Accepted Article

MISS VAISHNAVI AMARR REDDY (Orcid ID : 0000-0003-0553-272X)

Received Date : 27-Nov-2016

Revised Date : 15-Jan-2017

Accepted Date : 27-Jan-2017

Article type : Research Article

Spearmint R2R3-MYB transcription factor MsMYB negatively regulates monoterpene production and suppresses the expression of geranyl diphosphate synthase large subunit (*MsGPPS.LSU*)

Vaishnavi Amarr Reddy[1,2], Qian Wang[1#], Niha Dhar[1#], Nadimuthu Kumar[1], Prasanna Nori Venkatesh[1], Chakravarthy Rajan[1], Deepa Panicker[1], Vishweshwaran Sridhar[1], Hui-Zhu Mao[1], Rajani Sarojam[1*]

[1]Temasek Life Sciences Laboratory, 1 Research Link, National University of Singapore, Singapore 117604

[2]Department of Biological Sciences, National University of Singapore, Singapore 117543, Singapore

[#]Authors contributed equally

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/pbi.12701
This article is protected by copyright. All rights reserved.

Accepted Article

E-mail addressess:

Vaishnavi Amarr Reddy	vaishnavi@tll.org.sg
Qian Wang	emirate14@zju.edu.cn
Niha Dhar	niha@tll.org.sg
Nadimuthu Kumar	kumarnadi@tll.org.sg
Prasanna Nori Venkatesh	prasanna@tll.org.sg
Rajan Chakravarthy	biochackara@gmail.com
Deepa Panicker	nikappil@ntu.edu.sg
Vishweshwaran Sridhar	vishwesh@tll.org.sg
Hui-Zhu Mao	maohz@tll.org.sg

***Corresponding author**: Dr. Rajani Sarojam

Principal Investigator

Temasek Life Sciences Laboratory,

1 Research Link, National University of Singapore,

Singapore- 117604.

Email- rajanis@tll.org.sg

Fax- +65 68727517

Present address

Wang Qian - College of Biological and Environmental Sciences, Zhejiang Wanli University, Ningbo, 315100, Zhejiang, PR China

Chakravarthy Rajan - Singapore Centre on Environmental Life Sciences Engineering, Nanyang Technological University, Singapore, 637551

This article is protected by copyright. All rights reserved.

Accepted Article

Deepa Panicker - Singapore Centre on Environmental Life Sciences Engineering, Nanyang Technological University, Singapore, 637551

Running title: Spearmint *MsMYB* negatively regulates monoterpene biosynthesis

Keywords: transcription factor, R2R3-MYB, secondary metabolism, spearmint, GPPS, terpene

Summary

Many aromatic plants, like spearmint, produce valuable essential oils in specialized structures called peltate glandular trichomes (PGT). Understanding the regulatory mechanisms behind the production of these important secondary metabolites will help design new approaches to engineer them. Here, we identified a PGT-specific R2R3-MYB gene *MsMYB,* from comparative RNA-Seq data of spearmint, and functionally characterized it. Analysis of *MsMYB*-RNAi transgenic lines showed increased levels of monoterpenes and *MsMYB*-overexpressing lines exhibited decreased levels of monoterpenes. These results suggest that MsMYB is a novel negative regulator of monoterpene biosynthesis. Ectopic expression of *MsMYB,* in sweet basil and tobacco, perturbed sesquiterpene and diterpene derived metabolite production. In addition, we found that MsMYB binds to *cis*-elements of *MsGPPS.LSU* and suppresses its expression. Phylogenetic analysis placed MsMYB in subgroup 7 of R2R3-MYBs whose members govern phenylpropanoid pathway and are regulated by miR858. Analysis of transgenic lines showed that *MsMYB* is more

This article is protected by copyright. All rights reserved.

Accepted Article

specific to terpene biosynthesis as it did not affect metabolites derived from phenylpropanoid pathway. Further, our results indicate that *MsMYB* is probably not regulated by miR858, like other members of subgroup 7.

Introduction

Plants produce an overwhelming variety of specialized metabolites essential for ecological interactions, among which terpenes form the largest and structurally diverse class of natural products. They are involved in mediating plant defense responses and plant pollinator attraction, and facilitate plant-environment interactions (Bhargava et al., 2013; Pichersky and Gershenzon, 2002; Singh and Sharma, 2015). Apart from imparting ecological benefits to plants, terpenoids are of great economic importance to humans as well, as they are widely used in flavours, fragrances, cosmetics, pharmaceuticals, agricultural industries and chemical industries (Bouvier et al., 2005). Given the commercial and ecological importance of terpenoids, strategies to metabolically engineer them is of considerable interest. Monoterpenes form the C_{10} class of terpenoids and are generally colourless, lipophilic and volatile. They are responsible for the characteristic aromas and flavours of essential oils, floral scents and resin of aromatic plants (Loza-Tavera, 1999). The essential oil of spearmint (*Mentha spicata*) mainly consists of two monoterpenes, limonene and carvone, which are extensively exploited for their biological properties (Ringer et al., 2005). In spearmint, the essential oils are produced and stored in specialized structures called peltate glandular trichomes (Figure S1a). These so called “green

This article is protected by copyright. All rights reserved.

Accepted Article

biofactories" are widely found on aerial surfaces of many aromatic plants, and they actively produce and store large quantity of volatile metabolites (Champagne and Boutry, 2013; Lange and Turner, 2013; Turner et al., 2000).

Terpene biosynthesis pathway has been well studied in plants. The universal C_5 building blocks for all types of terpenes, isopentenyl diphosphate (IPP) and its isomer, dimethylallyl diphosphate (DMAPP), are synthesised either by the mevalonate (MVA) pathway in the cytosol, or by the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway in plastids (Vranova et al., 2013). The plastidial MEP pathway is largely responsible for producing C_5 precursors for monoterpenes and diterpenes production, whereas the cytosolic MVA pathway generates C_5 precursors for sesquiterpenes and triterpenes production (Dubey et al., 2003). However, few studies indicate that, under certain conditions, an exchange of precursor metabolites can occur between the cytosolic MVA and plastidial MEP pathways (Hemmerlin et al., 2012; Vranova et al., 2013). IPP and DMAPP undergo successive condensation reactions, catalysed by a class of enzymes called prenyltransferases, to form intermediates geranyl diphosphate (GPP; C_{10}), farnesyl diphosphate (FPP; C_{15}), and geranylgeranyl diphosphate (GGPP; C_{20}). These terpene diphosphates form the immediate precursors of monoterpenes, sesquiterpenes and diterpenes respectively (Kellogg and Poulter, 1997; Liang, 2009; Liang et al., 2002; Ogura and Koyama, 1998). Prenyltransferases are key to terpene production as they control the IPP flux into various branches of the terpene family (Liang et al., 2002; Oldfield and Lin, 2012).

This article is protected by copyright. All rights reserved.

Accepted Article

Geranyl diphosphate synthase (GPPS) is the prenyltransferase enzyme that is mainly responsible for the production of monoterpene precursor GPP in the plastids. It catalyzes a single condensation of terpene precursors DMAPP and IPP, to form GPP. Studies in few species of plants revealed that GPPS can function as homodimers or heterodimers (Nagegowda, 2010). Heteromeric GPPSs have been characterized so far only in angiosperms *Mentha piperita, Antirrhinum majus, Clarkia breweri*, and *Humulus lupulus*. All these plants produce large amounts of monoterpenes in specialized organs such as PGTs and flower petals (Burke et al., 1999; Tholl et al., 2004; Wang and Dixon, 2009). Heteromeric GPPS in mint consist of a large subunit (LSU) and a small subunit (SSU); both LSU and SSU are catalytically inactive alone (Burke and Croteau, 2002; Burke et al., 1999; Croteau et al., 2005). Interaction between the two subunits results in the formation of an active GPPS. Structural studies show that the LSU acts as the catalytic unit, whereas the SSU serves as the regulatory unit (Chang et al., 2010).

In the case of spearmint, the monoterpene biosynthetic pathway is well characterized (Champagne and Boutry, 2013; Croteau et al., 1991; Lange et al., 2011; Munoz-Bertomeu et al., 2008). Strategies to increase the yield of peppermint oil by manipulating genes that code for structural pathway enzymes have been reported (Diemer et al., 2001; Mahmoud et al., 2004). However, the developmental regulation of this secondary metabolite pathway still remains elusive. Transcription factors (TFs) can activate or repress multiple genes in a metabolic pathway, hence they are ideal targets for pathway engineering (Grotewold, 2008; Iwase et al., 2009). Only one TF *MsYABBY5* has been

This article is protected by copyright. All rights reserved.

Accepted Article

reported from spearmint that regulates monoterpene production (Wang et al., 2016) but several TFs have been identified from other plants like *Artemisia*, cotton, *Taxus*, rubber and rice that control terpene biosynthesis (Chen et al., 2012; Li et al., 2013; Miyamoto et al., 2014; Shen et al., 2016; Xu et al., 2004; Zhang et al., 2012).

The R2R3-MYBs represent one of the largest families of plant TFs (Kranz et al., 1998; Stracke et al., 2001) and can function as activators or repressors of genes (Adato et al., 2009; Legay et al., 2007). Many R2R3-MYBs have been characterized as transcriptional activators of flavonoid, phenylpropanoid and glucosinolate pathway genes in plants (Dubos et al., 2010; Zhao and Dixon, 2011; Zhong et al., 2010). R2R3-MYBs belonging to subgroup 4 are known to negatively regulate monolignol and flavonoid biosynthetic pathways (Fornalé et al., 2010; Jin et al., 2000; Legay et al., 2007; Preston et al., 2004; Tamagnone et al., 1998). Further, a single R2R3-MYB gene was also found to act as both repressor and activator of genes (Bhargava et al., 2010). Redundancy, dual function and opposite action as activators and repressors on common target genes by R2R3-MYBs help fine tune the regulation of plant secondary metabolite pathways. Despite several studies on R2R3 MYB-mediated regulation of plant phenylpropanoid pathway, knowledge of its role in terpene secondary metabolism is lacking (Matus, 2016). Besides TFs, microRNAs also are known regulators of various processes in plant growth and development. However, knowledge of their role in controlling secondary metabolic pathways is limited. MicroRNAs mainly target TFs to regulate various plant developmental processes (Sunkar et al., 2012; Wu et al., 2013; Zhang, 2015; Zhang and

This article is protected by copyright. All rights reserved.

Accepted Article

Wang, 2015). In plants, MYB TFs are potential targets for miR858 (Addo-Quaye et al., 2008; Guan et al., 2014; Jeong et al., 2013; Jia et al., 2015; Xia et al., 2012). Further, many studies have shown that miR858 can target R2R3-MYBs in apple, grapes, salvia and Arabidopsis (Li and Lu, 2014; Rock, 2013; Sharma et al., 2016; Xia et al., 2012). Presently, miR858 is considered as a potential regulator of plant phenylpropanoid metabolite pathway through its action on R2R3-MYBs (Fahlgren et al., 2007).

In this study, we identified a R2R3-MYB transcript which is preferentially expressed in spearmint PGTs from the comparative transcriptome data (Jin et al., 2014) of spearmint leaf, leaf stripped of PGT (L-PGT) and PGT. Phylogenetic analysis assigned MsMYB to subgroup 7 of R2R3-MYBs of Arabidopsis. Members of subgroup 7 are closely related R2R3-MYBs which control flavonoid accumulation in different parts of the Arabidopsis seedling, and are regulated by miR858 (Fahlgren et al., 2007; Sharma et al., 2016; Yang et al., 2012). Apart from Arabidopsis, several other R2R3-MYBs from different plants, similar to the members of subgroup 7, are also involved in flavonoid pathway (Adato et al., 2009). We functionally characterized the role of this PGT-specific R2R3-MYB and found it to be a negative regulator of monoterpene production in spearmint. Suppression of *MsMYB* expression, by RNAi approach, led to increase in monoterpene levels and its overexpression led to decrease in monoterpene levels in transgenic lines. This is the first report of regulation of monoterpene production by a R2R3-MYB. Ectopic expression of *MsMYB* in sweet basil (*Ocimum basilicum*), an aromatic plant similar to spearmint, and in tobacco (*Nicotiana sylvestris)*, affected the production of

This article is protected by copyright. All rights reserved.

Accepted Article

terpenes but not of flavonoids. We further found that MsMYB suppresses the expression of geranyl diphosphate synthase large subunit *(MsGPPS.LSU*) gene in spearmint. To investigate if *MsMYB* is regulated by miR858, similar to other R2R3-MYBs of subgroup 7, we performed small RNA-Seq of spearmint PGT and L-PGT. We found low amounts of miR858 in spearmint PGT but high amounts in L-PGT. Although *in silico* analysis predicted *MsMYB* as a target for miR858, we were unable to detect its cleaved products. This suggests that the differential expression of miR858 and *MsMYB* in separate tissues prevents *MsMYB* regulation by miR858. Collectively, the above results provide insight into the production of essential oils in a valuable crop-plant, like spearmint, and expand our knowledge of R2R3-MYB-mediated regulation of plant secondary metabolism.

Results

MsMYB shows preferential expression in spearmint PGTs

From the transcriptome data of three tissues of spearmint namely leaves, PGTs, and leaves devoid of PGTs (Jin et al., 2014), we annotated 45 R2R3-MYBs (Figure 1; Supporting Data S1). Out of these only two R2R3-MYBs, *MsMYB* and *MsMYB112* showed high expression in PGTs. *MsMYB* was chosen for further characterization. RNA-Seq expression of *MsMYB* along various tissues was further validated by quantitative RT-PCR (qRT-PCR) (Figure 2a). Full length open reading frame (ORF) of *MsMYB* was amplified from PGT cDNA and consisted of 813 base pairs (bp) encoding a polypeptide of 271 amino acids (Figure 3). It contained a R2, R3 repeat and five tryptophan residues within the

This article is protected by copyright. All rights reserved.

Accepted Article

R2 and R3 repeats which together forms a helix-turn-helix motif at the N-terminus. These features are generally conserved among all the R2R3-MYBs (Dubos et al., 2010). MsMYB showed highest sequence similarity (~60%) to *Salvia miltiorrhiza* MYB-related transcription factor in BLAST analysis. A phylogenetic tree was constructed using the amino acid sequences of known *Arabidopsis thaliana* (At) R2R3-MYBs (Figure S2). MsMYB fell under subgroup 7 along with AtMYB111, AtMYB11 and AtMYB12 (Dubos et al., 2010). All these three Arabidopsis MYBs function as activators of flavonoid pathway (Stracke et al., 2007). To examine subcellular localization of MsMYB, the ORF was fused with yellow fluorescent protein (YFP) and expressed under the control of CaMV 35S promoter in *Nicotiana benthamiana* leaves by agro infiltration. As shown in Figure 2b, the recombinant protein specifically localized to the nucleus.

