

RESEARCH ARTICLE

Open Access



# Comparative transcriptome analysis of oil palm flowers reveals an EAR-motif-containing R2R3-MYB that modulates phenylpropene biosynthesis

Ran Li<sup>1,6†</sup>, Vaishnavi Amarr Reddy<sup>1,2†</sup>, Jingjing Jin<sup>1</sup>, Chakaravarthy Rajan<sup>1,7</sup>, Qian Wang<sup>1,8</sup>, Genhua Yue<sup>1</sup>, Chin Huat Lim<sup>3</sup>, Nam-Hai Chua<sup>4</sup>, Jian Ye<sup>1,5\*</sup> and Rajani Sarojam<sup>1\*</sup>

## Abstract

**Background:** Oil palm is the most productive oil crop and the efficiency of pollination has a direct impact on the yield of oil. Pollination by wind can occur but maximal pollination is mediated by the weevil *E. kamerunicus*. These weevils complete their life cycle by feeding on male flowers. Attraction of weevils to oil palm flowers is due to the emission of methylchavicol by both male and female flowers. In search for male flowers, the weevils visit female flowers by accident due to methylchavicol fragrance and deposit pollen. Given the importance of methylchavicol emission on pollination, we performed comparative transcriptome analysis of oil palm flowers and leaves to identify candidate genes involved in methylchavicol production in flowers.

**Results:** RNA sequencing (RNA-Seq) of male open flowers, female open flowers and leaves was performed using Illumina HiSeq 2000 platform. Analysis of the transcriptome data revealed that the transcripts of methylchavicol biosynthesis genes were strongly up-regulated whereas transcripts encoding genes involved in lignin production such as, *caffeic acid O-methyltransferase (COMT)* and *Ferulate-5-hydroxylase (F5H)* were found to be suppressed in oil palm flowers. Among the transcripts encoding transcription factors, an EAR-motif-containing R2R3-MYB transcription factor (*EgMYB4*) was found to be enriched in oil palm flowers. We determined that *EgMYB4* can suppress the expression of a monolignol pathway gene, *EgCOMT*, in vivo by binding to the AC elements present in the promoter region. *EgMYB4* was further functionally characterized in sweet basil which also produces phenylpropenes like oil palm. Transgenic sweet basil plants showed significant reduction in lignin content but produced more phenylpropenes.

**Conclusions:** Our results suggest that *EgMYB4* possibly restrains lignin biosynthesis in oil palm flowers thus allowing enhanced carbon flux into the phenylpropene pathway. This study augments our understanding of the diverse roles that EAR-motif-containing MYBs play to fine tune the metabolic flux along the various branches of core phenylpropanoid pathway. This will aid in metabolic engineering of plant aromatic compounds.

**Keywords:** MYB transcription factor, Phenylpropene, Lignin, Oil palm, Basil, Phenylpropanoid pathway

\* Correspondence: jianye@im.ac.cn; rajanis@tll.org.sg

†Equal contributors

<sup>1</sup>Temasek Life Sciences Laboratory, National University of Singapore, Singapore 117604, Singapore

Full list of author information is available at the end of the article



## Background

Oil palm (*Elaeis guineensis*) is one of the most important monocot cash crop of Southeast Asia [1, 2]. Palm oil derived from its fruits is the largest source of edible vegetable oil in the world [3]. It was first introduced in Southeast Asia in 1848 and was planted on a commercial scale around 1917. The subsequent accelerated planting of palm trees and expansion of oil palm plantations boosted the economies of the developing Southeast Asian countries [4]. Oil palm is monoecious, producing male and female flowers on the same tree but at different times. To produce fruits, the plant needs to attract efficient pollinators to its flowers. *Elaeidobius kamerunicus*, a type of weevil which originated from West Africa, is considered the most competent and dominant insect pollinator species of oil palm. Introduction of this weevil into Malaysia during the 1980s enhanced the pollination rate of oil palm trees leading to 20–30% increase in fruit production [5]. To attract the weevil for pollination, both male and female flowers release a volatile phenylpropene compound called methylchavicol (also known as estragole) [6].

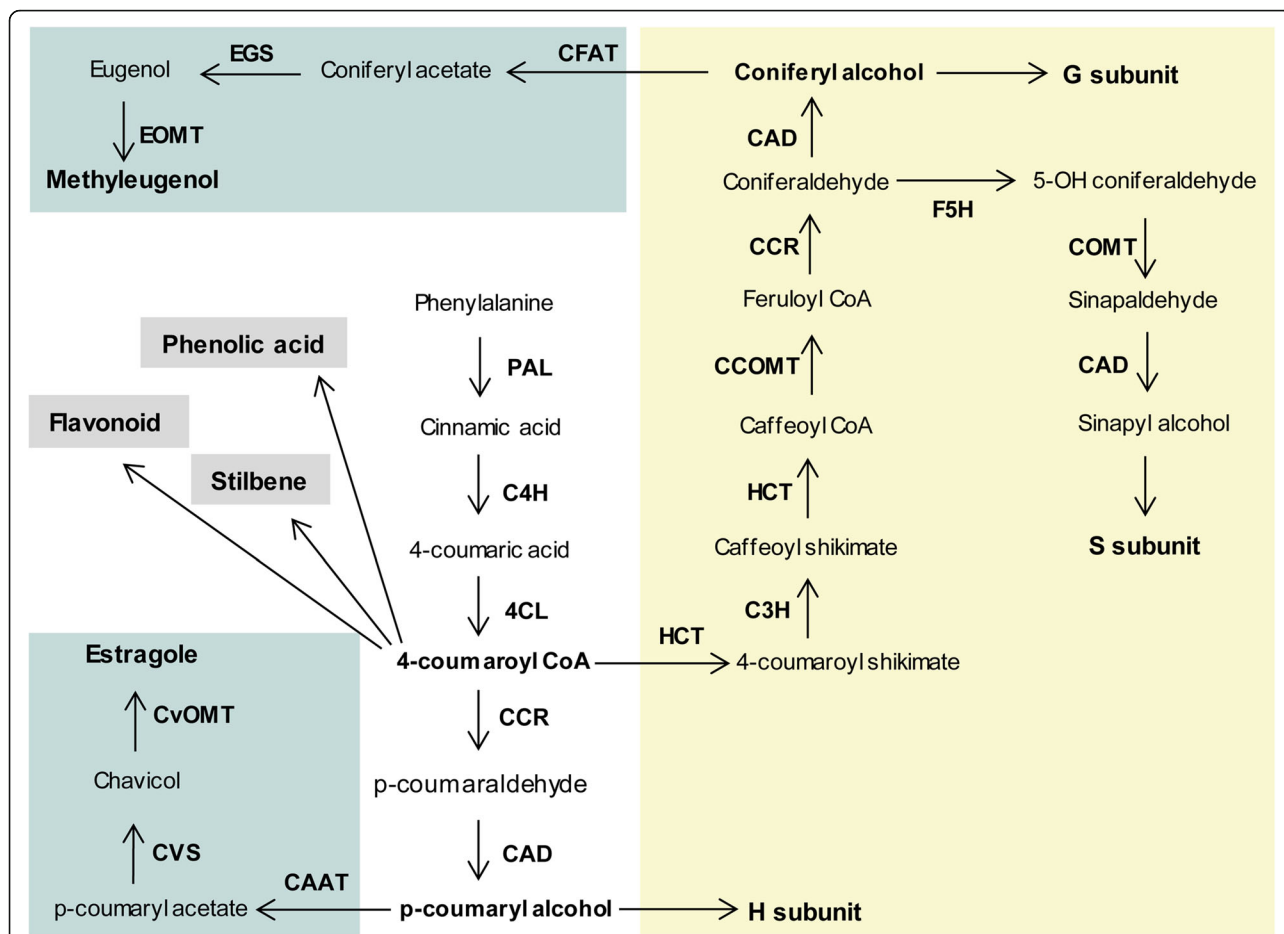
Phenylpropenes (C6-C3 carbon skeleton) are a class of volatile organic compounds (VOCs) produced by plants that serve as pollinator attractors and aid in pathogen defense [7, 8]. Commonly produced phenylpropenes include chavicol and eugenol, and their derivatives methylchavicol, methyleugenol and isoeugenol. Phenylpropenes are produced by the general phenylpropanoid pathway in plants [9]. This pathway is also responsible for the production of lignins, flavonoids, phenolic acids and stilbenes ([10]; Fig. 1). Among these phenylpropanoid derived metabolites, lignin is vital for plant growth and development. It enhances cell wall strength and is required for water transport and mechanical strength [11]. Lignin is a polymer derived from three monolignols, p-coumaryl, coniferyl and sinapyl alcohol. These three precursors produce the hydroxyphenyl (H), guaiacyl (G) and syringyl (S) units of lignin respectively (Fig. 1; [12]). Phenylpropenes are also produced from monolignols; for instance, chavicol is made from p-coumaryl alcohol, whereas, eugenol is derived from coniferyl alcohol (Fig. 1; [13]). Hence, plant tissues producing phenylpropenes, such as flowers, need to prudently regulate carbon flux more into scent production than towards lignin formation.

Lignin biosynthesis is extensively controlled by R2R3-MYB transcription factors. AC elements, which serve as the binding sites for R2R3-MYBs, have been found to be enriched in the promoter regions of many lignin biosynthesis genes [14–16]. Several R2R3-MYBs function as activators of lignin synthesis. However, members of the R2R3-MYB subgroup 4 that contain an EAR motif act as transcriptional repressors and some have been identified

as negative regulators of lignin production [10, 17, 18]. Perturbations in the expression levels of many of these repressors not only affected the final lignin content but also changed the flux along the various phenylpropanoid pathways. The first R2R3-MYBs characterized to down regulate lignification were *AmMYB308* and *AmMYB330* from *Antirrhinum majus*. Overexpression of *AmMYB308* or *AmMYB330* in tobacco repressed both lignin and phenolic acid metabolism [19]. From Arabidopsis, *AtMYB4* and *AtMYB32* were identified as repressors of lignin biosynthesis genes [20, 21]. The main target of *AtMYB4* was shown to be *cinnamate 4-hydroxylase* (*C4H*), a core enzyme of the phenylpropanoid pathway, and *AtMYB4* mutants showed enhanced accumulation of sinapate esters. *AtMYB32* has been suggested to be a repressor of Arabidopsis *caffeic acid O-methyltransferase* (*COMT*) gene. *Eucalyptus gunnii* *EgMYB1* affected lignin formation and repressed the expression of two key lignin genes *EgCCR* (cinnamoyl-CoA reductase) and *EgCAD2* (cinnamyl alcohol dehydrogenase) [22]. Further, heterologous expression of chrysanthemum *CmMYB1*, reduced lignin and flavonoid production in Arabidopsis [23].

From monocots, a few subgroup 4 R2R3-MYBs have been characterized as repressors of lignin genes. Maize *ZmMYB31* and *ZmMYB42* downregulate both maize and Arabidopsis *COMT* genes [24]. Ectopic expression of maize *ZmMYB42* in Arabidopsis decreased lignin content and suppressed flavonoid production as well [25]. But ectopic expression of maize *ZmMYB31* in Arabidopsis led to decreased lignin content and increased anthocyanin levels due to the redirection of carbon flux towards anthocyanin pathway. Apart from the *COMT* promoter, it was also shown that *ZmMYB31* could bind to the promoter of maize *Ferulate-5-hydroxylase* (*F5H*) gene in vivo [26]. Recently it was found that *COMT* is a common target of MYB31 and MYB42 in the mature leaves of maize, sorghum and rice. MYB31 and MYB42 were shown to target other genes involved in lignin pathway as well but in a more species-specific manner [27]. *ZmMYB11* from maize was identified based on the sequence similarity with *ZmMYB31* and *ZmMYB42*, and shown to regulate lignin metabolism by binding to *COMT* promoter [18]. A switchgrass (*Panicum virgatum*) ortholog of *AtMYB4*, *PvMYB4* was shown to negatively regulate lignin formation in both tobacco and switchgrass [16]. Recently from banana, *MusaMYB31* was identified as a repressor of lignin and polyphenols. It was able to down regulate many genes involved in lignin and general phenylpropanoid pathway [28].

