

# Transcriptome and metabolomic profiling: the effect of LED light quality on morphological traits, and phenylpropanoid-derived compounds accumulation in *Sarcandra glabra* seedlings

**Dejin Xie**

Fujian Agriculture and Forestry University <https://orcid.org/0000-0003-4879-4981>

**Lingyan Chen**

Fujian Agriculture and Forestry University

**Chengcheng Zhou**

Fujian Agriculture and Forestry University

**Muhammad Waqqas Khan Tarin**

Fujian Agriculture and Forestry University

**Deming Yang**

Fujian Agriculture and Forestry University

**Ke Ren**

Fujian Agriculture and Forestry University

**Tianyou He**

Fujian Agriculture and Forestry University

**Jundong Rong**

Fujian Agriculture and Forestry University

**Yushan Zheng** (✉ [zys1960@163.com](mailto:zys1960@163.com))

---

## Research article

**Keywords:** *Sarcandra glabra*, LED lights, Transcriptome, Metabolomic, Phenylpropanoid biosynthesis

**Posted Date:** May 31st, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-26873/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at BMC Plant Biology on October 15th, 2020.  
See the published version at <https://doi.org/10.1186/s12870-020-02685-w>.

# Abstract

## Background

*Sarcandra glabra* is an evergreen and traditional Chinese herb, having medicinal significance as anti-oxidant, anti-bacterial, anti-inflammatory, and anti-tumor. Recently, China has initiated to establish cultivation of this plant in greenhouse under artificial light-emitting diodes (LED). However, little is known regarding the effects of the different LED lights on plant growth, accumulation pattern of secondary metabolites, and the molecular mechanism of *Sarcandra glabra*.

## Results

Compared to white light (WL), the red light (RL) increased the height and decreased the stem diameter and leaf area, while blue light (BL) suppressed the height and leaf area. According to our transcriptome profiling, some differentially expressed genes (DEGs) were enriched in the phenylpropanoid biosynthesis. We identified 46 unigenes encoding for almost all known enzymes involved in phenylpropanoid biosynthesis, while the expression level of RNA-seq and qPCR were largely consistent. Meanwhile, we found 53 unigenes encoding R2R3-MYB proteins and 53 unigenes encoding bHLH proteins that several of them were related to flavonoids biosynthesis. Based on metabolomic profiling, a total of 454 metabolites were detected and the distribution of chemicals varied significantly. While flavonoids, phenolic acids, and tannins were mainly located in leaves; Organic acids, lignans and coumarins, and terpenoids were much more abundant in WG (root tissue under WL). Meanwhile, the yields of most flavonoids from BY (leaf tissue under BL) and the synthesis of primarily targeted compounds was lower than in WY (leaf tissue under WL) and RY (leaf tissue under RL). Instead, the leaves grown under RL exhibited a greater production of bioactive phytochemicals such as esculetin, fraxetin, esculin, and scopoletin.

## Conclusion

These results provide further insight into the molecular mechanism of metabolites accumulation patterns in *S. glabra* under different light conditions, enabling the development of optimum breeding conditions for this plant.

## Background

*Sarcandra glabra* (Thunb.) Nakai (Caoshanhu or Zhongjiefeng in Chinese medicine) belongs to the Chloranthaceae family, a widely distributed taxonomic group in southern of China and throughout Southeast Asia. *S. glabra* is an evergreen ornamental plant and traditional Chinese herb that uses seeds or division propagation to reproduce [1, 2]. The Chinese pharmacopoeia records that the gross dry plant has been used as medicine [3], which contains many plentiful bioactive phytochemicals that not only contribute to protect the plant against abiotic or biotic stress [4], but also confer anti-oxidant [5], anti-

bacterial [6], anti-inflammatory [7], and anti-tumor [8] properties. Previous research on this plant mainly focused on isolating and extracting its phytochemical components, which can be divided into six groups: organic acids [9, 10], flavonoids [9], coumarins [2], terpenoids [11], phenolic acids [5, 12], and polysaccharides [1]. Among these components, flavonoids, coumarins (e.g., isofraxidin and scopoletin), and phenolic acids (e.g., rosmarinic acid) are closely associated with the phenylpropanoid pathway, and their production and subsequent accumulation varies in different plant tissues and with different seasons.

To address the limited natural resources and meet consumer demand for safe and high-quality herbal products, China has initiated the project of Good Agricultural Practice (GAP) based on Chinese materia medica to improve herbs resources, and quality of cultivation. Thus, the production of *S. glabra* is gradually being shifted from the old cultivation model to more modern and sustainable practices, such as stereoscopic cultivation [13] or greenhouse cultivation [14]. Greenhouse cultivation is considered as a particularly effective agricultural method to produce large quantities of herb resources and ensuring product quality [15]. Light is a pivotal element in greenhouse internal environment, especially new-generation of LEDs that provide favorable conditions for plant growth and development [16, 17].

Nowadays, a large number of studies have shown that different LED lights trigger a cascade of responses to plant morphology, physiology, and metabolite accumulation in plants [18, 19]. Reportedly, stem length and leaf area in tomato seedlings were both restricted under the higher BL to RL proportion [20, 21]. Similarly, the hypocotyl length and leaf area were suppressed when the cucumber seedlings were exposed to a high percentage of BL [22]. The primary and species-specific secondary metabolites ingredients may be another factor of concern. For instance, the BL significantly increased the concentration of phenolic compounds in pea sprouts and basil [23, 24], flavonoids production in *Cyclocarya paliurus* [25], and the amount of salidroside in callus culture of *Rhodiola imbricata* [26]. By contrast, the red and far-red parts of the light spectrum induce the production of salicylic acid, which in turn increased flavonoid accumulation in *Ginkgo biloba* [27]. Furthermore, it was found that yellow LED light had beneficial effects on flavonoids concentration in *Epimedium pseudowushanense* [28]. Collectively, these findings demonstrated that flavonoids and other phenolic compounds derived from the phenylpropanoid pathway accumulated in response to the monochromatic LED light irradiation.

The phenylpropanoid pathway is also responsible for the production of coumarins, lignins, and phenolic acids [29]. Several studies have been conducted to describe phenylpropanoid biosynthesis, the various mechanisms involved in this complex pathway, as well as their influencing factors, such as enzymatic interactions, transcription factors (TFs), and the environmental factors [29, 30]. For TFs, MYB, bHLH, WDR, and the MBW complex (comprising of MYB, bHLH, and WD) are well conserved in plants and play a vital role in the regulation of the phenylpropanoid pathway [31–35].

This study aimed to experimentally compare the effects of WL, RL, and BL on *S. glabra* seedlings, and analyze the transcriptome and metabolomic profiling within these treatments. We also aimed to identify the structural and TFs (R2R3-MYB and bHLH) genes involved in the biosynthesis of flavonoids,

rosmarinic acid, and some coumarins (scopoletin, scopolin, and fraxetin), as well as their expression pattern were also discussed. The findings of this study will enable a better understanding of the effects of the different monochromatic LED lights on plant growth and secondary metabolite accumulation, and promote the quantity and quality of *S. glabra* cultivation by fine-tuning the optimum light condition.

## Results

### Effects of different light qualities on plant growth

At the final harvest (day 60), the leaf area was significantly lower in the RL and BL treatments than in the WL treatment (Fig. 1C), while plant height was significantly higher and lower in the RL and BL treatments, respectively, compared to the WL (Fig. 1D). Finally, on day 60, the stem diameter of samples in the WL and BL treatments was significantly larger than those in RL treatment (Fig. 1E). All statistical data were shown in Additional file 1: Table S1.

### Illumina sequencing, de novo assembly, and annotation of the reference transcriptome

In total, 1,025 million raw reads were generated from the 15 cDNA libraries with a total of 1,009 million high-quality clean reads. To be specific, each library produced about 55.9–92.7 million raw reads (CRA002486). The relevant parameters of *de novo* assembly were shown in Additional file 2: Table S2. The results of the functional annotation from the seven public databases showed that 46.9% of unigenes were annotated in at least one database, while only 3.0% of unigenes were predicted from all the databases (Additional file 3: Table S3). A total of 33,626 unigenes were categorized into 56 functional groups among three categories: biological process, cellular component and molecular function (Additional file 4: Fig. S1).

### DEGs analysis

To identify DEGs in the 15 samples, a pairwise comparison was performed and filed with  $\text{padj} < 0.05$  and  $|\log_2\text{FoldChange}| > 1$ . The results showed that 861 (BY vs RY), 378 (BY vs WY), 47 (RY vs WY), 10,033 (WG vs WY), 7917 (WJ vs WG), and 6379 (WJ vs WY) unigenes were significantly expressed.

The enrichment analysis was performed using the GO ontology analysis (Additional file 5: Fig. S2a-e). A large number of DEGs were predominant in five terms, namely: metabolic process, single-organism metabolic process, catalytic activity, ion binding, and transferase activity. Meanwhile, all DEGs were mapped in the KEGG database (Additional file 6: Fig. S3a-f). Interestingly, in almost every group (apart from RY vs WY) the DEGs were enriched in the biosynthesis of phenylpropanoids (including flavonoids). The phenylpropanoid-derived compounds are just originated from this biosynthetic pathway. Therefore, the follow-up experiment and analysis were performed to describe this regulatory mechanism, based on transcriptome and mebabolomic profiling.

### Metabolism analysis

Based on results obtained from the UPLC-MS/MS platform and self-built database, a total of 454 metabolites were detected, including various classes of lignans and coumarins, flavonoids, phenolic acids and other, and the majority of them are phenylpropanoid-derived compounds (Table 1 and Additional file 7: Table S4). Among them, fraxetin, scopolin, isofraxidin, fraxidin and esculin from the lignans and coumarins class and rosmarinic acid, rosmarinyl glucoside, chlorogenic acid, esculetin from the phenolic acids category were isolated from the samples, which are the major ingredients of *S. glabra*.

Table 1  
Classes of the detected metabolites in *S. glabra*

Metabolite class	Number of metabolites	Metabolic class	Number of metabolites
Flavonoids	90	Organic acids	32
Alkaloids	39	Lipids	60
Phenolic acids	66	Tannins	13
Lignans and Coumarins	11	Terpenoids	10
Amino acids and derivatives	52	Others	53
Nucleotides and derivatives	28		

PCA (Fig. 2A) and HCA (Fig. 2B) analyses aimed to evaluate the temporal accumulation patterns of metabolites identified in our samples. The PCA results showed that the samples tested were significantly different and that replication was generally consistent with that of others in the same group. The HCA presented metabolic profiling and distributed them into 10 classes in all of the samples. As shown in Fig. 2B, most flavonoids from BY were higher than in WY and RY. In contrast to WY, the contents of tannins in BY and RY were significantly different. However, the chemicals distribution was significantly different between the root (WG) and leaves (WY, RY, and BY). Leaf tissues contained a higher abundance of compounds such as flavonoids, phenolic acids, and tannins, while lipids, amino acids and derivative alkaloids, organic acids, nucleotides and derivatives, lignans and coumarins, terpenoids were much more abundant in the root tissue.