MsMYB **promoter shows trichome specific expression**

To identify the *cis*-acting regulatory elements in the promoter of *MsMYB,* we cloned a 615-bp genomic DNA fragment upstream of the translation start site by genome walking. Apart from the common TATA and CAAT box, several other *cis*-acting regulatory elements were identified (Figure 4a) using PlantCARE tool (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). Many of these elements were for hormones like ABRE-motif which is known for abscisic acid responsiveness (Basu et al., 2014), AuxRE-motif which is a auxin-responsive element (Ulmasov et al., 1995), CGTCA-motif and TGACG-motif which are known for MeJA-responsiveness (Zhou et al., 2012). Interestingly an AC-II element was also found. In general, AC-rich regions are known to be bound by R2R3-MYBs (Koschmann et al., 2012; Prouse and Campbell, 2012). This

This article is protected by copyright. All rights reserved.

Accepted Article

suggests that *MsMYB* might be regulated by other R2R3-MYBs. To check for the expression pattern, the 615-bp promoter fragment was fused with a β-glucuronidase (GUS) reporter gene and transformed into *N. benthamiana* plants. Trichome specific staining was observed in leaves and stems of tobacco transgenic plants (Figure 4b,c).

Manipulation of *MsMYB* expression levels affects monoterpene production in spearmint

To characterize the function of *MsMYB,* we generated transgenic spearmint lines harbouring RNAi construct and overexpression construct where *MsMYB* was under the control of CaMV 35S promoter. Four independent lines each confirmed by southern blot were selected for further characterization (Figure S1b,c). qRT-PCR analysis showed significant reduction in levels of *MsMYB* transcripts in RNAi lines (Figure 5a) and higher levels of *MsMYB* transcripts in overexpression lines when compared to wild type (WT) (Figure 5b). Both overexpression and RNAi transgenic plants appeared phenotypically similar to WT plants. Scanning electron microscopy was performed on these plants to take a closer look at leaf cells and PGTs. No phenotypical changes were observed.

Further, gas chromatography-mass spectrometry (GC-MS) analysis was performed to evaluate any changes in volatiles production. Major secondary metabolites of spearmint are monoterpenes; limonene and carvone (Figure S3a). GC-MS analysis revealed significant changes in the amount of these secondary metabolites in transgenic lines. Compared to WT plants, total

This article is protected by copyright. All rights reserved.

Accepted Article

monoterpene production in *MsMYB*-RNAi lines was higher; about 2.3 – 4.5 fold increase was observed (Figure 5c,d). Whereas, in *MsMYB*-overexpression lines a decrease of 0.5 - 0.7 fold was seen (Figure 5e,f). The above results indicated that *MsMYB* might be a negative regulator of secondary metabolism in spearmint.

MsMYB negatively regulates *MsGPPS.LSU*

To understand how MsMYB affects terpene metabolism, it is essential to know its downstream target genes. Towards this, we checked the expression levels of several transcripts encoding enzymes in various terpene precursor pathways, transporters and flavonoid pathway (Table S2), but none of them were significantly changed except for *MsGPPS.LSU*. It showed increased levels in *MsMYB*-RNAi lines and decreased levels in *MsMYB*-overexpressing lines (Figure 6a,b). GPPSs are known to localize to both non-green plastids and chloroplasts of photosynthetic cells where MEP pathway is active (Bouvier et al., 2000). We checked the localization of MsGPPS.LSU and found it to be localized in the chloroplast (Figure 9b). To check if the promoter of *MsGPPS.LSU* has putative MYB binding sites we cloned a 549-bp genomic DNA fragment upstream of the translation start site by genome walking. Using PlantCARE tool we identified two MYB binding sites (MBS) along with several elements for hormones and light responsiveness (Figure 7a).

To verify MsMYB interaction with *GPPS.LSU* promoter, yeast one-hybrid assays were performed. The whole *GPPS.LSU* promoter along with two truncated regions viz. site1 and site 2 which had the conserved MYB binding

This article is protected by copyright. All rights reserved.

Accepted Article

domains and mutated versions of site1 and site2 were used as bait to generate five Y1HGold [Bait/AbAi] strains i.e. pAbAi-GPPS-full, pAbAi-site1, pAbAi-site2, pAbAi-mut-site1 and pAbAi-mut-site2. After co-transformation with functional *MsMYB* gene in pGADT7 vector (pGADT7-MYB), three Y1HGold strains i.e. pAbAi-GPPS-full, pAbAi-site1 and pAbAi-site2 could grow on SD minus leucine with Aureobasidin (SD/-Leu/AbA, 800ng of AbA) auxotrophic medium but not the two mutant strains pAbAi-mut-site1 and pAbAi-mut-site2, confirming interaction of MsMYB with *GPPS.LSU* promoter *in vivo* (Figure 8). TFs can act as activators or repressors of transcription. Transgenic studies indicated that *MsMYB* is a negative regulator. To further determine the transcriptional repression activity of MsMYB in planta, transient expression assay in *N. benthamiana* was performed. We used $35S_{pro}$*:GFP* or $35S_{pro}$*:MsMYB* as effectors and $MsGPPS.LSU_{pro}$*:GUS* as reporter. Promoter activity of $MsGPPS.LSU_{pro}$*:GUS* was significantly reduced in $35S_{pro}$*:MsMYB* expressing leaves when compared to $35S_{pro}$*:GFP* expressing leaves (Figure 9a). These results show that *MsMYB* binds to the promoter of *MsGPPS.LSU* and suppresses its activity.

Terpene biosynthesis is known to be regulated by light in many plants (Cordoba et al., 2009). As we found light responsive elements in *MsGPPS.LSU* promoter, we checked if *MsGPPS.LSU* is also regulated by light. WT spearmint plants were grown in dark and light conditions. Leaf samples were collected at 0, 24, 48 and 72hrs interval to quantify *MsGPPS.LSU* transcript levels. Expression of *MsGPPS.LSU* was highly reduced in the plants grown under dark conditions suggesting that *MsGPPS.LSU* is regulated by light (Figure 7b). As GPPS is a

This article is protected by copyright. All rights reserved.

Accepted Article

key enzyme in the monoterpene pathway, multiple layers of gene regulation to fine tune the pathway is plausible.

Ectopic expression of *MsMYB* affects secondary terpene metabolism in sweet basil and *Nicotiana sylvestris*

Similar to spearmint, sweet basil too produces its essential oil in PGTs. The constituents of sweet basil essential oil are derived from terpene and phenylpropanoid pathways (Figure S3b). It is a mixture of monoterpenes, sesquiterpenes and the main phenylpropanoid derivative is eugenol. Ectopic expression of *MsMYB* in sweet basil was pursued to evaluate if *MsMYB* can affect sesquiterpene pathway and phenylpropanoid pathway as suggested by phylogenetic analysis. Three independent lines (line 5, line 7 and line 10) confirmed by southern blot were further selected for characterization (Figure S1d). Ectopic expression of *MsMYB* in these lines was confirmed by qRT-PCR (Figure 10a). GC-MS analysis on T-2 plants revealed an overall 73-85% decrease in total terpenes, both monoterpenes and sesquiterpenes (Figure 10b,c,d). However, the amount of eugenol derived from phenylpropanoid pathway remained unaltered (Figure 10e). Additionally, the leaves of transgenic lines were smaller in size when compared to WT (Figure S4a,b). From the NCBI database we obtained the sequences of sweet basil geranyl diphosphate synthase large subunit (*ObGPPS.LSU*) and farnesyl diphosphate synthase (*ObFPPS*). To investigate whether similar to spearmint, the decrease in terpenes in sweet basil is due to the reduction in transcripts of *ObGPPS.LSU* or because of reduction in *ObFPPS* transcript, we measured their levels by qRT-PCR. When compared to WT, *ObGPPS.LSU* expression was significantly

This article is protected by copyright. All rights reserved.

Accepted Article

reduced in transgenic plants (Figure 6c) but levels of *ObFPPS* were unaltered. Additionally, we measured the total flavonoid content in the leaves of sweet basil transgenic plants and found the total flavonoid content to be unaltered (Figure 10f). This suggests that MsMYB is a R2R3-MYB specific to terpene metabolism.

The sweet basil data showed that *MsMYB* can affect both monoterpene and sesquiterpene synthesis. In order to investigate whether it has a role in diterpene synthesis, we ectopically expressed *MsMYB* in *N. sylvestris*. Glandular trichomes of *N. sylvestris* mainly produce diterpenes; cembranoids (CBT-diol) which are generally derived from the same MEP pathway as the monoterpenes. Three independent lines were chosen for further characterization and ectopic expression of *MsMYB* was confirmed by qRT-PCR (Figure S5a). Analysis of CBT-diols production in transgenic plants showed a 42-50% reduction (Figure S5b). Similar to sweet basil plants, the leaves of tobacco transgenic lines were smaller in size when compared to WT (Figure S4c). The above results indicate that *MsMYB* can affect the flux in various branches of terpene metabolism

MsMYB is probably not a target of microRNA858

MicroRNA858 is known to regulate several R2R3-MYBs, specifically the subgroup 7 R2R3-MYBs, *AtMYB11, AtMYB111* and *AtMYB12* (Fahlgren et al., 2007; Sharma et al., 2016). As, MsMYB fell under subgroup 7 we were interested to know if it is regulated by microRNA858. We generated small RNA sequencing data from PGTs and L-PGT and identified ms-miR858. The

This article is protected by copyright. All rights reserved.

Accepted Article

sequence similarity of ms-miR858 is shown in Figure 11a. Expression analysis from the sequencing data showed that ms-miR858 had higher expression in L-PGT tissue when compared with PGT. This was validated by TaqMan® qPCR assay using miR396 as internal control due to its stable expression in L-PGT and PGT tissues (Figure 11b). The Plant Small RNA Target Analysis Server (Dai and Zhao, 2011) predicted *MsMYB* as a probable target of ms-miR858 (Figure 11c). To test if low levels of miR858 expression can target *MsMYB*, we looked for cleavage products through RNA ligase-mediated 5' amplification of cDNA ends (RLM-RACE) and Poly(A) Polymerase-Mediated (PPM)-RACE. However, no cleavage products of *MsMYB* were found in our conditions, thus reducing the possibility of it being regulated by miR858. *MsMYB* and miR858 appears to be differentially regulated at transcriptional level making their interaction less feasible.

Discussion

Extensive exploitation of aromatic plants for their commercially valued secondary metabolites calls for novel strategies for metabolic engineering to increase yield. A complete understanding of the secondary metabolism pathway by unravelling the enzymes and TFs controlling the pathway is indispensable for the same. PGTs are found on the aerial parts of many aromatic plants but very few studies have focused on identifying regulators of secondary metabolite production from these organs (Wang, 2014). In this study, we identified and characterized a R2R3-MYB gene, *MsMYB* from spearmint which is highly

This article is protected by copyright. All rights reserved.

Accepted Article

expressed in PGTs and acts as a negative regulator of monoterpene synthesis. GPPS enzyme is responsible for the formation of monoterpene precursor C_{10}-GPP in plastids and regulates flux into this pathway. Details about their function and regulation in plants are very scarce. The heteromeric GPPS in spearmint consists of a small and large subunit and interaction between them results in an active GPPS. Structural and mutagenic studies of these two subunits have shown that the large subunit is catalytically active and presumably through protein-protein interactions, the SSU shapes the active site cavity of LSU towards producing C_{10}-GPP and restricts the production of long chain C_{15} or C_{20} prenylphosphates (Chang et al., 2010). In spearmint, both the subunits of GPPS are preferentially expressed in PGTs (Turner and Croteau, 2004). MsMYB was found to interact with *MsGPPS.LSU* promoter and repress its activity. The increased production of monoterpene in *MsMYB*-RNAi lines can be attributed to enhanced production of GPP through greater production of MsGPPS.LSU. Misexpression of either the regulatory or catalytic subunit are known to alter enzyme activity in plants and animals (Gergs et al., 2004; He et al., 2004; Hu et al., 2016; Tiessen et al., 2002). In snapdragon and hop where heteromeric GPPS are found, only GPPS.SSU is closely correlated to monoterpene production whereas the GPPS.LSU is more ubiquitous. But spearmint *GPPS.LSU* expression is more PGT specific (Jin et al., 2014). Unlike the purified mint GPPS.LSU, which is inactive by itself, the potential GPPS.LSU from snapdragon and hops are active alone and functions as GGPPS (Tholl et al., 2004; Wang and Dixon, 2009). This suggests function of GPPS subunits from different plants might vary and also the mechanisms that regulate them.

This article is protected by copyright. All rights reserved.

Accepted Article

Ectopic expression of *MsMYB* in sweet basil and tobacco revealed that it could also perturb metabolites derived from sesquiterpene and diterpene pathways. Sesquiterpenes and diterpenes are derived from C_{15}-FPP and C_{20}-GGPP which are produced in cytoplasm and plastids through MVA and MEP pathway respectively. Studies have demonstrated that exchange of precursor metabolites can occur between the cytosolic MVA and plastidial MEP pathway. Metabolic intermediates like IPP, GPP and GGPP can be transported across plastidial membranes through transporters (Hemmerlin et al., 2012; Vranova et al., 2013). FPPS catalyses the sequential condensation of two molecules of IPP with DMAPP to form C_{15}-FPP the immediate precursor for all sesquiterpenes (Cornish, 1993; Lange et al., 2000). But FPPS is also know to accept C_{10}-GPP as an allylic substrate to generate FPP (Guo et al., 2015; Hemmerlin et al., 2003). Further sesquiterpenes synthesized in cytoplasm can be derived from plastidial GPP too (Adam et al., 1999). Hence an alteration in GPP produced in plastids can have an effect on sesquiterpene production. Ectopic expression of *MsMYB* in sweet basil resulted in reduced expression of *ObGPPS.LSU*. The decrease in sesquiterpene production observed in sweet basil indicates that GPP produced might be used towards synthesizing sesquiterpene especially since no change was observed on sweet basil FPPS. The monoterpene and diterpene pathway are both largely localized to plastids. GGPPS is responsible for producing GGPP for diterpenes (C_{20}) and tetraterpenes (C_{40}) synthesis. Apart from IPP and DMAPP, *in vitro* data show that GPP or FPP can also be used by GGPPS as allylic substrates to generate GGPP (Hefner et al., 1998; Okada et al., 2000; Takaya et al., 2003). The decrease in diterpenes produced

This article is protected by copyright. All rights reserved.

