Research Journal of
Botany

Identification of Terminal Flowering1 (TFL1) Genes Associated with the Teak (*Tectona grandis*) Floral Development Regulation Using RNA-seq

[1,3]D.S. Diningrat, [1]S.M. Widiyanto, [1]A. Pancoro, [1]Iriawati, [2]D. Shim, [2]B. Panchangam, [2]N. Zembower and [2]J.E. Carlson

[1]School of Life Sciences and Technology, Institute of Technology Bandung, Bandung, 40132, Indonesia

[2]Department of Ecosystem Science and Management, Pennsylvania State University, University Park, PA 16802, United States of America

[3]Department of Biology, Faculty of Mathematics and Natural Sciences, Medan State University, Medan, 20221, Indonesia

Corresponding Authors: S.M. Widiyanto and D.S. Diningrat, School of Life Sciences and Technology, Institute of Technology Bandung, Bandung, 40132, Indonesia

ABSTRACT

Teak is woody plants; a member of the Lamiaceae family. Teak is a plant that has a very high quality timber. Teak has constraints due to low reproductive rates and slow growth of the wood after entering the reproductive phase. Teak genetic engineering efforts by delaying flowering time was facing difficulties due to the lack of information about the role of genes regulating flowering identity in teak. Teak has indeterminate inflorescence same as the model plant Arabidopsis. In Arabidopsis, the role of Terminal Flowering 1 (TFL1) gene as a member of the Floral Meristem Identity (FMI) in regulating the vegetative to generative transition is by down regulation, so that, the downstream of the FMI genes up-regulation which resulted in the development towards the formation of flowers. In teak, this mechanism is not well known. The development of NGS technology-transcriptome analysis has allowed us to identify specific interest genes from non-model plant rapidly and cheaply relative. To determine the activity of the interest genes *in silico* can be undertaken with RNA-seq and QRT-PCR analysis approaches. In this study, it is identified that, TFL1 genes in teak with NGS transcriptome analysis approach that is annotated with *S. lycopersicum*. The TFL1 genes obtained from EST teak derived from vegetative and generative shoots buds RNA. The TFL1 genes activities on the tissues are done with RNA-seq analysis approach in order to obtain Digitally Gene Expression (DGE) of TFL1. The TFL1 gene activity was then validated *in silico* by QRT-PCR analysis. The results of the analysis showed that the TFL1-14 gene activity equivalent to the TFL1 gene activity in the model plant.

Key words: Terminal flowering 1, NGS-transcriptome analysis, DEG, QRT-PCR, EST, floral meristem identity genes

INTRODUCTION

Teak normally begin flowering at the age of 6-8 years after planting but in the artificial forest is reported to have early flowering at the age of two years after planting (Norwati *et al.*, 2011; Khanduri, 2012). At the beginning of flowering teak controlled by genetic and environmental factors (Palupi *et al.*, 2010). Early flowering in the terminal causes the main axis forking in the first
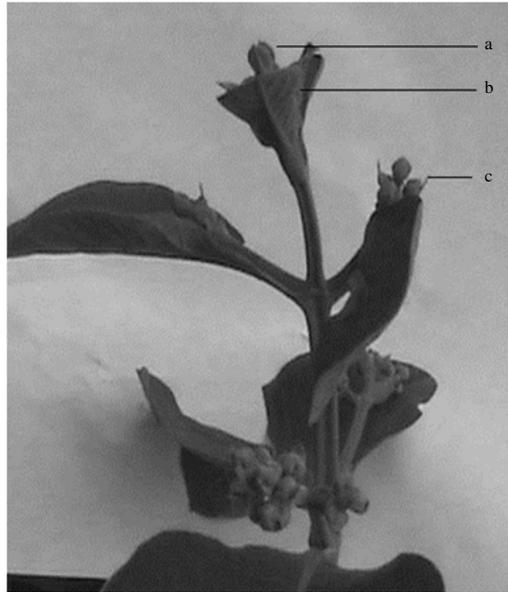
Fig. 1: Teak inflorescence, including, (a) Shoots generative apex (Generative Apical Bud/AB), (b) Lateral generative shoots (Generative Lateral Bud/LB2) and (c) Lateral generative shoots (Generative Lateral Bud/LB4)

year of flowering and forking on other shoots occur in the next flowering season (Palupi *et al.*, 2010; Norwati *et al.*, 2011; Khanduri, 2012). Figure 1 is shows the forking form of teak flowering. Forking in the major axis at the earliest stage of the inflorescence can reduce the growth of timber which can damage the quality of the wood (Widiyanto *et al.*, 2009; Palupi *et al.*, 2010; Norwati *et al.*, 2011; Khanduri, 2012).

Flowering reduces vegetative growth rate due to the utilization of energy for the flowering process (Widiyanto *et al.*, 2009). Molecular biology approaches becomes important to implement to understand the function and interaction of genes involved in the flowering process in teak (Ansari *et al.*, 2012). Understanding the role of each gene in regulating flowering in teak will facilitate the conduct of engineering to improve the quality of teak (Widiyanto *et al.*, 2009; Palupi *et al.*, 2010; Norwati *et al.*, 2011; Ansari *et al.*, 2012).

Flowering in teak can be divided into four sequential stages; (1) Activation of flowering time genes (flowering time genes) both by environmental and endogenous signals, (2) Activation of meristem identity genes (meristem identity genes) by some flowering time signals through various pathways that determine the identity of interest, (3) The identity of the gene activation floral organs (flower organ identity genes) by meristem identity genes that specify floral organs and (4) Activation of genes involved in organ builder of four floral organ formation (Rosli *et al.*, 2009). These stages are equivalent to the stages that occur in the model plant (Levy and Dean, 1998b). In the model plant, TFL1 is one of the meristem identity gene that play a role in regulating the other floral meristem identity genes (Larsson *et al.*, 1998; Levy and Dean, 1998a; Olsen *et al.*, 2002; Jack, 2004). Down regulation of TFL1 gene will result in up-regulation of downstream genes are LFY and AP1 (Olsen *et al.*, 2002; Ordidge *et al.*, 2005). Up regulation of LFY and AP1 were resulting in AGL4 induced flowering organs (Olsen *et al.*, 2002).

