J. Trop. Agric. and Fd. Sc. 47(2)(2019):37 – 54

Transcriptome analysis of floral scent biosynthesis during *Michelia alba* (cempaka putih) flower development

[Analisis transkrip biosintesis aroma bunga semasa perkembangan bunga *Michelia alba* (cempaka putih)]

V. Maheswary[1], S. Sanimah[2], H.N. Khairun[2], Y. Nurul Aishah[2], O. Qistina[2] and S. Vasanthi[2]

[1]Strategic Planning and Innovation Management Centre, MARDI Headquarters, Persiaran MARDI-UPM, 43400 Serdang, Selangor
[2]Biotechnology Research Centre, MARDI Headquarters, Persiaran MARDI-UPM, 43400 Serdang, Selangor
[3]Food Science Research Centre, MARDI Headquarters, Persiaran MARDI-UPM, 43400 Serdang, Selangor

Abstract

Despite a long history of cultivation of *Michelia alba* (cempaka putih) and the commercial importance of its essential oils, no information is currently available on potential changes in the expression of genes involved in the synthesis of individual volatile organic compounds (VOCs) during flower development. Here we describe the combined use of Digital Gene Expression (DGE) and chemical analysis (from a previous study) to discover genes for the VOCs biosynthesis pathways and related genes. From the DGE sequencing data, a total of 50 genes (including isoforms) were identified to be involved directly with fragrance biosynthesis from the pooled early stages (S1 – S6) and pooled late stages (S7 – S11) of flower development. Our transcriptome data revealed that the bulk of floral scent genes (29) in this species belonged to the isoprenoids biosynthesis pathway of which 15 transcripts belonged to the monoterpeoids biosythesis pathway, seven belonged to the sesquiterpenoids biosythesis pathway while the remaining seven belonged to the terpenoids backbone biosynthesis pathway. Fifteen genes were also identified to be involved in the biosynthesis of aromatic compounds in the phenylpropanoids/benzenoids biosynthesis pathway while six were found to belong to the fatty acid derivatives biosythesis pathway. Four other other genes were also identified to be related to floral scent biosynthesis in other metabolic pathways.

Keywords: *Michelia alba*, floral scent biosynthesis pathways, next generation sequencing, gene ontology

Introduction

Michelia alba, commonly called cempaka putih, belongs to the family Magnoliaceae and includes about 50 species. It is a tropical evergreen woody tree species which can grow to a height of 3 – 45 m and is widely cultivated in the tropical and subtropical South and Southeast Asian regions for its fragrant and ivory coloured fragrant flowers. *M. champaca* L. is native to India, Java, and the Philippines while *M. doltsopa* is a large tree growing to about 30 m bearing clusters of creamy white flowers in winter and is native to the eastern Himalayas and Meghalaya subtropical forests. *Michelia aenea* Dandy is native to China and

Article history
Received: 21.06.2017
Accepted: 17.07.2019

Authors' full names: Maheswary Vellupillai, Sanimah Simoh, Khairun Hisam Nasir, Nurul Aishah Yahya, Nur Qistina Othman dan Vasanthi Sivarajah
E-mail: maheswarykavitha@gmail.com

©Malaysian Agricultural Research and Development Institute 2019

Vietnam. Essential oils obtained from mature fresh *M. alba* flowers have already been used in perfumery and the cosmetic industry as major components of perfumes and fragrances such as *Joy* and *Jadore* (Anon. 2000).

The transition from bud development to full bloom is a critical phase switch from vegetative growth to reproductive growth in flowering plants for subsequent generation development. Endogenous and environmental signals initiate a complex network of genetic pathways to activate biosynthesis of floral scents which are assumed to play a role in both attraction of pollinators and defence against pathogens, parasites and herbivores (Knudsen et al. 2006; Muhlemann et al. 2014). The diversity of these floral scents results from different volatile organic compounds (VOCs) derived from different biosynthetic routes in plants such as terpenoids, phenylpropanoids/benzenoids and volatile fatty acid derivatives (Muhlemann et al. 2014). These compounds are synthesised during the development of the whole plant and the quantity and diversity of these VOCs are subject to changes in response to environmental stimuli (Dudareva et al. 2013). The unique and distinct floral scents of many ornamental plants such as *Clarkia breweri* (Dudareva and Pichersky 2000), *Rosa hybrid* (Hendel-Rahmanim et al. 2007), *Lilium* spp. (Zhang et al. 2013), tree peony (Zhao et al. 2012), *Syringa oblata* (Li et al. 2006; Zheng et al. 2015), *Prunus mume* (Hao et al. 2014), *Jasminum sambac* (Li et al. 2015), *Cananga odorata* (ylang ylang) (Jin et al. 2015), *Hedychium coronarium* (Yue et al. 2015), and *Lillium* 'Yelloween'(Wang et al. 2015) have already been investigated.

Terpenoids, also referred to as isoprenoids, are the largest and most diverse class of VOCs in plants (Dudareva et al. 2013; Knudsen et al. 2006). Terpenes are synthesised from two distinct and compartmentally separated pathways, the mevalonate (MVA) and 2-C-methyl-d-erythritol 4-phosphate (MEP) pathways (McGarvey and Croteau 1995). The distinct expression patterns of MEP and MVA pathway genes of different plants might result in differences in biosynthesis and emission of floral terpenoids. The biosynthesis and emission of terpenes have been investigated in snapdragon (Dudareva et al. 2003), *Clarkia breweri* (Dudareva et al. 1996), *A. thaliana* (Chen et al. 2003), *Lavandula angustifolia* (Guitton et al. 2010) and *Syringa oblata* (Zheng et al. 2015).

The phenylpropanoid/benzenoid class of metabolites is derived primarily from the carbon skeleton of phenylalanine, which is produced by the shikimate pathway (Orlova et al. 2006; Vogt 2010). They constitute a large class of secondary metabolites which are not usually volatile. However, several phenylpropanoids whose carboxyl group at C9 is reduced (to either the aldehyde, alcohol, or alkane/alkene) are volatiles. In addition, phenylpropanoids which contain alkyl additions to the hydroxyl groups of the benzyl ring or to the carboxyl group (i.e. ethers and esters) are also volatiles.

The chemical composition of floral VOCs produced by *M. alba* has been reported previously (Euyama et al. 1992; Sanimah et al. 2002; 2008; Shang et al. 2002). Sanimah et al. (2008) showed the presence of volatile isoprenoids or terpenoids, phenylpropanoids/benzenoids and fatty acid derivatives in this flower at development stages S5 – S11 (Sanimah et al. 2002). They also found that the primary components were monoterpenes, linalool and dihydrocarveol. However, two molecular studies using the subtractive hybridisation approach at mid-flower development stage S6 (Maheswary et al. 2008) and full bloom development stage (S10) (Maheswary et al. 2011) could not identify the key enzyme, linalool synthase (LIS), which was reported to produce the major fragrance compound linalool in *M. alba*. However, the studies did reveal several other genes involved in the biosynthesis of aromatic volatiles such as monoterpene glucosyltransferase (MGT), cytochrome P450, cytochrome

P450 monoxygenase, 10-hydroxygeraniol oxidoreductase, pulegone reductase, germacrene-D synthase and sesquiterpene cyclase in the terpenoids biosynthesis pathway, benzoyl CoA: benzyl alcohol benzoyl transferase, salicyclic acid carboxyl methyltransferase (SAMT), aromatic-L-amino acid decarboxylase, orcinol O-methyltransferase and eugenol synthase 2 in the phenylpropanoids/benzenoids pathway, and methyl jasmonate esterase and carboxylesterase (CXE) in the fatty acid derivatives pathway.