Accepted Article

in *MsMYB* overexpression tobacco plants can be due to a reduction in endogenous pool of GPP. Tobacco generally produces low levels of monoterpenes and there is no information currently about endogenous GPPS from it. Although introduction of monoterpene synthases from other plants does result in the emission of volatile monoterpenes in tobacco, suggesting some amount of GPP availability (Lücker et al., 2004). Apart from alteration in terpene derived secondary metabolites, overexpression of *MsMYB* in tobacco and sweet basil generated smaller leaves. Several major plant hormones are produced through terpene pathway such as gibberellin, cytokinin, abscisic acid and strigolactone which control various aspects of plant development (Gomez-Roldan et al., 2008; Umehara et al., 2008). Further experiments need to be done to identify perturbation in which hormone/hormones leads to smaller leaf size. This phenotype was not observed in spearmint. TFs can have additional targets when ectopically expressed. For example maize ZmMYB31 cannot bind to the promoter of maize *4CL* and *C3H* genes; however, it can repress the expression of Arabidopsis *4CL* and *C3H* when overexpressed in Arabidopsis (Fornalé et al., 2010). There is a possibilty that additional target genes related to primary terpene metablism are regulated when MsMYB is overexrpessed heterologously in these plants. Ectopic expression of *MsMYB* did not lead to any observable defect in chlorophyll formation which is a primary metabolite derived from diterpenes. An earlier study showed that GGPP used in chlorophyll biosynthesis is formed directly from DMAPP and IPP rather than from GPP and IPP. Thus any alterations in the amount of GPP formed does not affect chlorophyll biosynthesis (Van Schie et al., 2007).

This article is protected by copyright. All rights reserved.

Accepted Article

TFs are known to regulate multiple genes. *MsGPPS.LSU* is probably one of the many downstream targets regulated by MsMYB. Hence we cannot negate the fact that effects seen on terpene metabolism in spearmint, sweet basil and tobacco can be due to other genes as well. Transcriptome analysis of *MsMYB*-RNAi lines can provide us with more candidate genes that are potentially regulated by *MsMYB*. We found that *MsGPPS.LSU* is further regulated by light. Several genes of MEP pathway are known to be regulated by light (Cordoba et al., 2009). Secondary metabolite production is mainly for plants biotic and abiotic stress responses, so stringent multiple regulatory mechanisms involving activators and repressors must exist to determine the timing, amount and patterning of these compounds.

Majority of R2R3-MYBs characterized in plants are involved in the regulation of different classes of phenylpropanoid derived compounds (Liu et al., 2015). Phylogenetic analysis showed that MsMYB is very similar to subgroup 7 of Arabidopsis R2R3-MYBs which include MYB11, MYB12, and MYB111. These MYBs have been characterized as activators of biosynthetic enzymes of flavonoid pathway which is derived from the general phenylpropanoid pathway. Analysis of major flavonoid pathway genes from spearmint showed no change in *MsMYB* transgenic lines. Further no change was seen in either eugenol production which is derived from phenylpropanoid pathway or in total flavonoid content of leaves from sweet basil *MsMYB* overexpression lines. These results suggest *MsMYB* as a regulator of terpene metabolism and not affecting flavonoid production. This indicates towards diversification of function within the same subgroup. Though the N-terminus of R2R3-MYBs are conserved, C-

This article is protected by copyright. All rights reserved.

Accepted Article

terminus amino acids are very diverse and unique providing opportunity for divergence. Comparison of MYB proteins from different species will furnish new insights into the evolution of this important family of TF. In the past few years numerous studies have demonstrated miR858 can target R2R3-MYBs from various plants including the members of subgroup 7 of Arabidopsis R2R3-MYBs (Addo-Quaye et al., 2008; Rock, 2013; Sharma et al., 2016; Xia et al., 2012). Small RNA-Seq data of spearmint PGTs and L-PGTs showed an inverse pattern of expression for miR858 and *MsMYB*. High expression was seen in L-PGTs and low in PGTs for miR858 which is opposite to *MsMYB* expression, indicating a differential transcriptional regulation. *MsMYB* promoter is able to drive trichome specific expression as seen in tobacco. As genome of spearmint is not available, we cannot verify the promoter activity of miR858. Although *in silico* analysis predicts *MsMYB* to be a potential target for miR858, no cleavage products could be observed. This suggests that the PGT specific expression of *MsMYB* is governed mainly by the *cis*-elements present in its promoter and may be by other *trans*-factors but not by miR858.

In conclusion, *MsMYB* is the first R2R3-MYB gene identified from spearmint that is related to terpene secondary metabolism. MsMYB represses *MsGPPS.LSU* and is able to modify terpene production. This opens up new avenues and targets to metabolically engineer plants for altering pathways.

This article is protected by copyright. All rights reserved.

Accepted Article

Experimental procedures

Plant material and transformation

Secondary metabolites of commercial spearmint variety (*M. spicata*) and sweet basil (*O. basilicum*) were analysed by GC-MS and grown in green house under natural light conditions. *Agrobacterium* mediated transformation of spearmint was performed according to previously published protocol (Niu et al., 2000; Niu et al., 1998) with several modifications. The complete details of transformation are provided in supplementary data. *Agrobacterium* mediated transformation of sweet basil was performed as previously described by (Wang et al., 2016). Tobacco transformation was done as described previously (Gallois and Marinho, 1995).

RNA isolation and Quantitative real time PCR (qRT-PCR)

For PGT RNA isolation, initially, PGTs were isolated from 2-3 cm leaves as described previously (Jin et al., 2014). Later, total RNA was extracted from PGT using the Spectrum Plant total RNA kit from Sigma (UK). Total RNA from other tissues was extracted using an RNeasy® Plus Mini kit from Qiagen (Germany). 500ng of RNA was reverse transcribed to cDNA using iScript™ cDNA Synthesis kit form Bio-Rad (CA, USA). Expression levels of genes along various tissues were analysed using qRT-PCR that was carried out as described previously (Jin et al., 2014). In current study, elongation factor 1 (*ef1*) was used as internal control for both spearmint and sweet basil, due to its stable expression in plant (Nicot et al., 2005). For tobacco, elongation factor TuB (EF-TuB) was used as internal control. Error bars represent mean ± SE.

This article is protected by copyright. All rights reserved.

Accepted Article

TaqMan® qPCR assay

7ng of total RNA was reverse transcribed using the TaqMan® microRNA reverse transcription kit (Applied Biosystems, USA) according to manufacturer's protocol. Custom made primers were used specifically for miR858 and miR396 cDNA synthesis. Custom TaqMan® small RNA assays (Applied Biosystems, USA) were done for microRNA qPCR validation using custom made TaqMan® probe for miR858 and miR396. Expression values were calculated relative to the internal control miR396.

Promoter cloning and analysis

Genomic DNA was isolated from young leaves of spearmint using CTAB method. The flanking sequences of *MsMYB* and *MsGPPS.LSU* genes were amplified using a GenomeWalker™ Universal kit (Clontech, USA) and later ligated to pGEM®-T vector. The resulting product was transformed into *Escherichia coli* (*E. coli*) XL1-Blue cells and sequenced. The promoter was amplified with Phusion® High-Fidelity DNA Polymerase (NEB, USA) and subcloned into a gateway donor vector pENTR™/D-TOPO® (Invitrogen, Germany). Further, the recombinant plasmid was introduced into destination vector pBGWFS7 by LR recombination. The destination plasmid was further transformed into *Agrobacterium* EHA105 by heat shock. The recombinant *Agrobacterium* EHA105 strain was used to generate transgenic tobacco lines.

Gene amplification and plasmid construction

Full length ORF encoding MsMYB was obtained by performing 3' and 5' RACE using the SMARTer™ RACE cDNA amplification kit from Clontech (CA, USA).

This article is protected by copyright. All rights reserved.

Accepted Article

For sequencing of ORF, purified fragments were ligated with pGEM®-T vector and transformed into *E. coli* XL1-Blue cells. To overexpress *MsMYB*, firstly, the sequences were amplified with Phusion® High-Fidelity DNA Polymerase (NEB, USA). For *MsMYB* RNAi, four primers with restriction enzymes located at flanking region were used to amplify the fragment showing low similarity to other *MYB* genes. The purified PCR products were then cloned into the donor vector and subsequently introduced into the destination vector pK7WG2D via LR recombination. The *MsMYB* gene was driven by 35S promoter in both overexpression and RNAi plants. All destination plasmids harbouring target gene were transformed into *Agrobacterium* EHA105 by heat shock. The recombinant *Agrobacterium* EHA105 strains were used for plant transformation. Sequences of primers used in this study are listed in Table S1.

Subcellular localization of TFs

The full-length cDNA of *MsMYB* without the stop codon was cloned into the gateway vector pENTR/D-TOPO (Invitrogen, Germany), and then subsequently transferred into the destination vector pBA-DC-YFP (Zhang et al., 2005) which contains the CaMV 35S promoter and C-terminal in frame with YFP, to generate *MsMYB-YFP*. The construct was then introduced into *Agrobacterium* strain EHA105 by heat shock. 4′, 6-diamidino-2-phenylindole (DAPI) was used as maker to stain nucleus. *N. benthamiana* seeds were germinated on MS plate and transferred into soil. 3 weeks after growing in the greenhouse, the seedlings were used for agro-infiltration. Subcellular localization was performed as described by (Wang et al., 2016).

This article is protected by copyright. All rights reserved.

Accepted Article

Selection of transgenic lines

Initially, visual screening using GFP filter was pursued to isolate GFP-positive plants. Later, DNA was isolated from GFP-positive plants and insertion of gene of interest was checked using PCR. DNA-positive lines were then subjected to southern blot using DIG wash and block buffer set from Roche (IN, USA). DNA probe against CaMV 35S promoter was generated using PCR DIG probe synthesis kit from Roche (IN, USA) (Hart and Basu, 2009). Concentration of probe was quantified by doing a Dot plot using DIG nucleic acid detection kit from Roche (IN, USA) as described previously (Javelle et al., 2011). A total of 15 µg genomic DNA was digested overnight with *NdeI* at 37°C. Next day, digested product was electrophoresed on a 1% (w/v) agarose gel at 40 V for 5 h. After that, the gel was transferred to a nylon membrane and hybridized with CaMV 35S promoter probe followed by antibody binding and chemiluminescent reaction development by CDP star. Finally the membrane was exposed in a film and the number of T-DNA insertions were analysed. Spearmint plants were clonally propagated thrice before analysis. Basil plants were analysed in T2 generation and a minimum of 6 plants were sampled for each line. Tobacco plants were analysed in T1 generation. For each line, ~32 plants were used for phenotypical screening and 6 plants for GC-MS analysis.

GC-MS and total flavonoid content analysis

For GC-MS studies, spearmint plants were clonally propagated three times before collecting about 4-6 leaves of 2-3 cm which were ground to a fine powder using liquid nitrogen and homogenised using 500 µl ethyl acetate.

This article is protected by copyright. All rights reserved.

Accepted Article

Camphor was added as an internal control. For sweet basil, the analysis was performed in T-2 plants. From bottom, leaves of 3-4 cm at 4th node were used. For tobacco, 4-5 cm leaves were used from T1 plants. Diethyl sebacate was added as an internal standard in sweet basil and tobacco samples. Samples were incubated for 10 min at room temperature with vigorous shaking followed by a centrifugation for 10 min at 13,000 rpm. The top organic layer was transferred to a new tube and dehydrated using anhydrous Na_2SO_4. The samples were analysed using GC-MS (7890A with 5975C inert MSD with triple axis detector, Agilent Technologies, USA). 2 µl of samples were injected and separation was achieved with a temperature program of 50°C for 1 min and increased at a rate of 8°C/min to 300°C and held for 5 min, on a 30 m HP-5 MS column (Agilent Technologies, USA). Estimation of total flavonoid content was done using the aluminium chloride ($AlCl_3$) colorimetric method as described previously (Vaidya. et al., 2014).

Yeast one-hybrid assay

For ascertaining the interaction between MsMYB and *MsGPPS.LSU* promoter, a Matchmaker Gold yeast one-hybrid library screening system (Clontech, USA) was used. As *MsGPPS.LSU* promoter encompasses two different MYB binding sites, therefore, two promoter regions approximately 40-45 bp in size harbouring the two sites namely site 1 and site 2 were cloned into the MCS of pAbAi vector independently in addition to two mutant versions of these two sites. Furthermore, complete promoter was also cloned in pAbAi vector and used as bait for further confirmation. The *MsMYB* gene was cloned into the pGADT7 vector to be used as prey in the form of functional fusion protein along

This article is protected by copyright. All rights reserved.

Accepted Article

with GAL4 transcription activation domain (GAL4 AD). As a positive control, p53 bait and prey provided in the kit was used. Mutated site 1 and mutated site 2 were used as negative control. Yeast transformation and validation of positive interactions were implemented as described in the Matchmaker Gold Yeast one-hybrid system user manual (www.clontech.com/xxclt_ibcGetAttachment.jsp?cItemId=17599). The primers used are listed in Table S3.

Transactivation activity assay

The 549-bp promoter region of *MsGPPS.LSU* was amplified and inserted into pENTR™/D-TOPO®. The resulting plasmid was transformed into pBGWFS7 by LR recombination and further introduced into *Agrobacterium* EHA105. Leaves of *N. benthamiana* were agroinfiltrated with effector and reporter at a ratio of 1:1. Two days after infiltration, leaves were harvested to isolate crude protein. GUS quantitative assay was performed in triplicates as described in (Li et al., 2014).

RLM-RACE and PPM-RACE

For RLM-RACE 5' RACE adapter was ligated to 1 µg of total RNA and for PPM-RACE, poly(A) tailing of 1 µg of total RNA was done using poly(A) polymerase (Invitrogen, Germany). RNA was then reverse transcribed and amplified for 5' and 3' ends using FirstChioce RLM-RACE kit (Thermo Fisher, USA) according to manufacturer's protocol.

This article is protected by copyright. All rights reserved.

Accepted Article

Small RNA library construction and high-throughput sequencing

The RNA libraries were prepared using the TruSeq RNA library prep Kit v2, set A (Illumina Inc., USA) according to manufacturer's instructions. The quality and size of DNA libraries for sequencing were checked using the Agilent 2200 TapeStation system. Three libraries were run on single lanes on HiseqTM 2000 (Illumina Inc., USA), individually.

Bioinformatics analysis of sequencing data

After removing adapters by Cutadapt tool (Martin, 2011), modified sequences from 18 nt to 30 nt were used for further analysis. To begin with, sequences of rRNAs, tRNAs, snRNAs and snoRNAs available in Rfam11.0 database (ftp://ftp.ebi.ac.uk/pub/databases/Rfam/11.0/) were removed from the small RNA sequence reads. Deposited miR858 sequences from miRbase database (http://www.mirbase.org/search.shtml) were mapped to the remaining small RNA sequence reads by bowtie tool (Langmead and Salzberg, 2012) to identify ms-miRNA858. Maximum one mismatch per sequence was set as a parameter.

Phylogenetic analysis

Phylogenetic tree was constructed using MEGA6 software by Neighbour-joining method with bootstrap values of 1000 replicates and edited in FigTree. A. *thaliana* R2R3-MYB sequences were obtained from TAIR website.

Statistical analysis

Data are indicated as “mean ± SE” of three to six biological replicates each performed in triplicates. Statistical significance between transgenic plants and

This article is protected by copyright. All rights reserved.