The question arises, whether the regulation of TFL1 in model plants also occur in teak? This research will attempt to answer that question. NGS-transcriptome analysis performed to identify genes TFL1 (Liu *et al.*, 2013; Zhang *et al.*, 2013). To further, analyze the gene TFL1 in teak by the application of phylogenetic analysis. Phylogenetic analysis was performed on all TFL1 genes that exist in teak. In addition, phylogenetic analysis was also performed on selected TFL1 based on the lowest E-value compared with TFL1 gene in the other plants that are the result of BLASTX to NCBI nr protein database. The RNA-seq analysis is performed to produce DEG of TFL1 on both tissues, the vegetative and generative tissues of teak (Feng *et al.*, 2012; Mutasa-Gottgens *et al.*, 2012). The QRT-PCR analysis was performed to validate the results of DEG, the TFL1 gene activity in vegetative and generative tissues *in silico* (Brunner *et al.*, 2004; Barakat *et al.*, 2012). In this study, it is identified that TFL1 genes in teak with NGS transcriptome analysis approach that is annotated with *S. lycopersicum*.

MATERIALS AND METHODS

Teak tissues materials and RNA isolation: Vegetative and generative shoot buds of teak were collected from a 12 year old teak plant in Institute of Technology Bandung, Indonesia for RNA isolation. The following vegetative tissues were sampled from vegetative apical shoots. Generative tissues were sampled from lateral (nodal) floral-Buds 2nd of generative stage shoots. Both of teak tissue samples were frozen in liquid nitrogen immediately upon collection and put in dry shipper for shipping from ITB-lndonesia to Pennsylvania State University (PSU)-USA. Samples were immediately frozen at -80°C upon arrival at PSU until use. Total RNA was obtained by using the method for RNA isolation protocol that developed by Chang *et al.* (1993). Frozen tissue were ground to a fine powder under liquid nitrogen and dispersed in CTAB buffer. Following 2 chloroform extractions, RNA was precipitated with $LiCI_2$, again extracted with chloroform and precipitated with ethanol. The resulting RNA pellet was resuspended in 20-100 µL of DEPC-treated water. RNA concentration analysis on a Qubit^{TM} fluorometer (www.invitrogen.com/qubit) to show a total yield of RNA sample (Barakat *et al.*, 2009). The RNA concentrations are 555 and 206 ng µL^{-1} for vegetative and generative sample, respectively. The integrity of RNA was assessed with the Agilent 6000 RNA Nano Chip Kit on 2100 Bioanalyzer (Agilent Technologies) (Barakat *et al.*, 2012).

Paired-end cDNA library preparation and MiSeq Illumina sequencing: Total RNA of teak was extracted from the two tissues using the protocol described previously. The double-stranded cDNA was synthesized using the cDNA synthesis system using random hexamer primers (illumina) according to manufacturer's instructions (Li *et al.*, 2012; Lulin *et al.*, 2012; Fu *et al.*, 2013). The paired-end library was developed according to the protocol of the paired-end sample preparation kit (lllumina, USA) (Mizrachi *et al.*, 2010; Lulin *et al.*, 2012; Liu *et al.*, 2013). The resulting library was sequenced at Penn State University using Illumina MiSeq^{TM} 2000 (Illumina Inc., USA).

Transcript assembly and annotation: The FASTQ data file of two sequence computed with CLCbio for transcript assembly strategy (Angeloni *et al.*, 2011; Annadurai *et al.*, 2013). The paired-end reads were trimmed for quality score and the presence of repeated sequences >50 bp using the modified Mott-trimming algorithm present (default parameters) in CLCbio (Fu *et al.*, 2013). We assembled *de novo* the Illumina-trimmed paired-end reads into transcript contigs using the software 'CLC Genomics Workbench' by setting minimum 95% identity, minimum 40% overlap and 200 bp as minimum contig length (Liu *et al.*, 2013). The quality of the *de novo* assembly was

assessed with a local BLASTN (e-value $<10^{-6}$) alignment of all the contigs against *S. lycopersicum* (www.phytozome.com) using CLCbio workbench (Wang *et al.*, 2010; Zhang *et al.*, 2013). After teak TFL1 sequences obtained, then phylogenetic analysis performed on the TFL1-14 sequences to determine the TFL1 gene diversity that exist in teak. Phylogenetic analysis was also conducted to determine the teak TFL1-14 position compared with TFL1 of the other plant using BLASTX analysis approach to NCBI nr protein database. Phylogenetic analysis is performed using the ClustalW2 (http://www.ebi.ac.uk/) (Larkin *et al.*, 2007).

RNA-seq analysis: Comparison of digitally gene expression (DEG-seq) between TFL1 in vegetative and generative tissues was done using RNA-seq analysis software test developed by CLCbio genomic work bench (Eveland *et al.*, 2010; Guo *et al.*, 2011; Barakat *et al.*, 2012). DEG-seq analysis was used to identify TFL1 genes in transcript abundance because it integrates several statistical methods (Feng *et al.*, 2012; Huang *et al.*, 2012). The number of reads per contig for each TFL1 gene was compared between vegetative stage and generative tissues in teak separately (Guo *et al.*, 2011; Pestana-Calsa *et al.*, 2012; Sweetman *et al.*, 2012). RNA-seq employs a random sampling model based on the read count in vegetative and generative tissues libraries and performs a hypothesis test based on that model (Mutasa-Gottgens *et al.*, 2012). Further analysis of the DEG results should be validated by QRT-PCR Jian *et al.* (2008), Barakat *et al.* (2012) and Zhang *et al.* (2013).