The results of the comparative analysis are shown in Table 2. In the WY vs RY group, compared with WY, 11 flavonoids and six phenolic acids were significantly down-regulated in RY. Among them, only three flavonols (kaempferol, kaempferin, and cynaroside) were annotated in the flavonoid biosynthesis. In the WY vs BY group, 40 flavonoids, 11 phenolic acids, two lignans and coumarins, and three terpenoids were significantly differences. Only eight out of 40 flavonoids were mapped in the flavonoid biosynthesis, six of which increased by 3.2- to 6.2-fold in BY, but the content of cynaroside and protocatechuic aldehyde was approximately reduced by 4.5-fold in BY. For phenolic acids, 7 metabolites were down-regulated, including caffeic acid, esculetin, and sinapinaldehyde, while others increased by 2.5- to 3.3-fold. As for the two lignans and coumarins, daphnetin production decreased by 9.5-fold, while that of oxypeucedanin increased by 2.1-fold in BY. In the RY vs BY group, the flavonoids (49 compounds) showed similar accumulation patterns to those of the WY vs BY group. For phenolic acids (21 compounds), the

production of 12 compounds in BY was higher than that of RY by 2- to 5-fold, while that of the remaining compounds significantly decreased. Especially, in the WY vs WG group, compounds derived from flavonoid biosynthesis mainly accumulated in WY compared to WG. In contrast, some coumarins (esculin, isofraxidin, and fraxidin) were more abundant in WG, but other coumarins (scopolin and oxypeucedanin) and most phenolic acids were significantly more abundant in WY. Furthermore, we selected the main active constituents and compared the metabolic profiles of different samples (Fig. 3). It should be noted that BL inhibited the production of rosmarinic acid, esculetin, isofraxidin, scopolin, scopoletin, fraxidin, fumaric acid, and maleic acid, while RL stimulated the accumulation of esculetin, fraxetin, esculin, scopoletin, fumaric acid, and maleic acid in leaves.

Table 2

The quantitative statistics for the differential analysis of identified metabolites in different groups

Group name	Identified Metabolites	Identified Metabolites (significant)	Up (Down)-regulated	Annotated in KEGG	Annotated in KEGG (significant)
WY vs RY	450	44	29 (15)	203	14
WY vs BY	450	87	34 (53)	203	40
RY vs BY	449	114	45 (69)	202	51
WY vs WG	454	296	128 (168)	203	123

## Candidate genes involved in flavonoid, coumarin, and phenolic acid biosynthesis

We have identified 46 candidate unigenes encoding most of enzymes involved in the flavonoids, coumarins, and phenolic acid biosynthesis (Table 3 and Additional file 8: Table S5). Moreover, five unigenes relating to rosmarinic acid biosynthesis were also detected. All of these compounds are related in one way or another to the core phenylpropanoid pathway, comprising the three committed steps catalyzed by phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase (C4H), and 4-coumaroyl CoA ligase (4CL) enzymes. In this study, four PAL unigenes, four 4CL unigenes, and one C4H unigenes were identified, of which the PAL unigene (Cluster-22883.50216), 4CL unigene (Cluster-22883.47381), C4H unigene (Cluster-22883.50271) exhibited the highest expression level compared to other members. These also showed higher expression in BY and RY treatments. Five unigenes encoding hydroxycinnamoyltransferases were found, and the expression of unigene (Cluster-22883.48487) was detected in all samples, but it was most highly expressed in BY and RY. Furthermore, the unigene (Cluster-22883.21708) was only highly expressed in WG compared to other treatments. The scopoletin biosynthesis is derived from a key precursor - feruloyl CoA, produced by the action of caffeic acid 3-O-methyltransferase (COMT) and 4CL on caffeic acid or by Caffeoyl CoA 3-O-methyltransferase (CCoAOMT) reacting with caffeoyl CoA. In this pathway, only one COMT unigene (Cluster-22883.56077)

and one CCoAOMT unigene (Cluster-22883.54879) were detected, which were both significantly up-regulated in WG tissues and down-regulated in RY and BY compared to WY. Furthermore, five unigenes belonging to 2-oxoglutarate and Fe(II)-dependent dioxygenases were characterized. The unrooted phylogenetic tree showed that these were highly similar to feruloyl-CoA 6'-Hydroxylase (F6'H) family proteins. Among them, Cluster-22883.51443 and Scopoletin 8-hydroxylase protein (S8H, XP\_030962790.1) from *Quercus lobata* showed maximum similarity, Cluster-22883.11702 and Cluster-22883.33018 clustered with other F6'H1 enzymes that catalyzes feruloyl CoA to scopoletin (Additional file 9: Fig. S4). The pathway to the product of rosmarinic acid consists of four consecutive enzyme reactions, starting from the precursor (tyrosion) and three enzymes [tyrosine aminotransferase (TAT), hydroxyphenylpyruvate reductase (HPPR), and Cytochrome P450 reductase (CYP)] encoded by their respective unigenes (Cluster-22883.49301, Cluster-22883.50853, and Cluster-22883.50532). As for the rosmarinic synthase (RAS) enzyme, it is thought to be encoded by the unigene (Cluster-22883.29589) although this has not been conclusively proven (Additional file 10: Fig. S5). The unigene of Cluster-22883.49301 showed no significant differential expression in WY, RY, and BY, but Cluster-22883.50853 were highly expressed in BY. Cluster-22883.50532 had a significantly higher expression in WG, whereas it was significantly lower in BY, leading to a lower production of rosmarinic acid.

Table 3  
Summary of identified unigenes in the transcriptome of *S. glabra*

Gene	Number of unigenes	Gene	Number of unigenes	Gene	Number of unigenes
PAL	4	CHS	2	ANR	1
4CL	4	CHI	2	ANS	1
C4H	1	F3H	2	TAT	1
C3H	2	F3'H	1	HPPR	1
COMT	1	FLS	1	HPPD	1
CCoAOMT	1	UFGT	1	RAS	1
F6'H	4	RT	2	CPR	1
S8H	1	DFR	1	Hydroxycinnamoyl-transferase	5
SGTF	1	LAR	1	O-MT	2

The flavonoids biosynthesis shared the core pathway incorporating chalcone synthase (CHS), chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H) enzymes. The expression level of the unigenes encoding the above mentioned enzymes was higher in RY and BY than in WG. When the intermediate compound (dihydrokaempferol) was produced by the core enzymatic reactions, it would be flowed into several

branches to generate anthocyanins, proanthocyanidins, flavonols, etc. The biosynthesis of kaempferol is catalyzed by flavonol synthase (FLS) (Cluster-22883.47725) enzyme on dihydrokaempferol, and its production varied significantly with changing gene expression level under different LED lights and in different tissues. Rutin also originates from the precursor of dihydrokaempferol through the enzymatic reactions by flavonoid 3'-hydroxylase (F3'H, Cluster-22883.55303), FLS (Cluster-22883.47725), UDPG flavonoid O-glucosyltransferase (UFGT, Cluster-22883.49005) and rhamnosyl transferase (RT, Cluster-22883.85233 and Cluster-22883.49258). These unigenes (except Cluster-22883.85233) had similar expression patterns that BL had the potential to induce the higher expression quantity, while FLS was the limiting factor resulting in a lower production of flavonol compounds in the RL environment. Interestingly, the unigene of Cluster-22883.85233 was mainly expressed in root tissues to synthesize rutin, but Cluster-22883.49258 showed higher expression levels in leaves and stem tissues. A series of enzymes [dihydroflavonol 4-reductase (DFR), leucoanthocyanidin reductase (LAR), anthocyanin synthase (ANS), and anthocyanidin reductase (ANR)] are responsible for the production of catechin and epicatechin, which eventually synthesize the product of proanthocyanidins (PAs), and the unigenes encoding these enzymes had low expression levels in WG compared to other samples. Additionally, catechin was produced in large amounts in the leaf tissues, catalyzed by the combined action of these highly-expressed structural genes under the BL and RL environment. Overall, the BL and RL cultivation increased the expression level in a majority of structural genes, such as some isoforms in the PAL gene family, 4CL gene family, CHS, CHI, F3H, and C4H. The phenylalanine metabolic pathways were shown in Fig. 4A, and the heat map diagram of the aforementioned unigenes with the FPKM values was presented in Fig. 4B.

## Candidate transcription factors of MYB and bHLH

A total of 142 raw MYB (MYB and MYB-related) unigenes and 61 raw bHLH unigenes were collected. For MYB proteins, after screening, 92 MYB unigenes had different numbers of highly conserved DNA-binding domains and were divided into 4 classes: 1R-MYB, R2R3-MYB, 3R-MYB, and 4R-MYB proteins including 35, 53, 3, and 1 unigenes, respectively. Since the *Arabidopsis* genome was published, it has provided the first insight into the description and classification of plant MYB TFs. In the plants, R2R3-MYB proteins are the largest class in MYB family. They are involved in an extensively set of primary and secondary metabolism and have been divided into subgroups in *Arabidopsis* based on the conservation of the DNA-binding domain and of amino acid motifs in the C terminal domain [32]. Here, we have identified all the primary structures of 53 R2R3-MYB domains in *S. glabra* (Additional file 11: Fig. S6). We also selected 53 unigenes encoding R2R3-MYB in *S. glabra* to construct a phylogenetic tree with 125 full-length R2R3-MYB amino acid sequences from *Arabidopsis thaliana* (Fig. 5A). Numerous studies have shown that some subgroups of R2R3-MYB proteins have been involved in the regulation of flavonoids biosynthesis to date, namely S4, S5, S6, and S7. In subgroup 7, AtMYB11, AtMYB12, and AtMYB111 were closely related to flavonols biosynthesis. The unigene SgMYB38 was clustered with AtMYB111. The unigene of SgMYB53 showed a high degree of functional similarity within the S6 (AtMYB75/90/113/114), regulating the anthocyanin pigments pathway. A subset of six unigenes of SgMYB8/14/28/43/27/39 was in the same branch of the PAs biosynthetic pathway with AtMYB123 in subgroup 5. Moreover, we identified two

unigenes: SgMYB6 with a high amino acid sequence similarity to AtMYB38 (RAX2, S14), and SgMYB32 clustered with members of subgroup 16. The heat map showed the expression patterns of the 10 unigenes mentioned earlier according to the FPKM values (Fig. 5C). Based on the transcriptome results, the relative expression quantity of SgMYB38 was up-regulated in BY. In subgroup 5, the six unigenes had a significantly different expression pattern among different samples. SgMYB28, SgMYB8, and SgMYB39 exhibited the highest expression level in WY, compared to RY and BY. Their expression was also significantly lower in the WJ and WG. Meanwhile, SgMYB43 displayed a dramatic increase in expression in BY. Remarkably, SgMYB27 presented relatively higher expression in WJ, while both of SgMYB6 and SgMYB32 were strongly expressed in WG.