Accepted Article

WT was analysed using a two-tailed Student's *t*-test and indicated by asterisks. * indicates $p < 0.05$; ** indicates $p < 0.01$; *** indicates $p < 0.001$.

Accession Numbers

Sequence data of *A. thaliana* R2R3-MYBs used in Figure S2 can be found in TAIR library under accession numbers listed in supplemental data set 2. Sequence data of *ObGPPS.LSU* and *ObFPPS* can be found in NCBI database under the sequence ID DY340136 and DY332783 respectively. Sequence data of *MsMYB* has been deposited in GenBank under the accession number KY081780.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This research was funded by a grant from Singapore National Research Foundation (Competitive Research Programme Award No: NRF-CRP8-2011-02).

References

Adam, K.-P., Thiel, R. and Zapp, J. (1999) Incorporation of 1-[1-13C]Deoxy-d-xylulose in Chamomile Sesquiterpenes. *Archives of Biochemistry and Biophysics* **369**, 127-132.

This article is protected by copyright. All rights reserved.

Accepted Article

Adato, A., Mandel, T., Mintz-Oron, S., Venger, I., Levy, D., Yativ, M., Dominguez, E., Wang, Z.H., De Vos, R.C.H., Jetter, R., Schreiber, L., Heredia, A., Rogachev, I. and Aharoni, A. (2009) Fruit-Surface Flavonoid Accumulation in Tomato Is Controlled by a SlMYB12-Regulated Transcriptional Network. *PLoS Genet.* **5**, 23.

Addo-Quaye, C., Eshoo, T.W., Bartel, D.P. and Axtell, M.J. (2008) Endogenous siRNA and miRNA targets identified by sequencing of the Arabidopsis degradome. *Current Biology* **18**, 758-762.

Basu, S., Roychoudhury, A. and Sengupta, D.N. (2014) Deciphering the Role of various cis-acting regulatory elements in controlling SamDC gene expression in Rice. *Plant Signaling & Behavior* **9**, e28391.

Bhargava, A., Ahad, A., Wang, S.C., Mansfield, S.D., Haughn, G.W., Douglas, C.J. and Ellis, B.E. (2013) The interacting MYB75 and KNAT7 transcription factors modulate secondary cell wall deposition both in stems and seed coat in Arabidopsis. *Planta* **237**, 1199-1211.

Bhargava, A., Mansfield, S.D., Hall, H.C., Douglas, C.J. and Ellis, B.E. (2010) MYB75 Functions in Regulation of Secondary Cell Wall Formation in the Arabidopsis Inflorescence Stem. *Plant Physiology* **154**, 1428-1438.

Bouvier, F., Rahier, A. and Camara, B. (2005) Biogenesis, molecular regulation and function of plant isoprenoids. *Progress in Lipid Research* **44**, 357-429.

Bouvier, F., Suire, C., d'Harlingue, A., Backhaus, R.A. and Camara, B. (2000) Molecular cloning of geranyl diphosphate synthase and compartmentation of monoterpene synthesis in plant cells. *Plant Journal* **24**, 241-252.

Burke, C. and Croteau, R. (2002) Interaction with the small subunit of geranyl diphosphate synthase modifies the chain length specificity of geranylgeranyl diphosphate synthase to produce geranyl diphosphate. *Journal of Biological Chemistry* **277**, 3141-3149.

Burke, C.C., Wildung, M.R. and Croteau, R. (1999) Geranyl diphosphate synthase: Cloning, expression, and characterization of this prenyltransferase as a heterodimer. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 13062-13067.

Champagne, A. and Boutry, M. (2013) Proteomic snapshot of spearmint (Mentha spicata L.) leaf trichomes: A genuine terpenoid factory. *PROTEOMICS* **13**, 3327-3332.

Chang, T.H., Hsieh, F.L., Ko, T.P., Teng, K.H., Liang, P.H. and Wang, A.H.J. (2010) Structure of a Heterotetrameric Geranyl Pyrophosphate Synthase from Mint (Mentha piperita) Reveals Intersubunit Regulation. *Plant Cell* **22**, 454-467.

Chen, Y.Y., Wang, L.F., Dai, L.J., Yang, S.G. and Tian, W.M. (2012) Characterization of HbEREBP1, a wound-responsive transcription factor gene in laticifers of Hevea brasiliensis Muell. Arg. *Molecular Biology Reports* **39**, 3713-3719.

Cordoba, E., Salmi, M. and Leon, P. (2009) Unravelling the regulatory mechanisms that modulate the MEP pathway in higher plants. *J. Exp. Bot.* **60**, 2933-2943.

This article is protected by copyright. All rights reserved.

Accepted Article

Cornish, K. (1993) The separate roles of plant cis and trans prenyl transferases in cis-1,4-polyisoprene biosynthesis. *European Journal of Biochemistry* **218**, 267-271.

Croteau, R., Karp, F., Wagschal, K.C., Satterwhite, D.M., Hyatt, D.C. and Skotland, C.B. (1991) Biochemical-characterization of a spearmint mutant that resembles peppermint in monoterpene content. *Plant Physiology* **96**, 744-752.

Croteau, R.B., Davis, E.M., Ringer, K.L. and Wildung, M.R. (2005) (–)-Menthol biosynthesis and molecular genetics. *Naturwissenschaften* **92**, 562.

Dai, X.B. and Zhao, P.X. (2011) psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Research* **39**, W155-W159.

Diemer, F., Caissard, J.C., Moja, S., Chalchat, J.C. and Jullien, F. (2001) Altered monoterpene composition in transgenic mint following the introduction of 4S-limonene synthase. *Plant Physiology and Biochemistry* **39**, 603-614.

Dubey, V.S., Bhalla, R. and Luthra, R. (2003) An overview of the non-mevalonate pathway for terpenoid biosynthesis in plants. *Journal of Biosciences* **28**, 637.

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C. and Lepiniec, L. (2010) MYB transcription factors in Arabidopsis. *Trends in Plant Science* **15**, 573-581.

Fahlgren, N., Howell, M.D., Kasschau, K.D., Chapman, E.J., Sullivan, C.M., Cumbie, J.S., Givan, S.A., Law, T.F., Grant, S.R., Dangl, J.L. and Carrington, J.C. (2007) High-Throughput Sequencing of Arabidopsis microRNAs: Evidence for Frequent Birth and Death of MIRNA Genes. *Plos One* **2**.

Fornalé, S., Shi, X., Chai, C., Encina, A., Irar, S., Capellades, M., Fuguet, E., Torres, J.-L., Rovira, P., Puigdomènech, P., Rigau, J., Grotewold, E., Gray, J. and Caparrós-Ruiz, D. (2010) ZmMYB31 directly represses maize lignin genes and redirects the phenylpropanoid metabolic flux. *The Plant Journal* **64**, 633-644.

Gallois, P. and Marinho, P. (1995) Leaf Disk Transformation Using Agrobacterium tumefaciens-Expression of Heterologous Genes in Tobacco. In: *Plant Gene Transfer and Expression Protocols* (Jones, H. ed) pp. 39-48. Springer New York.

Gergs, U., Boknik, P., Buchwalow, I., Fabritz, L., Matus, M., Justus, I., Hanske, G., Schmitz, W. and Neumann, J. (2004) Overexpression of the catalytic subunit of protein phosphatase 2A impairs cardiac function. *Journal of Biological Chemistry* **279**, 40827-40834.

Gomez-Roldan, V., Fermas, S., Brewer, P.B., Puech-Pages, V., Dun, E.A., Pillot, J.-P., Letisse, F., Matusova, R., Danoun, S., Portais, J.-C., Bouwmeester, H., Becard, G., Beveridge, C.A., Rameau, C. and Rochange, S.F. (2008) Strigolactone inhibition of shoot branching. *Nature* **455**, 189-194.

Grotewold, E. (2008) Transcription factors for predictive plant metabolic engineering: are we there yet? *Current Opinion in Biotechnology* **19**, 138-144.

This article is protected by copyright. All rights reserved.

Accepted Article

Guan, X.Y., Pang, M.X., Nah, G., Shi, X.L., Ye, W.X., Stelly, D.M. and Chen, Z.J. (2014) miR828 and miR858 regulate homoeologous MYB2 gene functions in Arabidopsis trichome and cotton fibre development. *Nature Communications* **5**.

Guo, D., Li, H.-L. and Peng, S.-Q. (2015) Structure Conservation and Differential Expression of Farnesyl Diphosphate Synthase Genes in Euphorbiaceous Plants. *International Journal of Molecular Sciences* **16**, 22402-22414.

Hart, S.M. and Basu, C. (2009) Optimization of a Digoxigenin-Based Immunoassay System for Gene Detection in Arabidopsis thaliana. *Journal of Biomolecular Techniques : JBT* **20**, 96-100.

He, X., Anderson, J.C., Pozo, O.d., Gu, Y.-Q., Tang, X. and Martin, G.B. (2004) Silencing of subfamily I of protein phosphatase 2A catalytic subunits results in activation of plant defense responses and localized cell death. *The Plant Journal* **38**, 563-577.

Hefner, J., Ketchum, R.E.B. and Croteau, R. (1998) Cloning and Functional Expression of a cDNA Encoding Geranylgeranyl Diphosphate Synthase fromTaxus canadensisand Assessment of the Role of this Prenyltransferase in Cells Induced for Taxol Production. *Archives of Biochemistry and Biophysics* **360**, 62-74.

Hemmerlin, A., Harwood, J.L. and Bach, T.J. (2012) A raison d'être for two distinct pathways in the early steps of plant isoprenoid biosynthesis? *Progress in Lipid Research* **51**, 95-148.

Hemmerlin, A., Rivera, S.B., Erickson, H.K. and Poulter, C.D. (2003) Enzymes encoded by the farnesyl diphosphate synthase gene family in the big sagebrush Artemisia tridentata ssp spiciformis. *Journal of Biological Chemistry* **278**, 32132-32140.

Hu, R., Zhu, Y., Wei, J., Chen, J., Shi, H., Shen, G. and Zhang, H. (2016) Overexpression of PP2A-C5 that encodes the catalytic subunit 5 of protein phosphatase 2A in Arabidopsis confers better root and shoot development under salt conditions. *Plant, Cell & Environment*, n/a-n/a.

Iwase, A., Matsui, K. and Ohme-Takagi, M. (2009) Manipulation of plant metabolic pathways by transcription factors. *Plant Biotechnology* **26**, 29-38.

Javelle, M., Marco, C.F. and Timmermans, M. (2011) In Situ Hybridization for the Precise Localization of Transcripts in Plants. *Journal of Visualized Experiments : JoVE*, 3328.

Jeong, D.H., Schmidt, S.A., Rymarquis, L.A., Park, S., Ganssmann, M., German, M.A., Accerbi, M., Zhai, J.X., Fahlgren, N., Fox, S.E., Garvin, D.F., Mockler, T.C., Carrington, J.C., Meyers, B.C. and Green, P.J. (2013) Parallel analysis of RNA ends enhances global investigation of microRNAs and target RNAs of Brachypodium distachyon. *Genome Biology* **14**.

Jia, X.Y., Shen, J., Liu, H., Li, F., Ding, N., Gao, C.Y., Pattanaik, S., Patra, B., Li, R.Z. and Yuan, L. (2015) Small tandem target mimic-mediated blockage of microRNA858 induces anthocyanin accumulation in tomato. *Planta* **242**, 283-293.

This article is protected by copyright. All rights reserved.

Accepted Article

Jin, H.L., Cominelli, E., Bailey, P., Parr, A., Mehrtens, F., Jones, J., Tonelli, C., Weisshaar, B. and Martin, C. (2000) Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. *Embo Journal* **19**, 6150-6161.

Jin, J., Panicker, D., Wang, Q., Kim, M.J., Liu, J., Yin, J.-L., Wong, L., Jang, I.-C., Chua, N.-H. and Sarojam, R. (2014) Next generation sequencing unravels the biosynthetic ability of Spearmint (Mentha spicata) peltate glandular trichomes through comparative transcriptomics. *Bmc Plant Biology* **14**.

Kellogg, B.A. and Poulter, C.D. (1997) Chain elongation in the isoprenoid biosynthetic pathway. *Current Opinion in Chemical Biology* **1**, 570-578.

Koschmann, J., Machens, F., Becker, M., Niemeyer, J., Schulze, J., Bülow, L., Stahl, D.J. and Hehl, R. (2012) Integration of bioinformatics and synthetic promoters leads to the discovery of novel elicitor-responsive cis-regulatory sequences in Arabidopsis. *Plant Physiol* **160**.

Kranz, H.D., Denekamp, M., Greco, R., Jin, H., Leyva, A., Meissner, R.C., Petroni, K., Urzainqui, A., Bevan, M., Martin, C., Smeekens, S., Tonelli, C., Paz-Ares, J. and Weisshaar, B. (1998) Towards functional characterisation of the members of theR2R3-MYBgene family fromArabidopsis thaliana. *The Plant Journal* **16**, 263-276.

Lange, B.M., Mahmoud, S.S., Wildung, M.R., Turner, G.W., Davis, E.M., Lange, I., Baker, R.C., Boydston, R.A. and Croteau, R.B. (2011) Improving peppermint essential oil yield and composition by metabolic engineering. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 16944-16949.

Lange, B.M., Rujan, T., Martin, W. and Croteau, R. (2000) Isoprenoid biosynthesis: The evolution of two ancient and distinct pathways across genomes. *Proceedings of the National Academy of Sciences* **97**, 13172-13177.

Lange, B.M. and Turner, G.W. (2013) Terpenoid biosynthesis in trichomes-current status and future opportunities. *Plant Biotechnology Journal* **11**, 2-22.

Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357-U354.

Legay, S., Lacombe, E., Goicoechea, M., Briere, C., Seguin, A., Mackay, J. and Grima-Pettenati, J. (2007) Molecular characterization of EgMYB1, a putative transcriptional repressor of the lignin biosynthetic pathway. *Plant Sci.* **173**, 542-549.

Li, C.L. and Lu, S.F. (2014) Genome-wide characterization and comparative analysis of R2R3-MYB transcription factors shows the complexity of MYB-associated regulatory networks in Salvia miltiorrhiza. *Bmc Genomics* **15**.

Li, R., Weldegergis, B.T., Li, J., Jung, C., Qu, J., Sun, Y.W., Qian, H.M., Tee, C., van Loon, J.J.A., Dicke, M., Chua, N.H., Liu, S.S. and Ye, J. (2014) Virulence Factors of Geminivirus Interact with MYC2 to Subvert Plant Resistance and Promote Vector Performance. *Plant Cell* **26**, 4991-5008.

This article is protected by copyright. All rights reserved.

Accepted Article

Li, S., Zhang, P., Zhang, M., Fu, C. and Yu, L. (2013) Functional analysis of a WRKY transcription factor involved in transcriptional activation of the DBAT gene in Taxus chinensis. *Plant Biology* **15**, 19-26.