Validation tests of TFL1 by quantitative real-time PCR: Quantitative real-time RT-PCR (QRT-PCR) tests were conducted to determine the extent to which the number of EST reads per gene obtained by shotgun sequencing accurately reflected transcript levels in the source tissues (Brunner *et al.*, 2004; Jian *et al.*, 2008). The QRT-PCR estimates of transcript abundance were conducted on RNA from vegetative and generative bud tissues from teak (Heid *et al.*, 1996). The QRT-PCRs were prepared using the SYBR Green Master Mix kit (Applied Biosystems) and run in an Applied Biorad CFX 96 Fast Real-Time PCR system with default parameters (Livak and Schmittgen, 2001). Primers were designed using Primer3 software (Koressaar and Remm, 2007). The parameters used are the default parameters of Primer3 (Untergasser *et al.*, 2012). The parameters are set as follows: Number to return = 5, max stability = 9, max repeat mispriming = 12, pair max repeat mispriming = 24, max template mispriming = 12 and pair max template mispriming = 24. Parameters for thermodynamic also using the default parameters consisting of primer size optimum = 20 (18-27), primer tm optimum = 60 (57-63), max tm difference = 100%, primer gc minimum 20 and maksimum 80 (Http://bioinfo.ut.ee/primer3-0.4.0/input-help.htm) (Koressaar and Remm, 2007; Untergasser *et al.*, 2012). A gene encoding 18S rRNA was used as an endogenous standard to normalize template quantity.

The QRT-PCR analyses were performed to confirm the expression of TFL1 using *in silico* expression analysis (Barakat *et al.*, 2012). For each TFL1 gene, three biological replicates and three technical replicates were performed. Statistical analyses used to estimate the significance of the differences (Livak and Schmittgen, 2001; Brunner *et al.*, 2004; Barakat *et al.*, 2012).

RESULTS

NGS-transcriptome analysis of vegetative and generative teak shoots: The RNA isolation was using a modified method from Chang *et al.* (1993) performed to isolate RNA from teak tissue of vegetative and generative shoots buds. The RNA was checked for the quality using qubits and

Table 1: Summary statistics of sequencing and *de novo* assembly results

Term	Value
Input sequence	
Vegetative tissue	3,701,878
Generative tissue	3,778,316
Total bases	42,435,728
Contigs number	87,365
Minimum length of contigs	225
Maximum length of contigs	4,361
Average length of contigs	486
N75	359
N50	498
N25	805

bioanalyzer. Only RNA with best RNA Integrity Number (RIN) values further analyzed using Illumina NGS-Miseq platform (Collins *et al.*, 2008; Li *et al.*, 2012; Liu *et al.*, 2013). The Illumina Miseq sequencing platform generates 3,701,878 sequences for vegetative tissues and 3,778,316 sequences for generative tissues. These sequences were further analyzed, using CLC-bio workbench for trimming analysis to determine the quality of the sequence (Collins *et al.*, 2008; Wu *et al.*, 2010). The trimming results showed that the sequence has good quality. The following analysis also using CLC-bio workbench is *de novo* assembly (Annadurai *et al.*, 2013). The results of the *de novo* assembly are 87.365 contigs those resulted from the combination of vegetative and generative tissue sequences. Contigs quality was also tested by trimming using CLC-bio workbench (Wu *et al.*, 2010; Barakat *et al.*, 2012; Annadurai *et al.*, 2013). The trimming of the contigs result can be seen in Table 1.

 S. lycopersicum was used for contigs annotation. Results from BLASTN and annotations of teak contigs against *S. lycopersicum* cds database which produces 14 contigs hit clicking TFL1 gene. All TFL1 contigs then we call TFL1 unigene. The BLASTN results can be seen in Table 2. The TFL1 unigene produced had the different E-value and identity (%). It is decided to choose TFL1 for further analysis because it has the lowest E-value (Huang *et al.*, 2012; Barakat *et al.*, 2012). TFL1 unigenes then further analyzed by phylogenetic analysis using the ClustalW2 (http://www.ebi.ac.uk/) (Larkin *et al.*, 2007).

TFL1 in teak: The TFL1 genes hit by fourteen contigs. The range of number of hits is from 6-529 and the E-value range is 1.376E-07 up to 3.538. The greatest identities of the fourteenth TFL1 genes are entirely 100%. The range of greatest hits with a length is from 16-32. Greatest bit scores ranged 32.21-55.999 (Table 2). Teak-D-LB2_12_L001_R1_001 (paired) contig 81549, TFL1-14 chosen as gene for further analysis because it has the lowest E-value is 1.376E-07. Phylogenetic analysis results showed that there were ten groups of TFL1 in teak (Fig. 2).

 It can be classified into three major groups of genes TFL1 namely: Major group I consists of two minor groups, TFL1-1 and TFL1-7 clustered in first minor group while TFL1-4 and TFL1-10 clustered in the second minor group and TFL1-2 and TFL1-11 clustered in the third minor group. Major group II consists of four minor groups, TFL1-3 and TFL1-12 clustered in the 1st minor group. The second minor group consists of only one member i.e., TFL1-8. TFL1-5 and TFL1-14 clustered in the 3rd minor group, while TFL1-9 being the only member of the 4th minor group of major group II. Major group III consists of only a minor group consisting of TFL1-6 and TFL1-13. TFL1-14 selection based only on the lowest E-value, if we observe the phylogenetic analysis, there is no significant difference from TFL1-14 compared with other TFL1.