The stage of flower development has been shown to affect the nature and quantity of the VOCs produced and accumulated in many flowering species such as rose cultivars, where volatile emissions were found to be low in early flower development and later increased to a peak as the flowers reached full bloom (Guterman et al. 2002). The amount of scented products has also been shown to peak in the petals of roses at anthesis (Bergougnoux et al. 2007) and also in snapdragon *Antirrhinum* (Dudareva et al. 2003). However, not much has been investigated in the expression of floral scent genes at different flower stages in *M. alba*. At least four-fold differences have been observed in the expression of floral scent genes in *Jasminum sambac* between budding and blooming stages (Li et al. 2015). Therefore, in this study, two different pooled stages of flower development i.e. early stages (S1 – S6) and late stages (S7 – S11) were used to study the expression of genes involved in the biosynthesis of floral scents in *M. alba* using Digital Gene Expression (DGE) high-throughput sequencing technology.

Materials and methods

Plant materials

Fresh samples of *M. alba* at eleven flower developmental stages (*Plate 1*) were collected seasonally from trees around Selangor in Malaysia. Between S1 to S4 stages, flower buds are very small and closed with green petals and no fragrance. At S5, the buds are still closed with swollen yellowish petals. At S6, the still unopened buds are longer and greenish-cream in colour with a slight fragrance while at S7, the buds have cream-coloured fully opened bracts. S8 represents the initial flowering stage with quarterly open outer whorl of white petals while S9 represents the half flowering stage with semi-open outer and middle whorl of white petals. At S10, the flower is in full bloom with completely open outer, middle and inner whorl of white petals emitting a strong fragrance and at

S1: Flower buds very small, closed and petals green wth no fragrance
S5: Buds yellowish (colour break) and swollen
S6: Buds turned greenish-cream, swollen and elongated
S7: Buds turned full cream colour and bracts opened
S8: Quartes bloom, outer colour and bracts opened
S9: Half bloom, outer and middle whorl of petals opened
S10: Full bloom, outer, middle and inner whorl of petals opened
S11: Stamen turned brown, some petals fallen off
S12: Fruit formation

Plate 1. Flower development stages in Michelia alba

S11, the flowering stage ends with fully matured withering petals and brown stamen. Approximately 5 g samples of each flower stage were immediately frozen in liquid nitrogen prior to storing at –80 °C for RNA isolation.

RNA isolation

Total RNA was isolated from each of the frozen flower stages, S1 to S11, by first grinding the samples into a fine powder in a pre-chilled mortar and pestle before adding Cetyl Trimethyl Ammonium Bromide (CTAB) reagent. Subsequent steps were performed according to the method described by Chang et al. (1993). The total RNA was resuspended in RNase-free water and the RNA aliquots were quantified on a ND-1000 nanodrop (Nano Drop Technologies Inc, USA). Gel electrophoresis was performed on a 1% formaldehyde agarose gel to check the integrity of the RNA extracted. The RNA was then purified using the RNeasy Plant Mini Kit (Qiagen, Germany) and the quantity and quality of the purified RNA were checked again using the above methods. Equal amounts of purified total RNA from S1 to S6 flower developmental stages were pooled together to represent the early stages of flower development. The same was carried out for S7 to S11 flower developmental stages. A total of 20 μg of purified total RNA (500 ng/ μl) from each of the pooled samples were outsourced to the Illumina company (USA) to generate two DGE libraries consisting of 21-bp sequences using the second Illumina Genome Analyser platform (GA II). The RNA integrity and concentration of the two pooled samples were checked again by Illumina before deep sequencing.

BLAST analysis and identification of DGE sequence tags

The 21-bp sequence tags from the early and late flowering DGE libraries were matched against the TIGR Plant Transcript Assembly (plantTA) database (http://plantta.tigr.org/) (Childs et al. 2007) using the Basic Local Alignment Sequencing Tool (BLAST) program. The best hit out of ten hits was then selected for each sequence tag based on genes identified to be associated with fragrance biosynthesis pathways or with the highest scoring pair (hsp) of nucleotides.

Functional analysis using Gene Ontology (GO)

PlantTA sequences that mapped to the 21-bp sequence tags underwent BLAST against the SwissProt database with a cut-off E-value of $<10^{-6}$ before a functional analysis was performed on the resulting peptide sequences using the Blast2GO program (http://www.blast2go.de/) (Conesa et al. 2005; Gotz et al. 2008). Sequences were annotated according to the three main functional categories of Biological Process (BP), Molecular Function (MF) and Cellular Component (CC) (Ashburner et al. 2000; Botton et al. 2008). The sequences are in the process of being deposited in the NCBI Sequence Read Archive.

Results and discussion

Raw sequence analysis of DGE sequence tags

A total of about 5 million 21-bp raw sequence tags were generated by single pass (5’- sense strand) sequencing from each DGE library. The total number of sequence tags in the early and late stages of flower development libraries with frequencies of 1, 2 and ≥3 were 444,509 and 431,505 respectively (*Table 1*). After the elimination of sequence tags with frequencies ≤2 (poor quality sequencing scores due to sequencing error, giving rise to runs of the same base artefacts), the total number of sequence tags with frequencies of ≥3 remaining in the early and late libraries for blast analysis were 105,339 and 100,559 respectively.

Blast analysis of DGE sequence tags

Blast analysis results indicated that 96.2% of the raw sequence tags with frequencies of ≥3 matched to the annotated sequences in the plantTA database (*Table 1*) while

Table 1. Digital gene expression analysis of sequence tags in *Michelia alba* early (S1 – S6) and late (S7 – S11) flower development libraries before and after blast to PlantTA database

	No. of sequence tags	
Before blast	Early stages (S1 – S6)	Late stages (S7– S11)
Total no. of sequence tags with frequencies of 1, 2 and ≥3	444,509	431,505
No. of sequence tags with frequencies of ≥3	105,339	100,559
After blast		
No. of sequence tags with frequencies of ≥3	101,976	96,088
No. of sequence tags with no match	3,313	4,424
No. of sequence tags missing after blast	49	46

3.8% had no matches. An additional 0.05% were missing after blast analysis. These sequence tags could have been recognised by the program as possible artefacts (runs of the same base), and therefore omitted from analysis (Philippe et al. 2013).

Identification of differentially regulated genes

- **Isoprenoids biosynthesis pathway**

A majority of the genes (60%) belonged to the main isoprenoid biosynthesis pathway (*Table 2*). This finding is in agreement with previous GC-MS profiling studies in *M. alba* by Sanimah et al. (2008) who showed that 33 of the 78 compounds identified belonged to the isoprenoids and this comprised 30 – 50% of the total volatiles detected during stages S5 to S11. The most abundant (15) belonged to the monoterpenoids while seven belonged to the sesquiterpenoids. The remaining seven belonged to the terpenoid backbone biosynthesis pathway. In contrast to our findings, Jin et al. (2015) observed that more than 90% of the total identified terpenes in ylang ylang flowers were sesquiterpenes consisting of α-farnesene, α-bergamotene, germacrene-D, β-caryophyllene, humulene, farnesol, trans-β-farnesene and β-ylangen whereas monoterpenes were quantitatively less (10%), comprising mainly *cis*-β-ocimene, *trans*-β-ocimene and β-linalool.

i. Monoterpenoid biosynthesis

The highest transcript accumulation occurred for cytochrome P450 followed by cytochrome P450 monoxygenase during early flower development stages and decreased slightly towards the late stages (-0.89 and -0.96-fold respectively) (*Table 2*). This may be due to a high turn-over rate of linalool to the non-volatile linalool oxide and geraniol products to 2,6-dimethyl-octa-2,6-dene-1,8-diol respectively (*Figure 1*). Several cytochrome P450 genes (*71B36, 64213-66051, 71A4, CYPC, 78A3, 51, CYPA, 72A1, 71D7, 82A1*) and cytochrome P450 monoxygenase genes (*CYP81E11, CYP72A59, CYP83A, CYP51G1, CYP716A12, CYP72B, CYP74A3*) were identified which could be involved in the metabolism of terpenoid volatiles. Cytochrome P450s are found in all kingdoms and show extraordinary diversity in their chemical reactions. They are important in the oxidative, proxidative and reductive metabolisms of numerous and diverse endogenous compounds including terpenoids. Zheng et al. (2015) also found one cytochrome P450 gene (*CYP77A2*) that was up-regulated significantly at the bud stage and down-regulated at the flowering stage in *S. oblata*. Although cytochrome P450s genes are one of the largest gene families in plants, their functions in flowers are still largely unknown.