In comparison to lignin biosynthesis, very little is known about the regulation of volatile phenylpropenes in plants. Only few TFs have been reported to regulate the production of phenylpropenes. In *Petunia*, R2R3-MYBs, *ODORANT1* (*ODO1*), *EMISSION OF BENZENOIDS II* (*EOBII*)



**Fig. 1** The phenylpropanoid pathway in plant. CAAT, coumaryl alcohol acetyl transferase; CAD, cinnamyl alcohol dehydrogenase; CFAT, coniferyl alcohol acetyl transferase; CCOMT, caffeoyl-CoA O-methyltransferase; CCR, cinnamoyl-CoA reductase; C3H, p-coumaroylshikimate 3'-hydroxylase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumaroyl CoAligase; COMT, caffeic acid O-methyltransferase; CvOMT, chavicol O-methyltransferase; CVS, chavicol synthase; EGS, eugenol synthase; EOMT, eugenol O-methyltransferase; F5H, ferulate-5-hydroxylase; HCT, hydroxycinnamoyl-CoA shikimate/ quinate hydroxycinnamoyl transferase; PAL, phenylalanine ammonia lyase

and *PhMYB4*, were identified as regulators of volatile benzenoid/phenylpropanoid compounds in petals [29–33]. Both *PhODO1* and *PhEOB1* function as positive regulators of various shikimate and phenylpropanoid pathway genes while *PhMYB4* acts as a repressor of *C4H* gene and indirectly affects the formation of petunia floral volatiles. *Production of Anthocyanin Pigment1 (PAP1)* MYB transcription factor from Arabidopsis when ectopically expressed in *Petunia* or rose increased the production of phenylpropenes [34, 35]. Recently, *FaEOB1* was identified to control the production of eugenol in ripe strawberry fruits. In *FaEOB1*-silenced plants, the expression of *CAD* and *Eugenol synthase 2 (FaEGS2)* was down-regulated [36]. However, regulation of biosynthesis of volatile phenylpropenes in monocot plants have not been investigated.

In this study, we performed transcriptome sequencing of oil palm leaves, male flowers and female flowers and did a comparative study to identify the genes involved in

methylchavicol production. From the differentially expressed TFs, we identified an EAR-motif-containing R2R3-MYB gene, *EgMYB4 (Elaeis guineensis MYB4)*, which was highly enriched in open flowers. Expression of *EgMYB4* coincided with the spatial and developmental production of methylchavicol. Promoter studies showed that *EgMYB4* was able to bind to the promoter of oil palm *EgCOMT (caffeic acid O-methyltransferase)* and suppress its expression. To further characterize the function of *EgMYB4*, it was ectopically expressed in sweet basil (*Ocimum basilicum*) plants which produce high quantities of phenylpropenes similar to oil palm. Overexpression of *EgMYB4* decreased lignin content and enhanced the production of phenylpropenes in transgenic sweet basil plants. This suggests that, *EgMYB4* presumably suppresses lignin production in oil palm flowers and redirects the carbon flux to phenylpropene production to promote successful pollination.

## Results

### Methylchavicol is produced at open flower stage of oil palm

In oil palm plantations, open flowers give a strong odor like aniseed which is attributed to the emission of methylchavicol [37, 38]. Analysis of male flower, female flower and leaf samples by GC-MS confirmed that the strong odor from open flowers was due to methylchavicol emission (Additional file 1). In leaves, only one green leaf volatile, 2-hexenal, was detected. To investigate methylchavicol emission during different stages of flower development, volatile compounds from three developmental stages of both male and female flowers were extracted and analyzed by GC-MS. The developmental stages analyzed were; before anthesis stage (closed bud), at anthesis stage (newly opened) and after anthesis stage (old flowers). Our results showed that methylchavicol is strongly emitted in male and female flowers at the open flower stage (Fig. 2a and b).

### Sequencing, de novo assembly and annotation of transcriptome

To elucidate methylchavicol biosynthesis in oil palm, we performed RNA-sequencing (RNA-Seq) of male open flowers, female open flowers and leaves using Illumina HiSeq 2000. More than 22 million high quality reads of 101 base pairs from different tissues were generated. The quality of Illumina sequencing outputs was high, as evaluated by FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) (Additional file 2). Using the Trinity method (default parameter) [39], the sequence reads were finally assembled into 59,078 non-redundant unigenes with 96,062 isoforms (with N50 = 1884 bp and mean size = 1104 bp). All unigenes were longer than 200 bp. The unigenes were annotated by performing a BLASTX search against various protein databases including TAIR and UniProt (e-value bigger than 1e-3). Functions were assigned by top hit candidates. Expression levels for assembled transcripts were calculated by mapping the reads onto them using Bowtie [40]. Later, by using RSEM (RNA-Seq by Expectation-Maximization), the abundance of assembled transcripts was estimated and the expression levels were measured [41].

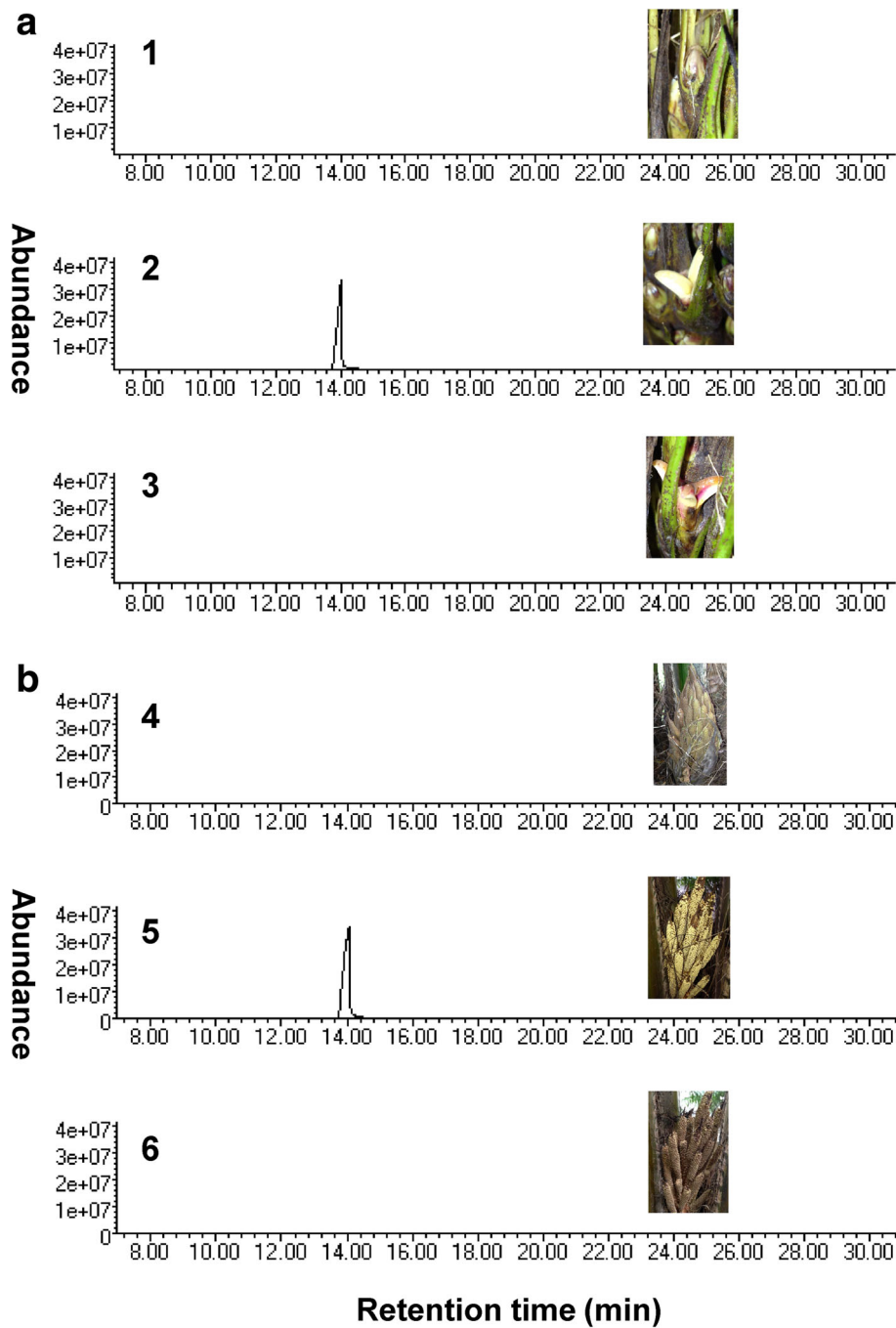
### EgMYB4 is an EAR-motif-containing R2R3-MYB TF that is enriched in flowers

Transcription factors (TFs) play an important role in regulating plant secondary metabolism [10, 42]. Seventy-nine upregulated and 219 downregulated TF transcripts were found in male and female open flowers when compared to leaves (Additional file 3A). The top 10 differentially expressed TFs between leaves and flowers are listed in Additional file 3B. Among them, the four MADS box TFs have been proposed to be involved in flower development [43]. The MYB-like TF was shown to be involved in

regulating flowering time [44]. Out of the four MYB TFs identified, two were similar to *Arabidopsis MYB21* and one each to *Arabidopsis MYB108* and *MYB4* respectively. Studies on *AtMYB21* showed that it functions in floral organ development particularly stamen development [45, 46], and *MYB108* was shown to regulate jasmonate-mediated stamen development [47]. The other differentially expressed MYB TF identified was *AtMYB4* which is implicated in regulating lignin pathway [20]. Analysis of *AtMYB4* and other R2R3-MYB belonging to subgroup 4 revealed that perturbations in lignin production was able to affect flux along the various phenylpropanoid pathways altering the formation of secondary metabolites [20, 26]. Since the production of both lignin and volatile phenylpropenes share common precursors of phenylpropanoid pathway, we decided to focus on this TF. Full length coding sequence of this MYB was obtained by reverse transcription PCR and named as *EgMYB4*. Amino acid sequence analysis indicated that it encodes a typical R2R3 domain and has an EAR motif (C2 motif) at the C-terminal similar to other members of R2R3-MYB subgroup 4 (Fig. 3a). Phylogenetic analysis based on amino acid sequences with other characterized R2R3-MYBs of subgroup 4 showed that *EgMYB4* is closely related to *Eucalyptus gunnii MYB1* (Fig. 3b). Expression profile analysis revealed high expression of *EgMYB4* in oil palm female and male open flowers when compared to leaves (Fig. 3c). Among the three developmental stages of flowers analyzed, *EgMYB4* expression increased in open flowers as compared to before anthesis stage flowers correlating with the emission of methylchavicol. In female flowers, the expression of *EgMYB4* decreased in after anthesis flowers while in male flowers it increased (Fig. 3c). To investigate the sub-cellular localization of *EgMYB4*, 35S: *EgMYB4-YFP* was transiently expressed in *N. benthamiana* leaves. Fluorescence analysis showed that *EgMYB4* exclusively localize in the nuclear bodies as confirmed by 4, 6-diamidino-2-phenylindole (DAPI) staining (Fig. 3d).