Table 2: Contigs related TFL1 genes results from BLASTN to the *S. lycopersicum* CDS database (www.phytozome.com)

| Query	Gene name	Number of hits	Lowest E-value	Accession (E-value)	Description (E-value)	Greatest identity (%)	Greatest hit length	Greatest bit score
Teak-D-LB2_12_L001_R1_001 (paired) contig82699	PEBP (phosphatidyl ethanolamine-binding protein) family protein (TFL1-1)	36	0.047	Solyc10g024470.1.1	PACid: 27281756	100	27	38.158
Teak-D-LB2_12_L001_R1_001 (paired) contig7980	PEBP (phosphatidyl ethanolamine-binding protein) family protein (TFL1-2)	6	3.538	Solyc05g055660.1.1	PACid: 27299647	100	16	32.210
Teak-D-LB2_12_L001_R1_001 (paired) contig12866	PEBP (phosphatidyl ethanolamine-binding protein) family protein (TFL1-3)	31	0.323	Solyc09g009560.1.1	PACid: 27310917	100	25	38.158
Teak-D-LB2_12_L001_R1_001 (paired) contig25376	PEBP (phosphatidyl ethanolamine-binding protein) family protein (TFL1-4)	24	1.929	Solyc01g009580.1.1	PACid: 27302368	100	23	32.210
Teak-D-LB2_12_L001_R1_001 (paired) contig31465	PEBP (phosphatidyl ethanolamine-binding protein) family protein (TFL1-5)	18	0.123	Solyc01g009580.1.1	PACid: 27302368	100	22	36.175
Teak-D-LB2_12_L001_R1_001 (paired) contig32476	PEBP (phosphatidyl ethanolamine-binding protein) family protein (TFL1-6)	16	0.052	Solyc09g005060.1.1	PACid: 27310154	100	23	38.158
Teak-D-LB2_12_L001_R1_001 (paired) contig39924	PEBP (phosphatidyl ethanolamine-binding protein) family protein (TFL1-7)	7	0.899	Solyc11g008660.1.1	PACid: 27295983	100	21	34.193
Teak-D-LB2_12_L001_R1_001 (paired) contig50579	PEBP (phosphatidyl ethanolamine-binding protein) family protein (TFL1-8)	44	0.119	Solyc09g005060.1.1	PACid: 27310154	100	19	36.175
Teak-D-LB2_12_L001_R1_001 (paired) contig51004	PEBP (phosphatidyl ethanolamine-binding protein) family protein (TFL1-9)	23	0.046	Solyc09g009560.1.1	PACid: 27310917	100	24	38.158
Teak-D-LB2_12_L001_R1_001 (paired) contig51065	PEBP (phosphatidyl ethanolamine-binding protein) family protein (TFL1-10)	20	0.045	Solyc01g009580.1.1	PACid: 27302368	100	22	36.175
Teak-D-LB2_12_L001_R1_001 (paired) contig51944	PEBP (phosphatidyl ethanolamine-binding protein) family protein (TFL1-11)	21	0.64	Solyc09g005060.1.1	PACid: 27310154	100	20	34.193
Teak-D-LB2_12_L001_R1_001 (paired) contig66149	PEBP (phosphatidyl ethanolamine-binding protein) family protein (TFL1-12)	39	0.539	Solyc09g005060.1.1	PACid: 27310154	100	32	55.999
Teak-D-LB2_12_L001_R1_001 (paired) contig80807	PEBP (phosphatidyl ethanolamine-binding protein) family protein (TFL1-13)	9	0.214	Solyc03g063100.1.1	PACid: 27289687	100	36	36.175
Teak-D-LB2_12_L001_R1_001 (paired) contig81549	PEBP (phosphatidyl ethanolamine-binding protein) family protein (TFL1-14)	529	1.38E-07	Solyc11g008660.1.1	PACid: 27295983	100	21	34.193
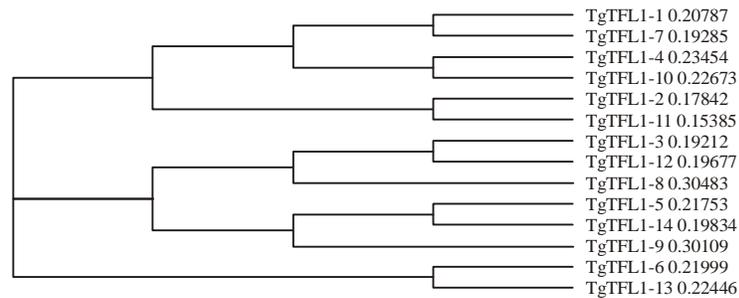
Fig. 2: Phylogram of teak TFL1 genes results from BLASTN to *S. lycopersicum* CDS database (http://www.phytozome.com)

The other phylogenetic analysis results of TFL1 gene teak against the NCBI nr protein sequence database (Fig. 3) showed that TFL1-14 in one group with TFL1-14 is one group with TFL1 of Arabidopsis and TFL1 of sunflowers. For the analysis of gene expression of TFL1 in the vegetative to generative transition of teak, we compare the results of TFL1-14 DEG with TFL1-14 QRT-PCR results. Based on these considerations, we design primer of TFL1-14 using Primer3 software (Koressaar and Remm, 2007; Untergasser *et al.*, 2012). The primer sequence of TFL1-14 is Left Primer (L): TTCTCTTTACGGGCTTCGA, Right Primer (R): CCGACGTGACAGCTTTTGT and L: AATTGTTGGTCTTCAACGAGGAA, R: AAAGGGCAGGGACGTAGTCAA for 18S. The 18S is used as a reference gene to be used for the QRT-PCR analysis.