Linalool synthase and (3R)-linalool synthase were also highly expressed during the early stages. However, the former was down-regulated (0.66-fold) while the latter was up-regulated 1.21-fold towards the late stages. The up-regulation of (3R)-linalool synthase may coincide with a strong floral scent and the rapid increase in the quantity of linalool released during the late stages from S9 to S11 as shown by Sanimah et al.

Table 2. Differential expression of putative genes identified in *Michelia alba* during pooled early stages (S1 – S6) and pooled late stages (S7 – S11) of flower development (E-value ≤1.0 x 10^{-4})

Metabolic pathways and putative genes identified (including isozymes)	Frequency of 21-bp sequence tags		Fold increase (+)/decrease (-)	Blast X protein ID	Organism
	Early stages (S1 – S6)	Late stages (S7 – S11)			
Isoprenoids biosynthesis pathway					
i. Monoterpenoids biosynthesis pathway					
Linalool synthase (LIS)	1,707	1,126	–0.66	CX544248	*Arabidopsis thaliana* (Mouse-ear cress)
(3R)-linalool synthase	1,303	1,589	+1.21	CF085915	*Artemisia annua* (Sweet wormwood)
R-linalool synthase	7	-	suppressed	TA509_39350	*Ocimum basilicum* (Sweet basil)
(-)-linalool synthase	447	562	+1.25	TA20643_3330	*Picea abies* (Norway spruce)
Monoterpene glucosyltransferase (MGT)	42	11	–0.26	DY391589	*Eucalyptus perriniana*
Cytochrome P450	35,193	31,657	–0.89	TA18525_3694	*Arabidopsis thaliana* (Mouse-ear cress)
Cytochrome P450 monoxygenase	9,379	9,094	–0.96	TA41491_3847	*Glycine max* (Soybean)
(-)-limonene synthase	12	11	–0.91	BW986931	*Picea abies* (Norway spruce)
(+)-alpha pinene synthase	3	27	+9.0	TA28042_3330	*Pinus taeda* (Loblolly pine)
(-)-alpha/beta pinene synthase	213	14	–0.07	TA1946_373101	*Picea abies* (Norway spruce)
E-beta-ocimene synthase	29	10	–0.34	CF505428	*Citrus unshiu* (Satsuma orange)
(-)-isopiperitenol dehydrogenase	278	44	–0.16	TA8782_4097	*Mentha piperita* (Peppermint)
Geraniol 10-hydroxylase	272	232	–0.85	DR573344	*Catharanthus roseus* (Rosy periwinkle)
10-hydroxygeraniol oxidoreductase	465	662	+1.42	AJ801875	*Catharanthus roseus* (Rosy periwinkle)
Gamma-terpinene synthase	177	186	+1.05	DT766392	*Citrus unshiu* (Satsuma orange)
ii. Sesquiterpenoids biosynthesis pathway					
(E)-beta-/Farnesene synthase	150	274	+1.83	TA180_34256	*Mentha piperita* (Peppermint)
Farnesyl pyrophosphate synthase (FPPS)	2,037	1,973	–0.97	CV469844	*Lycopersicon esculentum* (Tomato)
Farnesyl pyrophosphate synthetase½ (FPS2)	24	20	–0.83	TA39485_3702	*Arabidopsis thaliana* (Mouse-ear cress)

(cont.)

Table 2. Cont.

Metabolic pathways and putative genes identified (including isozymes)	Frequency of 21-bp sequence tags		Fold increase (+)/decrease (-)	Blast X protein ID	Organism
	Early stages (S1 – S6)	Late stages (S7 – S11)			
Germacrene A synthase/short/long form/ LTC1/LTC2	39	37	-0.94	TA3492_4236	*Cichorium intybus* (Chicory)
(-)-/(+)-Germacrene D synthase	350	317	-0.91	DY335632	*Ocimum basilicum* (Sweet basil)
Sesquiterpene synthase/2/A1/B1	308	342	+1.11	CV476674	*Lycopersicon hirsutum*
Sesquiterpene cyclase/1	397	331	-0.83	CN145031	*Oryza sativa* (japonica cultivar-group)
iii. Terpenoids backbone biosynthesis pathway					
Terpene synthase (TPS)	2,346	7,962	+3.39	DT477167	*Vitis vinifera* (Grape)
Terpenoid synthase	41	17	-0.41	TA19249_338618	*Vitis vinifera* (Grape)
Geranylgeranyl pyrophosphate synthase (GGPS)	383	246	-0.64	DW169459	*Helianthus annuus* (Common sunflower)
1-D/deoxyxylulose 5-phosphate synthase (DXPS)	1,679	1,464	-0.87	DR172445	*Catharanthus roseus* (Rosy periwinkle)
Geranyl/geranyl diphosphate synthase/ large subunit(GPPS)	432	339	-0.79	BF263001	*Arabidopsis thaliana* (Mouse-ear cress)
3-hydroxy-3-methylglutaryl coenzyme A reductase/isoform 1L (HMGR)	3,318	2,436	-0.73	AY741133	*Ginkgo biloba* (Ginkgo)
1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXPR)	474	539	+1.14	TA5739_4236	*Stevia rebaudiana* (Stevia)
Phenylpropanoids/benzenoids biosynthesis pathway					
Benzoyl coenzyme A: benzyl alcohol benzoyl transferase	205	177	-0.86	CF436641	*Nicotiana tabacum* (Common tobacco)
Benzoyl CoA benzoic acid benzoyltransferase	283	286	+1.01	DR543072	Verbena x hybrida
Caffeic acid-O-methyltransferase/2/II	4,455	2,803	-0.63	AJ568905	*Nicotiana tabacum* (Common tobacco)
Caffeic acid 3-O-methyltransferase1/2/3	1,542	1,933	+1.25	TA8853_29729	*Eucalyptus gunnii* (Cider tree)

(cont.)

Table 2. Cont.

Metabolic pathways and putative genes identified (including isozymes)	Frequency of 21-bp sequence tags		Fold increase (+)/decrease (-)	Blast X protein ID	Organism
	Early stages (S1 – S6)	Late stages (S7 – S11)			
Catechol O-methyltransferase	1,157	957	-0.83	BG417465	*Oryza sativa* (Japonica cultivar-group)
Caffeoyl-CoA O-methyltransferase/1/2/3/4/5/6	4,830	7,535	+1.56	BQ655308	*Pinus taeda* (Loblolly pine)
Caffeoyl-CoA 3-O-methyltransferase/1	153	110	-0.72	BP670429	*Linum usitatissimum* (Flax) (Linseed)
Cinnamate 4-hydroxylase/ CYP73	223	49	-0.22	CI641486	*Oryza sativa* (japonica cultivar-group)
Trans-cinnamate 4-hydroxylase	572	503	-0.88	TA432_3197	*Pinus taeda* (Loblolly pine)
Cinnamic acid 4-hydroxylase	200	165	-0.82	TA1224_73275	*Capsicum annuum* (Bell pepper)
Cinnamoyl CoA reductase/2	2,339	1,898	-0.81	TA12192_3332	*Oryza sativa* (japonica cultivar-group)
Cinnamoyl alcohol dehydrogenase	3	0	suppressed	CN206240	*Plantago major* (Common plantain)
Cinnamyl-alcohol dehydrogenase/1/2b/ CAD1	4,412	2,579	-0.58	TA3572_94328	*Fragaria ananassa* (Strawberry)
Salicyclic acid carboxyl methyltransferase (SAMT)	181	120	-0.66	BJ564811	*Ipomoea alba* (Moonflower)
4-coumarate coenzyme A ligase	190	169	-0.88	CA192636	*Zea mays* (Maize)
Fatty acid derivatives biosynthesis pathway					
Methyl jasmonate esterase	306	340	+1.11	TA54109_29760	*Solanum tuberosum* (Potato)
Jasmonate O-methyltransferase	249	168	- 0.67	TA24122_4081	*Arabidopsis thaliana*
Methylesterase	9,120	9,579	+1.05	TA16471_3993	*Lycopersicon esculentum* (Tomato)
O-methyltransferase/ZRP4/family 2	1,920	2,387	+1.24	CA502532	*Oryza sativa* (japonica cultivar-group)
Carboxylesterase (CXE)	568	627	+1.10	BQ582671	*Malus pumila*
24-methylenesterol C-methyltransferase2	581	400	-0.69	EB822518	*Oryza sativa* (Rice)

(*cont.*)

Table 2. Cont.

