3 according to their functional classes. The identity score of the highlighted sequences were divided into three classes: (1) High, BLASTX score > 80% identity; (2) Moderate, BLASTX score 60 - 80% identity and (3) Low, BLASTX score 40 - 60% identity.

DISCUSSION

ESTs generation and analysis were carried out as a preliminary approach for gene discovery from a cDNA library constructed from the floral meristem-bearing TIS4 tissue. Generally, ESTs generation permits rapid and high-volume screening of a cDNA library which could lead to the discovery of genes expressed specifically in the selected tissue. Although the number of clones sequenced in this study was small, this was the first genomic approach made towards gene discovery in teak especially in flowering related processes. To date, no publications or database of teak ESTs has been reported or submitted to the public domain. The analysis of the generated ESTs will also provide a general indication of the floral gene expression pattern by the TIS4 tissue. As a start, only 1125 fragments were sequenced in this study. Higher volume screening by sequencing more clones can be carried out in the future from the same cDNA library in the hope discovering more floral related genes.

ESTs offer a rapid and inexpensive way to gene discovery. However, because of the high-volume and high-throughput single-pass nature of generating ESTs, the data contain high error rates and low quality raw sequences. An EST also represents a small portion of an entire gene. Through single-pass DNA sequencing, normally only 300 - 500 readable bases are produced from each sequence read, but a full gene transcript may be several thousands of bases long. Reverse transcriptase also generates various lengths of cDNA from each mRNA template because they fall off the template at different places. Therefore, several ESTs could represent various parts of the same gene. In this study, redundancy in the identity search was minimized by clustering the ESTs using STACK_PACK software prior to BLASTX search. Clustering aims to place fragmented ESTs in the correct context and indexed by gene such that all fragmented data concerning a single gene is in a single index class (Hide et al., 1999). Through clustering, redundant fragments that share common sequences were grouped and longer contigs or consensus sequences were generated for better results in identity search. By sequencing more clones and adding more raw data, the clustering can be repeated to generate longer or more useful consensus sequences. Therefore, in this study we did not report on the redundancy analysis on the generated ESTs. For ESTs that were not clustered, they remained as singletons. Consensus sequences from clusters and singletons are known as nonredundants which are then searched for their identities.

Since no publication or database of teak ESTs has been reported or submitted to the public domain to date, we could not compare the expression pattern of the genes expressed in the TIS4 tissue with other tissues of teak plant. In future, ESTs could be generated from vegetative shoots, TIS1, TIS2, TIS3 and flower of teak in order to monitor and to compare the expression patterns of genes among the tissues. Generation of ESTs could also be used as a gene discovery tool from the respective tissues especially for floral genes. In addition to the classification of the generated ESTs by functional categories, further characterization of all the nonredundants was carried out using the InterProScan software where ESTs were grouped based on protein signatures that may be defined as families, domains, repeats or sites. The function of sequences within any group may be confined to a single biological process or it may have a diverse range of functions. After BLASTX analysis, only 107 sequences did not show any significant identities or no hit, whereas in the InterProScan analysis 381 sequences were not grouped into any of the GO categories. This is because InterProScan analysis and grouping were based on protein signatures (families, domains, repeats or sites) present in each submitted sequence. Although the identities of many sequences were known through BLASTX, they did not possess any domains, repeats or sites and therefore they could not be grouped based on

Table 3. Highlights of ESTs generated from TIS4 shoots of *T. grandis* based on the functional categories. Identity score; High, BLASTX score > 80% identity; Moderate, BLASTX score 60 - 80% identity; Low, BLASTX score 40 - 60% identity.