Expression profiles TFL1-14 in the regulation of vegetative to generative transition on teak: TFL1-14 gene level expression results of the DEG and QRT-PCR analysis can be seen in Fig. 4a. These results of DEG need to be confirmed in vegetative tissues and teak generative tissue with QRT-PCR analysis (Brunner *et al.*, 2004; Jian *et al.*, 2008; Guenin *et al.*, 2009; Howe *et al.*, 2013; Barakat *et al.*, 2012). Expression profile results of QRT-PCR analysis of TFL1-14 gene in the generative and vegetative tissue of teak can be seen in Fig. 4b. Figure 4b shows that the TFL1-14 gene expression profile results of QRT-PCR equivalent to the expression profile results of DEG analysis (Guenin *et al.*, 2009). The TFL1-14 expression profiles in down regulation during the formation of floral organs (Fig. 4a).

DISCUSSION

Flower formation is a crucial stage of plant development, because it determines the maturity of the plant (Torti *et al.*, 2012; Blazquez, 2000). Flowering plants that have been successful in generating flowers indicated that the plant is ready to produce offspring (Putterill *et al.*, 2004). The next stage after the forming of flower is the formation of seeds. In teak, flower formation occurs after the age of 6-8 years (Orwa *et al.*, 2009; Ansari *et al.*, 2012). This is a long time and it is a serious concern in the development of teak. Teak is very low reproductive rate if compared to other woody plants that live in the same habitat. Low reproductive rate is also a serious concern in the development of teak (Orwa *et al.*, 2009; Lyngdoh *et al.*, 2010). In teak wood production, the reproductive stage of teak is known to inhibit the growth of wood, so that, the teak will have a long time to harvest (Widiyanto *et al.*, 2009).
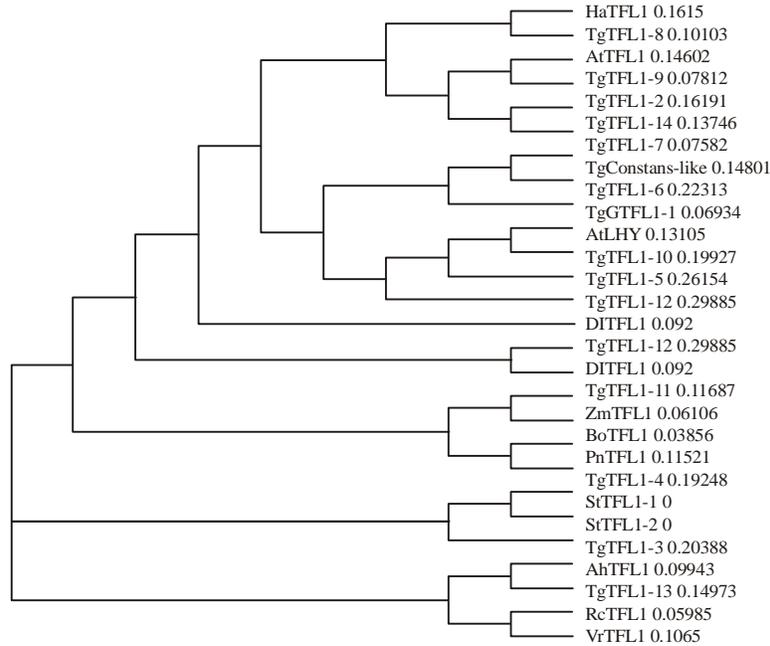
Query	Gene name	Species
AtTFL1	gi|20563264|gb|AF466812.1| *Arabidopsis thaliana* terminal flower 1 gene, complete cds	*Arabidopsis thaliana*
StTFL1-1	>gi|83583662|gb|DQ307621.1| *Solanum tuberosum* terminal flower 1 protein mRNA, complete cds	*Solanum tuberosum*
StTFL1-2	gi|568214624|ref|NM_001288549.1| *Solanum tuberosum* terminal flower 1 protein (LOC102577915), mRNA	*Solanum tuberosum*
DlTFL1	gi|635543644|gb|KJ480957.1| *Dimocarpus longan* cultivar Honghezi terminal flower 1 (TFL1) gene, complete cds	*Dimocarpus longan*
VrTFL1	gi|295148808|gb|GU947826.1| *Vitis riparia* terminal flower 1 (TFL1) gene, complete cds	*Vitis riparia*
TgCONSTANS-like	gi|662170386|gb|KF425509.1| *Tectona grandis* CONSTANS-like protein 9 mRNA, complete cds	*Tectona grandis*
ZmTFL1	gi|115498266|gb|DQ925416.1| *Zea mays* terminal flower 1 mRNA, complete cds	*Zea mays*
AhTFL1	gi|399207838|gb|JQ071508.1| *Arachis hypogaea* terminal flower 1 (TFL1) gene, complete cds	*Arachis hypogaea*
RcTFL1	gi|385866430|gb|JQ008813.1| *Rosa chinensis* cultivar Old Blush terminal flower 1 mRNA, complete cds	*Rosa chinensis*
BoTFL1	gi|335335969|gb|HM641253.1| *Bambusa oldhamii* terminal flower 1 (TFL1) mRNA, complete cds	*Bambusa oldhamii*
HaTFL1	gi|309257245|gb|GU985601.1| *Helianthus annuus* bio-material PI 578872 terminal flower 1 gene, complete cds	*Helianthus annuus*
AtLHY	gi|334182204|ref|NM_001197953.1| *Arabidopsis thaliana* protein LHY mRNA, complete cds	*Arabidopsis thaliana*

Fig. 3: Phylogram of teak TFL1-14 genes results from BLASTX to nr protein sequence database NCBI