Others related metabolic pathways	Frequency of 21-bp sequence tags		Fold increase (+)/decrease (-)	Blast X protein ID	Organism
	Early stages (S1 – S6)	Late stages (S7 – S11)			
Lipoxygenase (LOX)	7,231	9,718	+1.35	BI095683	*Ananas comosus* (Pineapple)
Peroxidase	8,172	8,163	-0.99	CD028586	*Spinacia oleracea* (Spinach)
Heat shock protein	29,834	30,605	+1.03	TA1382_263995	*Triticum aestivum* (Wheat)
Myb transcription factor (R2R3-MYB, MYB1)	2,785	2,611	-0.94	CA518555	*Nicotiana tabacum* (Common tobacco)

(2008) although linalool was not detected at all during the early stages (S5 to S8). This major fragrance compound subsequently became the most abundant component of essential oils at mature stages of flower development. (-)-linalool synthase was moderately expressed during early stages but up-regulated 1.25-fold towards the late stages. R-linalool synthase was expressed in very low amounts during the early stages and completely suppressed during the late stages of flower development. This could be due to the conversion of R-linalool to the non-volatile S-linalyl beta-D-glucopyranoside by monoterpene glucosyltransferase (MGT) during the early part of flower development. Lucker et al. (2001) also observed that S-linalool was converted to the non-volatile S-linalyl beta-D-glucopyranoside in transgenic petunia plants.

10-hydroxygeraniol oxidoreductase, gamma-terpinene synthase, (-)-alpha/beta pinene synthase, (-)-isopiperitenol dehydrogenase and geraniol 10-hydroxylase were also moderately expressed during the early stages. However, during the late stages only the former two were up-regulated 1.42- and 1.05-fold respectively, while the latter three were down-regulated 0.07-, 0.16- and 0.85- fold respectively. Monoterpene glucosyltransferase, (-)-limonene synthase, (+)-alpha pinene synthase and E-β-ocimene synthase also showed very low expression during early stages. All of these enzymes were down-regulated during the late stages of flowering except for (+)-alpha pinene synthase. These findings are in agreement with Sanimah et al. (2008) who also found the release of ocimene, limonene and alpa-pinene very low throughout the stages of flower development. In contrast to our findings, Jin et al. (2015) found that most of the monoterpene enzymes in ylang ylang flowers, such as β-ocimene and β-linalool, were undetectable during the early stages but highly induced during flower maturation. In addition, when the MEP and MVA pathways (*Figure 2*) are stimulated, genes

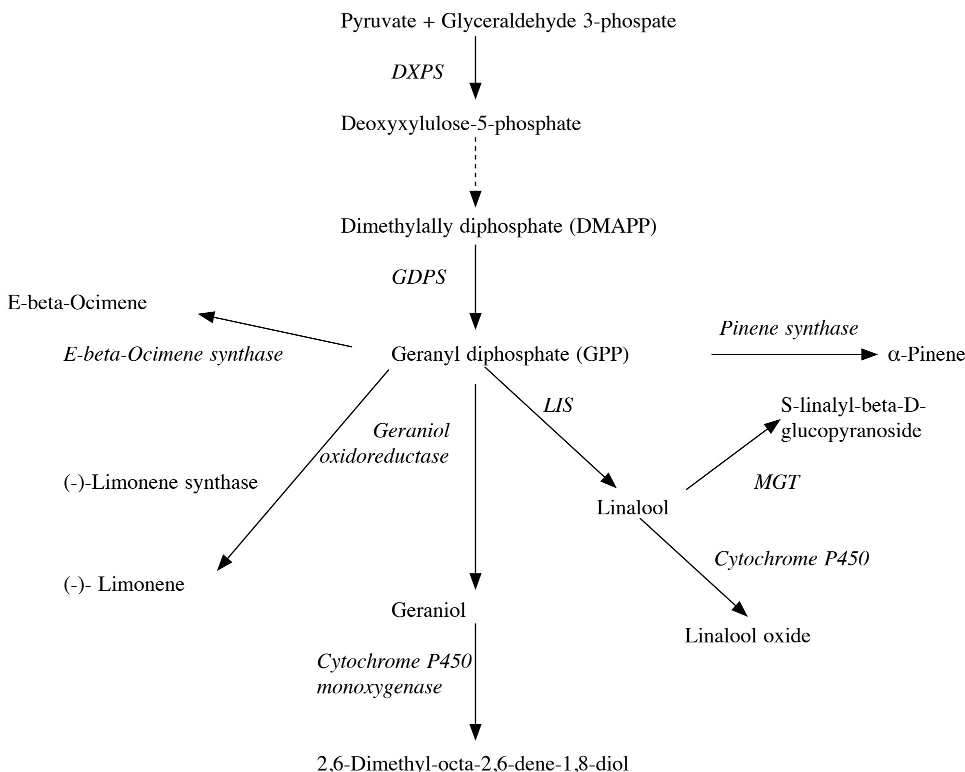

Figure 1. Putative metabolic pathway from Pyruvate and Glyceraldehyde-3-phosphate leading to monoterpene scent biosynthesis, and its related enzymes in Michelia alba. DXPS: Deoxyxylulose-5-phosphate synthase; GDPS: Geranyl diphosphate synthase; LIS: Linalool synthase; MGT: Monoterpene glucosyltransferase

for terpene-degrading enzymes such as limonene synthase and alpha/beta pinene synthase have been shown to have lower expression levels at the late stages, thus maintaining the release of a strong floral scent during blooming.

ii. Sesquiterpenoids biosynthesis

In the sesquiterpenoids biosynthesis pathway, farnesyl pyrophosphate synthase (FPPS) showed the highest transcript accumulation during the early stages but was slightly down-regulated 0.97-fold towards the late stages (*Table 2*). E-beta-farnesene synthase, germacrene D synthase, sesquiterpene synthase and sesquiterpene cyclase were all moderately expressed during the early stages, and their levels slightly decreased during open flower development except for E-beta farnesene synthase and sesquiterpene synthase which increased 1.83- and 1.11-fold respectively. Our finding is in agreement with Sanimah et al. (2008) who also found that the release of farnesene and germacrene-D was low during the early stages of flower development and not detected at all during the late stages. Jin et al. (2015) also found that germacrene-D detected in ylang ylang (*Cananga odorata*) flowers at all stages of flower development was retained or slightly decreased during open flower development. However, they also found that although major sesquiterpenes such as α-farnesene were present in a low quantity during the early stages, they subsequently became the most abundant during the mature stages of flower development in ylang ylang flowers. In