Liang, P.-H. (2009) Reaction Kinetics, Catalytic Mechanisms, Conformational Changes, and Inhibitor Design for Prenyltransferases. *Biochemistry* **48**, 6562-6570.

Liang, P.-H., Ko, T.-P. and Wang, A.H.J. (2002) Structure, mechanism and function of prenyltransferases. *European Journal of Biochemistry* **269**, 3339-3354.

Liu, J., Osbourn, A. and Ma, P. (2015) MYB Transcription Factors as Regulators of Phenylpropanoid Metabolism in Plants. *Molecular Plant* **8**, 689-708.

Loza-Tavera, H. (1999) Monoterpenes in Essential Oils. In: *Chemicals via Higher Plant Bioengineering* (Shahidi, F., Kolodziejczyk, P., Whitaker, J.R., Munguia, A.L. and Fuller, G. eds), pp. 49-62. Boston, MA: Springer US.

Lücker, J., Schwab, W., van Hautum, B., Blaas, J., van der Plas, L.H.W., Bouwmeester, H.J. and Verhoeven, H.A. (2004) Increased and Altered Fragrance of Tobacco Plants after Metabolic Engineering Using Three Monoterpene Synthases from Lemon. *Plant Physiology* **134**, 510-519.

Mahmoud, S.S., Williams, M. and Croteau, R. (2004) Cosuppression of limonene-3-hydroxylase in peppermint promotes accumulation of limonene in the essential oil. *Phytochemistry* **65**, 547-554.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011* **17**.

Matus, J.T. (2016) Transcriptomic and Metabolomic Networks in the Grape Berry Illustrate That it Takes More Than Flavonoids to Fight Against Ultraviolet Radiation. *Frontiers in Plant Science* **7**.

Miyamoto, K., Matsumoto, T., Okada, A., Komiyama, K., Chujo, T., Yoshikawa, H., Nojiri, H., Yamane, H. and Okada, K. (2014) Identification of Target Genes of the bZIP Transcription Factor OsTGAP1, Whose Overexpression Causes Elicitor-Induced Hyperaccumulation of Diterpenoid Phytoalexins in Rice Cells. *Plos One* **9**.

Munoz-Bertomeu, J., Ros, R., Arrillaga, I. and Segura, J. (2008) Expression of spearmint limonene synthase in transgenic spike lavender results in an altered monoterpene composition in developing leaves. *Metabolic Engineering* **10**, 166-177.

Nagegowda, D.A. (2010) Plant volatile terpenoid metabolism: Biosynthetic genes, transcriptional regulation and subcellular compartmentation. *Febs Letters* **584**, 2965-2973.

Nicot, N., Hausman, J.F., Hoffmann, L. and Evers, D. (2005) Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *Journal of Experimental Botany* **56**, 2907-2914.

Niu, X., Li, X., Veronese, P., Bressan, R.A., Weller, S.C. and Hasegawa, P.M. (2000) Factors affecting Agrobacterium tumefaciens-mediated transformation of peppermint. *Plant Cell Rep.* **19**, 304-310.

Niu, X., Lin, K., Hasegawa, P.M., Bressan, R.A. and Weller, S.C. (1998) Transgenic peppermint (Mentha x piperita L.) plants obtained by

This article is protected by copyright. All rights reserved.

Accepted Article

cocultivation with Agrobacterium tumefaciens. *Plant Cell Reports* **17**, 165-171.

Ogura, K. and Koyama, T. (1998) Enzymatic Aspects of Isoprenoid Chain Elongation. *Chemical Reviews* **98**, 1263-1276.

Okada, K., Saito, T., Nakagawa, T., Kawamukai, M. and Kamiya, Y. (2000) Five Geranylgeranyl Diphosphate Synthases Expressed in Different Organs Are Localized into Three Subcellular Compartments in Arabidopsis. *Plant Physiology* **122**, 1045-1056.

Oldfield, E. and Lin, F.-Y. (2012) Terpene Biosynthesis: Modularity Rules. *Angewandte Chemie (International ed. in English)* **51**, 1124-1137.

Pichersky, E. and Gershenzon, J. (2002) The formation and function of plant volatiles: perfumes for pollinator attraction and defense. *Current Opinion in Plant Biology* **5**, 237-243.

Preston, J., Wheeler, J., Heazlewood, J., Li, S.F. and Parish, R.W. (2004) AtMYB32 is required for normal pollen development in Arabidopsis thaliana. *The Plant Journal* **40**, 979-995.

Prouse, M.B. and Campbell, M.M. (2012) The interaction between MYB proteins and their target DNA binding sites. *Biochim. Biophys. Acta* **1819**, 67-77.

Ringer, K.L., Davis, E.M. and Croteau, R. (2005) Monoterpene metabolism. Cloning, expression, and characterization of (-)-isopiperitenol/(-)-carveol dehydrogenase of peppermint and spearmint. *Plant Physiology* **137**, 863-872.

Rock, C.D. (2013) Trans-acting small interfering RNA4: key to nutraceutical synthesis in 1 grape development? *Trends in plant science* **18**, 10.1016/j.tplants.2013.1007.1006.

Sharma, D., Tiwari, M., Pandey, A., Bhatia, C., Sharma, A. and Trivedi, P.K. (2016) MicroRNA858 Is a Potential Regulator of Phenylpropanoid Pathway and Plant Development. *Plant Physiology* **171**, 944-959.

Shen, Q., Yan, T., Fu, X. and Tang, K. (2016) Transcriptional regulation of artemisinin biosynthesis in Artemisia annua L. *Science Bulletin* **61**, 18-25.

Singh, B. and Sharma, R.A. (2015) Plant terpenes: defense responses, phylogenetic analysis, regulation and clinical applications. *3 Biotech* **5**, 129-151.

Stracke, R., Ishihara, H., Huep, G., Barsch, A., Mehrtens, F., Niehaus, K. and Weisshaar, B. (2007) Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the Arabidopsis thaliana seedling. *The Plant Journal* **50**, 660-677.

Stracke, R., Werber, M. and Weisshaar, B. (2001) The R2R3-MYB gene family in Arabidopsis thaliana. *Current Opinion in Plant Biology* **4**, 447-456.

Sunkar, R., Li, Y.-F. and Jagadeeswaran, G. (2012) Functions of microRNAs in plant stress responses. *Trends in Plant Science* **17**, 196-203.

Takaya, A., Zhang, Y.-W., Asawatreratanakul, K., Wititsuwannakul, D., Wititsuwannakul, R., Takahashi, S. and Koyama, T. (2003) Cloning, expression and characterization of a functional cDNA clone encoding geranylgeranyl diphosphate synthase of Hevea brasiliensis. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1625**, 214-220.

This article is protected by copyright. All rights reserved.

Accepted Article

Tamagnone, L., Merida, A., Parr, A., Mackay, S., Culianez-Macia, F.A., Roberts, K. and Martin, C. (1998) The AmMYB308 and AmMYB330 transcription factors from antirrhinum regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. *Plant Cell* **10**, 135-154.

Tholl, D., Kish, C.M., Orlova, I., Sherman, D., Gershenzon, J., Pichersky, E. and Dudareva, N. (2004) Formation of monoterpenes in Antirrhinum majus and Clarkia breweri flowers involves heterodimeric geranyl diphosphate synthases. *Plant Cell* **16**, 977-992.

Tiessen, A., Hendriks, J.H.M., Stitt, M., Branscheid, A., Gibon, Y., Farré, E.M. and Geigenberger, P. (2002) Starch Synthesis in Potato Tubers Is Regulated by Post-Translational Redox Modification of ADP-Glucose Pyrophosphorylase: A Novel Regulatory Mechanism Linking Starch Synthesis to the Sucrose Supply. *The Plant Cell* **14**, 2191-2213.

Turner, G.W. and Croteau, R. (2004) Organization of Monoterpene Biosynthesis in Mentha. Immunocytochemical Localizations of Geranyl Diphosphate Synthase, Limonene-6-Hydroxylase, Isopiperitenol Dehydrogenase, and Pulegone Reductase. *Plant Physiology* **136**, 4215-4227.

Turner, G.W., Gershenzon, J. and Croteau, R.B. (2000) Distribution of Peltate Glandular Trichomes on Developing Leaves of Peppermint. *Plant Physiology* **124**, 655-664.

Ulmasov, T., Liu, Z.B., Hagen, G. and Guilfoyle, T.J. (1995) Composite structure of auxin response elements. *Plant Cell* **7**, 1611-1623.

Umehara, M., Hanada, A., Yoshida, S., Akiyama, K., Arite, T., Takeda-Kamiya, N., Magome, H., Kamiya, Y., Shirasu, K., Yoneyama, K., Kyozuka, J. and Yamaguchi, S. (2008) Inhibition of shoot branching by new terpenoid plant hormones. *Nature* **455**, 195-200.

Vaidya., N., B., Brearley., T.A. and Joshee., N. (2014) Antioxidant Capacity of Fresh and Dry Leaf Extracts of Sixteen Scutellaria Species. *Journal of Medicinally Active Plants* **2**.

Van Schie, C.C.N., Ament, K., Schmidt, A., Lange, T., Haring, M.A. and Schuurink, R.C. (2007) Geranyl diphosphate synthase is required for biosynthesis of gibberellins. *The Plant Journal* **52**, 752-762.

Vranova, E., Coman, D. and Gruissem, W. (2013) Network Analysis of the MVA and MEP Pathways for Isoprenoid Synthesis. In: *Annual Review of Plant Biology, Vol 64* (Merchant, S.S. ed) pp. 665-700.

Wang, G. (2014) Recent progress in secondary metabolism of plant glandular trichomes. *Plant Biotechnology* **31**, 353-361.

Wang, G.D. and Dixon, R.A. (2009) Heterodimeric geranyl(geranyl)diphosphate synthase from hop (Humulus lupulus) and the evolution of monoterpene biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 9914-9919.

Wang, Q., Reddy, V.A., Panicker, D., Mao, H.Z., Kumar, N., Rajan, C., Venkatesh, P.N., Chua, N.H. and Sarojam, R. (2016) Metabolic engineering of terpene biosynthesis in plants using a trichome-specific transcription factor MsYABBY5 from spearmint (Mentha spicata). *Plant Biotechnology Journal* **14**, 1619-1632.

This article is protected by copyright. All rights reserved.

Accepted Article

Wu, L., Liu, D.F., Wu, J.J., Zhang, R.Z., Qin, Z.R., Liu, D.M., Li, A.L., Fu, D.L., Zhai, W.X. and Mao, L. (2013) Regulation of FLOWERING LOCUS T by a MicroRNA in Brachypodium distachyon. *Plant Cell* **25**, 4363-4377.

Xia, R., Zhu, H., An, Y.Q., Beers, E.P. and Liu, Z.R. (2012) Apple miRNAs and tasiRNAs with novel regulatory networks. *Genome Biology* **13**.

Xu, Y.H., Wang, J.W., Wang, S., Wang, J.Y. and Chen, X.Y. (2004) Characterization of GaWRKY1, a cotton transcription factor that regulates the sesquiterpene synthase gene (+)-delta-cadinene synthase-A. *Plant Physiology* **135**, 507-515.

Yang, C.Q., Fang, X., Wu, X.M., Mao, Y.B., Wang, L.J. and Chen, X.Y. (2012) Transcriptional regulation of plant secondary metabolism. *J Integr Plant Biol* **54**.

Zhang, B.H. (2015) MicroRNA: a new target for improving plant tolerance to abiotic stress. *Journal of Experimental Botany* **66**, 1749-1761.

Zhang, B.H. and Wang, Q.L. (2015) MicroRNA-Based Biotechnology for Plant Improvement. *Journal of Cellular Physiology* **230**, 1-15.

Zhang, Q., Zhu, J., Ni, Y., Cai, Y. and Zhang, Z. (2012) Expression profiling of HbWRKY1, an ethephon-induced WRKY gene in latex from Hevea brasiliensis in responding to wounding and drought. *Trees* **26**, 587-595.

Zhang, X., Garreton, V. and Chua, N.-H. (2005) The AIP2 E3 ligase acts as a novel negative regulator of ABA signaling by promoting ABI3 degradation. *Genes & Development* **19**, 1532-1543.

Zhao, Q. and Dixon, R.A. (2011) Transcriptional networks for lignin biosynthesis: more complex than we thought? *Trends in Plant Science* **16**, 227-233.

Zhong, R.Q., Lee, C.H. and Ye, Z.H. (2010) Functional Characterization of Poplar Wood-Associated NAC Domain Transcription Factors. *Plant Physiology* **152**, 1044-1055.

Zhou, M., Wu, L., Liang, J., Shen, C. and Lin, J. (2012) Expression analysis and functional characterization of a novel cold-responsive gene CbCOR15a from Capsella bursa-pastoris. *Molecular Biology Reports* **39**, 5169-5179.

Figure Legends

Figure 1. Heat map analysis of spearmint R2R3-MYBs. Differential expression pattern of 45 annotated R2R3-MYBs along various tissues [leaf (L), leaf stripped of PGTs (L-T) and PGTs (T)]. MYB and MYB112 are highlighted in red which show high expression in PGTs. Sequence data and expression values are available in supplemental data set 1.

This article is protected by copyright. All rights reserved.

Accepted Article

Figure 2. *MsMYB* expression and localization. qRT PCR was done to analyse the expression of *MsMYB* along the various tissues [leaf (L), leaf stripped of PGTs (L-T), root (R) and PGTs (T)]. (a) Expression levels of *MsMYB* showing preferential expression in PGTs. (b) Nucleus-specific localization of MsMYB in *N. benthamiana* leaf cells.

Figure 3. DNA and protein sequence of MsMYB. R2 and R3 repeats are highlighted in blue and red colours respectively. Five conserved tryptophan residues are highlighted with grey boxes. Nuclear localization signal as predicted by cNLS mapper is shown by a black box.

Figure 4. *MsMYB* promoter analysis and expression pattern in tobacco. (a) Regulatory elements present in the promoter region (-615bp) of *MsMYB.* (b), (c) *N. benthamiana* plants transformed with $MsMYB_{pro}$*:GUS* showing trichome specific expression in leaves and stems. Black arrows point GUS stained trichome heads.

Figure 5. **Transcript levels and GC analysis of transgenic plants**. Reduced (a) and increased (b) levels of *MsMYB* in transgenic spearmint *MsMYB*-RNAi and *MsMYB*- overexpressing lines when compared to WT. (c), (d) Increased levels of limonene and carvone in *MsMYB*-RNAi lines when compared to WT. (e), (f) Reduced levels of limonene and carvone in *MsMYB*-overexpressing lines when compared to WT. GFP: GFP overexpressing line. Data are indicated as mean ± SE. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

Figure 6. Transcript levels of *GPPS.LSU* in transgenic plants. (a) Increased levels of *MsGPPS.LSU* in *MsMYB*-RNAi lines. (b) Decreased levels of

This article is protected by copyright. All rights reserved.