### Phenylpropene synthases are enriched while COMT and F5H transcripts are repressed in oil palm flowers

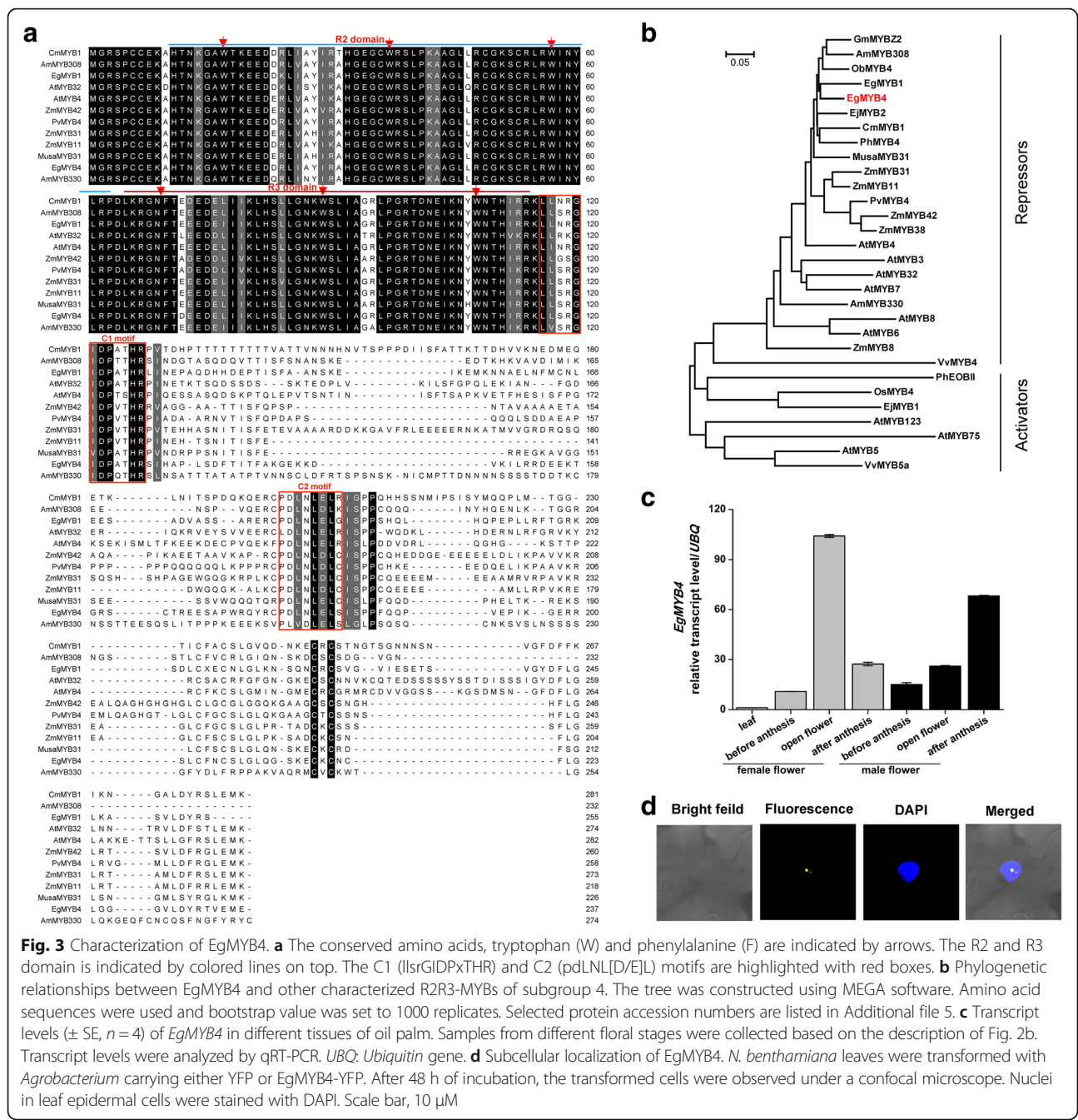
Among all the metabolic derivatives of phenylpropanoid pathway, lignin and phenylpropene biosynthesis share the most common precursors. To investigate the biosynthesis of lignin and methylchavicol in flowers, we first analyzed the expression of shared upstream biosynthesis genes from the RNA-Seq data. Many phenylpropanoid-pathway genes like *Phenylalanine ammonia-lyase (PAL)*, *4-coumarate:CoA ligase (4CL)*, *C4H*, *Cinnamoyl-CoA reductase (CCR1)* and *CAD* exist as a multigene family in various plant species. Studies have shown that they can have distinct or overlapping functions in phenylpropanoid metabolism depending on their substrate specificity



**Fig. 2** Analysis of methylchavicol in different flower stages. **a** GC-MS analysis of methylchavicol emission in female flower. 1, before anthesis; 2, open flower; 3, after anthesis. **b** GC-MS analysis of methylchavicol emission in male flower. 4, before anthesis; 5, open flower; 6, after anthesis

[48]. We indeed identified more than one transcript encoding these genes from our RNA-Seq data of leaves and flowers. Few transcripts showed constitutive expression while few other transcripts had tissue specific expression (Additional file 4). These genes can act in different branches of phenylpropanoid biosynthesis pathway, hence their function in the pathway cannot be

confirmed without the support of experimental data. To determine a clear expression pattern of the active metabolite pathway in flowers, genes specifically involved in latter steps of lignin and methylchavicol synthesis were investigated. Most differentially expressed transcripts between leaves and flowers were the genes encoding enzymes involved in the final two steps of methylchavicol



**Fig. 3** Characterization of EgMYB4. **a** The conserved amino acids, tryptophan (W) and phenylalanine (F) are indicated by arrows. The R2 and R3 domain is indicated by colored lines on top. The C1 (IlrGIDPxTHR) and C2 (pdLNL[D/E]L) motifs are highlighted with red boxes. **b** Phylogenetic relationships between EgMYB4 and other characterized R2R3-MYBs of subgroup 4. The tree was constructed using MEGA software. Amino acid sequences were used and bootstrap value was set to 1000 replicates. Selected protein accession numbers are listed in Additional file 5. **c** Transcript levels ( $\pm$  SE,  $n = 4$ ) of EgMYB4 in different tissues of oil palm. Samples from different floral stages were collected based on the description of Fig. 2b. Transcript levels were analyzed by qRT-PCR. UBQ: Ubiquitin gene. **d** Subcellular localization of EgMYB4. *N. benthamiana* leaves were transformed with *Agrobacterium* carrying either YFP or EgMYB4-YFP. After 48 h of incubation, the transformed cells were observed under a confocal microscope. Nuclei in leaf epidermal cells were stained with DAPI. Scale bar, 10  $\mu$ m

biosynthesis; chavicol synthase (CVS) and chavicol O-methyltransferase (CvOMT). They were significantly up-regulated in both male and female open flowers (Fig. 4b–f). CVS belongs to PIP family of NADPH-dependent reductases which also include eugenol synthases (EGS) and isoeugenol synthases (IGS). EGS and IGS have been identified and characterized from basil, *Petunia*, *Clarkia*, anise and recently from strawberry [7, 49–52]. A chavicol synthase (LtCES1) was also characterized from *Larrea tridentata* that can act on both coniferyl acetate and coumaryl acetate

substrates to form eugenol and chavicol, respectively [53]. The oil palm EgCVS showed high sequence similarity to these identified enzymes (Fig. 4c). Similarly, phenylpropene O-methyltransferases have been characterized from sweet basil and apple which are required for generating methylated phenylpropenes like, methyleugenol and methylchavicol [54, 55]. Two transcripts similar to chavicol O-methyltransferases (EgCvOMT1 and EgCvOMT2) were found in our RNA-Seq data, which shared high amino acid identity with sweet basil and apple CvOMTs (Fig. 4a). On the

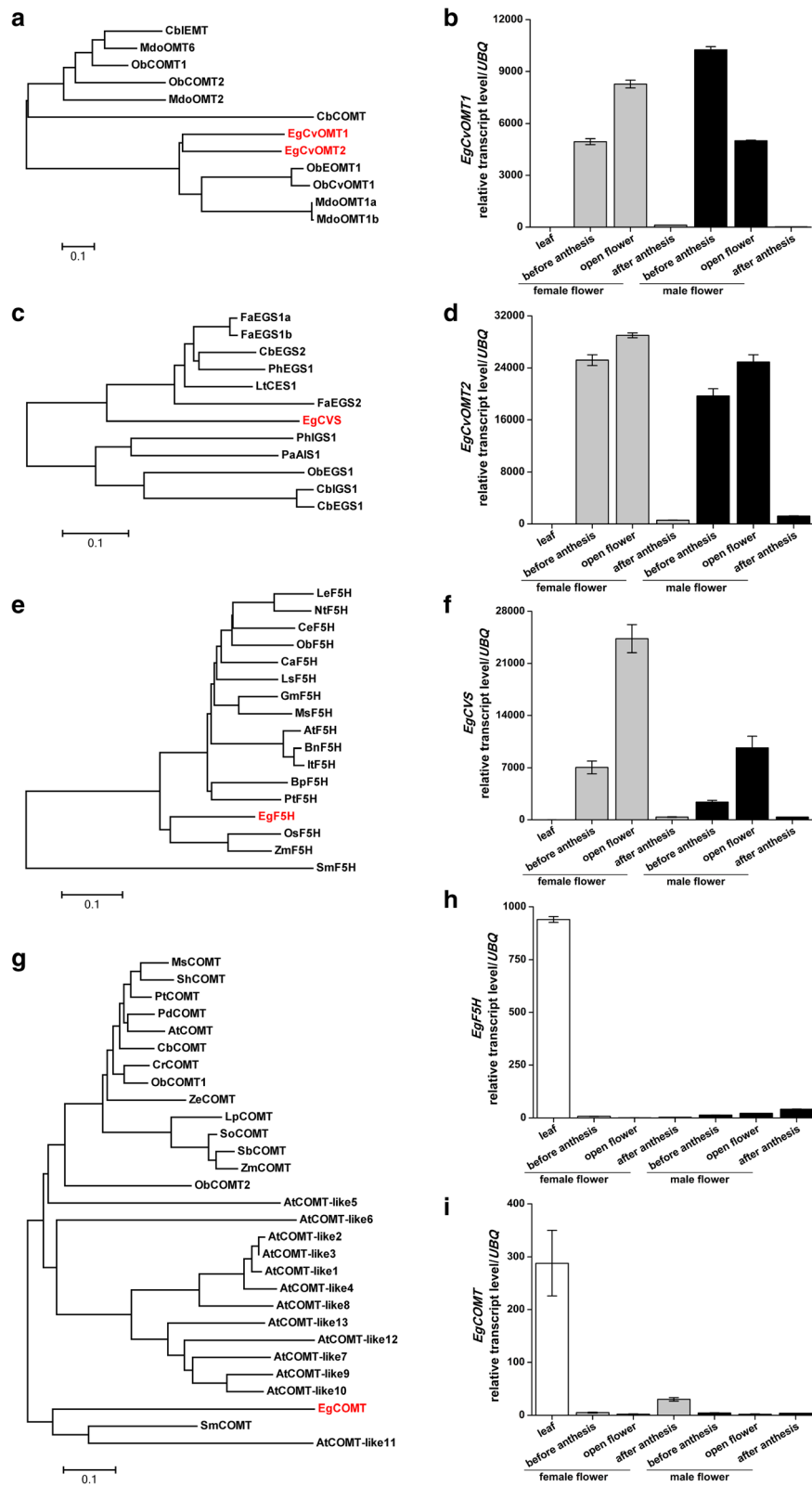


Fig. 4 (See legend on next page.)

(See figure on previous page.)