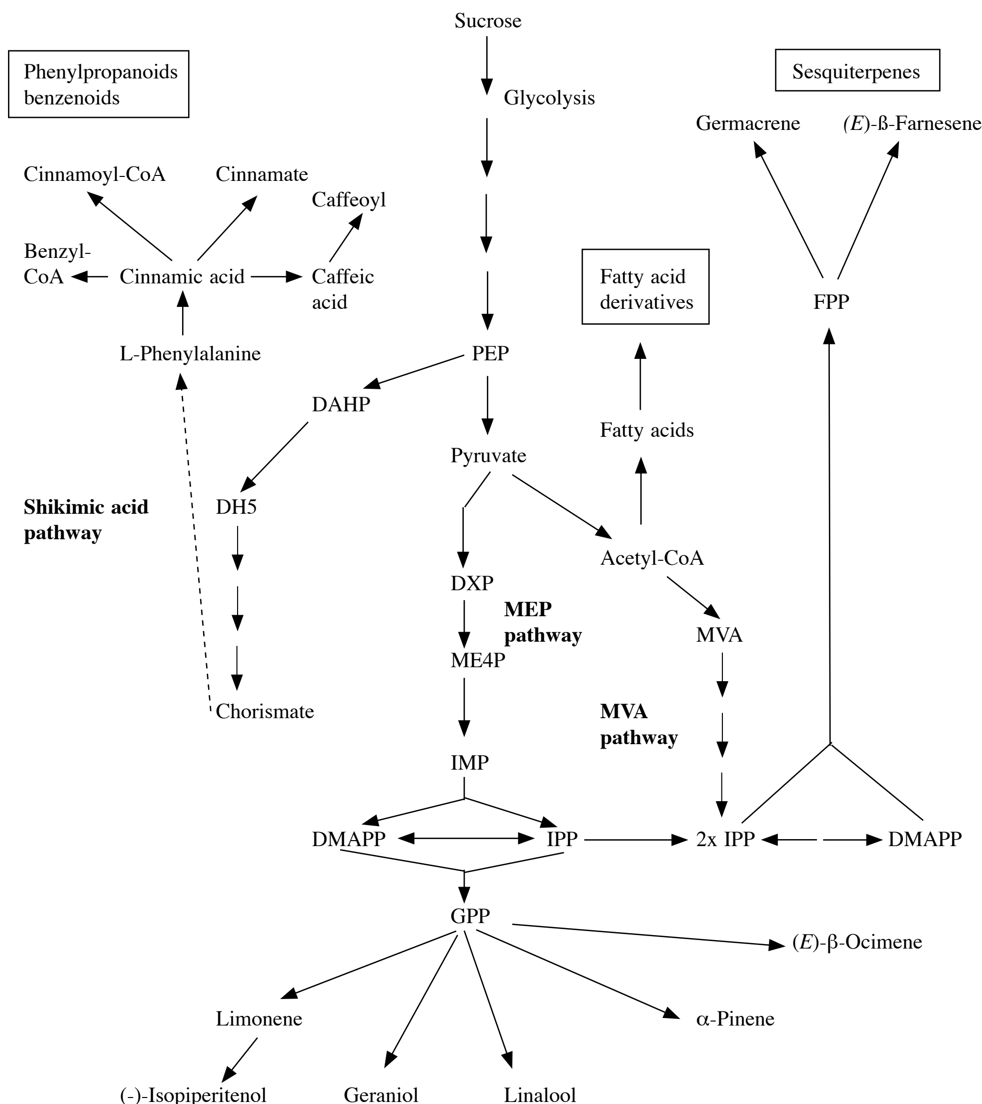

Figure 2. Proposed network of metabolic pathways from sucrose to branchways leading to biosynthesis of floral volatile compounds in Michelia alba flower including monoterpenes, sesquiterpenes, phenylpropanoids/benzenoids and fatty acid derivatives

M. alba, farnesyl pyrophosphate synthetase (FPS2) and germacrene A synthase were expressed in low amounts during the early stages and decreased slightly during late flower development stages.

iii. Terpenoids backbone biosynthesis pathway

We identified 3 transcripts in the terpenoids backbone biosynthesis pathway which showed high expression, i.e. 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGR), terpene synthase (TPS) and 1-deoxy-D-xylulose 5-phosphate synthase (DXPS) with the highest number of transcripts during the early stages of flower development (*Table 2*). During the late stages, terpene synthase was up-regulated 3.39-fold while HMGR and DXPS were down-regulated by 0.73- and 0.87-fold

respectively. In *L. Angustifolia* (Guitton et al. 2010) and snapdragon flowers (Dudareva et al. 2003), transcript levels of TPS were also observed to be upregulated during petal development. However, in ylang ylang (*Cananga odorata*) flowers, the level of terpene synthase detected at all stages was retained or slightly decreased during open flower development (Jin et al. 2015). Terpene synthases are responsible for generating the immense diversity in terpenes produced by plants (McGarvey and Croteau, 1995). Many TPSs have the ability to synthesise multiple products from a single prenyl diphosphate substrate (Degenhardt et al. 2009). Geranylgeranyl pyrophosphate synthase (GGPS) and geranyl diphosphate synthase (GPPS), the precursor for linalool and limonene compounds, were expressed in moderate amounts during the early stages and slightly down-regulated 0.64- and 0.79-fold respectively during the late stages. 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXPR) was also moderately expressed during the early stages of flowering but up-regulated slightly by 1.14-fold during the late stages. Terpenoid synthase was expressed in very low amounts during early stages and down-regulated 0.41-fold during the late stages. However, in *S. Oblata*, three terpenoid synthase genes, camphene synthase, 3S-LIS and TPS showed increased levels of expression during the flowering stage (Zheng et al. 2015).

- **Phenylpropanoids/benzenoids biosynthesis pathway**

A total of 15 enzymes (including isoforms) involved in the biosynthesis of aromatic volatile compounds were identified in this pathway. Six of these enzymes had the highest number of transcripts during the early stages of flowering including caffeoyl-CoA O-methyltransferase followed by caffeic acid O-methyltransferase, cinnamyl-alcohol dehydrogenase, cinnamoyl CoA reductase, caffeic acid 3-O-methyltransferase and catechol O-methyltransferase.
During the late stages of flowering, caffeoyl-CoA O-methyltransferase and caffeic acid-3-O-methyltransferase showed increased levels of expression by 1.56- and 1.25-fold respectively, while caffeic acid O-methyltransferase, catechol O-methyltransferase, cinnamoyl CoA reductase and cinnamyl-alcohol dehydrogenase showed decreased levels of expression by 0.63-, 0.83-, 0.81- and 0.58-fold respectively. In contrast to our findings, Jin et al. (2015) found these enzymes to be present almost exclusively in mature yellow ylang ylang (*Cananga odorata*) flowers. Eight other enzymes, namely benzoyl coenzyme A: benzyl alcohol benzoyl transferase, benzoyl CoA benzoic acid benzoyltransferase, caffeoyl-CoA 3-O-methyltransferase, cinnamate 4-hydroxylase, trans-cinnamate 4-hydroxylase, cinnamic acid 4-hydroxylase, salicyclic acid carboxyl methyltransferase (SAMT) and 4-coumarate coenzyme A ligase were moderately expressed during the early stages and slightly down-regulated during the late stages except for benzoyl CoA benzoic acid benzoyltransferase which showed a slight increase in expression by 1.01-fold. Cinnamoyl alcohol dehydrogenase showed a very low expression during the early stages and was completely suppressed during the late stages (*Table 2*). Salicyclic acid carboxyl methyltransferase which catalyses salicyclic acid (SA) to methyl salicyclate, has been isolated and characterised from many plants including *lilium* 'Yelloween' and shown to play an important role in floral scent methyl benzoate production and emission in lily flowers (Wang et al. 2015).

- **Fatty acid derivatives biosynthesis pathway**

In the fatty acid biosynthesis pathway, a total of six enzymes (including isoforms) were identified during the early flowering stages in *M. alba*. Two of these including methylesterase and O-methyltransferase were highly expressed during the early stages and up-regulated 1.05- and 1.24-

fold respectively during the late stages (*Table 2*) while the remaining four were moderately expressed during the early stages. Carboxylesterase (CXE) and methyl jasmonate esterase were up-regulated 1.10- and 1.11-fold respectively, towards the late stages while 24-methylenesterol C-methyltransferase and jasmonate O-methyltransferase were down-regulated 0.69- and 0.67-fold respectively, towards the late stages. In contrast to our findings, Jin et al. (2015) found that these aromatic compounds were also found almost exclusively in mature ylang ylang flowers.