Accepted Article

MsGPPS.LSU in *MsMYB*-OX lines. (c) Decreased levels of *ObGPPS.LSU* in sweet basil plants expressing *MsMYB.* Data are indicated as mean ± SE. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

Figure 7. Promoter analysis of *MsGPPS.LSU.* (a) Regulatory elements present in the promoter region (-549bp) of *MsGPPS.LSU.* (b) Decreased levels of *MsGPPS.LSU* in plants grown under dark when compared to plants grown under light. Data are indicated as mean ± SE.

Figure 8. Yeast one-hybrid assay between MsMYB and *GPPS.LSU* promoter.

(a) Bait strains of (1) pAbAi-GPPS-full, (2) pAbAi-site1, (3) pAbAi-site2, (4) pAbAi-p53 (positive control) and (5) auxotrophic Y1H Gold (negative control) growing on SD/-Ura.

(b) pGADT7-MYB prey expression in (1) pAbAi-GPPS-full, (2) pAbAi-site1, (3) pAbAi-site2; pGADT7-p53 prey expression in (4) pAbAi-p53 and empty pGADT7 in (5) pAbAi-p53 on SD/-Leu.

(c) DNA-protein interaction of (1) full *GPPS* promoter and MYB, (2) site 1 and MYB, (3) site 2 and MYB, (4) pAbAi-p53 and pGADT7-p53 (5) no interaction of pAbAi-p53 and empty pGADT7 on SD/-Leu/AbA.

(d) Absence of DNA-protein interaction of (1) pAbAi-mut-site1 and (2) pAbAi-mut-site2 with MYB.

SD/-Ura, SD medium without Ura; SD/-Leu, SD medium without Leu; SD/-Leu/AbA, SD medium without Leu but containing Aureobasidin A (800 ng).

This article is protected by copyright. All rights reserved.

Accepted Article

Figure 9. Transcriptional repression activity of MsMYB and localization of MsGPPS.LSU. (a) Reduced levels of GUS in *$35S_{pro}$:MsMYB* expressing leaves (b) Chloroplast specific localization of MsGPPS.LSU. Data are indicated as mean ± SE.

Figure 10. Ectopic expression of *MsMYB* in sweet basil alters the amount of secondary metabolites. (a) *MsMYB* expression in basil transgenic plants. (b), (c) Reduced levels of monoterpenes and sesquiterpenes. (d) Decreased levels of total terpenes. (e) Unaltered levels of phenylpropenes (eugenol). (f) Unaltered levels of total flavonoids. Data are indicated as mean ± SE. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

Figure 11. ms-miR858 analysis. (a) ms-miR858 sequence comparison with other known miR858 sequences from various species. miR858 sequences were obtained from miRBase. ms - *Mentha spicata*, ath - *Arabidopsis thaliana*, aly - *Arabidopsis lyrata*, mdm - *Malus domestica*, cme - *Cucumis melo*, ppe - *Prunus persica*. (b) Expression levels of ms-miR858 in leaf stripped of PGTs (L-T) and PGTs (T). (c) Probable binding of ms-miR858 to *MsMYB*.

Supplementary data

Figure S1. Scanning electron micrograph (SEM) of leaf surface and southern blot analysis of transgenic plants. (a) SEM of a spearmint leaf showing two kinds of glandular trichomes, (1) capitate glandular trichome and (2) peltate glandular trichome. (b), (c) Southern blot of *MsMYB*-RNAi and *MsMYB*-overexpressing spearmint lines showing a range of insertions. (d)

This article is protected by copyright. All rights reserved.

Accepted Article

Southern blot of transgenic sweet basil lines overexpressing *MsMYB* showing different T-DNA insertions. 15µg of DNA was digested with NdeI enzyme.

Figure S2. Phylogenetic tree showing the similarity of MsMYB to known *Arabidopsis thaliana* R2R3-MYBs. MsMYB is pointed with a red arrow. MsMYB falls under subgroup 7.

Figure S3. GC profiles of wild type plants. (a) Spearmint, 1. limonene; 2. carvone. (b) Sweet basil, 1. alpha-pinene; 2. beta-pinene; 3. eucalyptol; 4. linalyl acetate; 5. eugenol; 6. alpha-bergamotene; 7. germacrene D; 8.gamma-muurolene; 9. β-copaene.

Figure S4. Transgenic plants overexpressing *MsMYB* show smaller leaf size. Ectopic expression of *MsMYB* led to decreased leaf size in transgenic basil plants (a), (b) and tobacco plants (c).

Figure S5. Ectopic expression of *MsMYB* in tobacco. (a) *MsMYB* expression in tobacco transgenic plants. (b) Reduced levels of CBT-diols in transgenic tobacco plants expressing *MsMYB.* Data are indicated as mean ± SE. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

Table S1. Primers used in this study.

Table S2. List of genes analyzed in spearmint transgenic plants.

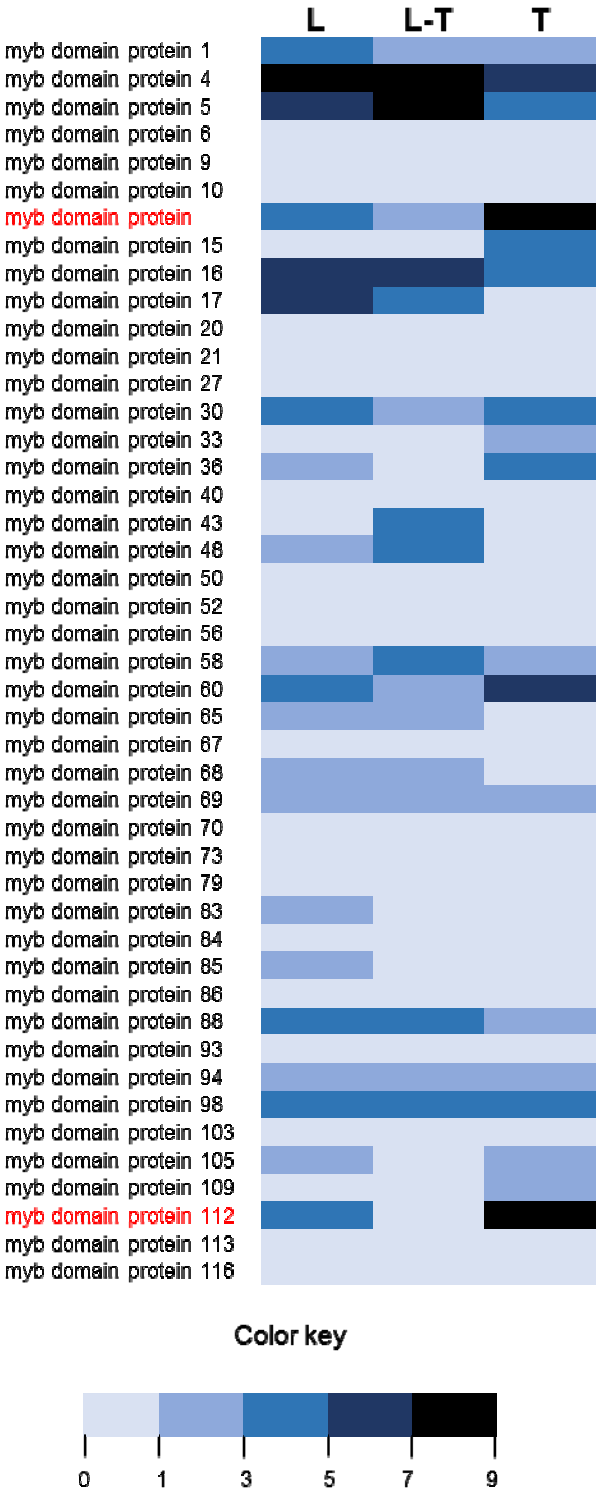
Table S3. Oligonucleotides used for the generation of bait sequences.

Data Set 1. Sequence data and expression values of spearmint R2R3-MYBs.

Data Set 2. Accession numbers of A.thaliana R2R3-MYBs.

This article is protected by copyright. All rights reserved.

Accepted Article



This article is protected by copyright. All rights reserved.

Accepted Article

Figure 1. Heat map analysis of spearmint R2R3-MYBs. Differential expression pattern of 45 annotated R2R3-MYBs along various tissues [leaf (L), leaf stripped of PGTs (L-T) and PGTs (T)]. MYB and MYB112 are highlighted in red which show high expression in PGTs. Sequence data and expression values are available in supplemental dataset 1.

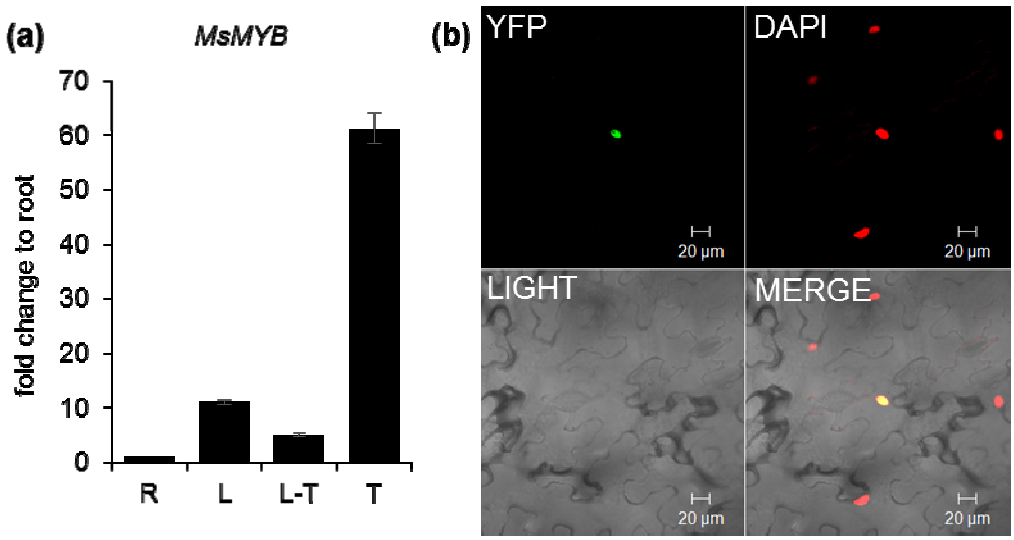


Figure 2. *MsMYB* expression and localization. qRT PCR was done to analyse the expression of *MsMYB* along the various tissues [leaf (L), leaf stripped of PGTs (L-T), root (R) and PGTs (T)]. (a) Expression levels of *MsMYB* showing preferential expression in PGTs. (b) Nucleus-specific localization of MsMYB in *N. benthamiana* leaf cells.

This article is protected by copyright. All rights reserved.

Accepted Article

> *MsMYB* (KY081780)

ATGGGAAGAGCGCCGTGCTGTGAGAAAGTTGGGTTGAAGAGAGGGAGATGGACTGCAGAAGAA
GATGAAAAGCTCAGAAAATATATTCAGGAAAATGGTGAAGGCTGCTGGCGATCATTGCCCAAG
AATGCAGGTTTACTTAGATGTGGAAAGAGTTGCAGACTGAGATGGATTAATTATTTGAGATCA
GATGTGAAGAGAGGGAATATTTCTTCTCAAGAAGAAGAAATCATCATTAATCTCCATGCATCT
ATGGGCAACAGGTGGTCCCTGATCGCCGCGCACTTGCCGGGTAGAACAGACAATGAAATCAAA
AATTACTGGAACTCCCATTTGAGCAGAAAATTCCACGGTTTCCGCCCTAATCCACAGTTCATT
CCGCCGCCGCCGCCTCCGCCGCCATCCTCCAAGCCCAAGAAGACAAAGAACAGTAATAAGAAG
GCAAAGGCGGCAGCTAAAACCGCCACTGCCGCCGTCGTCATGCCGACTACCCCCACTCCGGAG
AAAGAGTCCTCGGTGGGCAGACCTGGGAAGGAGAGAGAAAGTGAAGCGAGAGAGAGCGGCAGC
TCCATGGTGGGAGAGTTGGACGATTTAAATATGACGGAGGACTTGAGCGGCCTCTGGGGCCCC
ACACTGGACTTCGGGACCGGAAGTGAAATTTCGGATCCGGGTCTGGGCCAGCTAGATAACTCG
TCGTTGCAAATTTACGAAGAAACGCTTTCGTGGATTTGGAACGACGAGGACGACAAGTGGAAT
TCCAACGTGGACAACGGAGAAATGGATGGTGCAATGCTTTCTTGGTTGTTGTCATGA

> MsMYB

MGRAPCGEKVGLKRGRWTAEEDEKLRKYIQENGEGCWRSLPKNAGLLRCGKSCRLRWINYL

RSDVKRGNISSQEEEIIINLHASMGNRWSLIAAHLPGRTDNEIKNYWNSHLSRKFHGFRPN

PQFIPPPPPPPPSSKPKKTKNSNKKAKAAAKTATAAVVMPTTPTPEKESSVGRPGKERESE

ARESGSSMVGELDDLNMTEDLSGLWGPTLDFGTGSEISDPGLGQLDNSSLQIYEETLSWIW

NDEDDKWNSNVDNGEMDGAMLSWLLS

Figure 3. DNA and protein sequence of MsMYB. R2 and R3 repeats are highlighted in blue and red colours respectively. Five conserved tryptophan residues are highlighted with grey boxes. Nuclear localization signal as predicted by cNLS mapper is shown by a black box.

This article is protected by copyright. All rights reserved.