The bHLH proteins are another ancient and functionally diverse superfamily of TFs, which has been widely investigated in the past. In this study, we have identified 53 unigenes encoding bHLH proteins that have the distinguishable signature domain. We constructed the phylogenetic tree with these unigenes in *S. glabra* and 182 full-length AtbHLH protein sequences. The results of multiple sequences alignment of SgbHLH domains are displayed in Additional file 12: Fig. S7. As shown in Fig. 5B, all the bHLH proteins were divided into 21 subclasses indicating that the node values between different subclasses were low, but values within each subclass were high, suggesting that the latter have strong evolutionary relationships. A previous study classified bHLH proteins into 32 subfamilies [43], with subfamilies 2, 5, and 24 (corresponding to subclasses of 8, 7, and 15 of AtbHLH proteins) involved in the regulation of flavonoid biosynthesis. In a subclass of 8, we have identified a group of unigenes named Sg(bHLH) 22/24/25/43/45, which showed high sequence similarity with other AtbHLH protein members. Compared to the similar gene-expression change between the WY and RY groups, the relative expression level of these five unigenes was significantly lower in BY. Particularly, Sg45 was significantly up-regulated in WJ and WG. Within the subfamily of 7, Sg13 and Sg12 were clustered with the subset of At1g63650 (EGL), At5g41315 (GL3), At4g00480 (MYC1), and At4g09820 (TT8) that associated with some R2R3-MYB members to participate in the regulation of flavonoids metabolism, trichome formation, epidermal cell fate specification, and in the formation of hair and non-hair cells in root epidermis cells [43]. Furthermore, three unigenes of Sg(bHLH)4/27/37 encoding SPATULA (SPT), PIF7, and PIF1 belonged to subfamily of 15. The expression level of Sg4 was slightly reduced in BY and RY in comparison with WY, but Sg27 showed a lower level in BY, WJ, and WG (Fig. 5D).

## Verification of the transcript expression using qRT-PCR

To confirm the accuracy of the RNA-seq results, a total of 24 unigenes (five R2R3-MYB unigenes, five bHLH unigenes, and another 14 structural genes related to flavonoids biosynthesis) were selected (Additional file 8: Table S5). The qRT-PCR results for all unigenes were tested analyzed, revealing that the expression levels were mostly consistent with those obtained by RNA-seq. Furthermore, the correlation coefficient (R) for each unigene between RT-qPCR and RNA-seq for each unigene showed that RNA-seq data were accurate and hence can be used in future experiments (Fig. 6).

## Discussion

# Effects of different LED lights on plant morphological traits

In this study, we showed that LED lights had a significant influence on the morphological characteristics of *S. glabra*. Specifically, red light enhanced plant growth while blue light restricted it. Similar findings were also obtained in previous studies; for instance, radish stems were found to be thin and elongated under red light [38]. In experiments on the effects of blue and red light on stem elongation of tomato seedlings and other species, results showed that stem length decreased with increasing blue to red light ratio [39, 40]. However, another report on cucumber showed that seedlings showed increased height, hypocotyl, and epicotyl under blue light alone compared to other treatments [41]. Red and blue light are both significant components of the spectrum of photosynthetically active radiation (PAR) and hence play a fundamental role in plant photomorphogenesis. Plants sense and transmit light signals through photoreceptors such as phytochromes (Phy) [42, 43], cryptochromes (Cry) [44] and phototropins (Phot) [45], which in turn trigger plant growth and development by signal transduction. Phytochromes are R/FR light photoreceptors that can constantly perceive and respond to the surrounding environment [46]. Under irradiation by red light, phytochromes suppress the activity of CONSTITUTIVE PHOTOMORPHOGENICI 1 (COP1) and PHYTOCHROME INTERACTING FACTORS (PIFs) protein. The expression level of unigene (Cluster-22883.48222) encoding the COP1 protein is significantly reduced in RY compared to WY and BY, implying that red light accelerated Phys to break down COP1 and thus induced several photomorphogenesis-promoting TFs (e.g., HY5 and LAF1), as evidenced by their higher expression level [42]. In this study, we found that unigene (Cluster-22883.57986) encoding for HY5 and SgMYB32 encoding for LAF1 showed higher expression levels to various degrees.

The Cry family members in *Arabidopsis thaliana* are primarily involved in the inhibition of hypocotyl elongation, de-etiolation of seedlings, floral initiation, root development and other photoresponses [44], while those of other plants are involved in a more diverse range of functions [40]. The Phot family members are UV and blue-light photoreceptors. Phot1 and Phot2 are involved in the regulation of phototropism in different regulatory patterns: phot1 activation needs a low threshold of blue light intensity, while Phot1 and Phot2 are mediated by higher blue light intensities [40, 45]. In terms of hypocotyl elongation, previous research demonstrated that its inhibition was controlled by the combined action of Cry1, Cry2, and Phot1 during the de-etiolation stage of young seedlings [40, 44]. In this study, we found that four unigenes of Cluster-22883.51204, Cluster-22883.48125, Cluster-22883.53309, and Cluster-22883.50402 putatively encoded Phot1, Phot2, Cry1, and Cry2 proteins, respectively, and that these were involved in the perception and response to blue light irradiation (Additional file 14: Fig. S8a-b). Among these unigenes, the relative expression levels of Phot1, and Cry1 were up-regulated in BY, however, the Cry2 gene was down-regulated. As mentioned earlier, the higher expression level of Phot1 and Cry1 inhibited hypocotyl elongation in *S. glabra* seedlings, which were subsequently integrated with other signaling molecules to mediate the regulatory network of those blue light-regulated genes by transcriptional and post-translational mechanisms [44, 47]. In a previous study using a kinetic analysis method, it was reported that the rate of stem growth was determined by multiple environmental signals, such as hormones and endogenous rhythms [48]. The reason why different phenotypic and aerial

architecture development responses to red, blue, and blue: red light was different amongst different plants is complex and requires more investigation. Furthermore, the detailed mechanism of phytochrome-mediated regulation (e.g., stem elongation and bud outgrowth, also warrants further research [46].

## The conjoint analyses of the production of flavonoids and coumarins biosynthesis

All of the flavonoid compounds examined in this study are thought to be derived directly from the reaction between 4-coumaroyl CoA and three units of malonyl CoA catalyzed by CHS, CHI, F3H, F3'H, and FLS enzymes (encoded by early biosynthetic genes, EBGs). The intermediate products of this reaction are dihydroflavonols (e.g., dihydrokaempferol or dihydroquercetin) which are subsequently used as substrates in other pathways involved in flavonol, anthocyanin, and proanthocyanidin biosynthesis regulated by downstream genes [designated late biosynthetic genes (LBGs), Fig. 4A] [49–52]. Our analysis of metabolism profiling (Fig. 3) revealed that the yield of some flavonols (e.g., quercetin, quercitrin, kaempferol, and rutin) were significantly higher in BY. These play an important role in pollen fertility, flower color, and UV-B protection [53]. In flavonol metabolism, FLS, belonging to 2-oxoglutarate-dependant dioxygenase, is a branch-point enzyme with multiple isoforms that catalyzes the conversion from dihydrokaempferol to products in a tissue-specific and inducible manner [54, 55]. In this study, the unigene (Cluster-22883.47725) encoding FLS enzyme was identified. It was also shown that flavonol production varied according to its relative expression in WY, RY, BY, and WG. Given the higher relative expression of FLS in the blue light treatment, the production of quercetin, quercitrin, kaempferol, and rutin increased drastically in the leaves of *S. glabra*. Cinnamic acid and 4-coumaric acid are both upstream products of the first two steps of the core phenylpropanoid pathway, their production thus regulates the flow of precursors to all side branches in connection with the biosynthesis of lignans, flavonoids, and phenolic acids. Among these, lignins are synthesized by the conversion from 4-coumaric acid to caffeic acid catalyzed by C3H enzyme, and caffeic acid subsequently enters the lignin pathway. Meanwhile, 4CL catalyzes the conversion of 4-coumaric acid to 4-coumaroyl CoA, a substrate assigned into the chlorogenic pathway in combination with shikimate or quinic acid, and it also proceeds to all classes of flavonoids production together with malonyl CoA. Likewise, according to metabolite analysis, the evidence was emerging showing that the initial products of cinnamic acid and 4-coumaric acid contents accumulated significantly, but the bypass of caffeic acid production was inhibited in BY. Therefore, the increasing production of 4-coumaric acid led to more synthesis of 4-coumaroyl CoA, which in turn led to a higher production of flavonoids and chlorogenic (e.g., chalcone and naringenin). Generally, the accumulation of initial precursors is closely related to the expression level of biosynthetic genes, including PAL, C4H, and 4CL genes leading to varieties of flavonoids. In *A. thaliana*, the PAL gene family (AtPAL1-AtPAL4), 4CL gene family (At4CL1-At4CL4) and only one member of AtC4H have been identified and elaborated, some of which-like AtPAL1 and AtPAL2-are thought to have major role in the phenylpropanoids pathway [29, 56], while At4CL3 is thought to has preference for the flavonoids biosynthesis [57]. The unrooted phylogenetic tree showed that the unigenes of Cluster-22883.50216, Cluster-22883.50271, and Cluster-22883.47381 were highly homologous with PAL1/2, C4H, and 4CL respectively, which exhibited the up-regulated expression level under blue and red light condition,

suggesting that these consecutive genes served to the flavonoids biosynthesis (Additional file 15: Fig. S9a-b).

For coumarins, the compounds of fraxetin, esculin, isofraxidin, and fraxidin were most abundant in WG. The substantial accumulation of fraxetin was closely related to the higher expression level of the unigene (Cluster-22883.51443) encoding S8H in the roots. Furthermore, the variation in scopolin followed the gene expression pattern of Cluster-22883.45883 encoding for scopoletin glucosyltransferase (SGTF), which catalyzes scopolin formation. Other studies revealed that scopoletin could also be produced from the conversion of esculetin via the methylation of O-methyltransferase (O-MT). Through two branches of the core pathway, both feruloyl acid and esculetin contents decreased significantly with the reduction in scopoletin under blue light conditions. Similar to scopoletin, the biosynthetic pathway of esculetin originated from umbelliferone or caffeic acid on earth that are still unclear [4].