- **Other related metabolic pathways**

Several metabolic and scent-related enzymes indirectly involved in scent biosynthesis were also identified during the early stages of flowering in *M. alba* as in the preliminary studies (Maheswary et al. 2008; 2011) including lipoxygenase (LOX), peroxidase, heat shock protein and Myb transcription factor. Heat shock protein had the highest expression during the early stages of flowering and increased slightly (1.03-fold) towards the late stages followed by peroxidase which was slightly down-regulated (0.99-fold) towards the late stages. Lipoxygenase (LOX) which has been suggested to play a role in plant-pathogen interactions and their product, jasmonate, in resistance against insects and pathogens (Howe and Schilmiller 2002) also showed a high expression during the early stages and increased 1.35-fold towards the late stages. In addition, LOX genes have also been reported to be involved in converting storage lipids into subtrates for further oxidation to provide energy for scent emission as demonstrated by Hsiao et al. (2006) in *Phalaenopsis bellina*. This indicated the co-expression of enzymes in other metabolic pathways for the fragrance biosynthesis pathways to be complete.

- **Transcriptional regulation of floral scent biosynthesis**

In addition to the structural genes, we also found two genes belonging to the MYB transcription factor family, R2R3-MYB and MYB1 which showed high expression during the early stages but was slightly down-regulated (0.94-fold) during the late stages (*Table 2*). Zheng et al. (2015) also found six genes encoding nine transcription factors regulating fragrance biosynthesis in *Syringa oblata* Lindl inflorescence including R2R3-MYB and R2R3-MYB1which showed a higher expression in the flowering stage compared to the bud stages. In addition, Verdonk et al. (2005) reported that the suppression of ODORANT1 (ODO1) which belongs to the R2R3-MYB family, led to decreased levels of emitted volatile phenylpropanoids in *Petunia hybrida*. From this finding, it can be seen that we can manipulate the release of specific volatiles by the suppression or overexpression of specific genes encoding the release of different fragrances at different stages of flower development. However, the molecular regulation mechanism of floral scent biosynthesis depending on MYBs still needs further research.

Functional annotation using GO analysis

Figures 3a – 3c shows the gene ontology (GO) functional classification of the sequence tags in the 3 main categories, biological process (BP), molecular function (MF) and cellular component (CC). Many of the sequences are represented in more than one category.

In the BP category, the largest proportion of functionally assigned sequences fell into the metabolic processes sub-categories, followed by cellular processes, localisation, response to stimulus and biological regulation (*Figure 3a*) during the early stages and slightly down-regulated during the late stages. The large proportion of annotated sequences may suggest the high expression of genes involved in these processes throughout the flowering process.