Problems in teak flowering become important to learn because it is associated with the development of teak cultivation (Rosli *et al.*, 2009; Widiyanto *et al.*, 2009). Flowering mechanism that occurs in teak is still very limited information (Widiyanto *et al.*, 2009). In previous reports, we have learned about the role of LFY genes in regulating the transition of vegetative to generative of teak. In this report, we will be reported the role of other floral meristem identity genes, namely TFL1. The TFL1 role in regulating the teak transition of vegetative to generative will add information about the flowering mechanism of teak at the molecular level. This TFL1 expression profiles research on teak is expected to provide additional information on the mechanism of teak
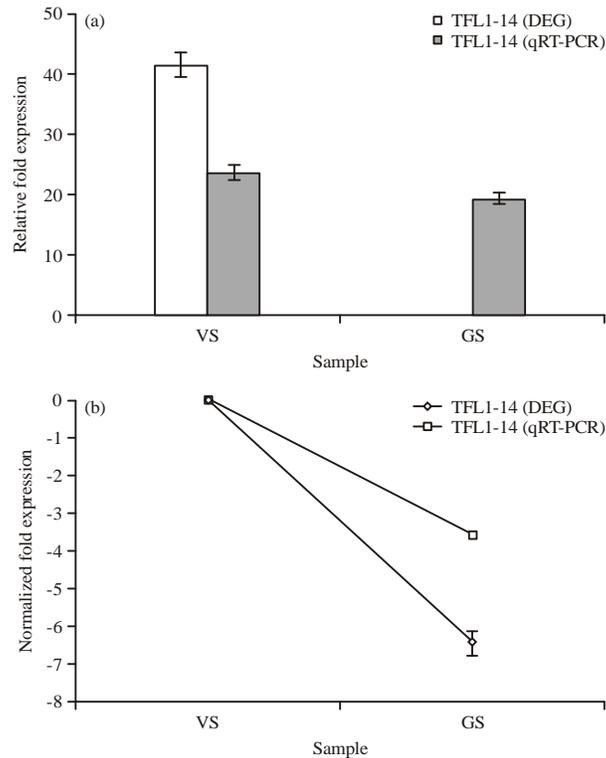
Fig. 4(a-b): Expressions (a) Level and (b) Profile TFL1-14 gene in the regulation of generative organ formation from vegetative shoots to generative shoot buds

flowering. In this study, the approach used NGS-transcriptome analysis to identify TFL1 genes in teak. The results of NGS-transcriptome analysis of the teak sequences obtained fourteen kinds TFL1 unigene which is annotated with *S. lycopersicum* (Olmstead, 2005; Lyngdoh *et al.*, 2010).

In Arabidopsis model plant, there are 14 TFL1 alleles that have been identified ((ABRC) www.arabidopsis.org) (Ordidge *et al.*, 2005). Although each allele has its own expression profile but the general profile of TFL1 expression is a gene that encodes a protein, is expressed in the cytoplasm (Ordidge *et al.*, 2005; Liu *et al.*, 2013). This gene controls the inflorescence meristem identity. This gene is involved in the initiation of flowering. These genes have an orthologous in Antirrhinum i.e., CENTRORADIALIS gene (CEN) (Jack, 2004; Putterill *et al.*, 2004). This gene is involved in protein trafficking to the protein storage in the vacuole (Olsen *et al.*, 2002). Genetic studies indicate that TFL1 acts in part by repressing the expression of LEAFY in the inflorescence strong conservation in the number, positioning and meristems (Olsen *et al.*, 2002; Ordidge *et al.*, 2005). Thus, down regulation of TFL1 leads to LFY expression and is one of the first steps in the genetic cascade that leads to flower formation (Olsen *et al.*, 2002).

In teak, flowering was also induced by environmental and endogenous factors (Rosli *et al.*, 2009; Palupi *et al.*, 2010). Both of these factors interact to induce flowering. In the model plant which is induced LFY gene as a floral meristem identity (William *et al.*, 2004; Widiyanto *et al.*, 2009). In this study we see the expression of teak TFL1-14 unigene in vegetative and generative shoot buds to induce flowering. The results of DEG and QRT-PCR gene expression analysis showed that TFL1-14 maintained in the down regulation trend in the regulation of teak floral organ formation. These

results indicate that the TFL1-14 expression profiles equivalent to the general pattern of TFL1 expression in the model plant. Based on the results we can assume that teak TFL1-14 is equivalent with model but we have more than one kind of TFL1. We need further analysis to identify other TFL1 unigene existing in the teak EST database that resulted by NGS-transcriptome analysis were performed. In order to further identify TFL1-14 genes in teak, we require advanced gene expression analysis, including *in situ* hybridization, gene over-expression and gene silencing.

However, this result is an initial study of TFL1 the other of the floral meristem identity gene expression in the teak flowering regulation. The authors hope that the results of this study may provide a basis for further research in understanding the regulatory mechanisms of vegetative to generative transition in teak.

CONCLUSION

This study shows that (1) *De novo* assembly result on the outcome of NGS-Transcriptome Analysis from teak vegetative and generative shoot buds sequence produce 87.365 contigs, (2) Identification and annotations results with *S. lycopersicum* CDS database obtained results 14 different unigene TFL1 in teak, (3) TFL1-14 has the smallest value of the E-value was analyzed further by DEG analysis and QRT-PCR analysis, (4) DEG expression profile results of TFL1-14 in equivalent with QRT-PCR results, (5) TFL1-14 has equivalent activity to the general TFL1 expression profile in the model plant and (6) Advanced research is needed to string up the understanding about the teak TFL1 gene. However, the results of this study are expected to provide the basis for research on the mechanism of flowering teak.

ACKNOWLEDGMENTS

REFERENCES

Angeloni, F., C.A.M. Wagemaker, M.S.M. Jetten, H.J.M. Op den Camp and E.M. Janssen-Megens *et al.*, 2011. *De novo* transcriptome characterization and development of genomic tools for *Scabiosa columbaria* L. using next-generation sequencing techniques. Mol. Ecol. Resour., 11: 662-674.