## **Some R2R3-MYB and bHLH proteins regulate expression of EBGs and LBGs**

To date, there have been various studies on TFs of R2R3-MYB, bHLH, WDR and MBW complexes as regulators of phenylpropanoid metabolism in plants [36, 51, 58–61]. These studies highlight the importance of these TFs and MBW complexes in the regulation of flavonoids biosynthesis in plant tissues and the seed coat. The whole flavonoids pathway relies on the coordinated expression of EBGs and LBGs. Moreover, in *A. thaliana*, the subclasses (S4, S5, S6, and S7) of R2R3-MYB proteins [32], the subgroups (S8, S7, and S15 in AtbHLH proteins) of bHLH proteins [34], and several WD40 proteins [35, 62] (e.g., TTG1) are also proved to be involved in the regulation of the expression of EBGs and LBGs in the flavonoids synthesis in a certain way. In flavonol biosynthesis, members of AtMYB11/12/111 from subgroup 7 regulate the EBG genes, in turn controlling flavonol production [50, 63]. Therefore, SgMYB38 was assigned to subgroup 7 and was likely to control the expression of CHS, CHI, F3H, F3'H, and FLS genes. The results of our transcriptome and RT-qPCR analyses showed that the expression level of these structural genes predominantly increased along with the accumulation of SgMYB38, and that the yield of some flavonol compounds increased by some degree in BY. The accumulation of PAs is mainly distributed in the innermost cell layer of the seed coat regulated by more than four MBW complexes, i.e., AtMYB123 (TT2)-TT8/GL3/EGL3-TTG1, and AtMYB5-TT8-TTG1, which target the different LBGs, such as ANS (LDOX), DFR, Anthocyanin biosynthesis via the regulation of the spatial and temporal expression of LBGs in vegetative tissues [64]. These results suggest that it is reasonable to assume that MBW complexes target gene specificity. Therefore, in this paper, we screened and identified some unigenes encoding R2R3-MYB, bHLH and WDR TFs. SgMYB8/14/27/28/43/39 belonging to subgroup 5 of R2R3-MYB proteins may encode the homologous protein TT2, while SgMYB53 showed a high degree of functional similarity within members in subgroup 6 of R2R3-MYB proteins. SgbHLH12/13 were classified in subgroup7 of the AtbHLH TFs predicted to encode TT8/GL3/EGL3/MYC1, while Cluster-22883.52183 was largely thought to encode TTG1. It is reasonable to assume that the assembled MYB complexes of SgMYB8/14/27/28/43/39-SgbHLH12/13-TTG1 may be involved in the synthesis of PAs, while SgMYB53-SgbHLH12/13-TTG1 are likely to be associated with anthocyanins biosynthesis. In this study,

we just identified these genes encoding TFs proteins and proposed potential MBW complexes in *S. glabra* according to the studied model plants. However, some aspects of the flavonoid biosynthesis still remain unknown: for instance, due to the unavailable genomic data for *S. glabra*, the copy numbers of important genes involved in the phenylpropanoid pathway have not yet been identified. An in-depth analysis of the MBW complexes structure, function, and regulation in *S. glabra* is still required.

## Conclusions

In conclusion, *S. glabra* seedlings cultured under different monochromatic LED lights (WL, RL, and BL) showed significant differences in plant growth. Additionally, a combined analysis of transcriptome and metabolomic profiling demonstrated that LED lights had a significant influence on metabolite accumulation and the expression patterns of structural genes and TFs (R2R3-MYB and bHLH) involved in phenylpropanoid biosynthesis. Compared with WL, BL showed higher expression levels in EBGs and some LBGs under the regulation of specific TFs, leading to a significant increase in flavonoid production in BY. FLS acted as a limiting factor, reducing the yield of flavonol compounds in the RL condition. Furthermore, the production of desired compounds (e.g., rosmarinic acid, esculetin, isofraxidin, scopolin, scopoletin, fraxidin, fumaric acid, and maleic acid) was negatively correlated with BL, while RL stimulated the accumulation of some vital compounds (e.g., esculetin, fraxetin, esculin, scopoletin, fumaric acid, and maleic acid) in leaves. This study gave valuable insights into the regulatory mechanism of the phenylpropanoid biosynthesis in *Sarcandra glabra*.

## Methods

### Treatments and samples collection

Seeds of *S. glabra* were collected from the Louyuan state-owned forest farm in Sanming City, Fujian Province, China. We have acquired an approval from forest farm owner, and Professor Xiaoxing Zou has identified the collected seeds propagated from *S. glabra*. The specimens were deposited in the herbarium of the GAP Engineering Research Center in Fujian Agriculture and Forestry University. The seeds were then cultivated at constant temperature and humidity incubator (25 °C, and 60% humidity) in the Bamboo Institute of the Fujian Agriculture and Forestry University, Fuzhou, Fujian Province, China (26°5' N, 119°13' E). To grow a pair of true leaves, the 90 seedlings were transplanted into black Poly Vinyl Chloride (PVC) rectangular pots (10 × 10 × 8 cm) using Pindstrup substrate (NO.5, pH 5.5) for cultivation. Seedlings were arranged randomly in the culture room under three different light treatments (WL, 380 ~ 760 nm; RL, 656 nm; BL, 450 nm). Each treatment had 10 pots with 3 replicates, making 30 pots in total. The LED lights, quantum sensor, and light meter (HR-450) were purchased from Hipoint Corporation. The photosynthetic photon flux density (PPFD) of the lights was set as 80  $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ , and the photoperiod was 16/8 h (day/night) (Fig. 1A). The respective intensities of the PPFD and light spectra were monitored using a HR-450 machine. After a 60-day period, 18 healthy seedlings from each treatment were uprooted, carefully washed and dried on filter paper (Fig. 1B). Leaves, stems, and roots were then separated and

triplicate composite samples were made from each treatment. In this study, we primarily focused on the effects of WL, RL, and BL on leaves, as well as the effects of WL on different tissues *S. glabra*. Therefore, fifteen samples [WY1/2/3, RY1/2/3, BY1/2/3, WJ1/2/3 (stem 1/2/3 under WL), and WG 1/2/3] were immediately frozen in liquid nitrogen and kept at  $-80^{\circ}\text{C}$  until further analysis.

## Determination of morphological traits

Plant height and stem diameter were measured by using a measuring tape and a vernier caliper, respectively throughout the cultivation period. The upper fully expanded leaves were measured using a portable leaf area meter (LI-3000C, Ecotek) 60 days after seeds were planted.

## RNA extraction and transcriptome profiling

Total RNA was extracted from the samples using RNeasy Pure Plant Kit (DP441, TIANGEN BIOTECH). The 15 RNA samples were then treated with DNase I (Takara) to remove genomic DNA contamination and sent to the Novogene Bioinformatics Technology Company (Beijing, China) for cDNA library construction and Illumina sequencing.

### Do novo assembly and annotation

For the quality of assembly, raw reads were filtered to remove adaptor sequences and reads with more than 10% of unknown bases. Low-quality reads, in which the percentage of quality value ( $\leq 20$ ) base was more than 50%, were also deleted. The Trinity method was employed to *de novo* assemble the clean reads [65]. Seven public databases or programs were used to annotate the unigenes, including the NCBI non-redundant protein (Nr), the NCBI nucleotide sequences (Nt), the protein family (Pfam), the euKaryotic Ortholog Groups (KOG), the Swiss-Prot database, the Kyoto Encyclopedia of Genes and Genomes (KEGG) and the Gene Ontology (GO).

## Gene expression analysis

The fragments per kilobase of transcript per million mapped transcript (FPKM) values were used to analyze gene expression. The differential expression analysis of any two sets of treatments was measured using the R package *DEGseq2* [66]. An adjusted p-value of 0.05 was set as the threshold to determine significant differences in DEGs. The enrichment analysis was performed with GO ontology analysis using the R package *GOseq* [67]. The KEGG pathway was annotated using the KEGG database. The corrected P-value  $< 0.05$  and  $|\log_2\text{FoldChange}| > 1$  were set as the threshold to determine the significant differences under the GO and KEGG enrichment analyzes. Lastly, a heat map diagram of the selected unigenes was constructed using heml software with the FPKM values.

## Candidate transcription factors of MYB and bHLH

Using BLASTx analysis against the PlnTFDB [68] (<http://planttfdb.cbi.pku.edu.cn/>) database to identify MYB and bHLH TFs according to the transcriptome data. Referring to reviews on the characteristics and classification of *Arabidopsis thaliana* R2R3-MYB proteins to screen and classify the unigenes encoding R2R3-type MYB proteins in *S. glabra* [32, 37]. Likewise, we referred to the previous researches on the

characteristics of *Arabidopsis thaliana* bHLH proteins to examine the distribution and predicted DNA binding features of bHLH proteins in *S. glabra* [34, 69].

## qRT-PCR confirmation of the gene expression profile

First-strand cDNA was generated from 1 µg of total RNA using the PrimeScript™ RT reagent Kit with gDNA Eraser (RR047A, Takara, Japan). Gene-specific primers and a reference gene (CAC, clathrin adaptor complexes) for qRT-PCR were designed to use by Primer Premier 5.0 software (Primer, Canada) and synthesized by the SunYa Company (Fuzhou, China) (Additional file 16: Table S7). All analyses were repeated twice for each biological replicate. The qRT-PCR was performed on the ABI 7500 Real-Time PCR system (Applied Biosystems, USA) using SYBR Green premix Ex Taq Kit (RR820A, Takara, Japan). The expression levels of the tested reference genes were determined from CT values and calculated using the  $2^{-\Delta\Delta CT}$  method.

## Phylogenetic analysis and statistical analysis

A multiple sequence alignment analysis was performed using the DNAMAN software. The phylogenetic trees were constructed by using MEGA7.0 with a neighbor-joining method and 1000 bootstrap replicates after alignment by ClustalW [70]. All values were presented as the means and standard errors ( $\pm$ ). Statistical analyses were performed by using the SPSS 20.0 software (IBM, USA), Duncan's multiple range test was applied to compare the means at  $P < 0.05$ .

## Metabolomic profiling analysis

### Sample preparation and extraction

Meanwhile, the 12 samples (WY1/2/3, RY1/2/3, BY1/2/3, WG1/2/3), other than WJ, were also sent to Metware Technology Company (Wuhan, China) for metabolomic profiling. The freeze-dried leaf and root were crushed using a mixer mill with a zirconia bead for 90 s at 30 Hz. 100 mg powder was weighted and extracted overnight at 4 °C using 0.6 ml 70% aqueous methanol. Following centrifugation at 10,000 g for 10 min, the extracts were absorbed and filtrated before UPLC-MS/MS analysis. The sample extracts were analyzed using a UPLC-ESI-MS/MS system (UPLC, Shim-pack UFLC SHIMADZU CBM30A system; MS, Applied Biosystems 4500Q TRAP).

### Quality control data analysis

To investigate the metabolic differences and the degree of variation between twelve samples and quality control sample (a mixer sample pooled together with equal mass from the aforementioned samples) with three replications, the unsupervised principal component analysis (PCA) was performed using the R package *prcomp*. Data was unit variance scaled before performing PCA. To investigate metabolic changes in different samples, a hierarchical cluster analysis (HCA) was carried out using the R package *heatmap*, and the raw data was disposed of unit variance (UV) scaling before HCA analysis.