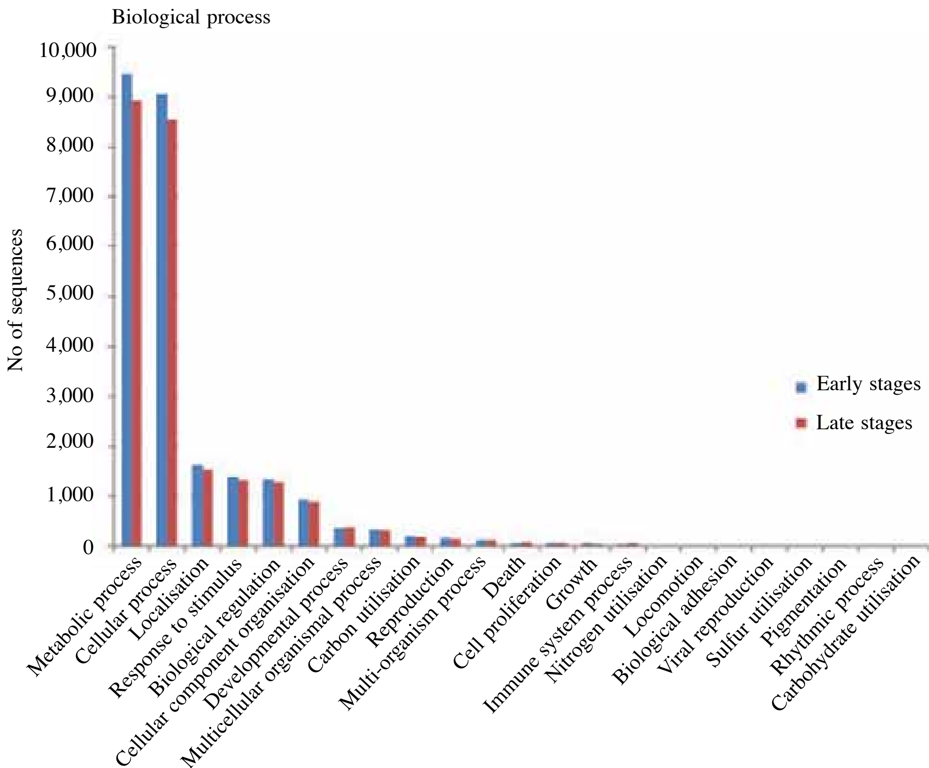

Figure 3a. Gene Ontology (GO) functional subcategories for biological process

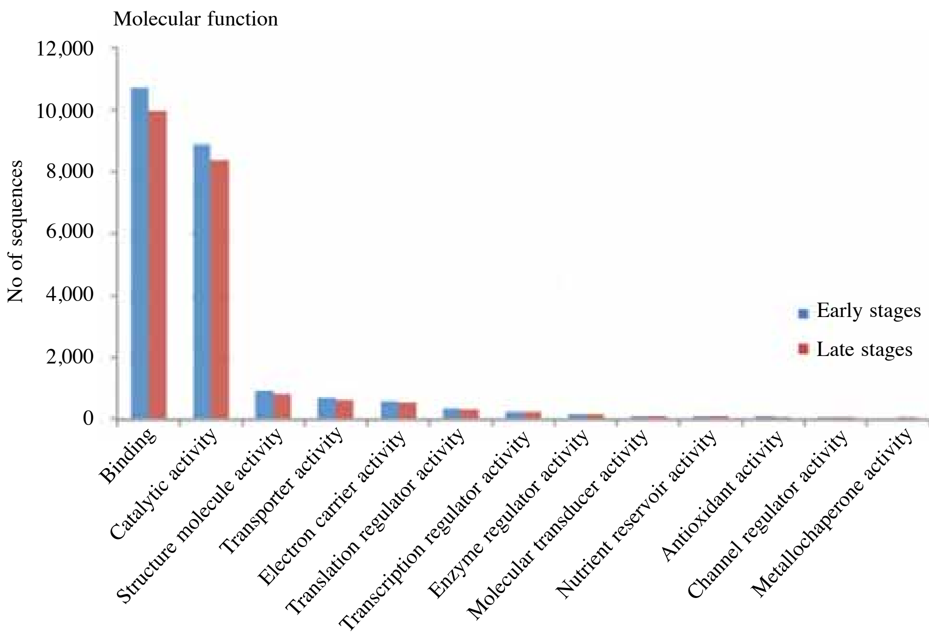

Figure 3b. Gene Ontology (GO) functional subcategories for molecular function

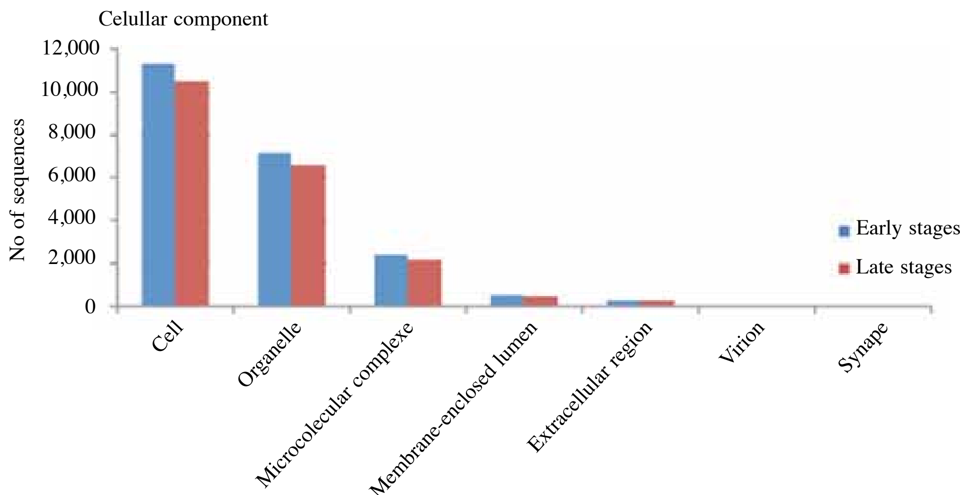

Figure 3c. Gene Ontology (GO) functional subcategories for cellular component

The number of sequences in the remaining subcategories were much lower with slight down-regulation except for developmental process, death, immune system process and viral reproduction which increased by 1.01- to 1.2-fold. Interestingly, the expression of genes in the carbohydrate utilisation subcategory was induced only during the late stages while there was no change in the expression of genes in the locomotion, sulfur utilisation, pigmentation and rhythmic process subcategories.

In the MF category (*Figure 3b*), the most highly represented groups are binding and catalytic activity putative functions. The number of 21-bp sequence tags was 0.50- to 0.96-fold less in all the molecular functions during the late stages of flowering compared to the early stages except for the metallochaperone activity group which was induced only during the late stages.

In the CC category, the majority of the sequences mapped to the cell, organelles and macromolecular complex putative functions. The number of 21-bp sequence tags was 0.89- to 0.92-fold less in all the putative functions during the late flowering stages except for the virion group which was 1.08-fold more during the late stages and the synapse group which was induced only during the late stages.

Conclusion

Our data has revealed several differentially expressed scent-related genes associated with the biosynthesis of VOCs such as terpenoids, phenylpropanoids/benzenoids and fatty acid derivatives during the early to late stages of flowering in *M. alba*. These findings indicate that different volatile compounds are emitted in different quantities at different stages of flower development. The release mechanism of specific floral scents at different stages of flower development may probably be to attract insects for the preparation of the reproduction process. These floral attractants may offer opportunity for further investigation of plant-insect interactions and pollination mechanisms. The availability of transcriptomic data for this flower can provide a useful database and valuable resource for future functional studies such as molecular genetics, variety breeding and biochemical characterisation of this species. It can also help us to better understand the biological and molecular mechanisms of floral scent production in this flower. However, functional characterization of the identified genes has to be carried out at each stage of flower development which can lead to the development of a molecular markers approach to improve and enhance

floral scents in flowers using genetic engineering techniques. A study on the changes in essential oils composition can also contribute to the cosmetic industry.

Acknowledgement
The authors would like to thank Mr Lee Weng Wah for assisting us in the bioinformatics analysis of the genomics data. Special thanks are also due to Ms Lina Rozano and Ms Al-Aida Quswa for assisting us in the bioinformatics analysis of the genomics data and Ms Nor Sufiah Sebaweh for technical assistance. The authors would also like to thank Mr Muhammad Fairuz Mohd Yusof for assisting us in submitting the DGE sequences to NCBI genebank. This work was part of the research supported by the Ministry of Science, Technology and Innovation, Malaysia (NBD Top Down 09-03-03-004, BTK/TD/007).

References

Anon. (2000). JOY perfume tree. Retrieved on 15 June 2000 from http://mgonline.com/michele.html.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S. and Eppig, J.T. (2000). Gene Ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25(2): 25 – 29

Bergougnoux, V., Caissard, J.C., Jullien, F., Magnard, J.L., Scalliet, G., Cock, J., Hugueney, P. And Baudino, S. (2007). Both the adaxial and abaxial epidermal layers of the rose petal emit volatile scent compounds. *Planta* 226: 853 – 866

Botton, A., Galla, G., Conesa, A., Bachem, C., Ramina, A. and Barcaccia, G. (2008). Large-scale Gene Ontology analysis of plant transcriptome-derived sequences retrieved by AFLP technology. *BMC Genomics* 9: 347

Chang, S.J., Puryean, J. and Cairney, J. (1993). A simple and efficient method for isolating RNA from pine trees. *Plant Mol. Biol.* 11:113 – 116

Chen, F., Tholl, D., D'Auria, J.C., Farooq, A., Pichersky, E. and Gershenzon, J. (2003). Biosynthesis and emission of terpenoid volatiles from *Arabidopsis* flowers. *Plant Cell*15(2): 481 – 494

Childs, K.L., Hamilton, J., Zhu, W., Ly, E., Cheung, F., Wu, H., Rabinowicz, P.D., Town, C.D., Buell, C.R. and Chan, A.P. (2007). The TIGR plant transcript assemblies database. *Nucleic Acids Res.* 35: D846 – D851 (Database issue doi:10.1093/nar/gkl785)

Conesa, A., Gotz, S., Garcia-Gomez, J.M., Terol, J., Talon, M. and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Oxford University Press.* 21(18): 3674 – 3676

Degenhardt, J., Köllner, T.G. and Gershenzon, J. (2009). Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry* 70(15-16): 1621-1637

Dudareva, N., Cseke, L., Blanc, V.M. and Pichersky, E. (1996). Evolution of floral scent in *Clarkia*: novel patterns of S-Linalool synthase gene expression in the *C. breweri* flower. *Plant Cell* 8(7): 1137 – 1148

Dudareva, N. and Pichersky, E. (2000). Biochemical and molecular genetic aspects of floral scents. *Plant Physiol* 122(3): 627 – 633

Dudareva, N., Martin, D., Kish, C.M., Kolosova, N., Gorenstein, N. and Faldt, J. (2003).*(E)-B*-ocimene and myrcene synthase genes of floral scent biosynthesis in snapdragon: function and expression of three terpene synthase genes of a new terpene synthase subfamily. *Plant Cell* 15(5): 1227 – 1241

Dudareva, N., Klempien, A., Muhlemann, J.K. and Kaplan, I. (2013). Biosynthesis, function and metabolic engineering of plant volatile organic compounds. *New Phytologist* 198: 16 – 32

Euyama, Y., Hashimoto, S., Nii, H. and Furukawa, K. (1992). The chemical composition of the flower oil and the leaf oil of *Michelia alba* D.C. *J. Ess. Oil Res* 4: 15 – 23

Gotz, S., Garcia-Gomez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talon, M., Dopazo, J. and Conesa, A. (2008). High throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res* 36: 3420 – 3435

Guitton, Y., Nicole, F., Moja, S., Valot, N., Legrand, S. and Jullien, F. (2010). Differential accumulation of volatile terpene and terpene synthase mRNAs during lavender (*Lavandula angustifolia* and *L. intermedia*) inflorescence development. *Physiol Plantarum* 138(2): 150 – 163

Guterman, I., Shalit, M., Menda, N., Piestun, D., Dafny-Yelin, M., Shalev, G., Bar, E., Davydov, O., Ovadis, M. and Emanuel, M. (2002). Rose scent: genomics approach to discovering novel floral fragrance-related genes. *Plant Cell* 14(10): 2325 – 2338

Hao, R.J., Zhang, Q., Yang, W.R., Wang, J., Cheng, T.R., Pan, H.T. et al. (2014). Emitted and endogenous floral scent compounds of *Prunus mume* and hybrids. *Biochem System Ecol*. 54: 23 – 30