**Fig. 4** Expression of phenylpropanoid pathway genes in oil palm flower. Transcript levels ( $\pm$  SE,  $n = 4$ ) of *EgCvOMT1* (**b**), *EgCvOMT2* (**d**), *EgCVS* (**f**), *EgF5H* (**h**) and *EgCOMT* (**i**) in different tissues of oil palm. Samples from different floral stages were collected based on the description of Fig. 2b. Transcript levels were analyzed by qRT-PCR. *UBQ*: *Ubiquitin* gene. **a** Phylogenetic relationships between *EgCvOMT* and O-methyltransferases from other plants. **c** Phylogenetic relationships between *EgCVS* and other NADPH dependent PIP reductases from plants. **e** Phylogenetic relationships between *EgF5H* and P450s from other plants. **g** Phylogenetic relationships between *EgCOMT* and O-methyltransferases from other plants. Neighbor-joining phylogenetic trees were constructed using MEGA5.1 based on multiple protein sequence alignments made with ClustalX

other hand, transcripts encoding genes involved in lignin biosynthesis genes such as *caffeic acid O-methyltransferase* (*COMT*) and *Ferulate-5-hydroxylase* (*F5H*) were highly down-regulated in oil palm male and female open flowers (Fig. 4h and i). *COMT* and *F5H* are key genes mainly committed to the formation of S-lignin subunit [12]. *EgCOMT* and *EgF5H* showed high similarity to the well-characterized genes from other species (Fig. 4e and g). Protein accession numbers and nucleotide sequences of the genes in Fig. 4 are listed out in Additional files 5 and 6 respectively.

#### EgMYB4 interacts with *EgCOMT* promoter

AC elements are present in the promoters of many lignin biosynthesis genes and are known to be bound by MYBs which regulates their expression. Additionally, *COMT* and *F5H* genes are known to be down-regulated by EAR motif-containing MYBs in monocots. Maize *ZmMYB11* was found to bind to the promoter of *ZmCOMT* in vivo; *ZmMYB31* and *ZmMYB42* was found to bind to the promoters of both *ZmCOMT* and *ZmF5H* genes and repress their expression [18, 26]. To investigate if *EgF5H* and *EgCOMT* are regulated by *EgMYB4*, we cloned and screened a 2 kb-promoter region of both genes. One AC-IV element (ACCAAAC) was found in both the promoters (Fig. 5a; [16]). Within 0.25 kb of promoter region of *EgCOMT*, we identified two new AC elements (AACAAACC) and named it as AC-V. To test the interaction of *EgMYB4* to these AC elements, an electrophoretic mobility shift assay (EMSA) was performed using the purified recombinant His-*EgMYB4* protein. His-*EgMYB4* bound to both AC-IV and AC-V elements but failed to interact with mutated versions of AC-IV (GAAGGGA) and AC-V (GGAGGAT) (Fig. 5b). DNA binding specificity was further confirmed by a competition experiment using 250-fold excess unlabeled cold probe which led to the disappearance of labeled DNA/protein complex.

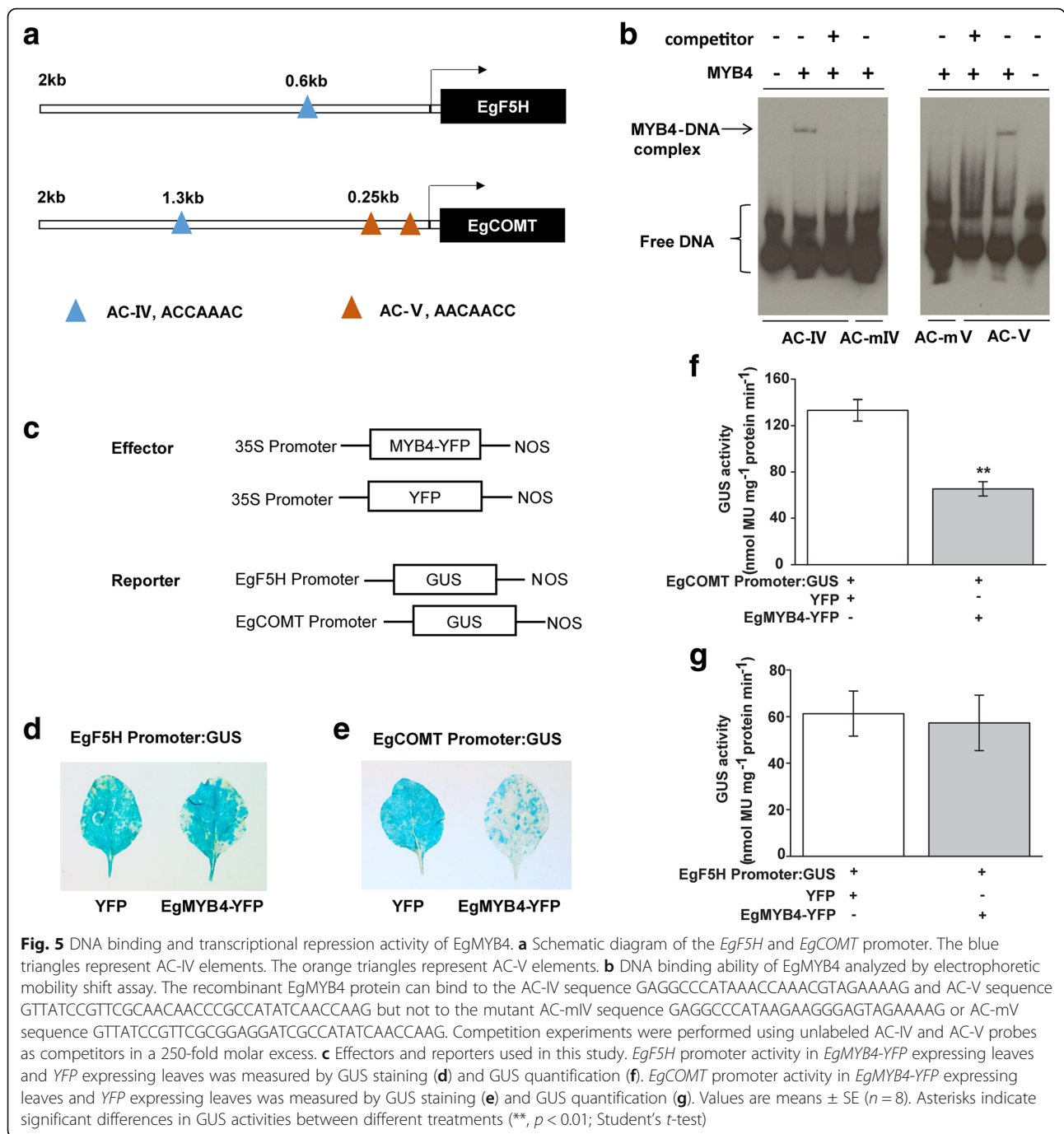
Further, the transcriptional repression activity of *EgMYB4* was determined by using *N. benthamiana* as a transient expression system [56]. Based on the distribution of AC elements, 1 kb promoter region of *EgF5H* and 2 kb promoter region of *EgCOMT* was introduced into pCAMBIA1391 vector. *EgF5H* promoter: *GUS* or *EgCOMT* promoter: *GUS* were used as reporters and 35S:*EgMYB4-YFP* was used as an effector (Fig. 5c). The

promoter activity of *EgF5H* had no difference between *EgMYB4-YFP* and only *YFP* expressing leaves (Fig. 5d and g). However, the promoter activity of *EgCOMT* was significantly suppressed in *EgMYB4-YFP* expressing leaves when compared to leaves expressing only *YFP* (Fig. 5e and f). These results indicate that *EgCOMT* is repressed by *EgMYB4*. Suppression of *COMT* activity is known to reduce lignin formation [57–60]. From the above results, we can postulate that increased expression of *EgMYB4* in oil palm flowers reduces lignin synthesis by repressing *EgCOMT*, which might indirectly affect the metabolic flux into methylchavicol synthesis.

#### Ectopic expression of *EgMYB4* in sweet basil decreases total lignin content and increases phenylpropene production

As oil palm is a non-model plant and transformation studies are not feasible, we decided to characterize *EgMYB4* in sweet basil plants. Similar to oil palm flowers, sweet basil also produces phenylpropenes in leaf glandular trichomes which form the main components of sweet basil essential oil (Additional file 7; [9]). To validate our hypothesis that *EgMYB4* mediated lignin suppression in oil palm flowers affects flux into phenylpropene synthesis, *EgMYB4* was ectopically expressed in sweet basil plants. The sweet basil variety used in our study produces two types of phenylpropenes, majorly eugenol/methyleugenol and small amounts of methylchavicol. Five independent transgenic lines were initially selected and of them two lines, MYB4–2 and MYB4–5, were advanced for further characterization. Both of these lines showed high expression of *EgMYB4* transgene (Fig. 6a). Plants overexpressing *EgMYB4* showed many phenotypic alterations. These plants were dwarfed, exhibited weak stem, had reduced leaf size and delayed flowering (Fig. 6b and c). These phenotypes are characteristic of lignin-deficient mutants [26, 61]. In plants, the cell walls of sclerenchyma and vascular tissues are lignified providing strength and rigidity. Transverse sections of mature stems from transgenic and wild type (WT) sweet basil plants were analyzed by toluidine blue and phloroglucinol staining to look for possible changes in these lignified tissues. There was considerable amount of reduction in the number of lignified cell layers forming the sclerenchyma and vascular tissues in the transgenic plants (Fig. 7a-d). The amount of total lignin in mature stems was measured by thioglycolic acid method and was found to be reduced by 34.2% and 31.4%

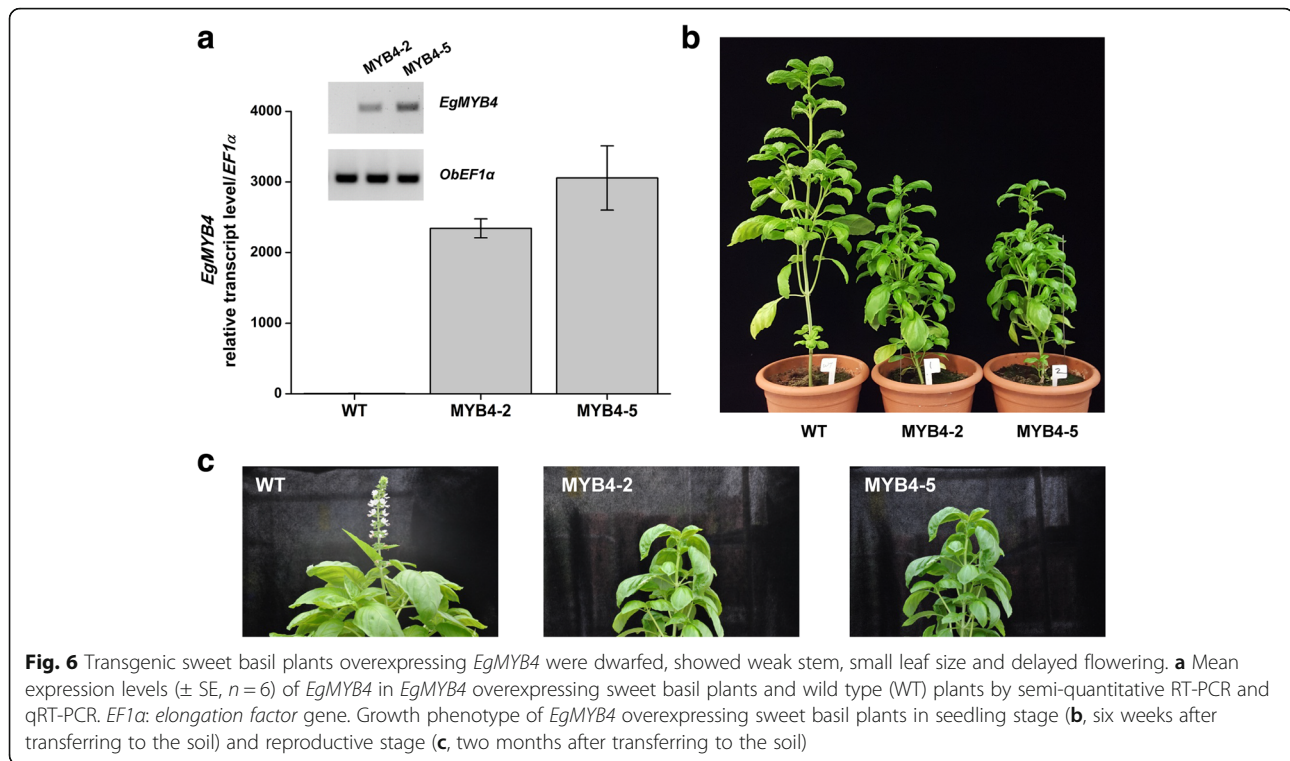




in *EgMYB4* overexpressing lines MYB4-2 and MYB4-5 respectively when compared to WT plants (Fig. 7e). These results indicate that overexpression of *EgMYB4* in sweet basil could reduce lignin biosynthesis in sweet basil plants.