## Abbreviations

LED: Light-emitting diodes; WL: White light; RL: Red light; BL: Blue light; DEGs: Differentially expressed genes; WG: Root tissue under WL; WJ: Stem tissue under WL; BY: Leaf tissue under BL; WY: Leaf tissue under WL; RY: Leaf tissue under RL; GAP: Good Agricultural Practice; TFs: Transcription factors; MBW: MYB-bHLH-WD40; Nr: Non-redundant protein; Nt: NCBI nucleotide sequences; Pfam: The protein family; KOG: The euKaryotic Ortholog Groups; KEGG: The Kyoto Encyclopedia of Genes and Genomes; GO: the Gene Ontology; FPKM: Fragments per kilobase of transcript per million mapped transcript; PVC: Poly Vinyl Chloride; PPF: Photosynthetic photon flux density; PCA: Principal component analysis; HCA: Hierarchical cluster analysis; UV: unit variance; PAL: Phenylalanine ammonia lyase; C4H: Cinnamate 4-hydroxylase; 4CL: 4-coumaroyl CoA ligase; C3H: Coumarate 3-hydroxylase; HCT: Hydroxycinnamoyl-CoA shikimate hydroxycinnamoyl transferase; HQT: Hydroxyl-cinnamoyl CoA quinate hydroxycinnamoyl transferase; COMT: Caffeic acid 3-O-methyltransferase; CCoAOMT (CCOMT): Caffeoyl CoA 3-O-methyltransferase; F6'H: Feruloyl-CoA 6'-Hydroxylase; S8H: Scopoletin 8-hydroxylase; SGTF: Scopoletin glucosyltransferase; O-MT: O-methyltransferase; CHS: Chalcone synthase; CHI: Chalcone isomerase; F3H: Flavanone 3-hydroxylase; F3'H: Flavonoid 3'-hydroxylase; FLS: Flavonol synthase; UFGT: UDPG flavonoid O-glucosyltransferase; RT: Rhamnosyl transferase; DFR: Dihydroflavonol 4-reductase; LAR: Leucoanthocyanidin reductase; ANR: Anthocyanidin reductase; ANS or (LDOX): Anthocyanin synthase or (leucoanthocyanidin dioxygenase); TAT: Tyrosine aminotransferase; HPPR: Hydroxyphenylpyruvate reductase; HPPD: 4-hydroxyphenylpyruvate dioxygenase; RAS: Rosmarinate synthase (rosmarinic acid synthase); CYP (CPR): Cytochrome P450 reductase. CA2H: Caffeic acid 2-hydroxylase, PAs: Proanthocyanidins; R: Correlation coefficient; PAR: Photosynthetically active radiation; Phy: Phytochromes; Cry: Cryptochromes; Phot: Phototropins; COP1: CONSTITUTIVE PHOTOMORPHOGENIC 1; PIFs: PHYTOCHROME INTERACTING FACTORS; EBGs: Early biosynthetic genes; LBGs: Late biosynthetic genes.

## Declarations

## Ethics approval and consent to participate

Not applicable

### Consent to publish

Not applicable

### Availability of data and materials

15 transcriptome raw data produced by Illumina NovaSeq 6000 have been deposited in the Genome Sequence Archive (SRA) database (<https://bigd.big.ac.cn/gsa/>) under the accession number CRA002486 (<http://bigd.big.ac.cn/gsa/s/2LVZ8h94>).

### Competing interests

The authors declare no conflict of interest.

## Funding

This work was supported by the Science and Technology Innovation Platform of Fujian province of China (2008Y2001). GAP Engineering Research Center of Fujian province of China. The funding body had no role in the design of study, analysis, and interpretation of data and in writing the manuscript.

## Authors' Contributions

YZ made the research plan. TH and JR collected the materials. DX, M.W.K, DY, and KR performed the experiments. DX, LC, CZ, and M.W.K analyzed the data and wrote the manuscript. All authors discussed the results and comment on the manuscript. All authors gave final approval for publication.

## Acknowledgements

We thank Louyuan state-owned forest farm owner to provide the seeds, and we also acknowledge Xiaoxing Zou at Fujian Agriculture and Forestry University to identify this plant material. We are very grateful to Shaoyan Shen, Junming Zheng, and Lianfeng Gu for providing useful suggestions, discussion to this study. We are very grateful to the editor and reviewers for evaluating our manuscript and providing comments for its improvement.

## References

1. Li H, Gong X, Wang Z, Pan C, Zhao Y, Gao X, Liu W. Multiple fingerprint profiles and chemometrics analysis of polysaccharides from *Sarcandra glabra*. *Int J Biol Macromol*. 2019; 123:957-967.
2. Zhou H, Liang J, Lv D, Hu Y, Zhu Y, Si J, Wu S. Characterization of phenolics of *Sarcandra glabra* by non-targeted high-performance liquid chromatography fingerprinting and following targeted electrospray ionisation tandem mass spectrometry/time-of-flight mass spectrometry analyses. *Food Chem*. 2013; 138(4):2390-2398.
3. The state pharmacopoeia commission of People's Republic of China. *Pharmacopoeia of the People's Republic of China*, vol. 1: Beijing: Chemical Industry Press; 2015.
4. Bourgaud F, Hehn A, Larbat R, Doerper S, Gontier E, Kellner S, Matern U. Biosynthesis of coumarins in plants: a major pathway still to be unravelled for cytochrome P450 enzymes. *Phytochem Rev*. 2006; 5(2):293-308.
5. Liu J, Li X, Lin J, Li Y, Wang T, Jiang Q, Chen D. *Sarcandra glabra* (Caoshanhu) protects mesenchymal stem cells from oxidative stress. A bioevaluation and mechanistic chemistry. *BMC Complem Altern M*. 2016; 16:423.

6. Cao H, Tan R, He R, Tang L, Wang X, Yao N, Duan W, Hu Y, Yao X, Kurihara H. *Sarcandra glabra* Extract Reduces the Susceptibility and Severity of Influenza in Restraint-Stressed Mice. *Evid Based Complement Alternat Med*. 2012; 2012:236539.
7. Wei S, Chi J, Zhou M, Li R, Li Y, Luo J, Kong L. Anti-inflammatory lindenane sesquiterpenoids and dimers from *Sarcandra glabra* and its upregulating AKT/Nrf2/HO-1 signaling mechanism. *Ind Crop Prod*. 2019; 137:367-376.
8. Guo X, Shen L, Tong Y, Zhang J, Wu G, He Q, Yu S, Ye X, Zou L, Zhang Z *et al*. Antitumor activity of caffeic acid 3,4-dihydroxyphenethyl ester and its pharmacokinetic and metabolic properties. *Phytomedicine*. 2013; 20(10):904-912.
9. Li X, Zhang Y, Zeng X, Yang L, Deng Y. Chemical profiling of bioactive constituents in *Sarcandra glabra* and its preparations using ultra-high-pressure liquid chromatography coupled with LTQ Orbitrap mass spectrometry. *Rapid Commun Mass Spectrom*. 2011; 25(17):2439-2447.
10. Zheng W, Wang S, Chen X, Hu Z. Analysis of *Sarcandra glabra* and its medicinal preparations by capillary electrophoresis. *Talanta*. 2003; 60(5):955-960.
11. Yaermaimaiti S, Wang P, Luo J, Li R, Kong L. Sesquiterpenoids from the seeds of *Sarcandra glabra* and the potential anti-inflammatory effects. *Fitoterapia*. 2016; 111:7-11.
12. Zhou H, Liang J, Lv D, Hu Y, Zhu Y, Si J, Wu S. Characterization of phenolics of *Sarcandra glabra* by non-targeted high-performance liquid chromatography fingerprinting and following targeted electrospray ionisation tandem mass spectrometry/time-of-flight mass spectrometry analyses. *Food Chem*. 2013; 138(4):2390-2398.
13. Liao P, Liu P, Wang Y, Huang C, Lan L, Yang Y, Cui X. Stereoscopic cultivation of *Panax notoginseng*: A new approach to overcome the continuous cropping obstacle. *Ind Crop Prod*. 2018; 126:38-47.
14. Ding J, Tu H, Zang Z, Huang M, Zhou S. Precise control and prediction of the greenhouse growth environment of *Dendrobium candidum*. *Comput Electron Agr*. 2018; 151:453-459.
15. Yano A, Cossu M. Energy sustainable greenhouse crop cultivation using photovoltaic technologies. *Renew Sust Energ Rev*. 2019; 109:116-137.
16. Tarakanov I, Yakovleva O, Konovalova I, Paliutina G, Anisimov A. Light-emitting diodes: On the way to combinatorial lighting technologies for basic research and crop production. *Acta Hortic*. 2012; 956:171-178.
17. Yeh N, Chung J. High-brightness LEDs—Energy efficient lighting sources and their potential in indoor plant cultivation. *Renew Sust Energ Rev*. 2009; 13(8):2175-2180.
18. De Keyser E, Dhooghe E, Christiaens A, Van Labeke M, Van Huylenbroeck J. LED light quality intensifies leaf pigmentation in ornamental pot plants. *Sci Hortic*. 2019; 253:270-275.
19. Berkovich YA, Konovalova IO, Smolyanina SO, Erokhin AN, Avercheva OV, Bassarskaya EM, Kochetova GV, Zhigalova TV, Yakovleva OS, Tarakanov IG. LED crop illumination inside space greenhouses. *Reach*. 2017; 6:11-24.
20. Thwe AA, Kasemsap P, Vercambre G, Gay F, Phattaralerphong J, Gautier H. Impact of red and blue nets on physiological and morphological traits, fruit yield and quality of tomato (*Solanum*