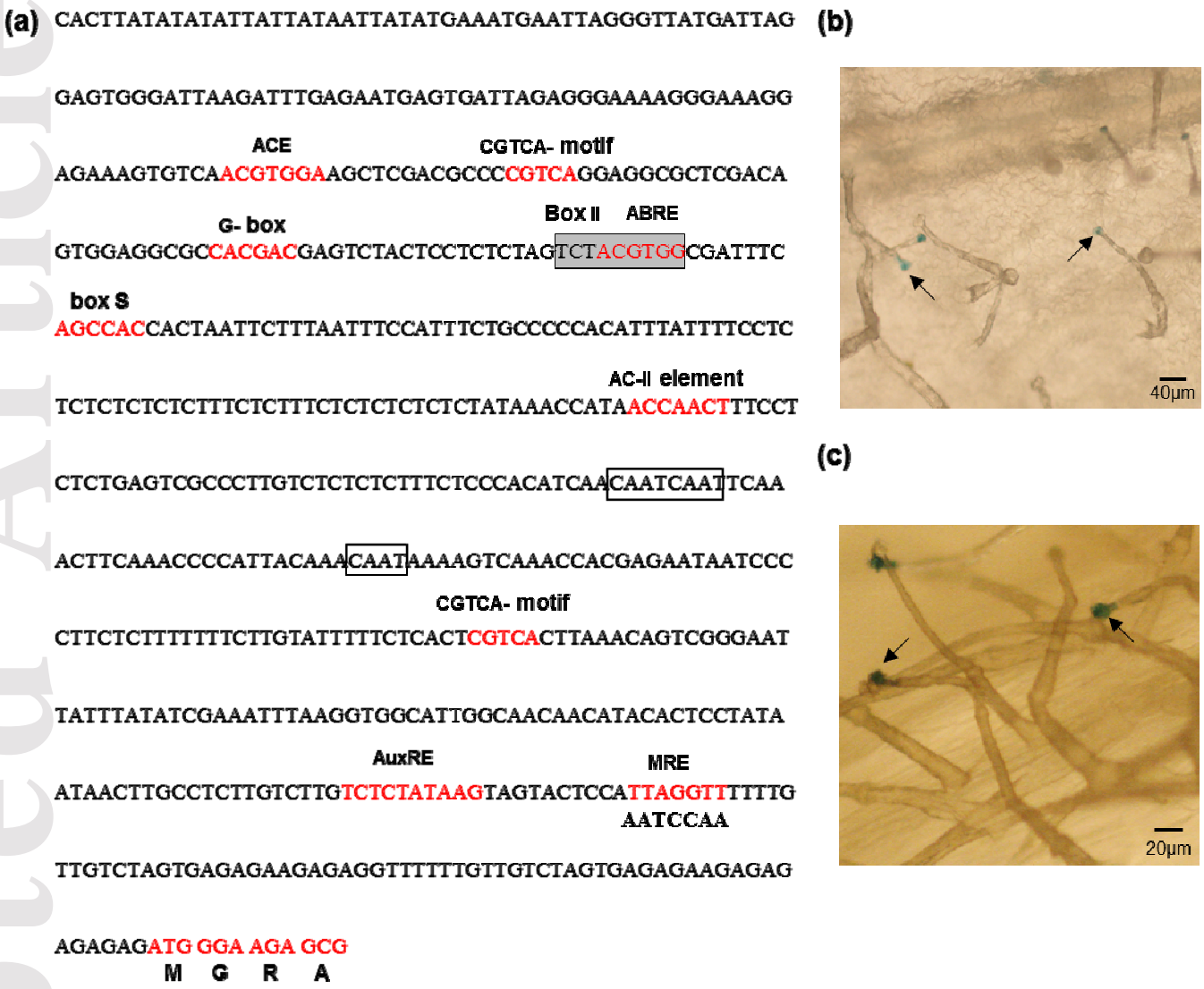


Figure 4. *MsMYB* promoter analysis and expression pattern in tobacco. (a) Regulatory elements present in the promoter region (-615bp) of *MsMYB.* (b), (c) *N. benthamiana* plants transformed with *MsMYB*$_{pro}$*:GUS* showing trichome specific expression in leaves and stems. Black arrows point GUS stained trichome heads.

This article is protected by copyright. All rights reserved.

Accepted Article

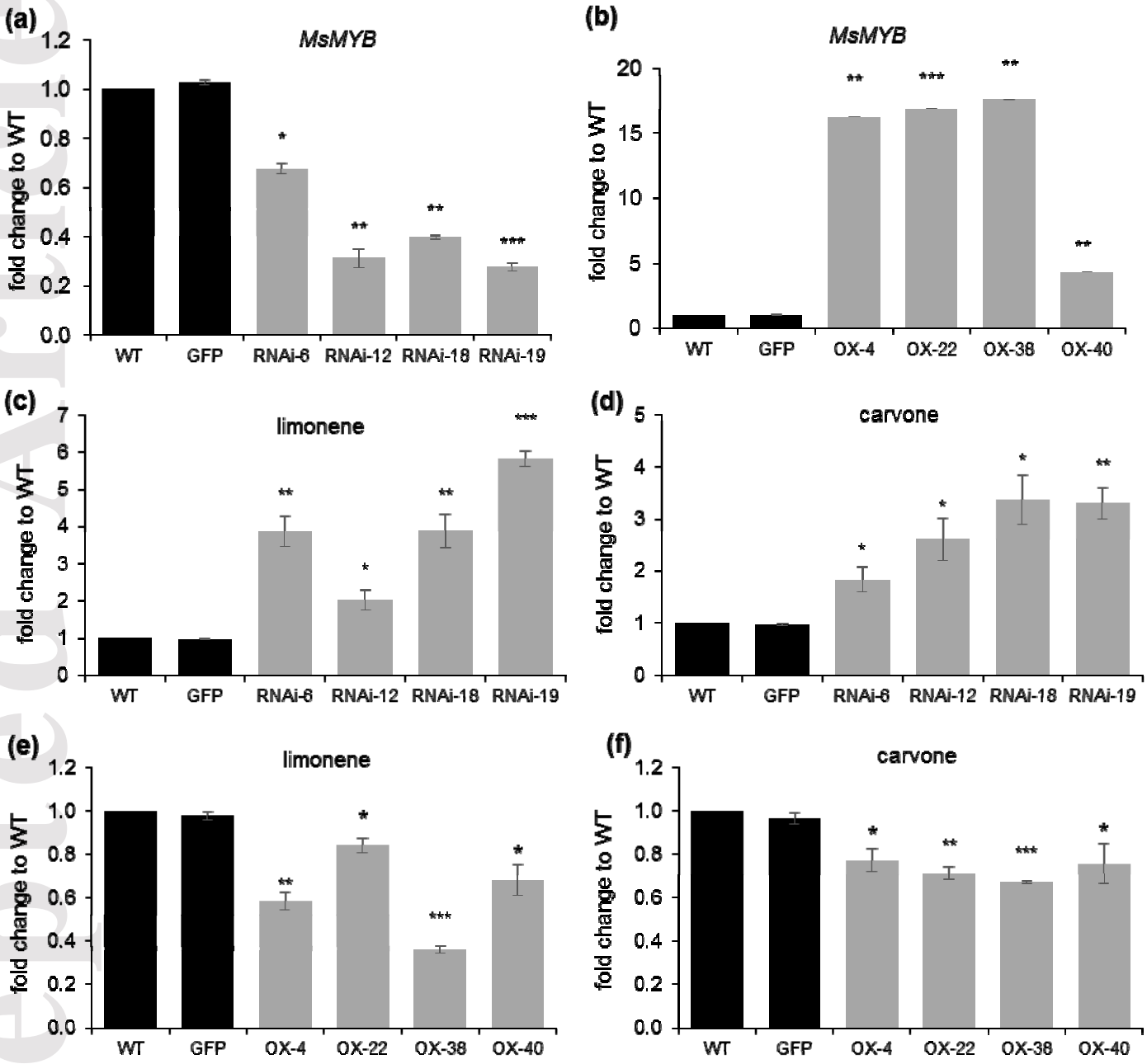
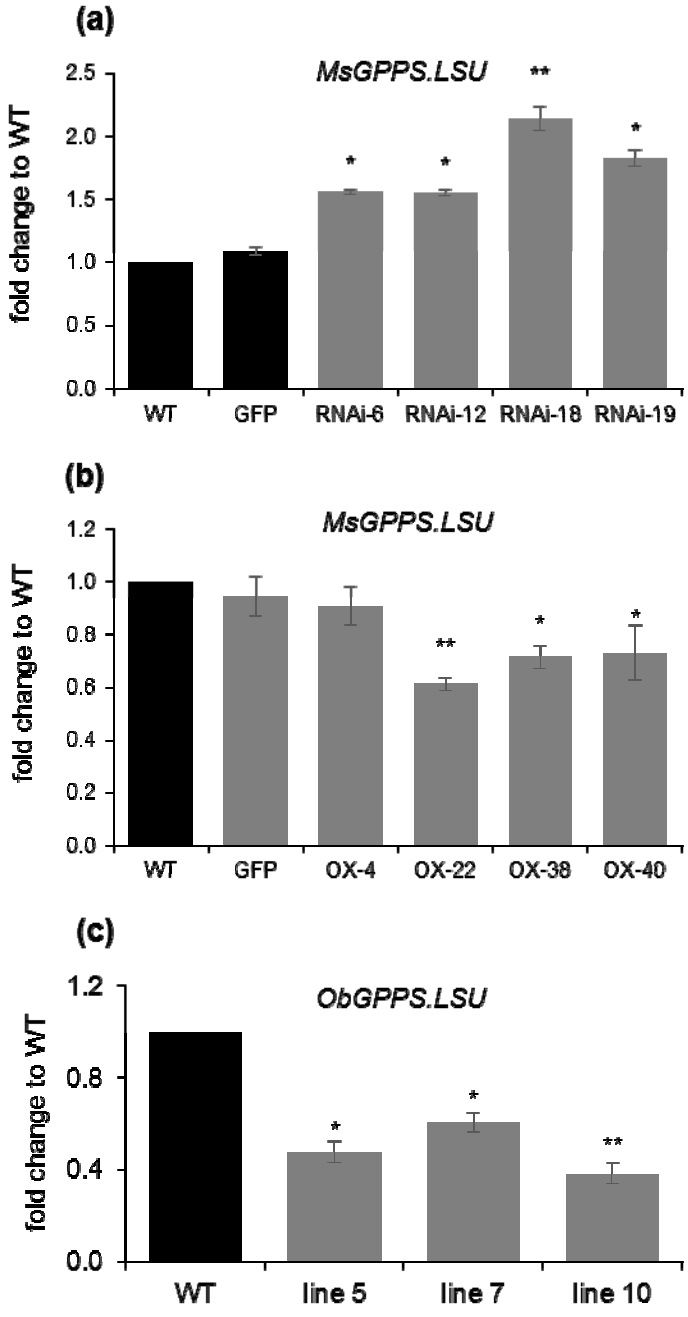


Figure 5. **Transcript levels and GC analysis of transgenic plants**. Reduced (a) and increased (b) levels of *MsMYB* in transgenic spearmint *MsMYB*-RNAi and *MsMYB*-overexpressing lines when compared to WT. (c), (d) Increased levels of limonene and carvone in *MsMYB*-RNAi lines when compared to WT. (e), (f) Reduced levels of limonene and carvone in *MsMYB*-overexpressing lines

This article is protected by copyright. All rights reserved.

when compared to WT. GFP: GFP overexpressing line. Data are indicated as mean ± SE. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.



This article is protected by copyright. All rights reserved.

Figure 6. Transcript levels of *GPPS.LSU* in transgenic plants. (a) Increased levels of *MsGPPS.LSU* in *MsMYB*-RNAi lines. (b) Decreased levels of *MsGPPS.LSU* in *MsMYB*-OX lines. (c) Decreased levels of *ObGPPS.LSU* in sweet basil plants expressing *MsMYB.* Data are indicated as mean ± SE. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

Accepted Article

This article is protected by copyright. All rights reserved.

Accepted Article

(a)

GT1- motif
TGA- box
ATCATGGTTAAACATATGAAAAAATTGACGTAAACTTTGATTTTATACTAA

ABRE
TCCGTCTAGGCCGTTTAGGCACCGCCTAGACTCCCGCCTAGGCTGCCTAGG

CGCTAATCGCTCCCCCACCGTTCTACTAGCGTCTAGCGACTTTTAGAACACT

GATCAACCGTCGTTGCAATAAAAGATGCAACACGGAAATGCAATTGAGGA

TCCGATTCGAGAATCCATTGATGGTGGTGCATAAACTAGCGATTCAGATCA

CGTCA- motif
TCTATTACATTTCCTTGTTACACCGTGGGATGATCGTCATGTATTTAATGTA

MBS (site2)
ATTGTGTTTATTGAATTATTTAAAGCAGTTTAGGGCTTTAAAAGCAGTTAG
GTCAAT

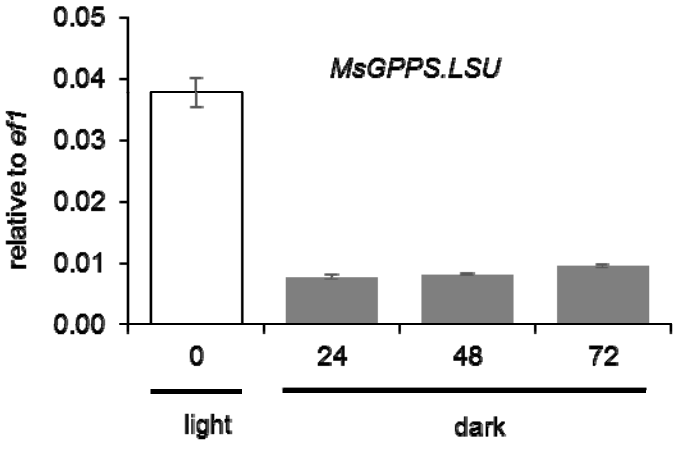
LRE
GCGTTAAACTGCCAAACTCTCCCAATCAAAATCTCCGCTCTTTTATTGCAAC

TGACG- motif
ATCACATATATTCATTCACCTAATTGCCAAAACAATTCTGACGTTGTTTCTT

MBS (site1)
CGCACAGATTTCCAGTTGCAGTGCTGGGTTGATATATATATAATCTCGTGT
GTCAAC

GARE- motif
ATTTATATATATATATATATATTTCTGTTGGTAAAAATG GGA AGA GCG
M G R A

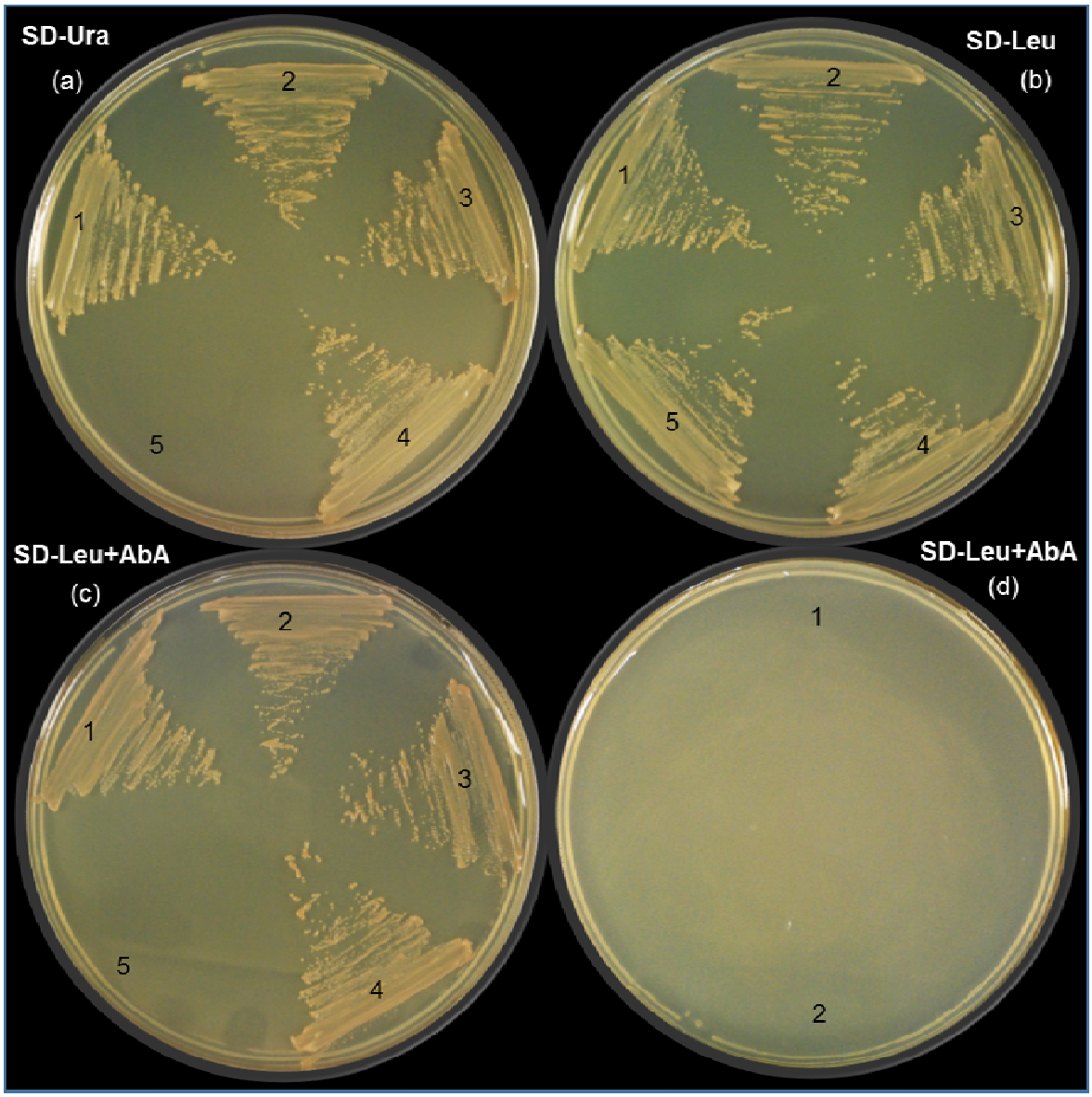
(b)



This article is protected by copyright. All rights reserved.