Annadurai, R.S., R. Neethiraj, V. Jayakumar, A.C. Damodaran and S.N. Rao *et al.*, 2013. *De novo* transcriptome assembly (NGS) of *Curcuma longa* L. rhizome reveals novel transcripts related to anticancer and antimalarial terpenoids. PloS One, Vol. 8.

Ansari, S.A., C. Narayanan, S.A. Wali, R. Kumar, N. Shukla and S.K. Rahangdale, 2012. ISSR markers for analysis of molecular diversity and genetic structure of indian teak (*Tectona grandis* L.F.) populations. Ann. For. Res., 55: 11-23.

Barakat, A., D.S. DiLoreto, Y. Zhang, C. Smith and K. Baier *et al.*, 2009. Comparison of the transcriptomes of American chestnut (*Castanea dentata*) and Chinese chestnut (*Castanea mollissima*) in response to the chestnut blight infection. BMC Plant Biol., Vol. 9. 10.1186/1471-2229-9-51

Barakat, A., M. Staton, C.H. Cheng, J. Park and N.B. Yassin *et al.*, 2012. Chestnut resistance to the blight disease: Insights from transcriptome analysis. BMC Plant Biol., Vol. 12. 10.1186/1471-2229-12-38

Blazquez, M.A., 2000. Flower development pathways. J. Cell Sci., 113: 3547-3548.

Brunner, A.M., I.A. Yakovlev and S.H. Strauss, 2004. Validating internal controls for quantitative plant gene expression studies. BMC Plant Biol., Vol. 4. 10.1186/1471-2229-4-14

Chang, S., J. Pryear and J. Cairney, 1993. A simple and efficient method for isolating RNA from pine trees. Plant Mol. Biol. Rep., 11: 113-117.

Collins, L.J., P.J. Biggs, C. Voelckel and S. Joly, 2008. An approach to transcriptome analysis of non-model organisms using short-read sequences. Genome Inform., 21: 3-14.

Eveland, A.L., N. Satoh-Nagasawa, A. Goldshmidt, S. Meyer and M. Beatty *et al.*, 2010. Digital gene expression signatures for maize development. Plant Physiol., 154: 1024-1039.

Feng, C., M. Chen, C.J. Xu, L. Bai and X.R. Yin *et al.*, 2012. Transcriptomic analysis of Chinese bayberry (*Myrica rubra*) fruit development and ripening using RNA-Seq. BMC Genomics, Vol. 13. 10.1186/1471-2164-13-19

Fu, N., Q. Wang and H.L. Shen, 2013. *De novo* assembly, gene annotation and marker development using Illumina paired-end transcriptome sequences in celery (*Apium graveolens* L.). PloS One, Vol. 8. 10.1371/journal.pone.0057686

Guenin, S., M. Mauriat, J. Pelloux, O.V. Wuytswinkel, C. Bellini and L. Gutierrez, 2009. Normalization of QRT-PCR data: The necessity of adopting a systematic, experimental conditions-specific, validation of references. J. Exp. Bot., 60: 487-493.

Guo, S., J. Liu, Y. Zheng, M. Huang and H. Zhang *et al.*, 2011. Characterization of transcriptome dynamics during watermelon fruit development: sequencing, assembly, annotation and gene expression profiles. BMC Genomics, Vol. 12. 10.1186/1471-2164-12-454

Heid, C.A., J. Stevens, K.J. Livak and P.M. Williams, 1996. Real time quantitative PCR. Genome Res., 6: 986-994.

Howe, G.T., J. Yu, B. Knaus, R. Cronn and S. Kolpak *et al.*, 2013. A SNP resource for douglas-fir: *De novo* transcriptome assembly and SNP detection and validation. BMC Genomics, Vol. 14. 10.1186/1471-2164-14-137

Huang, H.H., L.L. Xu, Z.K. Tong, E.P. Lin, Q.P. Liu, L.J. Cheng and M.Y. Zhu, 2012. *De novo* characterization of the Chinese Fir (*Cunninghamia lanceolate*) transcriptome and analysis of candidate genes involved in cellulose and lignin biosynthesis. BMC Genomics. 10.1186/1471-2164-13-648

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

Jian, B., B. Liu, Y. Bi, W. Hou, C. Wu and T. Han, 2008. Validation of internal control for gene expression study in soybean by quantitative real-time PCR. BMC Mol. Biol., Vol. 9. 10.1186/1471-2199-9-59

Khanduri, V.P., 2012. Annual variation in floral phenology and pollen production in a 25-year-old plantation of *Tectona grandis*. Nordic J. Bot., 30: 82-89.

Koressaar, T. and M. Remm, 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics, 23: 1289-1291.

Larkin, M.A., G. Blackshields, N.P. Brown, R. Chenna and P.A. McGettigan *et al.*, 2007. Clustal W and clustal X version 2.0. Bioinformatics, 23: 2947-2948.

Larsson, A.S., K. Landberg and D.R. Meeks-Wagner, 1998. The *TERMINAL FLOWER2* (*TFL2*) gene controls the reproductive transition and meristem identity in *Arabidopsis thaliana*. Genetics, 149: 597-606.

Levy, Y.Y. and C. Dean, 1998a. Control of flowering time. Curr. Opin. Plant Biol., 1: 49-54.

Levy, Y.Y. and C. Dean, 1998b. The transition to flowering. Plant Cell, 10: 1973-1989.

Li, D., Z. Deng, B. Qin, X. Liu and Z. Men, 2012. *De novo* assembly and characterization of bark transcriptome using Illumina sequencing and development of EST-SSR markers in rubber tree (*Hevea brasiliensis* Muell. Arg.). BMC Genomics, Vol. 13. 10.1186/1471-2164-13-192

Liu, T., S. Zhu, Q. Tang, P. Chen, Y. Yu and S. Tang, 2013. *De novo* assembly and characterization of transcriptome using illumina paired-end sequencing and identification of CesA gene in ramie (*Boehmeria nivea* L. Gaud). BMC Genomics, Vol. 14. 10.1186/1471-2164-14-125

Livak, K.J. and T.D. Schmittgen, 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods, 25: 402-408.