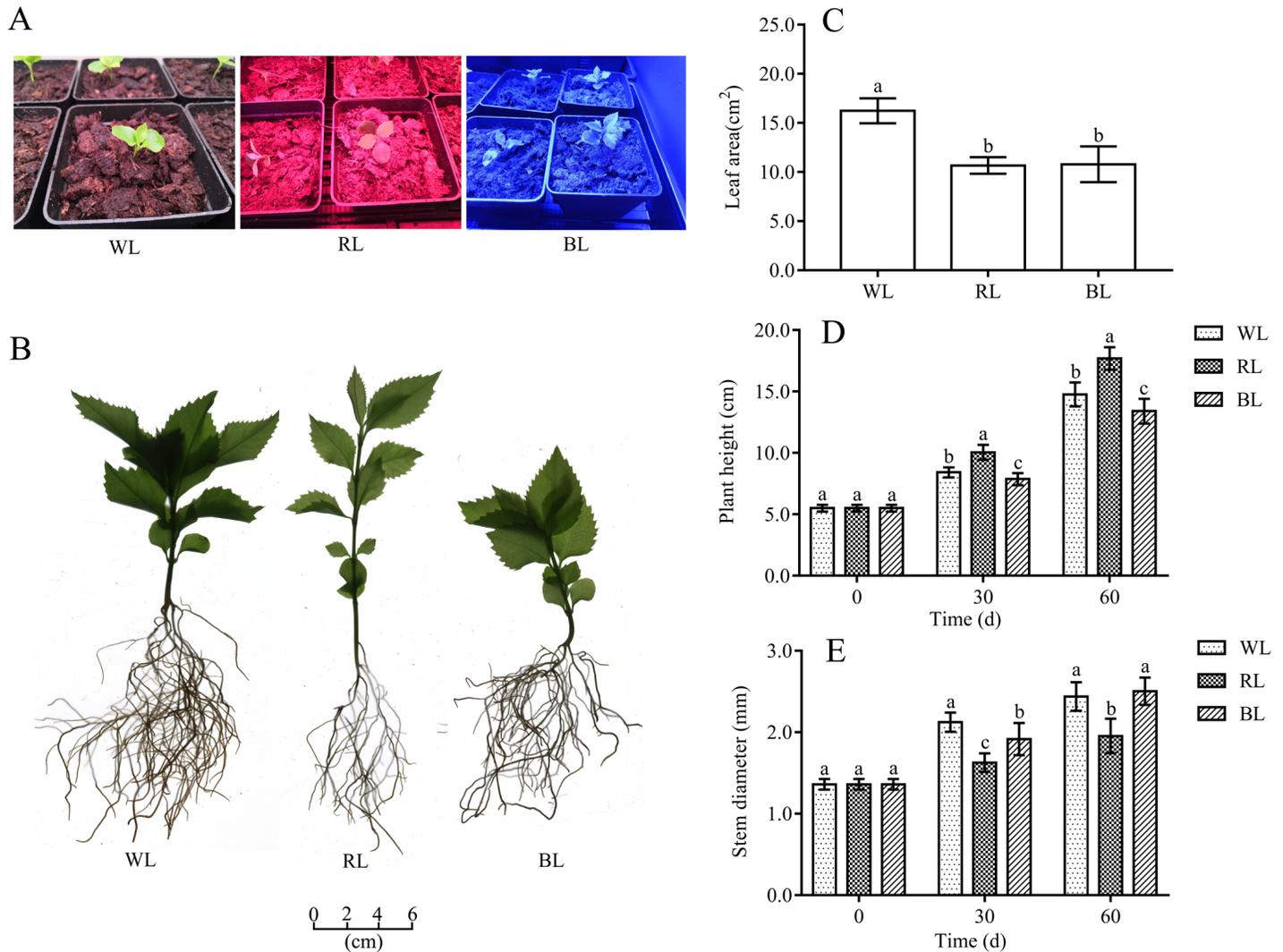
- lycopersicum* Mill.). *Sci Hortic.* 2020; 264:109185.
21. Nanya K, Ishigami Y, Hikosaka S, Goto E. Effects of blue and red light on stem elongation and flowering of tomato seedlings. *Acta Hortic.* 2012; 956:261-266.
  22. Hernández R, Kubota C. Physiological responses of cucumber seedlings under different blue and red photon flux ratios using LEDs. *Environ Exp Bot.* 2016; 121:66-74.
  23. Liu H, Chen Y, Hu T, Zhang S, Zhang Y, Zhao T, Yu H, Kang Y. The influence of light-emitting diodes on the phenolic compounds and antioxidant activities in pea sprouts. *J Funct Foods.* 2016; 25:459-465.
  24. Taulavuori K, Pyysalo A, Taulavuori E, Julkunen-Tiitto R. Responses of phenolic acid and flavonoid synthesis to blue and blue-violet light depends on plant species. *Environ Exp Bot.* 2018; 150:183-187.
  25. Liu Y, Fang S, Yang W, Shang X, Fu X. Light quality affects flavonoid production and related gene expression in *Cyclocarya paliurus*. *J Photoch Photobio B.* 2018; 179:66-73.
  26. Kapoor S, Raghuvanshi R, Bhardwaj P, Sood H, Saxena S, Chaurasia OP. Influence of light quality on growth, secondary metabolites production and antioxidant activity in callus culture of *Rhodiola imbricata* Edgew. *J Photoch Photobio B.* 2018; 183:258-265.
  27. Ni J, Dong L, Jiang Z, Yang X, Sun Z, Li J, Wu Y, Xu M. Salicylic acid-induced flavonoid accumulation in *Ginkgo biloba* leaves is dependent on red and far-red light. *Ind Crop Prod.* 2018; 118:102-110.
  28. Yang Q, Pan J, Shen G, Guo B. Yellow light promotes the growth and accumulation of bioactive flavonoids in *Epimedium pseudowushanense*. *J Photoch Photobio B.* 2019; 197:111550.
  29. M.Fraser C, Chapple C. The Phenylpropanoid Pathway in Arabidopsis. *The Arabidopsis Book.* 2011; 9:e152.
  30. Vogt T. Phenylpropanoid Biosynthesis. *Mol Plant.* 2010; 3(1):2-20.
  31. Xu W, Grain D, Bobet S, Gourrierc J, Thévenin J, Kelemen Z, Lepiniec L, Dubos C. Complexity and robustness of the flavonoid transcriptional regulatory network revealed by comprehensive analyses of MYB-bHLH-WDR complexes and their targets in Arabidopsis seed. *New Phytol.* 2014; 202(1):132-144.
  32. Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L. MYB transcription factors in Arabidopsis. *Trends Plant Sci.* 2010; 15(10):573-581.
  33. Zimmermann IM, Heim MA, Weisshaar B, Uhrig JF. Comprehensive identification of *Arabidopsis thaliana* MYB transcription factors interacting with R/B-like BHLH proteins. *Plant J.* 2004; 40(1):22-34.
  34. Toledo-Ortiz G, Huq E, Quail PH. The Arabidopsis Basic/Helix-Loop-Helix Transcription Factor Family. *Plant Cell.* 2003; 15(8):1749-1770.
  35. Smith TF, Gaitatzes C, Saxena K, Neer EJ. The WD repeat: a common architecture for diverse functions. *Trends Biochem Sci.* 1999; 24(5):181-185.
  36. Carretero-Paulet L, Galstyan A, Roig-Villanova I, Martínez-García JF, Bilbao-Castro JR, Robertson DL. Genome-Wide Classification and Evolutionary Analysis of the bHLH Family of Transcription Factors in Arabidopsis, Poplar, Rice, Moss, and Algae. *Plant Physiol.* 2010; 153(3):1398-1412.

37. Stracke R, Werber M, Weisshaar B. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr Opin Plant Biol.* 2001; 4(5):447-456.
38. Samuolienė G, Sirtautas R, Brazaitytė A, Sakalauskaitė J, Sakalauskienė S, Duchovskis P. The impact of red and blue light-emitting diode illumination on radish physiological indices. *Cent Eur J of Biol.* 2011; 6(5):821-828.
39. Nanya K, Ishigami Y, Hikosaka S, Goto E. Effects of blue and red light on stem elongation and flowering of tomato seedlings. *Acta Hort.* 2012; 956:261-266.
40. Huché-Théliier L, Crespel L, Gourrierc JL, Morel P, Sakr S, Leduc N. Light signaling and plant responses to blue and UV radiations—Perspectives for applications in horticulture. *Environ Exp Bot.* 2016; 121:22-38.
41. Hernández R, Kubota C. Physiological responses of cucumber seedlings under different blue and red photon flux ratios using LEDs. *Environ Exp Bot.* 2015; 29(1):66-74.
42. Li J, Li G, Wang H, Wang Deng X. Phytochrome signaling mechanisms. *The Arabidopsis Book.* 2011; 9:e148.
43. Franklin KA, Quail PH. Phytochrome functions in Arabidopsis development. *J Exp Bot.* 2010; 61(1):11-24.
44. Yu X, Liu H, Klejnot J, Lin C. The Cryptochrome Blue Light Receptors. *The Arabidopsis Book.* 2010; 8:e135.
45. Christie JM. Phototropin Blue-Light Receptors. *Annu Rev Plant Biol.* 2007; 58(1):21-45.
46. Demotes-Mainard S, Péron T, Corot A, Bertheloot J, Le Gourrierc J, Pelleschi-Travier S, Crespel L, Morel P, Huché-Théliier L, Boumaza R *et al.* Plant responses to red and far-red lights, applications in horticulture. *Environ Exp Bot.* 2016; 121:4-21.
47. Sellaro R, Hoecker U, Yanovsky M, Chory J, Casal JJ. Synergism of Red and Blue Light in the Control of Arabidopsis Gene Expression and Development. *Curr Biol.* 2009; 19(14):1216-1220.
48. Parks BM, Folta KM, Spalding EP. Photocontrol of stem growth. *Curr Opin Plant Biol.* 2001; 4(5):436-440.
49. Koes R, Quattrocchio F, Mol J. The flavonoid biosynthetic pathway in plants: Function and evolution. *Bioessays.* 1994; 16:123-132.
50. Falcone Ferreyra ML, Rius SP, Casati P. Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Front Plant Sci.* 2012; 3:222.
51. Xu W, Dubos C, Lepiniec L. Transcriptional control of flavonoid biosynthesis by MYB–bHLH–WDR complexes. *Trends Plant Sci.* 2015; 20(3):176-185.
52. Grotewold E. Plant metabolic diversity: a regulatory perspective. *Trends Plant Sci.* 2005; 10(2):57-62.
53. Wang Z, Wang S, Xiao Y, Li Z, Wu M, Xie X, Li H, Mu W, Li F, Liu P *et al.* Functional characterization of a HD-ZIP IV transcription factor NtHDG2 in regulating flavonols biosynthesis in *Nicotiana tabacum*. *Plant Physiol Bioch.* 2020; 146:259-268.

54. Owens D, Alerding A, Crosby K, Bandara A, Westwood J, Winkel B. Functional Analysis of a Predicted Flavonol Synthase Gene Family in Arabidopsis. *Plant Physiol.* 2008; 147(3):1046-1061.
55. Lukacin R, Wellmann F, Britsch L, Martens S, Matern U. Flavonol synthase from *Citrus unshiu* is a bifunctional dioxygenase. *Phytochemistry.* 2003; 62:287-292.
56. Huang J, Gu M, Lai Z, Fan B, Shi K, Zhou Y, Yu J, Chen Z. Functional analysis of the Arabidopsis PAL gene family in plant growth, development, and response to environmental stress. *Plant Physiol.* 2010; 153(4):1526-1538.
57. Ehling J, Büttner D, Wang Q, Douglas CJ, Kombrink E. Three 4-coumarate: coenzyme A ligases in Arabidopsis thaliana represent two evolutionarily divergent classes in angiosperms. *Plant J.* 1999; 19(1):9-20.
58. Ambawat S, Sharma P, Yadav NR, Yadav RC. MYB transcription factor genes as regulators for plant responses: an overview. *Physiol Mol Biol Pla.* 2013; 19(3):307-321.
59. Liu J, Osbourn A, Ma P. MYB transcription factors as regulators of phenylpropanoid metabolism in plants. *Mol Plant.* 2015; 8(5):689-708.
60. Bipei, Zhang, Martin, Hülkamp. Evolutionary Analysis of MBW Function by phenotypic rescue in *Arabidopsis thaliana*. *Front Plant Sci.* 2019; 10:375.
61. Feller A, Machemer K, Braun EL, Grotewold E. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant J.* 2011; 66(1):94-116.
62. Ramsay N, Glover B. MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. *Trends Plant Sci.* 2005; 10(2):63-70.
63. Stracke R, Jahns O, Keck M, Tohge T, Niehaus K, Fernie AR, Weisshaar B. Analysis of production of flavonol glycosides-dependent flavonol glycoside accumulation in *Arabidopsis thaliana* plants reveals MYB11-, MYB12- and MYB111-independent flavonol glycoside accumulation. *New Phytol.* 2010; 188(4):985-1000.
64. Gonzalez A, Zhao M, Leavitt JM, Lloyd AM. Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in Arabidopsis seedlings. *Plant J.* 2008; 53(5):814-827.
65. Grabherr M, Haas B, Yassour M, Levin J, Thompson D, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q *et al.* Full-Length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* 2011; 29:644-652.
66. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014; 15(12):550.
67. Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 2010; 11(2):R14.
68. Pérez-Rodríguez P, Riaño-Pachón DM, Corrêa LGG, Rensing SA, Kersten B, Mueller-Roeber B. PlnTFDB: updated content and new features of the plant transcription factor database. *Nucleic Acids Res.* 2010; 38:D822-D827.