We have performed RNA sequencing of sweet basil leaf glandular trichomes from which the orthologs of *F5H* and *COMT* genes were identified. Expression levels of these genes were measured in *EgMYB4* overexpressing and wild type sweet basil plants. The expression levels of *ObCOMT1*, *ObCOMT2* and *ObF5H* were significantly

decreased in the transgenic lines when compared to WT plants (Fig. 7f-h). These results suggest that ectopic expression of *EgMYB4* down regulates the expression of both *ObCOMT* and *ObF5H* genes thereby affecting lignin biosynthesis in sweet basil plants. Although in vivo studies *EgMYB4* did not suppress *EgF5H*, we did observe a decrease in *ObF5H* expression when *EgMYB4* was ectopically expressed. The direct interaction of *EgMYB4* with the promoters of *ObCOMTs* and *ObF5H* remains to be confirmed. However similar results have been observed



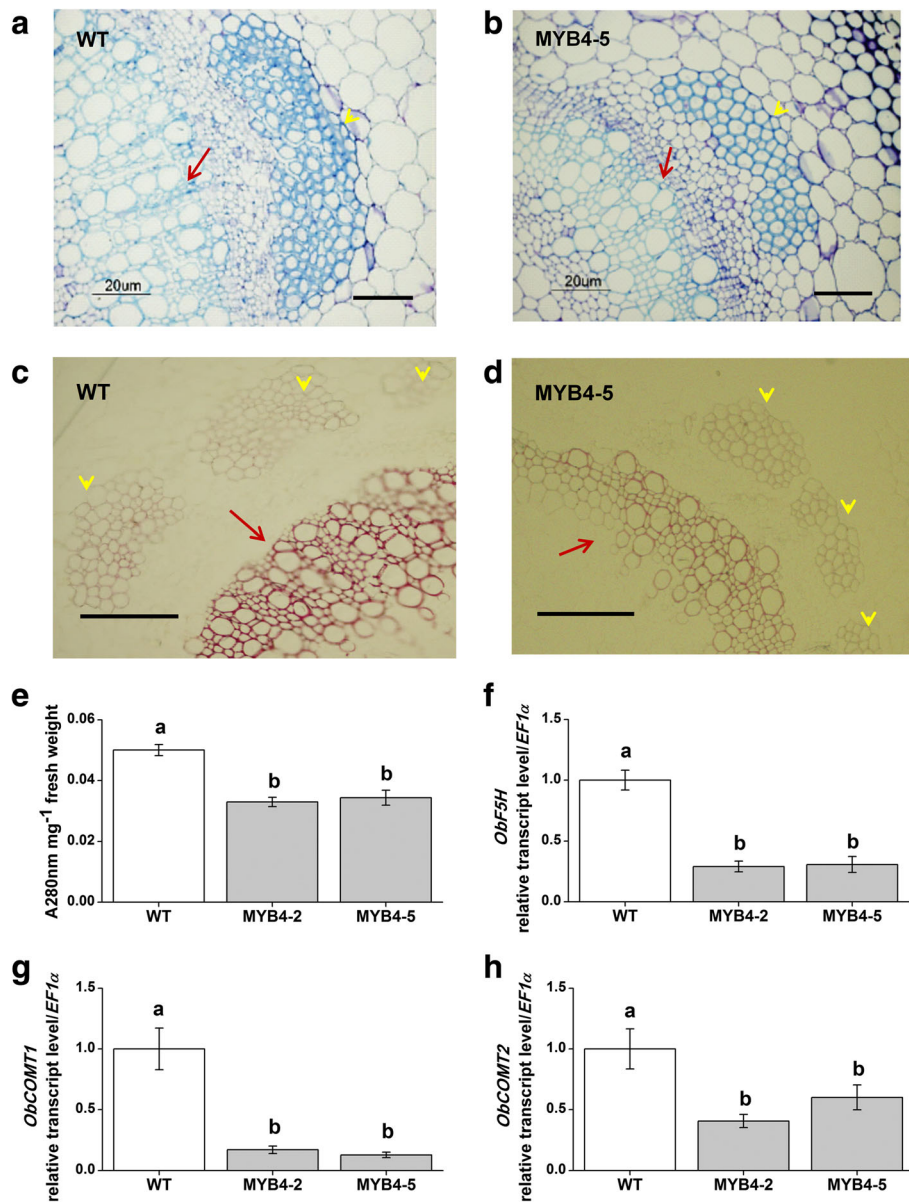
with other R2R3-MYBs like *ZmMYB31* and *AtMYB4* where overexpression leads to dose dependent selection of additional target genes in transgenic plants [20, 26].

To determine the effect of *EgMYB4* overexpression on phenylpropenes, volatile compounds were extracted and analyzed from transgenic lines and WT plants. The transgenic plants exhibited a significant increase in total phenylpropene levels (Fig. 8d). The amount of methylchavicol in MYB4-2 and MYB4-5 was ~15- and 10-folds higher than WT plants (Fig. 8a). Meanwhile, the amount of methyleugenol in MYB4-2 and MYB4-5 was ~ 4- and 3-folds higher than WT plants (Fig. 8c). However, we did not see any changes in eugenol levels (Fig. 8b). This result suggests that ectopic expression of *EgMYB4* can affect flux into phenylpropene pathway.

## Discussion

Efficient pollination is vital for plant's survival. The interaction between plant and its pollinator is considered to be mutualistic and the result of coevolution between plant and pollinator species. A stringent one-to-one coevolution is generally found between one plant species and its insect partner, for example fig plants and its specific pollinator [62]. In case of oil palm, average yields in Southeast Asia increased considerably only after introduction of *Elaeodobius kamarunicus* from Africa. This strongly suggests that oil palm is highly dependent on just a single species of insect for pollination suggesting a specialized mutualistic palm-pollinator relationship

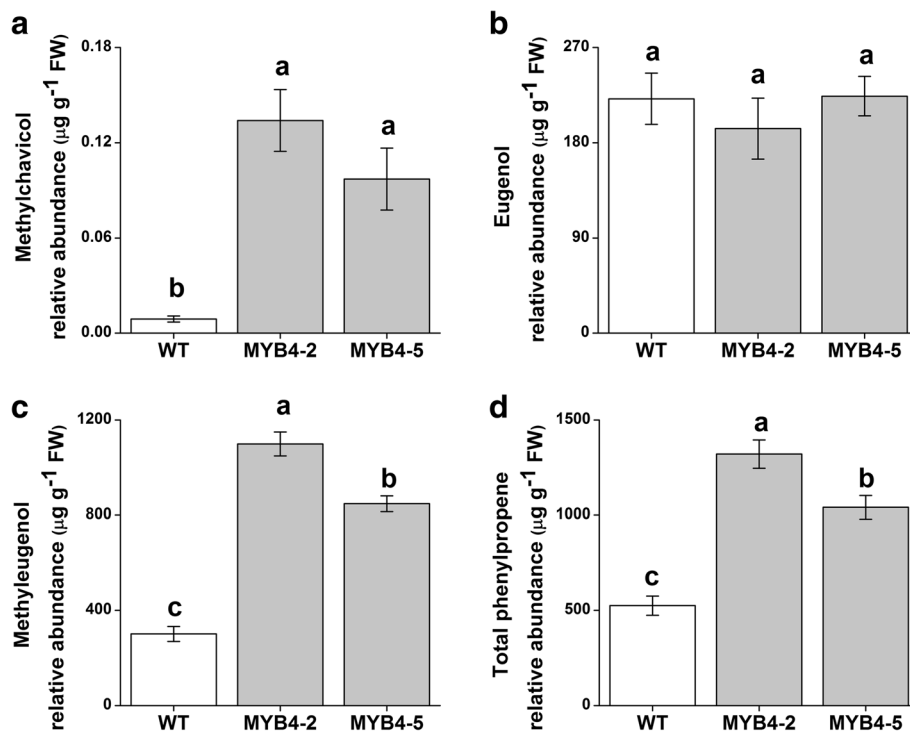
[63]. Phenylpropenes form one of the main components of floral scents that help attract pollinators. Oil palm flowers emit methylchavicol to attract insect pollinators. Phenylpropenes are derived from the general phenylpropanoid pathway. Pathway studies propose that both lignin and phenylpropene production share common substrates [13]. Hence within a cell, tight regulation of metabolic pathway enzymes is required to ensure that substrates used by competing pathways are regulated aptly to produce the desired metabolic outcome. The biosynthesis of lignin consumes large amounts of carbon resources. The biomass of the oil palm is lignocellulosic rich with lignin contents ranging from 18 to 23% among different varieties and tissues [64]. From a plant's perspective, suppression of lignin biosynthesis in flowers allows the plant to reallocate carbon sources towards a vital floral phenylpropene production, which will aid in attracting pollinators. Transcriptome data analysis of open oil palm flowers revealed that *EgCVS* and *EgCvOMT* are highly expressed in flowers whereas expression levels of *EgCOMT* and *EgF5H* are significantly reduced when compared to leaves (Additional file 4). By comparative transcriptome analysis and protein-DNA interaction assays, we identified a R2R3 EAR-motif-containing *EgMYB4* from oil palm which can bind to AC elements of *EgCOMT* and suppress its activity in vivo. No interaction with *EgF5H* was observed (Fig. 5). Although one AC element exists in the promoter of *EgF5H* which was cloned, we cannot completely rule out the possibility that the length



**Fig. 7** Transgenic sweet basil plants overexpressing *EgMYB4* show decreased lignin content. Transverse section of WT (a) and *EgMYB4*-5 line (b) stems stained with toluidine blue. Transverse section of WT (c) and *EgMYB4*-5 line (d) stems stained with phloroglucinol. Red arrows indicate the xylem tissue and yellow arrows indicate the sclerenchyma tissues. Scale bar, 20  $\mu$ m. **e** Thioglycolic acid quantification of total lignin content in wild type (WT) and *EgMYB4*-expressing sweet basil plants. Values are mean  $\pm$  SE ( $n = 6$ ). Mean expression levels ( $\pm$  SE,  $n = 6$ ) of *ObF5H* (f), *ObCOMT1* (g) and *ObCOMT2* (h) in *EgMYB4* overexpressing sweet basil plants and WT plants. Letters indicate significant differences among different lines ( $p < 0.05$ , Duncan's multiple range test)

of the promoter tested might not have been enough to observe the binding of *EgMYB4* to *EgF5H*. Additionally, there might be other EAR-motif-containing MYBs that repress *EgF5H* in oil palm flowers. Monocot transcription factors like *ZmMYB11*, *MusaMYB31*, *ZmMYB31*, *ZmMYB42* and *PvMYB4* which show high sequence identity with *EgMYB4* are reported to repress COMT along with other genes in the lignin biosynthesis pathway [25–28]. In our study, we tested the interaction of *EgMYB4*

with the promoters of *EgCOMT* and *EgF5H* only, we cannot dismiss the possibility that it might regulate other genes in phenylpropanoid pathway. qRT-PCR analysis of *EgMYB4* at different developmental stages of flowers showed high levels of *EgMYB4* in after anthesis stage male flowers. At this stage, flowers do not emit phenylpropanes suggesting additional role for *EgMYB4* in flowers. Perturbation in one branch of phenylpropanoid pathway is known to affect other branches as well. After examining the



**Fig. 8** Transgenic sweet basil plants overexpressing *EgMYB4* show increased phenylpropenes biosynthesis. Mean levels ( $\pm$  SE,  $n = 6$ ) of methylchavicol (a), eugenol (b), methyl eugenol (c) and total phenylpropenes (d) produced by wild type (WT) and *EgMYB4*-expressing sweet basil plants. Letters indicate significant differences among different lines ( $p < 0.05$ , Duncan's multiple range test)

transcripts of genes involved in flavonoid pathway from the RNA-Seq data of oil palm flowers, we were not able to observe a clear pattern that would suggest either up or downregulation of flavonoid pathway in flowers (Additional file 8). Further work is required to analyze other phenylpropanoid derived metabolites in oil palm flowers apart from methylchavicol to determine the effect of *EgMYB4* on their production.