Clone ID	Putative identity	Identity score	Functional class
052	ubiquinol-cytochrome C reductase	High	Energy
209	aldehyde dehydrogenase	High	
529	cinnamyl alcohol dehydrogenase	High	
873	thioredoxin	Low	
cn10	alcohol dehydrogenase	Moderate	
cn42	cytochrome P450	Moderate	
cn25	cytochrome c oxidase	Moderate	
294	photosystem I-N subunit	Low	
834	phenylcoumaran benzylic ether reductase	High	
870	NADH dehydrogenase	High	
081	S-adenosyl-L-methionine synthetase	High	Metabolism
106	fructose-bisphosphate aldolase	High	
305	malate dehydrogenase	High	
188	UDP-galactose 4-epimerase	High	
232	long chain acyl-CoA synthetase	High	
276	stearoyl-acyl carrier protein desaturase	High	
347	laccase	Moderate	
544	cytosolic glutamine synthetase	High	
847	caffeoyl-CoA O-methyltransferase	High	
cn31	glutaredoxin	Moderate	
047	dehydrin	High	Disease and Defense
063	peroxidase	Moderate	
303	defender against death (dad) 1	High	
059	galactinol synthase	High	
391	copper-binding protein	Moderate	
cn3	metallothionein	High	
cn5	copper-zinc superoxide dismutase	HIgh	
cn33	cysteine protease inhibitor	Moderate	
449	glutathione transferase	Moderate	
940	ozone-responsive stress-related protein	Moderate	
203	protein phosphatase type 1	High	Signal Transduction
246	serine/threonine protein kinase	High	
470	GTP-binding protein	High	
cn57	protein phosphatase 2C	Moderate	
937	acid phosphatase 1	Low	
1230	protein kinase	Low	
231	zinc-finger protein	Moderate	Transcription and Post-Transcription
236	AG-motif binding protein-4	Moderate	
376	S-ribonuclease binding protein SBP1	High	
401	Mpt5 family RNA-binding protein	High	
570	DNA-binding protein	Moderate	
638	yippee-like protein	Moderate	
858	arginine/serine-rich splicing factor	High	
125	RNA binding protein 3	Moderate	
129	small nuclear ribonucleoprotein	High	
995	RNA polymerase II subunit	High	

Table 3. contd.

070	60S ribosomal protein	High	Protein Synthesis
101	40S ribosomal protein	High	
226	chloroplast 50S ribosomal protein	Moderate	
372	ribosomal protein S27	High	
393	elongation factor 1 alpha	High	
506	eukaryotic initiation factor 4A	High	
522	translation initiation factor eIF-1A	High	
445	ribosomal protein L32	Moderate	
774	glutamyl-tRNA synthetase	Moderate	
1272	elongation factor 1-beta	Moderate	
104	copper chaperone	High	Protein Destination and Storage
159	ubiquitin-conjugating enzyme	High	
314	polyubiquitin	High	
398	ubiquitin-protein ligase	Moderate	
516	heat shock protein	Moderate	
585	proteasome subunit alpha-3	High	
cn32	class II small heat shock protein	Moderate	
cn68	cyclophilin	High	
984	cysteine proteinase	Moderate	
1141	4-coumarate:coenzyme A ligase	High	
067	arabinogalactan protein	Low	Cell Structure
091	alpha-tubulin	High	
339	histone H2B	High	
403	polygalacturonase	Low	
497	actin	High	
714	actin depolymerizing factor	Moderate	
871	pectin methyl esterase	High	
417	expansin	High	
437	histone H4	High	
782	oxidase	Moderate	Secondary Metabolism
378	isoflavone reductase	Low	
086	nucleoside transporter	Moderate	Transporters
225	copper-transporting ATPase	Low	
499	tonoplast water channel	Low	
559	sucrose transporter 1	High	
657	carnitine/acylcarnitine translocase	High	
839	sulfate transporter	Moderate	
cn28	aquaporin	High	
cn34	aquaglyceroporin	Moderate	
cn43	lipid transferase	Low	
381	importin alpha 1	High	
065	RAB5B	Low	Intracellular Traffic
069	phophate translocator	High	
265	clathrin	High	
400	annexin	Moderate	
473	ADP-ribosylation factor	High	
498	vesicle-associated membrane protein	Moderate	
572	small GTP-binding protein	High	
590	coatomer protein complex	Low	
207	transposases	Moderate	Transposons

their protein families, domains, repeats or sites. They also were not given GO number and were not grouped into any of the GO categories.