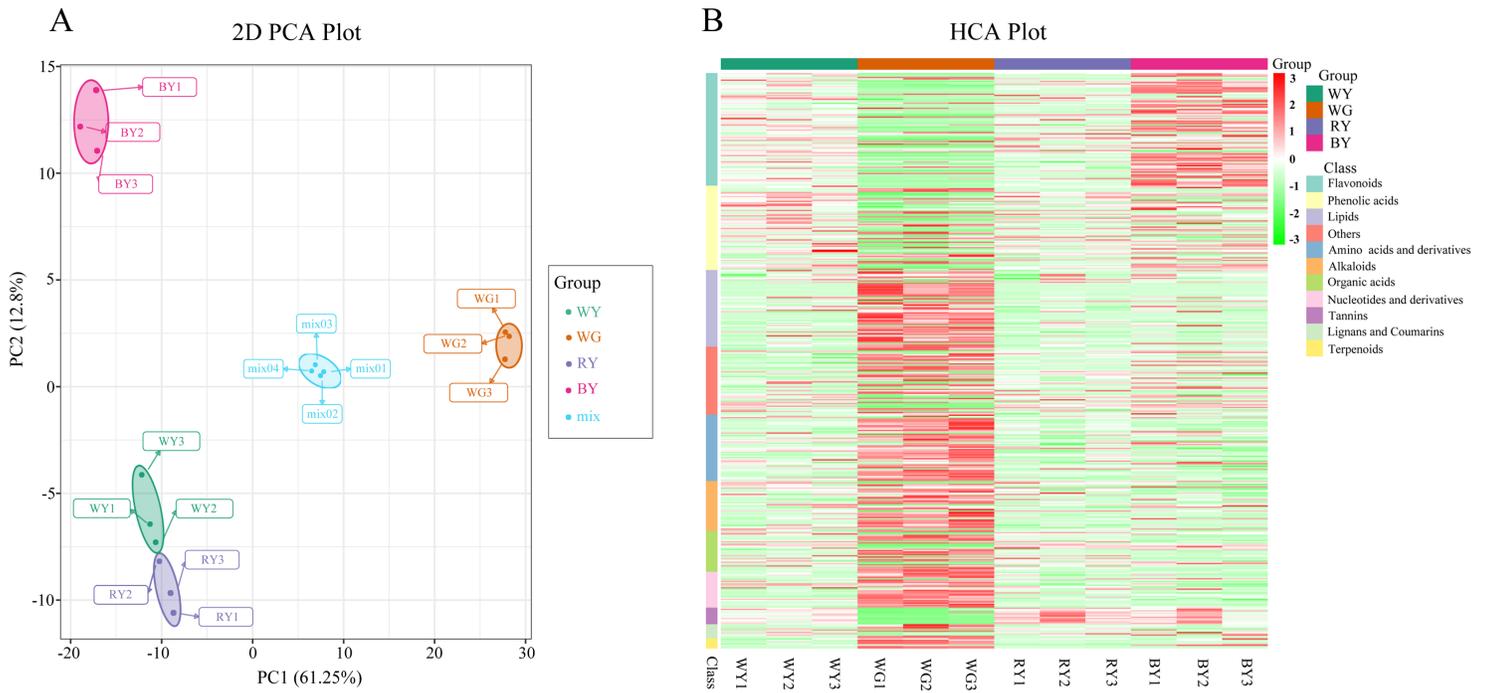
69. Atchley WR, Terhalle W, Dress A. Positional dependence, cliques, and predictive motifs in the bHLH protein domain. *J Mol Evol.* 1999; 48(5):501-516.
70. Sudhir K, Glen S, Li M, Christina K, Koichiro T. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol.* 2018; 35(6):1547-1549.

## Figures



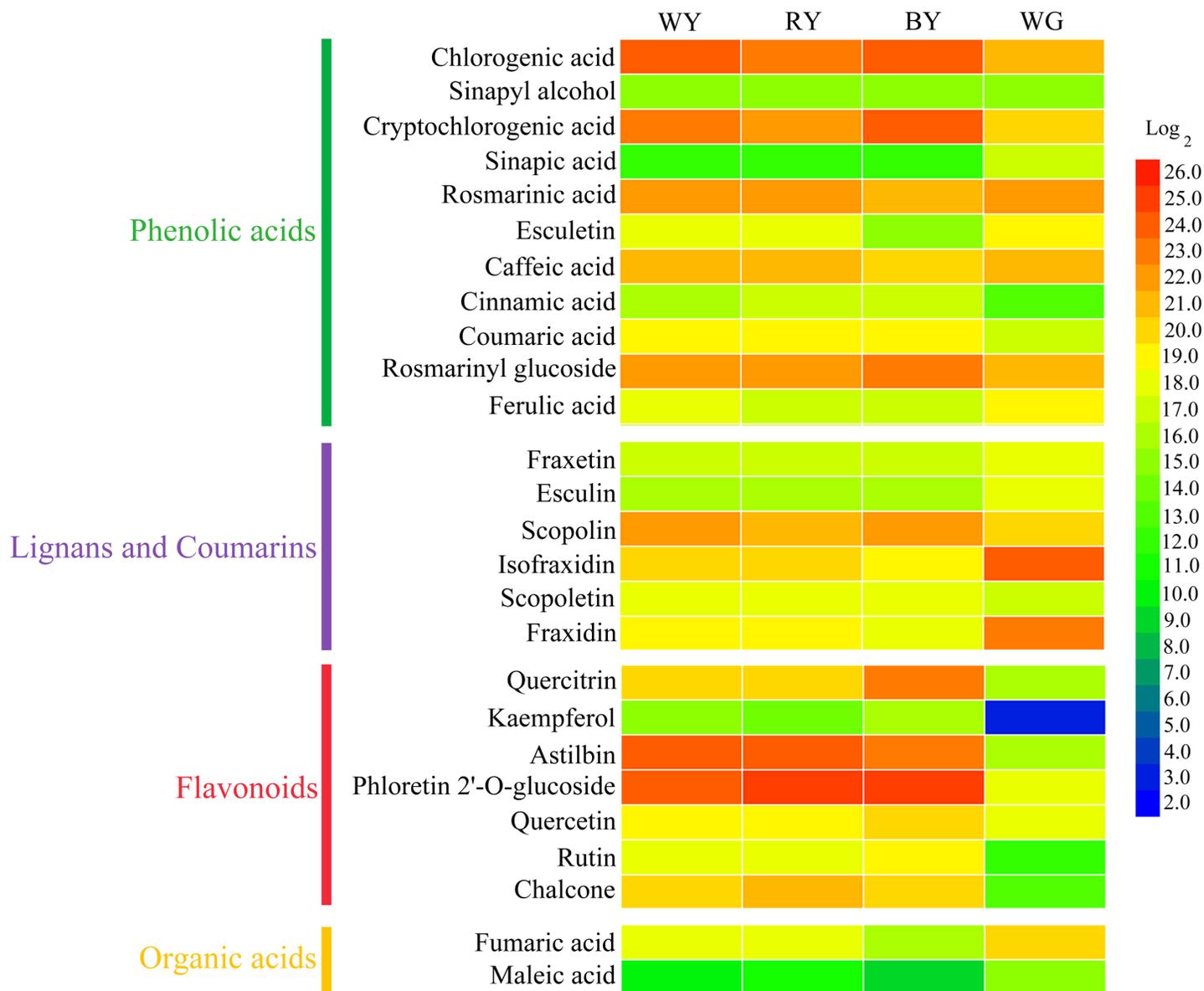
**Figure 1**

Images of the effect of LED lights on leaf area, plant height, and stem diameter. (A) the *S. glabra* seedlings grown under different LED lights; (B) the whole plant morphology under WL, RL, BL; (C) the leaf area of the seedlings on the 60th day; (D) the plant height; (E) the stem diameter; Different lower cases indicate significant differences ( $P < 0.05$ ) based on Duncan's multiple range test on days 0, 30, and 60, respectively; Bars indicate standard errors.



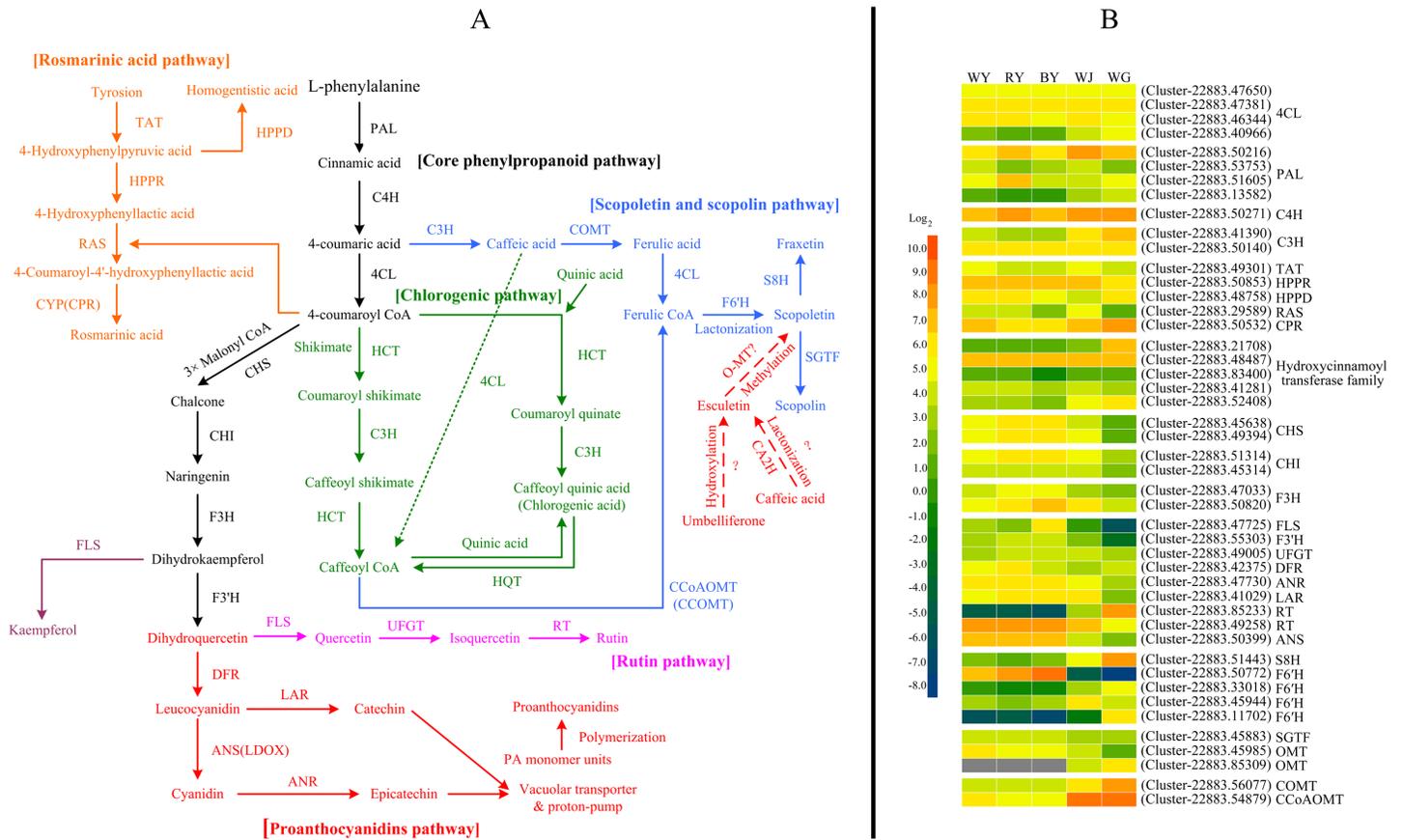
**Figure 2**

Results of PCA (A) and HCA (B) analyses of 12 samples from *S. glabra*. PC1 (61.25% variance), and PC2 (12.8% variance) were used to generate the scores plot. Samples of WY, RY, BY, WG, and mix are marked in green, purple, pink, brown, and blue, respectively (Fig. 2A). A heatmap was generated from the metabolism profiling data of the 12 samples, each rectangle representing the value of a correlation between samples, while the samples of WY, RY, BY, WG are also marked in the same colors, respectively (Fig. 2B).