Lulin, H., Y. Xiao, S. Pei, T. Wen and H. Shangqin, 2012. The first Illumina-based *de novo* transcriptome sequencing and analysis of safflower flowers. PloS One, Vol. 7. 10.1371/journal.pone.0038653.g005

Lyngdoh, N., G. Joshi, G. Ravikanth, R.U. Shaanker and R. Vasudeva, 2010. Influence of levels of genetic diversity on fruit quality in teak (*Tectona grandis* Lf). Curr. Sci., 99: 639-644.

Mizrachi, E., C.A. Hefer, M. Ranik, F. Joubert and A.A. Myburg, 2010. *De novo* assembled expressed gene catalog of a fast-growing eucalyptus tree produced by illumina mRNA-Seq. BMC Genomics, Vol. 11. 10.1186/1471-2164-11-681

Mutasa-Gottgens, E.S., A. Joshi, H.F. Holmes, P. Hedden and B. Gottgens, 2012. A new RNAseq-based reference transcriptome for sugar beet and its application in transcriptome-scale analysis of vernalization and gibberellin responses. BMC Genomics, Vol. 13. 10.1186/1471-2164-13-99

Norwati, A., B. Norlia, H. M. Rosli, M. Norwati and R. Abdullah, 2011. Development of transgenic teak (*Tectona grandis*) expressing a *cry*1Ab gene for control of the skeletonizer. AsPac. J. Mol. Biol. Biotechnol., 19: 149-156.

Olmstead, R., 2005. A synoptical classification of the lamiales. Lamiales-Synoptical Classification, 2: 1-16.

Olsen, K.M., A. Womack, A.R. Garrett, J.I. Suddith and M.D. Purugganan, 2002. Contrasting evolutionary forces in the *Arabidopsis thaliana* floral developmental pathway. Genetics, 1: 1641-1650.

Ordidge, M., T. Chiurugwi, F. Tooke and N.H. Battey, 2005. *LEAFY*, *TERMINAL FLOWER1* and *AGAMOUS* are functionally conserved but do not regulate terminal flowering and floral determinacy in *Impatiens balsamina*. Plant J., 44: 985-1000.

Orwa, C., A. Mutua, R. Jamnades and A. Simons, 2009. *Tectona grandis*. Agrofor. Database, 4: 1-6.

Palupi, E.R., J.N. Owens, S. Sadjad, S. Sudarsono and D.D. Solihin, 2010. The importance of fruit set, fruit abortion and pollination success in fruit production of teak (*Tectona grandis*). Can. J. For. Res., 40: 2204-2214.

Pestana-Calsa, M.C., C.M. Pacheco, R.C. de Castro, R.R. de Almeida, N. Patricia, V. de Lira and T. Calsa Jr., 2012. Cell wall, lignin and fatty acid-related transcriptome in soybean: Achieving gene expression patterns for bioenergy legume. Genet. Mol. Biol., 35: 322-330.

Putterill, J., R. Laurie and R. Macknight, 2004. It's time to flower: The genetic control of flowering time. BioEssays, 26: 363-373.

Rosli, H.M., N.M. Saleh, M.M. Noor, M. Norwati, B. Norlia and A. Norwati, 2009. Analysis of expressed sequence tags derived from inflorescence shoot of *Tectona grandis* (teak). Afr. J. Biotechnol., 8: 5633-5640.

Sweetman, C., D.C. Wong, C.M. Ford and D.P. Drew, 2012. Transcriptome analysis at four developmental stages of grape berry (*Vitis vinifera* Cv. Shiraz) provides insights into regulated and coordinated gene expression. BMC Genomics, Vol. 13. 10.1186/1471-2164-13-691

Torti, S., F. Fornara, C. Vincent, F. Andres and K. Nordstrom *et al.*, 2012. Analysis of the arabidopsis shoot meristem transcriptome during floral transition identifies distinct regulatory patterns and a leucine-rich repeat protein that promotes flowering. Plant Cell Online, 24: 444-462.

Untergasser A., I. Cutcutache, T. Koressaar, J. Ye, B.C. Faircloth, M. Remm and S.G. Rozen, 2012. Primer3-new capabilities and interfaces. Nucleic Acids Res., Vol. 40. 10.1093/nar/gks596

Wang, Z., B. Fang, J. Chen, X. Zhang and Z. Luo *et al.*, 2010. De novo assembly and characterization of root transcriptome using Illumina paired-end sequencing and development of cSSR markers in sweetpotato (Ipomoea batatas). BMC Genomics, Vol. 11. 10.1186/1471-2164-11-726.

Widiyanto, S.N., A. Sukmawan, N. Haro and H. Rahmania, 2009. Transient expression of β-glucuronidase reporter gene in Agrobacterium-inoculated shoots of various teak clones. Afr. J. Biotechnol., 8: 2143-2150.

William, D.A., Y. Su, M.R. Smith, M. Lu, D.A. Baldwin and D. Wagner, 2004. Genomic identification of direct target genes of LEAFY. Proc. Natl. Acad. Sci. USA., 101: 1775-1780.

Wu, T., Z. Qin, X. Zhou, Z. Feng and Y. Du, 2010. Transcriptome profile analysis of floral sex determination in cucumber. J. Plant Physiol., 167: 905-913.

Zhang, J., K. Wu, S. Zeng, J.A.T. da Silva, X. Zhao, C.E. Tian. H. Xia and J. Duan, 2013. Transcriptome analysis of cymbidium sinense and its application to the identification of genes associated with floral development. BMC Genomics, Vol. 14. 10.1186/1471-2164-14-279