Accepted Article

Figure 7. Promoter analysis of *MsGPPS.LSU.* (a) Regulatory elements present in the promoter region (-549bp) of *MsGPPS.LSU.* (b) Decreased levels of *MsGPPS.LSU* in plants grown under dark when compared to plants grown under light. Data are indicated as mean ± SE.



This article is protected by copyright. All rights reserved.

Accepted Article

Figure 8. Yeast one-hybrid assay between MsMYB and *GPPS.LSU* promoter.

(a) Bait strains of (1) pAbAi-GPPS-full, (2) pAbAi-site1, (3) pAbAi-site2, (4) pAbAi-p53 (positive control) and (5) auxotrophic Y1H Gold (negative control) growing on SD/-Ura.

(b) pGADT7-MYB prey expression in (1) pAbAi-GPPS-full, (2) pAbAi-site1, (3) pAbAi-site2; pGADT7-p53 prey expression in (4) pAbAi-p53 and empty pGADT7 in (5) pAbAi-p53 on SD/-Leu.

(c) DNA-protein interaction of (1) full *GPPS* promoter and MYB, (2) site 1 and MYB, (3) site 2 and MYB, (4) pAbAi-p53 and pGADT7-p53 (5) no interaction of pAbAi-p53 and empty pGADT7 on SD/-Leu/AbA.

(d) Absence of DNA-protein interaction of (1) pAbAi-mut-site1 and (2) pAbAi-mut-site2 with MYB.

SD/-Ura, SD medium without Ura; SD/-Leu, SD medium without Leu; SD/-Leu/AbA, SD medium without Leu but containing Aureobasidin A (800 ng).

This article is protected by copyright. All rights reserved.

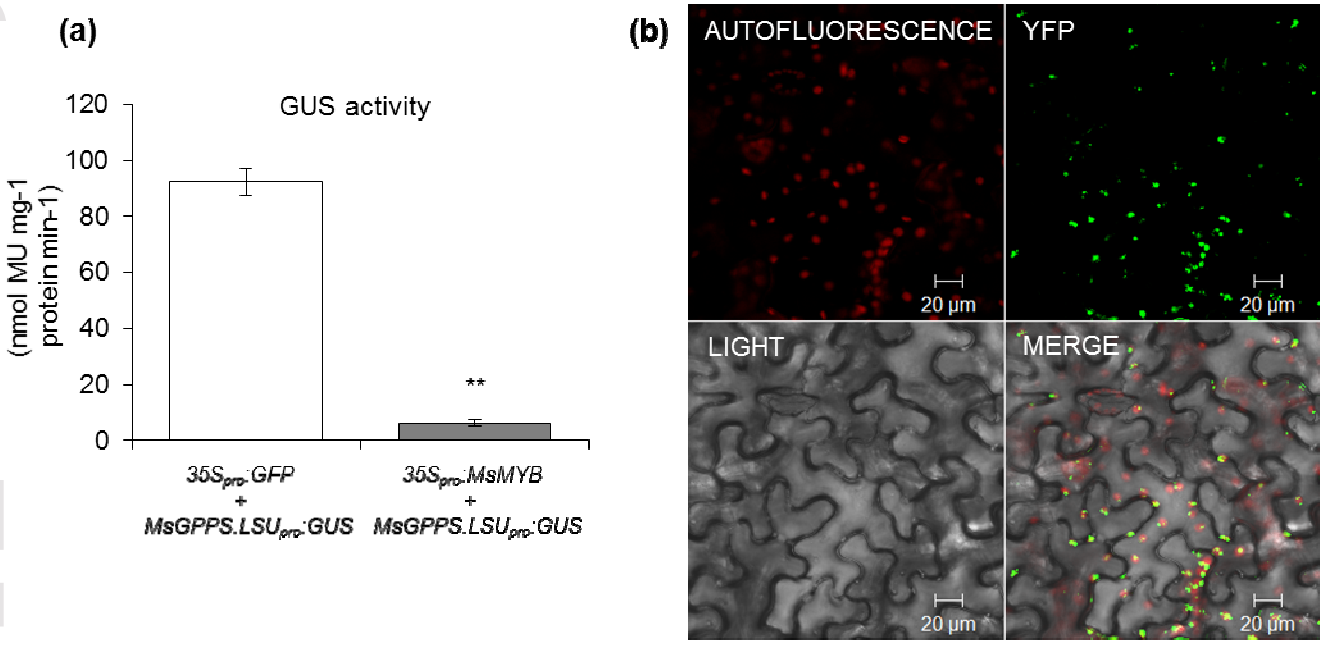


Figure 9. Transcriptional repression activity of MsMYB and localization of MsGPPS.LSU. (a) Reduced levels of GUS in *35S$_{pro}$:MsMYB* expressing leaves (b) Chloroplast specific localization of MsGPPS.LSU. Data are indicated as mean ± SE.

This article is protected by copyright. All rights reserved.

Accepted Article

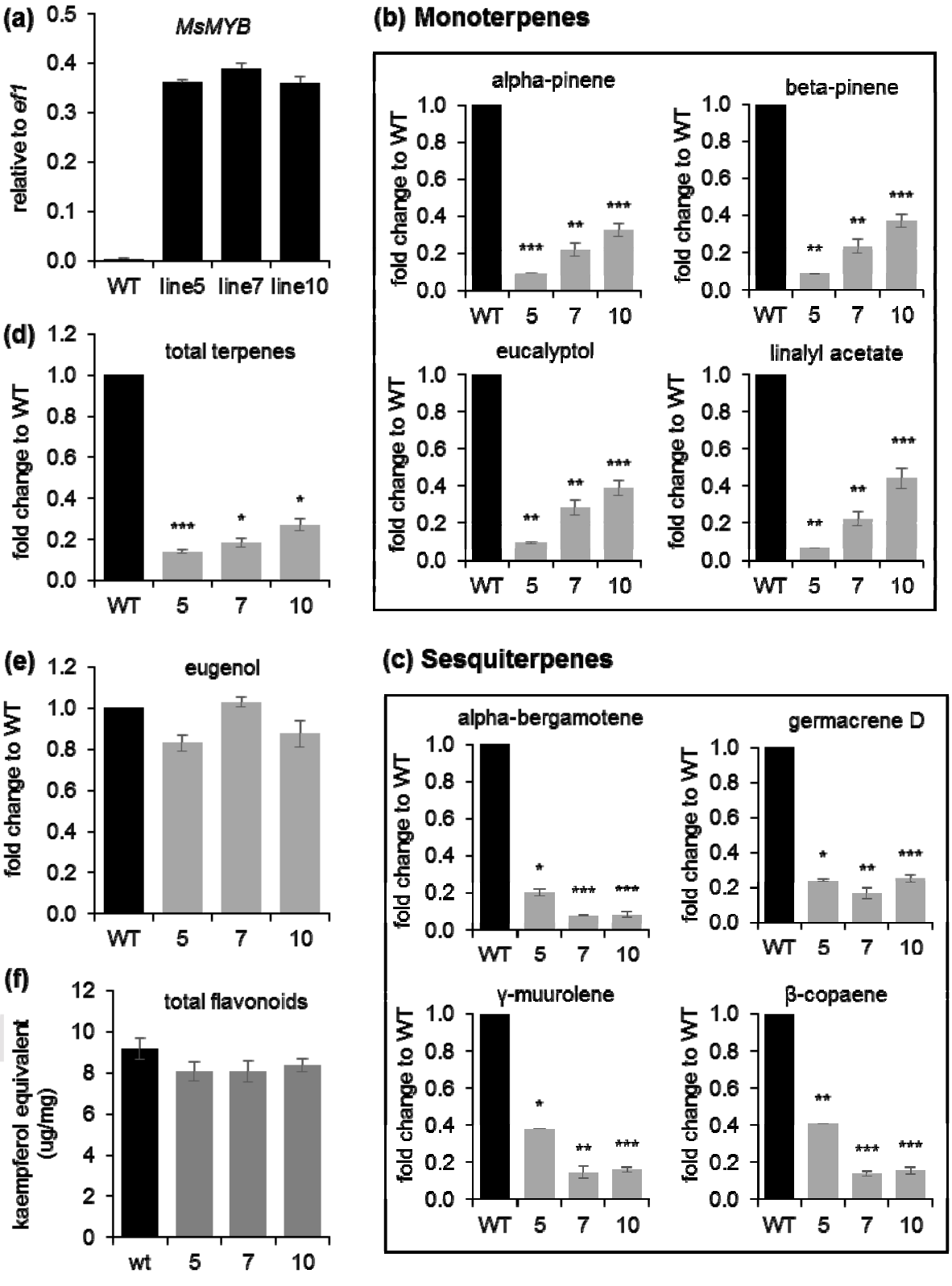


Figure 10. Ectopic expression of *MsMYB* in sweet basil alters the amount of secondary metabolites. (a) *MsMYB* expression in basil transgenic plants. (b), (c) Reduced levels of monoterpenes and sesquiterpenes. (d) Decreased

This article is protected by copyright. All rights reserved.

levels of total terpenes. (e) Unaltered levels of phenylpropenes (eugenol). (f) Unaltered levels of total flavonoids. Data are indicated as mean ± SE. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

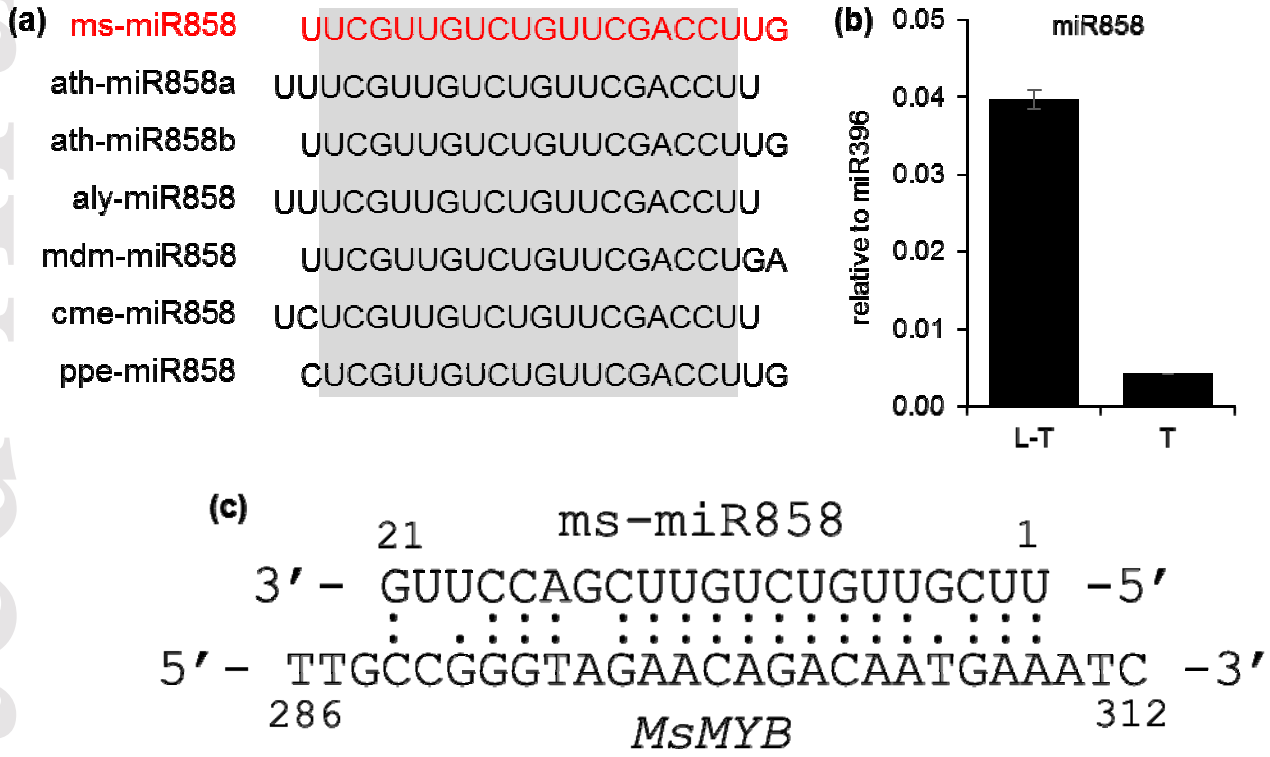


Figure 11. ms-miR858 analysis. (a) ms-miR858 sequence comparison with other known miR858 sequences from various species. miR858 sequences were obtained from miRBase. ms - *Mentha spicata*, ath - *Arabidopsis thaliana*, aly - *Arabidopsis lyrata*, mdm - *Malus domestica*, cme - *Cucumis melo*, ppe - *Prunus persica*. (b) Expression levels of ms-miR858 in leaf stripped of PGTs (L-T) and PGTs (T). (c) Probable binding of ms-miR858 to *MsMYB*.

This article is protected by copyright. All rights reserved.

Accepted Article