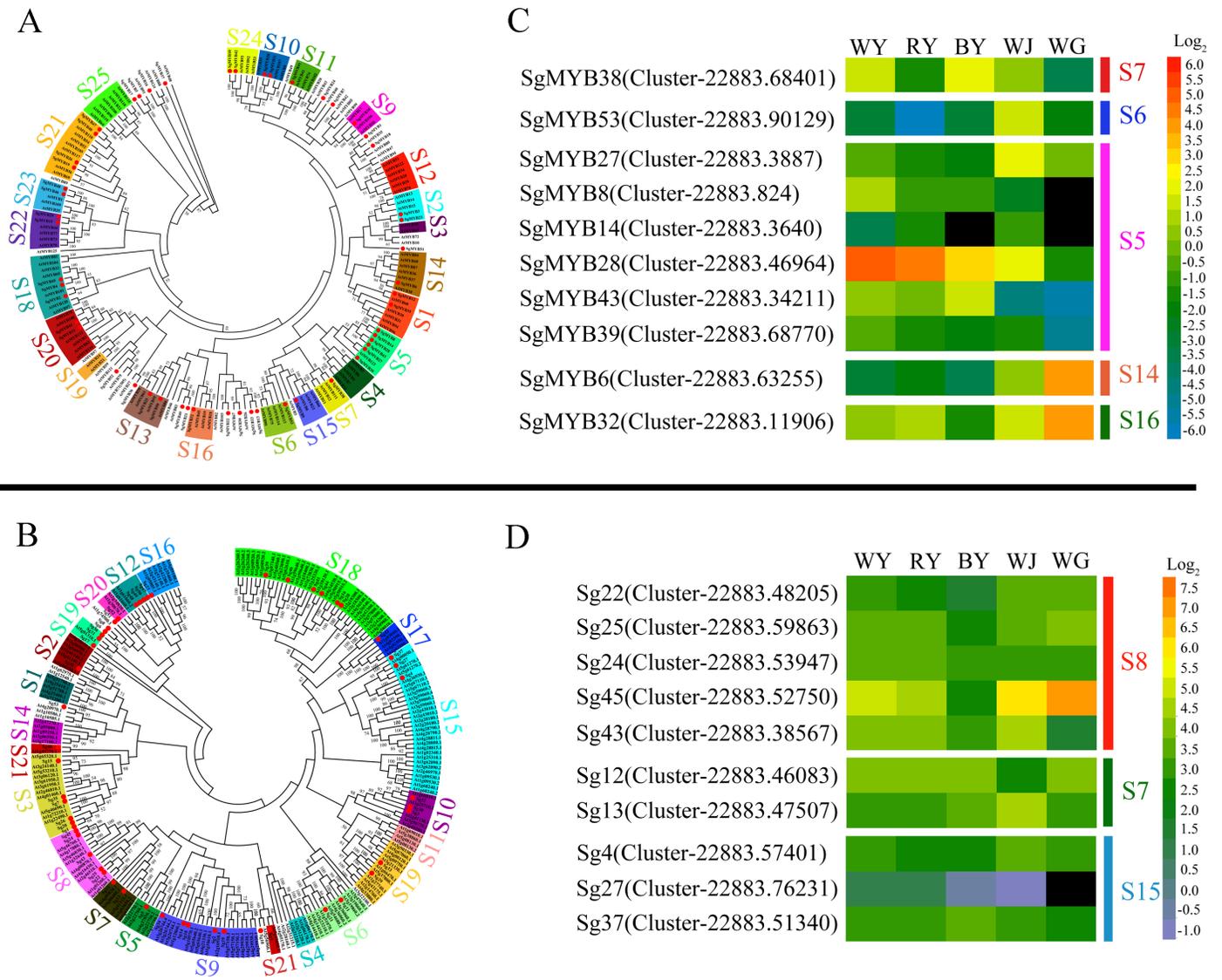
**Figure 3**

Heat map diagram of the relative concentration of the 24 principle active metabolites. The left legend represents different metabolites classes in different colors. The name of metabolites are listed on the vertical line, and the samples are listed horizontally. All of the relative content values of the metabolites are shown as logarithms.



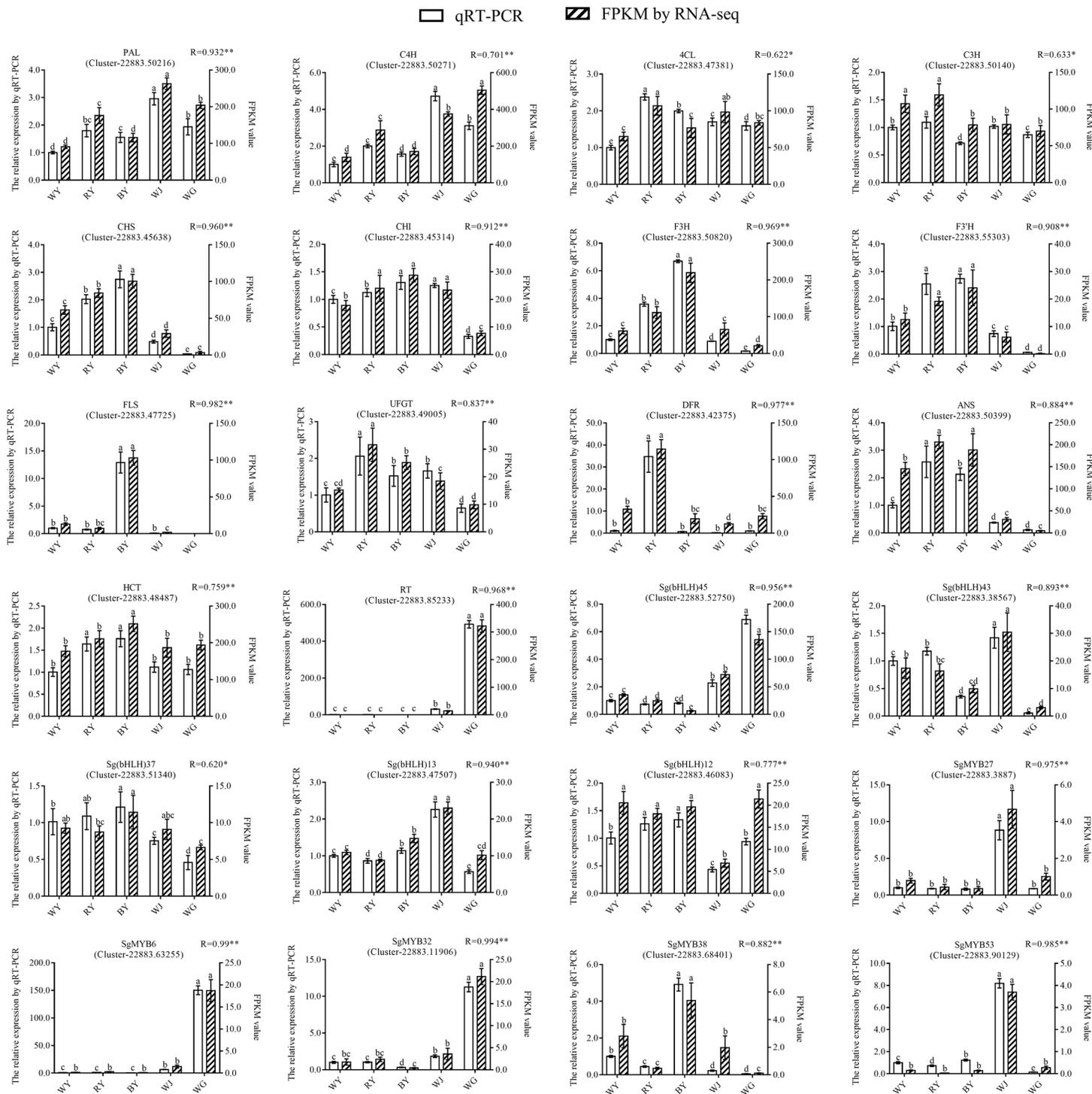
**Figure 4**

Phenylalanine metabolic pathways (A) and the heat map diagram of the unigenes (B). The name of unigenes are listed on the vertical line, and the samples are listed horizontally. The scale represents the logarithms of the FPKM values of these unigenes based on the transcriptome data.



**Figure 5**

Phylogenetic trees and heat map diagrams with R2R3-MYB and bHLH amino acid sequences. Bootstrap values are displayed as percentages (1000 replicates) when >50 %; (A) The subgroups (S1-S25) of R2R3-MYB highlighted with different colors were designated as previously reported [32, 37], and 53 unigenes encoding R2R3-MYB proteins in *S. glabra* were numbered randomly from 1 to 53. (B) Subgroups S1-S21 for bHLH highlighted with different colors were designated as previously reported [34], and 53 unigenes encoding bHLH proteins of *S. glabra* were numbered randomly from 1 to 53. All the labeled R2R3-MYB and bHLH amino acid sequences are given in Additional file 13: Table S6; (C and D) All the selected unigenes are listed on the vertical line. The scale represents the logarithms of the FPKM values of the unigenes based on the transcriptome data. The black rectangles mean that the FPKM value is zero.



**Figure 6**

Real-time quantitative PCR confirmation of the 24 selected unigenes. The left y-axis and white legend represent the relative expression levels determined by RT-qPCR, while the right-axis and slash legend represent the FPKM values obtained by RNA-seq. Expression values were adjusted by setting the expression of WY to 1 for each unigene. All RT-qPCRs for each unigene used triplicates with two repeats per experiment. Error bars indicate SD, and different lower case letters (a-e) represent a significant

difference among the five samples at  $p < 0.05$ . The correlation coefficient (R) between RT-qPCR and RNA-seq for each unigene is shown with its corresponding significance level (\*\* $p < 0.01$ , \* $p < 0.05$ ).

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1TableS1.xls](#)
- [Additionalfile2TableS2.doc](#)
- [Additionalfile8TableS5.xls](#)
- [Additionalfile7TableS4.xlsx](#)
- [Additionalfile4FigureS1.doc](#)
- [Additionalfile3TableS3.doc](#)
- [Additionalfile13TableS6.txt](#)
- [Additionalfile9FigureS4.doc](#)
- [Additionalfile10FigureS5.doc](#)
- [Additionalfile5FigureS2.doc](#)
- [Additionalfile11FigureS6.doc](#)
- [Additionalfile16TableS7.doc](#)
- [Additionalfile12FigureS7.doc](#)
- [Additionalfile14FigureS8.doc](#)
- [Additionalfile6FigureS3.doc](#)
- [Additionalfile15FigureS9.doc](#)