Overexpression of *EgMYB4* in sweet basil resulted in increased phenylpropene levels and decreased lignin content (Figs. 7 and 8). Sweet basil *EgMYB4* overexpressing plants exhibited growth phenotypes typical to lignin deficient plants, including dwarfism, weakened stem, reduced leaf size and delayed flowering [26]. These phenotypes of decreased lignin imply that ectopic expression of *EgMYB4* can affect lignin biosynthesis genes in sweet basil. Downregulation of *ObCOMTs* and *ObF5H* was observed in transgenic plants. However, in the in vivo studies *EgMYB4* did not suppress *EgF5H*. There is a possibility that *EgMYB4* might have different targets in different plants or the resulting perturbation in the lignin pathway might affect the steady state mRNA level of *ObF5H*. Additionally, this might also be due to the high level of *EgMYB4* expression in the transgenics. Similar results have been observed with *AtMYB4* and *ZmMYB31*. *AtMYB4* mainly regulates cinnamate 4-hydroxylase (*C4H*) to control

hydroxycinnamic acid metabolism. But when overexpressed it was able to down regulate the expression of additional genes (*CHS*, *4CL1* and *4CL3*) in the pathway. It was found that there is a dose dependent selection of target genes by *AtMYB4* [20]. Similarly, *ZmMYB31* interacts with *ZmCOMT* and *ZmF5H* promoters in vivo. However, it can repress the expression of Arabidopsis *4CL* and *C3H* when overexpressed in Arabidopsis [26]. The direct binding of *EgMYB4* to the promoters of sweet basil *COMTs*, *F5H* and other lignin pathway genes needs to be confirmed. Downregulation of *COMT* and *F5H* by transgenic approaches has been reported in many plants like Arabidopsis, tobacco, poplar, alfalfa, maize, switchgrass and fescue to manipulate lignin polymer. They mainly affect the formation of S-lignin subunits [57–60, 65–69]. We observed that both methylchavicol and methyleugenol was increased in *EgMYB4* overexpressing sweet basil lines which are derived from *p*-coumaryl and coniferyl alcohol respectively. Probably the downregulation of sinapyl alcohol (S-lignin) pathway increases flux towards the formation of *p*-coumaryl and coniferyl alcohol thus enhancing the formation of both methylchavicol and methyleugenol which are derived from these precursors in sweet basil.

In oil palm, transcripts encoding enzymes directly responsible for methylchavicol biosynthesis like *EgCvOMT*

and EgCVS show transcriptional activation only in flowers (Additional file 4). However, in addition to direct activation of enzymes that are responsible for metabolite production, activation and suppression of enzymes involved in flux leading to its formation also help to fine tune the amount of metabolite formation, in a particular tissue at a developmentally relevant time. For example, in *Petunia* R2R3-MYBs *ODORANT1* (*ODO1*), *EMISSION OF BENZENOIDS I*, *EMISSION OF BENZENOIDS II* (*EOBII*), are all identified as positive regulators of various genes involved in the production of floral volatiles [29–33]. But *PhMYB4* belonging to R2R3-MYB subgroup 4 acts as a repressor of *C4H* gene of phenylpropanoid pathway and indirectly affects the formation of petunia floral volatiles [33]. Similarly, production of methylchavicol in oil palm flowers can be regulated at various levels by different genes acting as positive or negative regulators. Multiple or single TFs might be involved in stage specific activation of EgCvOMT and EgCVS. Our work presented here shows the regulation imparted by one of the probably many transcription factors involved in temporal and spatial regulation of methylchavicol biosynthesis in oil palm flowers. Emission of methylchavicol forms an important agronomic trait in oil palm as it affects pollination. In our study, we identified a transcription factor potentially regulating methylchavicol emissions and enzymes catalyzing its production. These genes may be used as potential candidates in molecular breeding strategies of oil palm. Discovering genetic variations of these genes in different oil palm accessions and its effect on emission/yield can help design breeding of high yielding varieties.

## Conclusions

A volatile phenylpropene, methylchavicol is released from oil palm flowers to attract the pollinator weevil. RNA-Seq analysis of flowers showed increased expression of transcripts coding for enzymes involved in methylchavicol production but decreased expression of *COMT* and *F5H* transcripts which are specifically involved in lignin formation. We identified a R2R3-MYB enriched in flowers, *EgMYB4* that can bind to *EgCOMT* promoter and suppress it. Functional analysis of *EgMYB4* in sweet basil demonstrated the ability of *EgMYB4* to reduce lignin formation and enhance the formation of phenylpropene suggesting a similar role for *EgMYB4* in oil palm flowers. This study on methylchavicol emission in oil palm flowers will further enrich our understanding of transcriptional regulation of the complex phenylpropanoid pathways which produces an array of compounds in plants. Moreover, it will help in oil palm breeding and

metabolic engineering of phenylpropanoid pathway to produce aromatic compounds of interest.

## Methods

### Plant materials

The oil palm species used in this study was *Dura* [2]. Samples from different flower stages and leaves were collected from palm trees planted in Temasek Life Sciences Laboratory, Singapore and Wilmar International Plantation, Palembang, Indonesia. Commercial sweet basil (*O. basilicum*) was tested for its secondary metabolites by GC-MS and grown in green house under natural light conditions. *Agrobacterium* mediated transformation of sweet basil was performed as previously described by [70]. T<sub>0</sub> and T<sub>1</sub> transgenic plants were selected using GFP as visual marker. For all experiments, T<sub>1</sub> plants were used. *Nicotiana benthamiana* seeds were germinated on MS plate and transferred into soil. Twenty days after growing in the greenhouse, the seedlings were used for Agro-infiltration.

### Compound analysis

1.5 cm of one oil palm male flower bunch or 3 female flowers were placed in a 5 mL glass bottle containing 2 mL hexane and shaken for 10 min at room temperature. After centrifugation, 500  $\mu$ L of supernatant was transferred into a 2 mL GC vial and analyzed by gas chromatography-mass spectrometry (GC-MS). For sweet basil samples, leaves were ground in liquid nitrogen and 200 mg of powder was transferred to 1.5 mL eppendorf tubes. 500  $\mu$ L of hexane (containing 20  $\mu$ g of diethyl sebacate as internal standard) was added and vortexed for 2 min. After centrifugation, 400  $\mu$ L of supernatant was transferred into a 2 mL GC vial and analyzed by GC-MS. Six plants from each transgenic line were measured.

### RNA isolation and RNA sequencing

RNA from oil palm tissues was extracted as described previously [71]. Briefly, about 0.5 g of each sample was ground in liquid nitrogen and the powder was then transferred to a pre-chilled polypropylene (Falcon) tube. 5 mL of pre-heated (65 °C) CTAB extraction buffer (2% (w/v) CTAB, 2% (w/v) polyvinylpyrrolidone (PVP-40), 100 mM Tris-HCl (pH 8.0), 25 mM EDTA, 2 M NaCl, 0.1% spermidine and 2%  $\beta$ -mercaptoethanol) was added to each tube and samples were incubated for 30 min at 65 °C. Later, the samples were extracted with chloroform: isoamylalcohol (24:1) for 2 times. The supernatant (1.0 mL) was then transferred to RNase free 1.5 mL eppendorf tubes and 0.5 mL of 96–100% ethanol was added. The mixture was immediately loaded onto RNA binding columns (Qiagen RNA Mini extraction kit) and RNA was extracted according to manufacturer's protocol. RNA sequencing and assembly was performed as

described previously [72]. RNA from sweet basil samples were isolated by using RNA Mini extraction kit (Qiagen).

#### Quantitative real-time PCR (qRT-PCR) and reverse transcription PCR (RT-PCR)

Eight hundred ng of total RNA for each sample was reverse transcribed using the PrimeScript™ RT-PCR Kit (TaKaRa). qRT-PCR was performed on an ABI 7900 HT fast real time system (Life technologies) using SYBR Green Real-time PCR Master Mixes (Life technologies). For RT-PCR, genes were amplified in T100™ Thermal Cycler (Bio-Rad) by the following program, 95 °C for 2 min; 23 cycles of 95 °C for 40s, 60 °C for 40s, 72 °C for 15 s; 72 °C for 5 min. 10 µL of PCR products were analyzed by gel electrophoresis. The primers used for RNA detection of target genes by qRT-PCR and RT-PCR are listed in Additional file 9. Oil palm *Ubiquitin* gene (*EgUBQ*) or sweet basil *elongation factor* (*ObEF1α*) gene was used as internal controls.

#### Constructs

Full-length open reading frames encoding EgMYB4 without a stop codon were amplified by PCR using *Pfu* DNA polymerase (Thermo Scientific) with primers listed in Additional file 9. The obtained DNA fragments were cloned either into pBA-YFP vector [56] to generate YFP fused protein or pET28b (Novagene) to generate His-tag fused protein.

#### Purification of recombinant protein

The construct containing His-MYB4 was transformed into *E. coli* BL21 (DE3). Expression of His-MYB4 was induced by adding 0.4 mM isopropyl-β-thiogalactopyranoside (IPTG) followed by an incubation at 37 °C for 3 h. Later, the cells were collected and the recombinant protein was purified using His-Trap (GE healthcare) according to the manufacturer's instruction.

#### Electrophoretic mobility shift assay (EMSA)

The probes used in EMSA were all labelled by Biotin and are listed in Additional file 9. EMSA was performed using a LightShift Chemiluminescent EMSA Kit (Thermo) according to the manufacturer's instructions. Competition experiments were performed using unlabeled DNA as a competitor in a 250-fold molar excess.

#### Transcriptional repression assay

The 2 kb promoter region of *EgCOMT* and 1 kb promoter region of *EgF5H* were PCR-amplified (primers are listed in Additional file 9) and cloned into pCAM-BIA1391. All constructs were introduced into AGL1 *Agrobacterium*. Leaves of *N. benthamiana* were agroinfiltrated with the indicated constructs (Fig. 5) at a ratio of 1:1. Two days after infiltration leaves were harvested

and frozen in liquid nitrogen or stained with GUS staining buffer [56]. Each treatment was repeated eight times. GUS quantitative assay was performed as described previously [56].

#### Lignin measurement

Total lignin content was measured using thioglycolic acid method as described previously [61]. 100 mg of fresh sweet basil stems were used and six plants from each transgenic line were measured.

#### Histology

Stems from the basal portion of adult wild type sweet basil and transgenic plants were fixed in historesin and sectioned. Sections were stained either with toluidine blue or phloroglucinol for lignin analysis. Lignin analysis was performed as described previously [73].

#### Data analysis

Differences in total lignin and phenylpropene levels on different lines were determined by analysis of variance (ANOVA). Differences in GUS activities on different treatments were analyzed by using the Student's t-test. All tests were carried out with Statistica (Statistica, SAS Institute Inc., <http://www.sas.com/>).