Hendel-Rahmanim, K., Masci, T., Vainstein, A. and Weiss, D. (2007). Diurnal regulation of scent emission in rose flowers. *Planta* 226(6): 1491 – 1499

Howe, G.A. and Schilmiller, A.L. (2002). Oxylipinmetabolism in response to stress. *Curr OpinPlant Biol* 5: 230–236

Hsiao, Y.Y., Tsai, W.C., Kuoh, C.S., Huang, T.H., Wang, H.C., Wu, T.S., Leu, Y.L., Chen, W.H. and Chen, H.H. (2006). Comparison of transcripts in Phalaenopsis bellina and Phalaenopsis equestris (Orchidaceae) flowersto deduce monoterpene biosynthesis pathway. *BMC Plant Biology* 6: 14

Jin, J., Kim, M.J., Dhandapani, S., Tjhang, J.G., Yin, J.L., Wong, L., Sarojam, R., Chua, N.H. and Jang, I.C. (2015). The floral transcriptome of ylang ylang (*Cananga odorata* var. *Fruticosa*) uncovers biosynthetic pathways for volatile organic compounds and a multifunctional and novel sesquiterpene synthase. *Journal of Experimental Botany* Doi:10.1093/jxb/erv196

Knudsen, J.T., Eriksson, R., Gershenzon, J. and Stahl, B. (2006). Diversity and distribution of floral scent. *Botanical Review* 72(1): 1 – 120

Li, Z., Cao, H. Liu, L. and Li, B. (2006). Chemical constituents of aroma in fresh *Syringa Oblata* flowers. *J. Zhejiang For Coll*. 23(2): 159 – 162

Li, Y.H., Zhang, W. and Li, Y. (2015). Transcriptomic analysis of flower blooming in *Jasminum sambac* through *De Novo* RNA sequencing. *Molecules* 20: 10734 – 10747

Maheswary, V., Yong, S.H., Nurul Aishah, Y., Sew, Y.S., Khairun, H.N. and Hassan, M.D. (2008). Isolation of transcripts related to floral scent biosynthesis from Cempaka putih (*Michelia alba*) flower using subtractive hybridization. *Pertanika J. Trop. Agric. Sci*. 31(1): 107 – 116

Maheswary, V., Sanimah, S., Yong, S.H., Nurul Aishah, Y. and Vasanthi, S. (2011). Differentially expressed scent-related genes in early bloom *Micelia alba*. *J. Trop. Agric.* and *Fd. Sc*. 39(2): 191 – 201

McGarvey, D.J. and Croteau, R. (1995). Terpenoid metabolism. *The Plant Cell* 7: 1015 – 1026

Muhlemann, J.K., Klempien, A. and Dudareva, N. (2014). Floral volatiles: from biosynthesis to function. *Plant, Cell and Environment* 37: 1936 – 1949

Orlova, I., Marshall-Colón, A. and Schnepp, J. (2006). Reduction of benzenoid synthesis in petunia flowers reveals multiple pathways to benzoic acid and enhancement in auxin transport. *The Plant Cell* 18: 3458 – 3475

Philippe, N., Elias, B.S., Anthony, B., Alban, M., Florence, R., Qiang, B., John, De V., Eric, R. and Therese, C. (2013). Combining DGE and RNA-sequencing to identify new poly A^+ non-coding transcripts in the human genome. *Nucleic Acids Research* 42(5): 2820 – 2832

Sanimah, S., Mohd Shaib, J., Nor Afni, Z. and Siti Saleha, T. (2002). Preliminary study on *Michelia alba* (cempaka putih) floral development. *Proceedings of the 14th Malaysian Society of Plant Physiology Conference,* 10 – 12 Sept, 2002, Melaka (Jaafar, H.Z., Md. Shah, U.K. and Madon, M.S. eds.) p. 24 – 26

Sanimah, S., Suri, R., Nor Azizun, R., Hazniza, A., Radzali, M., Rusli, I. and Hassan, M.D. (2008). Volatile compounds of essential oil from different stages of *Michelia alba* (cempaka putih) flower development. *J. Trop. Agric. and Fd. Sc*. 36(1):109 – 119

Shang, C., Hu, Y., Deng, C. and Hu, K. (2002). Rapid determination of volatile constituents of *Michelia alba* flowers by gas chromatography-mass spectrometry with solid-phase microextraction. *Journal of Chromatography A* 942: 283 – 288

Verdonk, J.C., Haring, M.A., Tunen, A.Jv. and Schuurink, R.C. (2005). *ODORANT1* regulates fragrance biosynthesis in petunia flowers. *Plant Cell* 17(5): 1612 – 1624

Vogt, T. (2010). Phenylpropanoid biosynthesis. *Molecular Plant* 3: 2 – 20

Wang, H., Sun, M., Li, L.L., Xie, X.H. and Zhang, Q.X. (2015). Cloning and characterization of a benzoic acid/salicyclic acid carboxyl methyltransferase gene involved in floral scent production from lily (*Lillium* 'Yelloween'). *Genetics and Molecular Research* 14(4): 14510 – 14521

Yue, Y., Yu, R. and Fan, Y. (2015). Transcriptome profiling provides new insights into the formation of floral scent in *Hedychium coronarium*. *BMC Genomics* 16: 470

Zhang, H.X., Hu, Z.H., Leng, P.S., Wang, W.H., Xu, F. and Zhao, J. (2013). Qualitative and quantitative analysis of floral volatile components from different varieties of *Lillium* spp. *Scientia Agri Sin*. 46(4): 790 – 799

Zhao, J., Hu, Z.H., Leng, P.S., Zhang, H.X. and Cheng, F.Y. (2012). Fragrance composition in six tree peony cultivars. *Kor. J. Hortic. Sci. Technol*. 30(6): 617 – 625

Zheng, J., Hu, Z., Guan, X., Dou, D., Bai, G., Wang, Y., Guo, Y., Li, W. and Leng, P. (2015). Transcriptome analysis of *Syringa oblata* Lindl. Inflorescence identifies genes associated with pigment biosynthesis and scent metabolism. *PloS ONE 10*(*11*): e0142542. Doi:10.1371/journal.pone0142542

Abstrak

Walaupun *Michelia alb*a (cempaka putih) mempunyai sejarah lama dalam penanaman dan kepentingan komersial minyak patinya, tidak terdapat maklumat terkini potensi perubahan dalam ekspresi gen-gen yang terlibat dalam sintesis setiap bahan organik meruap (*VOC*) semasa perkembangan bunganya. Di sini kami menghuraikan penggunaan gabungan *Digitial Gene Expression (DGE)* dan analisis kimia (dari kajian terdahulu) untuk mengenal pasti gen-gen dalam tapak biosintesis *VOCs* dan gen-gen berkaitan. Dari data penjujukan *DGE* sebanyak 50 gen-gen (termasuk isoform) dikenal pasti terlibat secara langsung dengan biosintesis wangi daripada gabungan peringkat-peringkat awal (S1 – S6) dan gabungan peringkat-peringkat akhir (S7 – S11) semasa perkembangan bunga. Data transkriptom kami mendedahkan bahawa sebahagian besar gen-gen (28) yang menghasilkan wangi bunga dalam spesis ini didapati daripada tapak utama biosintesis isoprenoid di mana 15 transkrip tergolong dalam tapak biotensis monoterpenoid, 6 transkrip tergolong dalam tapak biosintesis monoterpenoid, 6 transkrip tergolong dalam tapak biosintesis tulang belakang terpenoid, manakala baki 7 transkrip adalah tergolong dalam tapak biosintesis tulang belakang terpenoid. Lima belas gen dikenal pasti terlibat dalam biosintesis bahan aromatik dalam tapak fenilpropanoid/benzenoid manakala 6 gen didapati tergolong dalam tapak biosintesis derivatif asid lemak. Empat gen lain juga dikenalpasti berkaitan dengan biosintesis aroma bunga di tapak metabolik lain