After BLASTX analysis, five light induction and floral related genes were discovered. Basic information about the five sequences were obtained from the GeneBank. However, only the early flowering *3* gene from *Mesembryanthemum crystallinum* was characterized before and further information was published (Boxall et al., 2005). The other four genes sequences were directly submitted to GeneBank and no further published information about the characterization. In future, these five genes can be further characterized by sequencing the whole fragments or used them as probes in cDNA library screening. In this preliminary investigation, the discovery of five light induction and floral related genes from 1,125 sequenced clones was very encouraging, although they were not fully characterized yet.

The application of modern biotechnological techniques including molecular biology and genetic engineering is an alternative approach that is very promising and worth looking into towards solving the flowering problem in planted teak trees. The timing of the first flowering has been reported as an important factor that determines the length of the clear bole for timber (Krishnapillay, 2000). Flowering in teak is terminal, both on the main stem and the branches. Juvenile teak trees are monopodium, which grow on a single leading shoot. As a result, the main stem can be forked and therefore, limiting the length of the clear bole. Repeated terminal flowering and branching will give the trees a broad-crown appearance with shorter clear bole. Although selection for late flowering individuals as planting materials and application of proper plantation management strategies could help to delay early flowering, the introduction of genes that could turn off or delay early flowering would be highly desirable for the production of planting materials for new teak plantations. Therefore, it is important to isolate and to understand the functions and interactions of some important genes involved in the flowering processes. Generation of ESTs from selected floral tissue of teak could serve as a rapid and inexpensive tool towards the discovery of these desired floral genes.

REFERENCES

Boonkird S (1966). Flowering of teak. Nat. His. Bull. Siam Soc. 21: 69-73.

Boxall SF, Foster JM, Bohnert HJ, Cushman JC, Nimmo HG, Hartwell J (2005). Conservation and divergence of circadian clock operation in a stress-inducible crassulacean acid metabolism species reveals clock compensation against stress. Plant Physiol. 137: 969-982.

Ewing B, Hillier L, Wendl MC, Green P (1998). Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. 8: 175-185.

Ewing B, Green P (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 8: 186-194.

Gram K, Larsen CS (1958). The flowering of teak (*Tectona grandis*) in aspects of tree breeding. Nat. His. Bull. Siam Soc. 19: p. 1

Hide W, Miler R, Ptitsyn A, Kelso J, Gopallakrishnan C, Christoffels A (1999). EST clustering tutorial. Retrieved from http://www.sanbi.ac.za/ stack.

Jack T (2004). Molecular and genetic mechanisms of floral control. Plant Cell, 16: S1-S17.

Kadambi K (1972). Silviculture and management of teak. Bulletin 24, School of Forestry, Stephen F. Austin State University, Texas, p. 136.

Krishnapillay B (2000). Silviculture and management of teak plantations. Unasylva, 201(51): 14-21.

Nanda KK (1962). Some observations on growth, branching behaviour and flowering of teak (*Tectona grandis* L. f.) in relation to light. Indian Forester, 1962: 207-218.

Rao YS (1997). Keynote address of International Teak Symposium. In Proceedings of the International Teak Symposium, Kerala, India, 2-4 December 1991, eds. Chand Basha S, Mohanan C and Sankar S. pp. 1-6.

Schultz DJ, Craig R, Cox-Foster DL, Mumma RO, Medford JI (1994). RNA isolation from recalcitrant plant tissue. Plant Mol. Biol. Reporter, 12: 310-316.

Zdobnov EM, Apweiler R (2001). InterProScan-an integration platform for the signature-recognition methods in InterPro. Bioinformatics, 17(9): 847-848.