#### Additional files

**Additional file 1:** Analysis of volatile compounds in different oil palm tissues. (DOCX 115 kb)

**Additional file 2:** Overview of RNA-seq result. (DOCX 302 kb)

**Additional file 3:** Transcription factors found in RNA-seq data. (DOCX 209 kb)

**Additional file 4:** Heat map of phenylpropanoid pathway transcripts expression in oil palm leaves and male and female open flowers generated by Heatmapper. (DOCX 490 kb)

**Additional file 5:** List of plant genes used in the phylogenetic analysis described in Figs. 3 and 4. (DOCX 25 kb)

**Additional file 6:** Nucleotide sequence information of phenylpropanoid pathway genes from RNA-seq. (DOCX 22 kb)

**Additional file 7:** GC-MS analysis of phenylpropenes in sweet basil leaves. (DOCX 74 kb)

**Additional file 8:** Expression levels of predicted flavonoid biosynthesis genes in oil palm leaves and flowers. (DOCX 20 kb)

**Additional file 9:** DNA primers used in this study. (DOCX 21 kb)

#### Abbreviations

C4H: cinnamate 4-hydroxylase; CAD: cinnamyl alcohol dehydrogenase; CCR: cinnamoyl-CoA reductase; COMT: caffeic acid O-methyltransferase; EMSA: electrophoretic mobility shift assay; GC-MS: gas chromatography-mass spectrometry; GFP: green fluorescent protein; GUS: β-glucuronidase; RNA-Seq: RNA sequencing; TF: transcription factors; YFP: yellow fluorescent protein

#### Acknowledgements

We thank KharMeng Ng for his invaluable assistance with experiments. This research was supported by funds from Wilmar International Ltd. and Temasek Life Sciences Laboratory.

### Funding

This research was supported by funds from Wilmar International Ltd. and Temasek Life Sciences Laboratory.

### Availability of data and materials

The RNA-Seq data supporting the result of this article is available in the DNA Data Bank of Japan (DDBJ: <http://www.ddbj.nig.ac.jp/>) with accession number DRA004319. Nucleotide sequences from this study were submitted to the National Center for Biotechnology Information (NCBI)/GenBank using BankIt with accession numbers: *EgCVS* (KT778617), *EgMYB4* (KT778616), *EgCOMT* promoter (KT778614), *EgF5H* promoter (KT778615). All the phylogenetic trees and associated matrices have been deposited to the TreeBASE (<http://treebase.org/treebase-web/home.html>) with the study number 21622. The TreeBASE generated URL to access the deposition is <http://purl.org/phylo/treebase/phyloids/study/TB2:S21622>.

### Authors' contributions

RL, JY and RS designed experiments. RL, VAR and QW performed the experiments. CHL assisted with oil palm flower sample preparation. JY conducted the experiments related with RNA-Seq. JJ analyzed the transcriptome data. CR performed the sweet basil transformation. GY supported oil palm genome information. RL, NHC, JY and RS wrote the article, which was reviewed and approved by all authors.

### Ethics approval and consent to participate

This research was a collaborative effort between Wilmar International and Temasek Life Sciences Laboratory. Oil palm (*Elaeis guineensis*) samples were collected from palm trees planted in Temasek Life Sciences Laboratory, Singapore and Wilmar International Plantation, Palembang, Indonesia. Sweet basil seeds were obtained from commercial suppliers and grown in green house at Temasek Life Sciences Laboratory. No specific permits were required.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Author details

<sup>1</sup>Temasek Life Sciences Laboratory, National University of Singapore, Singapore 117604, Singapore. <sup>2</sup>Department of Biological Sciences, National University of Singapore, Singapore 117543, Singapore. <sup>3</sup>R&D Department, Wilmar International Plantation, Palembang, Indonesia. <sup>4</sup>Laboratory of Plant Molecular Biology, Rockefeller University, New York, NY 10065, USA. <sup>5</sup>State Key Laboratory of Plant Genomics, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China. <sup>6</sup>Present Address: Department of Molecular Ecology, Max Planck Institute for Chemical Ecology, 07745 Jena, Germany. <sup>7</sup>Present Address: Singapore Centre on Environmental Life Sciences Engineering, Nanyang Technological University, Singapore, Singapore. <sup>8</sup>Present Address: College of Biological and Environmental Sciences, Zhejiang Wanli University, Ningbo, Zhejiang, China.

Received: 24 May 2017 Accepted: 13 November 2017

Published online: 23 November 2017

### References

- Singh R, Ong-Abdullah M, Low ETL, Manaf MAA, Rosli R, Nookiah R, Ooi LCL, Ooi SE, Chan KL, Halim MA. Oil palm genome sequence reveals divergence of interfertile species in old and new worlds. *Nature*. 2013;500(7462):335–9.
- Lee M, Xia JH, Zou Z, Ye J, Alfiko Y, Jin J, Lieando JV, Purnamasari MI, Lim CH, Suwanto A, Wong L, Chua NH, Yue GH. A consensus linkage map of oil palm and a major QTL for stem height. *Sci Rep*. 2015;5:8232.
- Morcillo F, Cros D, Billotte N, Ngando-Ebongue GF, Domonhédou H, Pizot M, Cuéllar T, Espéout S, Dhoubi R, Bourgis F. Improving palm oil quality through identification and mapping of the lipase gene causing oil deterioration. *Nat Commun*. 2013;4:2160.
- Corley RHV, Tinker P. The oil palm: John Wiley & Sons; 2008.
- Syed R, Pushparajah E, Soon CP. Insect pollination of oil palm: feasibility of introducing *Elaeidobius* spp. into Malaysia. In: *The oil palm in the eighties A report of the Proceedings of the International Conference on Oil Palm in Agriculture in the Eighties, Kuala Lumpur, 17–20 June 1981 Volume I: 1982*. Incorporated Society of Planters: 263–289.
- Hussein M, Lajis N, Ali J. Biological and chemical factors associated with the successful introduction of *Elaeidobius kamerunicus* Faust, the oil palm pollinator in Malaysia. *Acta Hort*. 1990;288:81–7.
- Koeduka T, Fridman E, Gang DR, Vassão DG, Jackson BL, Kish CM, Orlova I, Spassova SM, Lewis NG, Noel JP. Eugenol and isoeugenol, characteristic aromatic constituents of spices, are biosynthesized via reduction of a coniferyl alcohol ester. *Proc Natl Acad Sci U S A*. 2006;103(26):10128–33.
- Tan KH, Nishida R. Methyl eugenol: its occurrence, distribution, and role in nature, especially in relation to insect behavior and pollination. *J Insect Sci*. 2012;12(1):56.
- Gang DR, Wang J, Dudareva N, Nam KH, Simon JE, Lewinsohn E, Pichersky E. An investigation of the storage and biosynthesis of phenylpropenes in sweet basil. *Plant Physiol*. 2001;125(2):539–55.
- Liu J, Osbourn A, Ma P. MYB transcription factors as regulators of phenylpropanoid metabolism in plants. *Mol Plant*. 2015;8(5):689–708.
- Zhao Q, Dixon RA. Transcriptional networks for lignin biosynthesis: more complex than we thought? *Trends Plant Sci*. 2011;16(4):227–33.
- Bonawitz ND, Chapple C. The genetics of lignin biosynthesis: connecting genotype to phenotype. *Annu Rev Genet*. 2010;44:337–63.
- Muhlemann JK, Woodworth BD, Morgan JA, Dudareva N. The monolignol pathway contributes to the biosynthesis of volatile phenylpropenes in flowers. *New Phytol*. 2014;204(3):661–70.
- Hatton D, Sablowski R, Yung MH, Smith C, Schuch W, Bevan M. Two classes of cis sequences contribute to tissue-specific expression of a PAL2 promoter in transgenic tobacco. *Plant J*. 1995;7(6):859–76.
- Raes J, Rohde A, Christensen JH, Van de Peer Y, Boerjan W. Genome-wide characterization of the lignification toolbox in Arabidopsis. *Plant Physiol*. 2003;133(3):1051–71.
- Shen H, He X, Poovaiah CR, Wuddineh WA, Ma J, Mann DG, Wang H, Jackson L, Tang Y, Neal Stewart C. Functional characterization of the switchgrass (*Panicum virgatum*) R2R3-MYB transcription factor PvMYB4 for improvement of lignocellulosic feedstocks. *New Phytol*. 2012;193(1):121–36.
- Stracke R, Werber M, Weisshaar B. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr Opin Plant Biol*. 2001;4(5):447–56.
- Vélez-Bermúdez IC, Salazar-Henao JE, Fornalé S, López-Vidriero I, Franco-Zorrilla JM, Grotewold E, Gray J, Solano R, Schmidt W, Pagés M, MYB A. ZML complex regulates wound-induced lignin genes in maize. *Plant Cell*. 2015;27(11):3245–59.
- Tamagnone L, Merida A, Parr A, Mackay S, Culianez-Macia FA, Roberts K, Martin C. The AmMYB308 and AmMYB330 transcription factors from antirrhinum regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. *Plant Cell*. 1998;10(2):135–54.
- Jin H, Cominelli E, Bailey P, Parr A, Mehrtens F, Jones J, Tonelli C, Weisshaar B, Martin C. Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. *EMBO J*. 2000;19(22):6150–61.
- Preston J, Wheeler J, Heazlewood J, Li SF, Parish RW. AtMYB32 is required for normal pollen development in *Arabidopsis thaliana*. *Plant J*. 2004;40(6):979–95.
- Legay S, Sivadon P, Blervacq AS, Pavy N, Baghdady A, Tremblay L, Levasseur C, Ladouce N, Lapiere C, Séguin A. EgMYB1, an R2R3-MYB transcription factor from eucalyptus negatively regulates secondary cell wall formation in Arabidopsis and poplar. *New Phytol*. 2010;188(3):774–86.
- Zhu L, Shan H, Chen S, Jiang J, Gu C, Zhou G, Chen Y, Song A, Chen F. The heterologous expression of the chrysanthemum R2R3-MYB transcription factor CmMYB1 alters lignin composition and represses flavonoid synthesis in *Arabidopsis thaliana*. *PLoS One*. 2013;8(6):e65680.
- Fornalé S, Sonbol FM, Maes T, Capellades M, Puigdomènech P, Rigau J, Caparrós-Ruiz D. Down-regulation of the maize and *Arabidopsis thaliana*

- caffeic acid O-methyl-transferase genes by two new maize R2R3-MYB transcription factors. *Plant Mol Biol.* 2006;62(6):809–23.
25. Sonbol FM, Fornalé S, Capellades M, Encina A, Touriño S, Torres JL, Rovira P, Ruel K, Puigdomenech P, Rigau J. The maize ZmMYB42 represses the phenylpropanoid pathway and affects the cell wall structure, composition and degradability in *Arabidopsis thaliana*. *Plant Mol Biol.* 2009;70(3):283–96.
  26. Fornalé S, Shi X, Chai C, Encina A, Irar S, Capellades M, Fuguet E, Torres JL, Rovira P, Puigdomenech P. ZmMYB31 directly represses maize lignin genes and redirects the phenylpropanoid metabolic flux. *Plant J.* 2010; 64(4):633–44.
  27. Agarwal T, Grotewold E, Doseff AI, Gray J. MYB31/MYB42 Syntelogs exhibit divergent regulation of phenylpropanoid genes in maize, sorghum and rice. *Sci Rep.* 2016;6:28502.
  28. Tak H, Negi S, Ganapathi T. Overexpression of *MusaMYB31*, a R2R3 type MYB transcription factor gene indicate its role as a negative regulator of lignin biosynthesis in banana. *PLoS One.* 2017;12(2):e0172695.
  29. Verdonk JC, Haring MA, van Tunen AJ, Schuurink RC. ODORANT1 regulates fragrance biosynthesis in petunia flowers. *Plant Cell.* 2005;17(5):1612–24.
  30. Spitzer-Rimon B, Marhevka E, Barkai O, Marton I, Edelbaum O, Masci T, Prathapani NK, Shklarman E, Ovadis M, Vainstein A. EOBI1, a gene encoding a flower-specific regulator of phenylpropanoid volatiles' biosynthesis in petunia. *Plant Cell.* 2010;22(6):1961–76.
  31. Colquhoun TA, Schwieterman ML, Wedde AE, Schimmel BC, Marciniak DM, Verdonk JC, Kim JY, Oh Y, Gális I, Baldwin IT. EOBI1 controls flower opening by functioning as a general transcriptomic switch. *Plant Physiol.* 2011;156(2):974–84.
  32. Spitzer-Rimon B, Farhi M, Albo B, Cna'ani A, MMB Z, Masci T, Edelbaum O, Yu Y, Shklarman E, Ovadis M. The R2R3-MYB-like regulatory factor EOBI, acting downstream of EOBI1, regulates scent production by activating ODO1 and structural scent-related genes in petunia. *Plant Cell.* 2012;24(12):5089–105.
  33. Colquhoun TA, Kim JY, Wedde AE, Levin LA, Schmitt KC, Schuurink RC, Clark DG. PhMYB4 fine-tunes the floral volatile signature of *Petunia hybrida* through PhC4H. *J Exp Bot.* 2010;62(3):1133–43.
  34. Zvi MMB, Negre-Zakharov F, Masci T, Ovadis M, Shklarman E, Ben-Meir H, Tzifira T, Dudareva N, Vainstein A. Interlinking showy traits: co-engineering of scent and colour biosynthesis in flowers. *Plant Biotechnol J.* 2008;6(4):403–15.
  35. Zvi MMB, Shklarman E, Masci T, Kalev H, Debener T, Shafir S, Ovadis M, Vainstein A. PAP1 transcription factor enhances production of phenylpropanoid and terpenoid scent compounds in rose flowers. *New Phytol.* 2012;195(2):335–45.
  36. Medina-Puche L, Molina-Hidalgo FJ, Boersma MR, Schuurink RC, López-Vidriero I, Solano R, Franco-Zorrilla J-M, Caballero JL, Blanco-Portales R, Muñoz-Blanco J. A R2R3-MYB transcription factor (FaEOBI1) regulates eugenol production in ripe strawberry (*Fragaria x ananassa*) fruit receptacles. *Plant Physiol.* 2015;168:598–614.
  37. Misztal PK, Owen SM, Guenther AB, Rasmussen R, Geron C, Harley P, Phillips GJ, Ryan A, Edwards DP, Hewitt CN. Large estragole fluxes from oil palms in Borneo. *Atmos Chem Phys.* 2010;10:4343–58.
  38. Barcelos E, de Almeida Rios S, Cunha RN, Lopes R, Motoike SY, Babiychuk E, Skirycz A, Kushnir S. Oil palm natural diversity and the potential for yield improvement. *Front Plant Sci.* 2015;6:190.
  39. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* 2011;29(7):644–52.
  40. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. *Nat Methods.* 2012;9(4):357–9.
  41. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics.* 2011;12(1):323.
  42. Vom Endt D, Kijne JW, Memelink J. Transcription factors controlling plant secondary metabolism: what regulates the regulators? *Phytochemistry.* 2002;61(2):107–14.
  43. Smaczniak C, Immink RG, Angenent GC, Kaufmann K. Developmental and evolutionary diversity of plant MADS-domain factors: insights from recent studies. *Development.* 2012;139(17):3081–98.
  44. Fujiwara S, Oda A, Yoshida R, Niinuma K, Miyata K, Tomozoe Y, Tajima T, Nakagawa M, Hayashi K, Coupland G. Circadian clock proteins LHY and CCA1 regulate SVP protein accumulation to control flowering in *Arabidopsis*. *Plant Cell.* 2008;20(11):2960–71.
  45. Cheng H, Song S, Xiao L, Soo HM, Cheng X, Xie D, Peng J. Gibberellin acts through jasmonate to control the expression of MYB21, MYB24, and MYB57 to promote stamen filament growth in *Arabidopsis*. *PLoS Genet.* 2009;5(3):e1000440.
  46. Mandaokar A, Thines B, Shin B, Markus Lange B, Choi G, Koo YJ, Yoo YJ, Choi YD, Choi G, Browse J. Transcriptional regulators of stamen development in *Arabidopsis* identified by transcriptional profiling. *Plant J.* 2006;46(6):984–1008.
  47. Mandaokar A. MYB108 acts together with MYB24 to regulate jasmonate-mediated stamen maturation in *Arabidopsis*. *Plant Physiol.* 2009;149(2):851–62.
  48. Humphreys JM, Chapple C. Rewriting the lignin roadmap. *Curr Opin Plant Biol.* 2002;5(3):224–9.
  49. Koeduka T, Louie GV, Orlova I, Kish CM, Ibdah M, Wilkerson CG, Bowman ME, Baiga TJ, Noel JP, Dudareva N. The multiple phenylpropene synthases in both *Clarkia breweri* and *Petunia hybrida* represent two distinct protein lineages. *Plant J.* 2008;54(3):362–74.
  50. Koeduka T, Orlova I, Baiga TJ, Noel JP, Dudareva N, Pichersky E. The lack of floral synthesis and emission of isoeugenol in *Petunia axillaris* subsp. *parodii* is due to a mutation in the isoeugenol synthase gene. *Plant J.* 2009;58(6):961–9.
  51. Aragüez I, Osorio S, Hoffmann T, Rambla JL, Medina-Escobar N, Granell A, Botella MÁ, Schwab W, Valpuesta V. Eugenol production in achenes and receptacles of strawberry fruits is catalyzed by synthases exhibiting distinct kinetics. *Plant Physiol.* 2013;163(2):946–58.
  52. Koeduka T, Baiga TJ, Noel JP, Pichersky E. Biosynthesis of t-anethole in anise: characterization of t-anol/isoegenol synthase and an O-methyltransferase specific for a C7-C8 propenyl side chain. *Plant Physiol.* 2009;149(1):384–94.
  53. Vassão DG, Kim S-J, Milhollan JK, Eichinger D, Davin LB, Lewis NG. A pinoresinol-lariciresinol reductase homologue from the creosote bush (*Larrea tridentata*) catalyzes the efficient in vitro conversion of p-coumaryl/coniferyl alcohol esters into the allylphenols chavicol/eugenol, but not the propenylphenols p-anol/isoegenol. *Arch Biochem Biophys.* 2007;465(1):209–18.
  54. Gang DR, Lavid N, Zubieta C, Chen F, Beuerle T, Lewinsohn E, Noel JP, Pichersky E. Characterization of phenylpropene O-methyltransferases from sweet basil facile change of substrate specificity and convergent evolution within a plant O-methyltransferase family. *Plant Cell.* 2002;14(2):505–19.
  55. Yauk YK, Chagné D, Tomes S, Match AJ, Wang MY, Chen X, Maddumage R, Hunt MB, Rowan DD, Atkinson RG. The O-methyltransferase gene MdoOMT1 is required for biosynthesis of methylated phenylpropenes in ripe apple fruit. *Plant J.* 2015;82:937–50.
  56. Li R, Weldegergis BT, Li J, Jung C, Qu J, Sun Y, Qian H, Tee C, van loon JJ, Dicke M, Chua NH, Liu SS, ye J. Virulence factors of geminivirus interact with MYC2 to subvert plant resistance and promote vector performance. *Plant Cell.* 2014;26(12):4991–5008.
  57. Ni W, Paiva NL, Dixon RA. Reduced lignin in transgenic plants containing a caffeic acid-O-methyltransferase antisense gene. *Transgenic Res.* 1994;3(2):120–6.
  58. Guo D, Chen F, Inoue K, Blount JW, Dixon RA. Downregulation of caffeic acid 3-O-methyltransferase and caffeoyl CoA 3-O-methyltransferase in transgenic alfalfa: impacts on lignin structure and implications for the biosynthesis of G and S lignin. *Plant Cell.* 2001;13(1):73–88.
  59. Piquemal J, Chamayou S, Nadaud I, Beckert M, Barrière Y, Mila I, Lapiere C, Rigau J, Puigdomenech P, Jauneau A. Down-regulation of caffeic acid O-methyltransferase in maize revisited using a transgenic approach. *Plant Physiol.* 2002;130(4):1675–85.
  60. Chen L, Auh C-K, Dowling P, Bell J, Lehmann D, Wang ZY. Transgenic down-regulation of caffeic acid O-methyltransferase (COMT) led to improved digestibility in tall fescue (*Festuca arundinacea*). *Funct Plant Biol.* 2004;31(3):235–45.
  61. Bonawitz ND, Im Kim J, Tobimatsu Y, Ciesielski PN, Anderson NA, Ximenes E, Maeda J, Ralph J, Donohoe BS, Ladisch M. Disruption of mediator rescues the stunted growth of a lignin-deficient *Arabidopsis* mutant. *Nature.* 2014;509(7500):376–80.
  62. Shimizu A, Dohzono I, Nakaji M, Roff DA, Miller DG III, Osato S, Yajima T, Niitsu S, Utsugi N, Sugawara T. Fine-tuned bee-flower coevolutionary state hidden within multiple pollination interactions. *Sci Rep.* 2014;4
  63. Barfod AS, Hagen M, Borchsenius F. Twenty-five years of progress in understanding pollination mechanisms in palms (Arecaceae). *Ann Bot.* 2011;108(8):1503–16.
  64. Hashim R, Saari N, Sulaiman O, Sugimoto T, Hiziroglu S, Sato M, Tanaka R. Effect of particle geometry on the properties of binderless particleboard manufactured from oil palm trunk. *Mater Design.* 2010;31(9):4251–7.



65. Meyer K, Shirley AM, Cusumano JC, Bell-Lelong DA, Chapple C. Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in Arabidopsis. *Proc Natl Acad Sci U S A*. 1998;95(12):6619–23.
66. Franke R, McMichael CM, Meyer K, Shirley AM, Cusumano JC, Chapple C. Modified lignin in tobacco and poplar plants over-expressing the Arabidopsis gene encoding ferulate 5-hydroxylase. *Plant J*. 2000;22(3):223–34.
67. Huntley SK, Ellis D, Gilbert M, Chapple C, Mansfield SD. Significant increases in pulping efficiency in C4H-F5H-transformed poplars: improved chemical savings and reduced environmental toxins. *J Agr Food Chem*. 2003;51(21):6178–83.
68. Li L, Zhou Y, Cheng X, Sun J, Marita JM, Ralph J, Chiang VL. Combinatorial modification of multiple lignin traits in trees through multigene cotransformation. *Proc Natl Acad Sci U S A*. 2003;100(8):4939–44.
69. Fu C, Mielenz JR, Xiao X, Ge Y, Hamilton CY, Rodriguez M, Chen F, Foston M, Ragauskas A, Bouton J. Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proc Natl Acad Sci U S A*. 2011;108(9):3803–8.
70. Wang Q, Reddy VA, Panicker D, Mao HZ, Kumar N, Rajan C, Venkatesh PN, Chua NH, Sarojam R. Metabolic engineering of terpene biosynthesis in plants using a trichome-specific transcription factor MsYABBY5 from spearmint (*Mentha spicata*). *Plant Biotech J*. 2016;14:1619–32.
71. Sangha JS, Gu K, Kaur J, Yin Z. An improved method for RNA isolation and cDNA library construction from immature seeds of *Jatropha curcas* L. *BMC research notes*. 2010;3(1):126.
72. Jin J, Panicker D, Wang Q, Kim MJ, Liu J, Yin J-L, Wong L, Jang I-C, Chua N-H, Sarojam R. Next generation sequencing unravels the biosynthetic ability of spearmint (*Mentha spicata*) peltate glandular trichomes through comparative transcriptomics. *BMC Plant Biol*. 2014;14(1):292.
73. Liljegren SJ, Ditta GS, Eshed Y, Savidge B, Bowman JL, Yanofsky MF, MADS-box SHATTERPROOF genes control seed dispersal in Arabidopsis. *Nature*. 2000;404(6779):766–70.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

